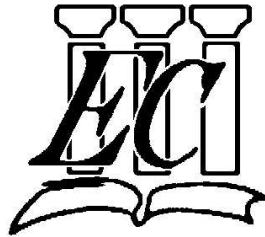


# Arquitectura del PC

## Volumen I. Microprocesadores

---

Manuel Ujaldón Martínez  
Doctor en Ingeniería Informática  
Profesor Titular de Universidad  
Departamento de Arquitectura de Computadores  
Escuela Técnica Superior de Ingeniería Informática  
Universidad de Málaga  
Marzo, 2003



Editorial Ciencia-3, S.L.

“Arquitectura del PC. Edición 2003. Volumen I. Microprocesadores.”

Autoedición y bicromía: Autor.  
Confeción de glosarios, índices, resúmenes y sumarios: Autor.  
Cuestionarios, ejercicios y solucionarios: Autor.  
Digitalización del material gráfico: Autor.  
Diseño de portada: Ana Serrano y autor.  
Formato, grafismos y diseño de página: Autor.  
Fotocomposición y fotografías: Autor <sup>1</sup>  
Ilustrador de diagramas y figuras: Autor.  
Impresión y encuadernación: Megaprint, S.L.  
Maquetación y estilos: Autor.

créditos

ISBN

Depósito legal:  
Impreso en España

ISBN para este volumen: 84-95391-86-4  
ISBN obra completa (4 volúmenes): 84-95391-90-2

Editorial y  
distribución

Editorial Megamultimedia, S.L.  
Camino San Rafael, 71  
Polígono Ronda Exterior.  
29006 Málaga.  
Teléfono: 95 236 31 43  
Fax: 95 236 41 01

Editorial Ciencia-3, S.L.  
C/ Comercio, 4. 28007 Madrid  
Teléfonos: 91 552 76 80  
91 552 86 17  
Fax: 91 501 57 01  
E-mail: [cienciatres@teleline.es](mailto:cienciatres@teleline.es)

copyright

© Manuel Ujaldón Martínez (2003)

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética u óptica, o cualquier sistema de almacenamiento y recuperación de información, sin la previa autorización por escrito del titular del Copyright.

objetivo

Los contenidos de este libro están orientados a la formación académica del alumnado universitario. El principal objetivo que el autor persigue con su publicación es el de facilitar su propia labor docente como Profesor Titular en la Escuela Técnica Superior de Ingeniería Informática de la Universidad de Málaga. Fuera del ámbito meramente educativo, este material no puede ser utilizado bajo ningún concepto.

fé de errores

Los contenidos e indicaciones técnicas de este libro han sido redactadas con sumo cuidado, pero sus muchas imperfecciones reflejan las carencias de un denodado esfuerzo personal. Por todo ello, el autor se ve obligado a declinar cualquier tipo de garantía acerca de la veracidad y exactitud de la información aquí publicada, comunicando a su vez que **no se hace responsable jurídico ni ofrece ningún tipo de cobertura por las consecuencias derivadas de posibles inconsistencias o errores.**

correcciones  
por Internet

El autor estará encantado de recibir cualquier corrección o aportación que se le desee hacer llegar por parte de docentes, estudiantes y lectores en general, para lo cual pueden utilizar libremente su dirección de correo electrónico, [ujaldon@ac.uma.es](mailto:ujaldon@ac.uma.es) (acompañando siempre como asunto ó subject “Comentario del libro”) o participar en el foro de erratas y opiniones que pondremos a su disposición en nuestra página Web:

☞ <http://www.ac.uma.es/~ujaldon/libro>

---

<sup>1</sup> A excepción de algunas fotos donadas por otras fuentes, hecho que se refleja a pié de foto donde corresponda.

# Índice general

---

Índice de figuras	XV
Índice de fotografías	XVII
Índice de tablas	XIX
Indice de ejemplos	XXIII
Agradecimientos	1
Editorial	3
<b>1. Génesis y evolución del PC</b>	<b>7</b>
1.1. La década de los 40: Desde el tubo de vacío al transistor . . . . .	8
1.2. La década de los 50: Desde el transistor al circuito integrado . . . . .	8
1.3. La década de los 60: Desde Fairchild a Intel . . . . .	12
1.4. La década de los 70: Desde el microprocesador al PC . . . . .	14
1.4.1. La carrera por el primer microprocesador . . . . .	14
1.4.2. Los modelos que tomaron el relevo . . . . .	20
1.5. El IBM PC . . . . .	25
1.5.1. La fábula de Intel y el IBM-PC . . . . .	26
1.5.2. La fábula de Microsoft y el IBM-PC . . . . .	27
1.5.3. La fábula de IBM y su PC . . . . .	28
1.6. El marco de las generaciones de computadores . . . . .	28
Resumen . . . . .	30
La anécdota: Intel, cadena hotelera . . . . .	30
Cuestionario de evaluación . . . . .	31

<b>PRIMERA PARTE: MICROPROCESADORES</b>	<b>33</b>
<b>2. El legado generacional</b>	<b>35</b>
2.1. La década de los 80: Intel frente a Motorola	37
2.1.1. Primera generación (1978-1982)	38
2.1.2. Segunda generación (1982-1985)	38
2.1.3. Tercera generación (1985-1989)	39
2.2. La década de los 90: Omnipresente Intel	40
2.2.1. Cuarta generación (1989-1993)	41
2.2.2. Quinta generación (1993-1997)	41
2.2.3. Sexta generación (1997-2000)	42
Resumen	43
La anécdota: IBM, ese celestino entre el chip y el procesador	45
Cuestionario de evaluación	46
<b>3. Las principales magnitudes</b>	<b>47</b>
3.1. Frecuencia de reloj	49
3.2. Tecnología de integración	50
3.2.1. Evolución y significado	50
3.2.2. Efectos directos sobre otras variables	54
3.2.3. Efectos laterales entre las variables afectadas	56
3.2.4. Cómo dar empleo a un ejército de transistores	58
3.3. Paralelismo a nivel de instrucción	58
3.3.1. Segmentación (pipelining)	59
3.3.2. Superescalaridad	60
3.3.3. Combinación de segmentación y superescalaridad	62
3.3.4. Supersegmentación	63
3.3.5. Dependencias: Las enemigas del paralelismo	64
3.4. Memoria caché integrada	69
3.4.1. Breve sinopsis histórica	72
3.4.2. Jerarquía	73
3.4.3. Optimizaciones	77
3.4.4. Proximidad al núcleo del procesador	81
3.4.5. Ubicación del controlador de caché	85
3.4.6. Velocidad	88

3.4.7. Análisis del rendimiento de caché en relación al procesador . . . . .	88
3.4.8. Análisis del coste asociado a una caché . . . . .	93
3.5. Conjunto de instrucciones . . . . .	94
3.5.1. CISC versus RISC . . . . .	95
3.5.2. Diseño RISC . . . . .	97
3.5.3. Diseños VLIW . . . . .	104
3.5.4. Instrucciones multimedia . . . . .	106
Resumen . . . . .	119
La anécdota: Magnitudes oficiales y oficiosas . . . . .	121
Cuestionario de evaluación . . . . .	122
<b>4. La quinta generación</b>	<b>133</b>
4.1. Intel . . . . .	134
4.1.1. Pentium . . . . .	134
4.1.2. Pentium MMX . . . . .	140
4.2. AMD . . . . .	143
4.2.1. K5 . . . . .	144
4.3. Cyrix . . . . .	145
Resumen . . . . .	146
Cuestionario de evaluación . . . . .	146
<b>5. La sexta generación</b>	<b>149</b>
5.1. Intel . . . . .	151
5.1.1. Pentium Pro: El esqueleto de la sexta generación . . . . .	151
5.1.2. Predicción de salto . . . . .	158
5.1.3. Pentium II: Un nuevo formato . . . . .	161
5.1.4. Deschutes: Pentium II Turbo . . . . .	164
5.1.5. Pentium II Xeon: Para servidores . . . . .	165
5.1.6. Celeron: Un paso en falso . . . . .	166
5.1.7. CeleronA: La rectificación . . . . .	167
5.1.8. Celeron Coppermine: Reconversión a Pentium III . . . . .	167
5.1.9. Pentium III: Más instrucciones multimedia . . . . .	168
5.1.10. Coppermine: Novedades en la integración . . . . .	173
5.1.11. Pentium III Xeon: Más para servidores . . . . .	175
5.2. AMD . . . . .	176

5.2.1.	K6: El primer problema para Intel . . . . .	176
5.2.2.	K6-2: Comienza el baile multimedia . . . . .	177
5.2.3.	K6-III: El ejemplo de configuración equilibrada . . . . .	179
5.3.	Comparativa: Intel versus AMD . . . . .	180
5.3.1.	Síntesis de la trayectoria de Intel . . . . .	181
5.3.2.	Síntesis de la trayectoria de AMD . . . . .	181
5.3.3.	Conclusiones . . . . .	182
5.4.	Cyrix . . . . .	183
	Resumen . . . . .	184
	Cuestionario de evaluación . . . . .	185
<b>6.</b>	<b>La séptima generación</b> . . . . .	<b>189</b>
6.1.	K7 de AMD . . . . .	192
6.1.1.	Frecuencia de reloj . . . . .	192
6.1.2.	El bus local . . . . .	192
6.1.3.	Tecnología de integración . . . . .	193
6.1.4.	Zócalo de la placa base . . . . .	193
6.1.5.	Memoria principal . . . . .	194
6.1.6.	Memoria caché . . . . .	194
6.1.7.	El entendimiento con el sistema de memoria . . . . .	194
6.1.8.	El conjunto de instrucciones . . . . .	195
6.1.9.	Segmentación . . . . .	197
6.1.10.	Superescalaridad . . . . .	197
6.1.11.	Resolución de dependencias . . . . .	197
6.1.12.	Instrucciones multimedia . . . . .	198
6.1.13.	Unidades de ejecución . . . . .	198
6.1.14.	Valoración final . . . . .	198
6.2.	Thunderbird y Duron: Las 0.18 micras en AMD . . . . .	199
6.2.1.	Thunderbird . . . . .	199
6.2.2.	Duron . . . . .	200
6.2.3.	Duron con aditivos . . . . .	201
6.3.	Athlon XP . . . . .	202
6.3.1.	Frecuencia de reloj . . . . .	202
6.3.2.	El bus local y sus diálogos con memoria principal . . . . .	204
6.3.3.	Tecnología de integración . . . . .	204

6.3.4.	El jeroglífico de la frecuencia . . . . .	206
6.3.5.	Paralelismo a nivel de instrucción . . . . .	206
6.3.6.	Memoria caché . . . . .	206
6.3.7.	Conjunto de instrucciones . . . . .	207
6.3.8.	Valoración final . . . . .	207
6.4.	Pentium 4 de Intel . . . . .	208
6.4.1.	Nueva arquitectura . . . . .	209
6.4.2.	Frecuencia de reloj . . . . .	210
6.4.3.	Tecnología de integración . . . . .	210
6.4.4.	El bus local . . . . .	211
6.4.5.	Diálogos con memoria principal . . . . .	212
6.4.6.	El conjunto de instrucciones y su decodificación . . . . .	215
6.4.7.	Memoria interna . . . . .	216
6.4.8.	Segmentación . . . . .	219
6.4.9.	Superescalaridad . . . . .	221
6.4.10.	Resolución de dependencias . . . . .	221
6.4.11.	Instrucciones multimedia . . . . .	222
6.4.12.	Unidades de ejecución . . . . .	224
6.4.13.	Valoración final . . . . .	224
6.5.	Northwood: Las 0.13 micras en Intel . . . . .	226
6.5.1.	Mejoras introducidas . . . . .	226
6.5.2.	Los diálogos con memoria principal . . . . .	227
6.5.3.	HyperThreading . . . . .	229
6.6.	Comparativa: K7 versus Pentium 4 . . . . .	231
6.6.1.	Ejecución de una instrucción . . . . .	232
6.6.2.	Frecuencia . . . . .	237
6.6.3.	Paralelismo a nivel de instrucción . . . . .	237
6.6.4.	Incidencia de las dependencias . . . . .	239
6.6.5.	Acceso a memoria . . . . .	241
6.6.6.	Valoración final . . . . .	242
6.6.7.	Comparativa a 0.13 micras . . . . .	244
6.7.	Otros fabricantes . . . . .	245
6.7.1.	Cyrix . . . . .	245
	Resumen . . . . .	245

La anécdota: El curso cíclico de la historia . . . . .	246
Cuestionario de evaluación . . . . .	247
<b>7. La octava generación</b> . . . . .	<b>255</b>
7.1. El K8 de AMD . . . . .	256
7.1.1. Etimología . . . . .	256
7.1.2. Frecuencia . . . . .	257
7.1.3. Tecnología de integración . . . . .	257
7.1.4. Paralelismo a nivel de instrucción . . . . .	258
7.1.5. Memoria interna . . . . .	260
7.1.6. Conjunto de instrucciones . . . . .	260
7.1.7. El interfaz del procesador . . . . .	266
7.1.8. Las diferencias entre Athlon 64 y Opteron . . . . .	271
Resumen . . . . .	271
La anécdota: Alpha 21464, o el primer procesador arácnido . . . . .	273
Cuestionario de evaluación . . . . .	273
<b>8. El futuro del procesador</b> . . . . .	<b>275</b>
8.1. Memoria histórica de la última década . . . . .	276
8.1.1. Lección 1. 1993-94: Frecuencia y algo más . . . . .	276
8.1.2. Lección 2. 1995-96: Apuesta prematura por las arquitecturas de 32 bits . . . . .	276
8.1.3. Lección 3. 1997-98: El falso reclamo de las instrucciones multimedia . . . . .	276
8.1.4. Lección 4. 1999-2000: El hardware esquivo los condicionamientos software . . . . .	277
8.1.5. Lección 5. 2001-2002: Ansiedad por los 64 bits . . . . .	277
8.2. Perspectivas para la próxima década . . . . .	277
8.2.1. La frecuencia . . . . .	278
8.2.2. La tecnología de integración . . . . .	279
8.2.3. El paralelismo a nivel de instrucción . . . . .	280
8.2.4. La memoria caché . . . . .	281
8.2.5. El conjunto de instrucciones . . . . .	281
8.2.6. El microprocesador del año 2010 . . . . .	282
8.2.7. El microprocesador del año 2015 . . . . .	282
8.3. El futuro de la computación . . . . .	283
8.3.1. Aritmética entera . . . . .	283
8.3.2. Aritmética de punto flotante . . . . .	283



8.3.3. El supercomputador más potente . . . . .	285
La anécdota: La ley de Moore . . . . .	286
<b>Ejercicios suplementarios</b>	<b>287</b>
Relacionar la dimensión RISC/CISC de un microprocesador . . . . .	287
Clasificar las mejoras del procesador Pentium . . . . .	287
<b>Para saber más</b>	<b>289</b>
Acerca de la historia del PC . . . . .	289
Sobre componentes en general . . . . .	290
Sobre microprocesadores . . . . .	291
Diccionarios de terminología informática . . . . .	293
<b>Soluciones a las cuestiones</b>	<b>295</b>
<b>Glosario conceptual</b>	<b>315</b>
<b>Glosario comercial</b>	<b>319</b>
<b>Glosario institucional</b>	<b>325</b>

---

## **VOLUMEN II. LA INFORMACION**

### **SEGUNDA PARTE: MEMORIAS**

#### **9. Evolución**

- 9.1. Semblanza histórica por décadas
- 9.2. Evolución por generaciones

#### **10. Memoria principal**

- 10.1. Etimología
- 10.2. La operación de refresco
- 10.3. El controlador de memoria principal
- 10.4. Parámetros de funcionalidad y rendimiento
- 10.5. Fiabilidad
- 10.6. Conexión a la placa base

- 10.7. Formato
- 10.8. Voltaje
- 10.9. Autoconfiguración
- 10.10. Descomposición
- 10.11. Entrelazado
- 10.12. Concurrencia
- 10.13. Arquitectura e interfaz
- 10.14. Etiquetado y especificaciones
- 10.15. Diez consejos para elegir la memoria principal del PC

## 11. Memoria caché

- 11.1. Arquitectura e interfaz
- 11.2. Organización
- 11.3. Aspectos eléctricos
- 11.4. Criterios de selección

## 12. Memoria de vídeo

- 12.1. El proceso de generación de imágenes
- 12.2. Elementos de la tarjeta gráfica
- 12.3. Evolución de la memoria de vídeo
- 12.4. Magnitudes de la memoria de vídeo

## 13. El futuro de la memoria

- 13.1. Evolución de la velocidad
- 13.2. Repercusiones sobre la arquitectura del PC
- 13.3. La futura memoria principal (DRAM)
- 13.4. La futura memoria caché (SRAM)
- 13.5. La futura memoria secundaria (disco)
- 13.6. Los límites del almacenamiento
- 13.7. Tecnologías emergentes

## TERCERA PARTE: BUSES

### 14. Caracterización de los buses

- 14.1. Líneas
- 14.2. Jerarquía

14.3. Protagonistas

14.4. El bus local

## 15. Buses de propósito general

15.1. El bus de expansión

15.2. Buses serie versus paralelo

15.3. El bus serie estándar: RS-232

15.4. El bus paralelo estándar: Centronics

## 16. Buses de propósito específico

16.1. El bus gráfico: AGP

16.2. Buses de disco: IDE y SCSI

16.3. Buses serie multimedia: USB y Fire Wire

16.4. Sonido y módem integrados: AC'97

# VOLUMEN III. LA ESTRUCTURA

## CUARTA PARTE: PLACAS BASE

### 17. Los elementos que componen la placa base

17.1. Elementos pasivos

17.2. Relojes y multiplicadores

17.3. El juego de chips (chipset)

17.4. El triángulo que circunda el procesador

17.5. Los buses

17.6. El triángulo que circunda la memoria

17.7. Firmware para la configuración del sistema

17.8. Puntos para la conexión de dispositivos

### 18. El juego de chips

18.1. Retrospectiva

18.2. Juegos de chips para Pentium

18.3. Juegos de chips para Pentium Pro

18.4. Juegos de chips para Pentium II y III

18.5. Juegos de chips para Pentium II y III Celeron

18.6. Juegos de chips para Pentium II y III Xeon

- 18.7. Juegos de chips para Pentium III Coppermine
- 18.8. Juegos de chips para Athlon
- 18.9. Juegos de chips para Athlon Thunderbird
- 18.10. Juegos de chips para Athlon XP
- 18.11. Juegos de chips para Pentium 4 Willamette
- 18.12. Juegos de chips para Pentium 4 Northwood
- 18.13. Juegos de chips para Pentium 4 Celeron
- 18.14. Juegos de chips para Pentium 4 Xeon

## 19. El formato de la placa base

- 19.1. Formato baby-AT
- 19.2. Formato AT
- 19.3. Formato ATX
- 19.4. Formato micro-ATX
- 19.5. Formato flex-ATX
- 19.6. Formato ATX12v

## QUINTA PARTE: MONTAJE

### 20. Adquisición de componentes

- 20.1. Cómo comprar
- 20.2. Dónde comprar
- 20.3. Cuándo comprar
- 20.4. Qué comprar
- 20.5. Cuatro reglas básicas para una certera selección de componentes

### 21. Elementos y riesgos del montaje

- 21.1. Elementos pasivos: Interconexiones
- 21.2. Elementos activos: Alimentación
- 21.3. Riesgos y cómo evitarlos
- 21.4. La carcasa

### 22. El montaje paso a paso

- 22.1. La secuencia de montaje y sus variantes
- 22.2. Preliminares
- 22.3. Desmonte de la carcasa

- 22.4. Inserción del microprocesador en su zócalo
- 22.5. Incorporación del disipador de calor
- 22.6. Inclusión de los módulos de memoria
- 22.7. Fijación de la placa base a la lámina de la carcasa
- 22.8. Configuración de los jumpers
- 22.9. Liberación del frontal de las bandejas
- 22.10. Inserción de los dispositivos de almacenamiento masivo
- 22.11. Dotación de corriente a la placa base
- 22.12. Introducción de la placa base en la carcasa
- 22.13. Inserción de tarjetas
- 22.14. Agregado de ventiladores adicionales
- 22.15. Conexión de discos
- 22.16. Acoplamiento de LED externos
- 22.17. Incorporación de componentes vía conectores externos
- 22.18. Validación del montaje: El encendido del equipo

## **VOLUMEN IV. EL TALLER**

### **SEXTA PARTE: CONFIGURACION**

#### **23. El sistema básico de entrada/salida (BIOS)**

- 23.1. Tipos de BIOS
- 23.2. Elementos ligados a la BIOS
- 23.3. Funciones de la BIOS

#### **24. Los menús de configuración de la BIOS**

- 24.1. Introducción
- 24.2. Entrada en el menú de configuración
- 24.3. Los submenús del menú de configuración
- 24.4. Las opciones del menú de configuración
- 24.5. Salida del menú de configuración
- 24.6. Las opciones de visita obligada en BIOS PnP

#### **25. Optimización del sistema**

- 25.1. El sistema de computación

- 25.2. La jerarquía de memoria
- 25.3. Las vías de comunicación
- 25.4. Los periféricos
- 25.5. El sistema de alimentación
- 25.6. El sistema de ventilación

## SEPTIMA PARTE: AVERIAS

### 26. Prevención de riesgos y averías

- 26.1. Condiciones medioambientales
- 26.2. Interferencias por ondas
- 26.3. Las ondas electromagnéticas
- 26.4. Deficiencias en el suministro eléctrico
- 26.5. Deficiencias en las líneas telefónicas y de datos
- 26.6. Elementos de corriente
- 26.7. El apagado incontrolado del equipo
- 26.8. Agresiones por virus informáticos
- 26.9. Realización de copias de seguridad

### 27. Mantenimiento del sistema

- 27.1. Mantenimiento de componentes
- 27.2. Mantenimiento de soportes de información y áreas de datos
- 27.3. Reglas básicas de ergonomía

### 28. Diagnóstico y reparación de averías

- 28.1. Preliminares
- 28.2. Sintomatología
- 28.3. Fallos hardware provenientes de la fase de montaje
- 28.4. Averías más frecuentes
- 28.5. Averías durante la secuencia de arranque
- 28.6. Problemas con la BIOS
- 28.7. Problemas con la contraseña del firmware
- 28.8. Problemas con el sistema de ventilación

## **VOLUMEN V. LOS CHIPS (sólo en la Web)**

### **OCTAVA PARTE: VARIABLES EXTERNAS**

#### **29. Velocidad: Limitaciones**

- 29.1. El transistor
- 29.2. El techo físico para la frecuencia de un microprocesador
- 29.3. El techo lógico para la frecuencia de un microprocesador

#### **30. Sobreaceleración: Oportunidades**

- 30.1. Ventajas e inconvenientes de la sobreaceleración
- 30.2. Las oportunidades que se nos ofrecen
- 30.3. Los elementos que debemos conocer
- 30.4. Las claves para sobreacelerar con éxito
- 30.5. El remarcado de chips

#### **31. Calor: Riesgos y soluciones**

- 31.1. Efectos térmicos
- 31.2. El calor para los no iniciados
- 31.3. Generación y transferencia de calor en el interior del PC
- 31.4. Elementos para la evacuación de calor: Disipación
- 31.5. Elementos para la medición de la temperatura del sistema
- 31.6. Elementos para la inyección de frío: Refrigeración

#### **32. Voltaje**

- 32.1. El papel del voltaje
- 32.2. Variables que influyen en el voltaje
- 32.3. Evolución histórica
- 32.4. Manipulación del voltaje del microprocesador
- 32.5. El regulador de voltaje
- 32.6. Fluctuaciones de corriente
- 32.7. El voltaje en el resto del equipo
- 32.8. APM: Gestión avanzada del consumo

#### **33. Estimaciones futuras**

- 33.1. Frecuencia

- 33.2. Temperatura
- 33.3. Voltaje
- 33.4. Integración

## NOVENA PARTE: ESTRUCTURA INTERNA

### 34. Tecnología de integración

- 34.1. El proceso de fabricación de chips
- 34.2. Los niveles de integración
- 34.3. El empaquetado de la circuitería
- 34.4. Envoltura externa

### 35. La celda básica de memoria

- 35.1. Arquitectura de un chip de memoria
- 35.2. La celda básica de una memoria dinámica
- 35.3. Mejoras en su implementación

### 36. Fabricación futura

- 36.1. Limitaciones
- 36.2. Alternativas de fabricación
- 36.3. EUV: Litografía láser

## DECIMA PARTE: TENDENCIAS FUTURAS

### 37. Física: Chips cuánticos

- 37.1. Qubits: Representación de la información
- 37.2. Rigidez: Procesamiento de la información
- 37.3. Entanglement: Teletransporte de la información
- 37.4. Restricciones
- 37.5. Previsiones

### 38. Nanotecnología: Chips moleculares

- 38.1. Métodos ascendentes y descendentes
- 38.2. Métodos descendentes

### 39. Biotecnología: Chips biológicos

- Neurochips, chips biónicos y biochips



# Índice de figuras

---

2.1. Diagrama de bloques de un microprocesador por generaciones . . . . .	44
3.1. La secuencia en el tratamiento de las principales magnitudes del procesador . . . . .	49
3.2. La señal de reloj de un microprocesador . . . . .	50
3.3. Un típico transistor MOS . . . . .	51
3.4. Sección transversal de un chip . . . . .	51
3.5. Ejecución segmentada y superescalar de instrucciones . . . . .	60
3.6. Evolución superescalar de los procesadores de Intel . . . . .	61
3.7. Cuatro técnicas de paralelismo a nivel de instrucción . . . . .	64
3.8. Operaciones y datos necesarios para completar una instrucción . . . . .	70
3.9. La ubicación de los niveles de memoria en el Pentium Pro . . . . .	76
3.10. Cachés y buses de un procesador de sexta generación . . . . .	77
3.11. Supersegmentación en los Alpha 21164 y 21264 de Digital . . . . .	78
3.12. Ubicación del área de datos de caché y su controlador . . . . .	86
3.13. La filosofía de diseño VLIW . . . . .	105
3.14. El formato de instrucción IA-64 del procesador Itanium . . . . .	106
3.15. El diseño VLIW sobre el código de instrucción del Crusoe . . . . .	107
3.16. Composición software y hardware del procesador Crusoe . . . . .	107
3.17. Diagrama de bloques de un computador SIMD . . . . .	108
3.18. Las dos vías para la programación de aplicaciones con instrucciones multimedia . . . . .	113
3.19. Los conjuntos de instrucciones multimedia de Intel y AMD . . . . .	115
3.20. Evolución de los conjuntos multimedia y su relación con DirectX . . . . .	119
4.1. El formato de instrucción de la familia de procesadores x86 de Intel . . . . .	139
4.2. Diagrama de bloques del procesador Pentium de Intel . . . . .	140
5.1. Derivación de las etapas de segmentación del Pentium en las del Pentium Pro . . . . .	154

5.2. Diagrama general de funcionamiento del Pentium Pro . . . . .	155
5.3. Las unidades funcionales del Pentium Pro para la ejecución de instrucciones . . . . .	157
5.4. Las unidades funcionales del Pentium Pro y sus puertos . . . . .	160
5.5. Las unidades funcionales del Pentium II . . . . .	164
5.6. Las unidades funcionales del Pentium III . . . . .	172
6.1. Relación entre el procesador K7 y sus diferentes niveles de memoria . . . . .	195
6.2. Diagrama de bloques del procesador K7 de AMD . . . . .	196
6.3. El perfecto entendimiento entre el Athlon XP y la memoria DDRAM de 2x133 MHz . . . . .	204
6.4. Diagrama de bloques del microprocesador Pentium 4 de Intel . . . . .	210
6.5. El perfecto entendimiento del Pentium 4 con la memoria RDRAM . . . . .	214
6.6. Los diferentes bancos de registros del Pentium 4 . . . . .	217
6.7. Inserción de la caché de traza en la arquitectura Pentium 4 . . . . .	218
6.8. Derivación del cauce de segmentación del Pentium 4 a partir de sus predecesores . . . . .	220
6.9. Las sucesivas extensiones multimedia realizadas por Intel hasta el Pentium 4 . . . . .	224
6.10. Interfaz de diálogo entre el Pentium 4 Northwood y la memoria DDRAM de 2x133 MHz . . . . .	228
6.11. Interfaz de diálogo entre el Pentium 4 Northwood y la memoria RDRAM de 32 bits . . . . .	229
6.12. El paralelismo a distintos niveles de una arquitectura: SMP, SMT y PNI . . . . .	230
6.13. K7 vs. Pentium 4. Fin de la primera fase: Búsqueda de instrucción . . . . .	232
6.14. K7 vs. Pentium 4. Fin de la segunda fase: Decodificación de instrucción . . . . .	233
6.15. K7 vs. Pentium 4. Fin de la tercera fase: Planificación y reordenación de instrucción . . . . .	234
6.16. K7 vs. Pentium 4. Fin de la cuarta fase: Búsqueda de operandos . . . . .	236
6.17. K7 vs. Pentium 4. Fin de la ejecución de una instrucción entera . . . . .	238
7.1. Diagrama de bloques de la arquitectura K8 de AMD . . . . .	259
7.2. Evolución de los bancos de registros en las arquitecturas PC de 16, 32 y 64 bits . . . . .	262
7.3. El banco de registros de la arquitectura K8 . . . . .	265
7.4. Los principales bloques funcionales de la arquitectura K8 . . . . .	266
7.5. La arquitectura K8 en su variante más general: Un multiprocesador con 4 CPU . . . . .	269
7.6. Evolución prevista del ancho de banda en el controlador de memoria DDRAM integrado del K8 . . . . .	271

# Índice de fotografías

---

1.1. Ubicación geográfica del Silicon Valley . . . . .	9
1.2. El paso del tiempo en Fairchild Semiconductor . . . . .	10
1.3. El padre del chip, Jack Kilby . . . . .	11
1.4. Retrospectiva de Intel y AMD . . . . .	12
1.5. El primer inversor y el primer anuncio de Intel . . . . .	13
1.6. Trabajadores de Intel en sus inicios y en nuestros días . . . . .	14
1.7. Fotos antiguas y actuales de T.Hoff, S.Mazor, F.Faggin y M.Shima . . . . .	16
1.8. Fabricación del microprocesador 4004 . . . . .	18
1.9. Reliquias del pasado: El computador ENIAC, la calculadora Busicom y el 4004 . . . . .	19
1.10. Algunos de los primeros microprocesadores . . . . .	23
1.11. El nacimiento del PC: El IBM-PC y su sistema operativo MS-DOS . . . . .	26
2.1. El PC en sus dos primeras generaciones . . . . .	39
2.2. Procesadores de las 4 firmas más conocidas en sus seis primeras generaciones . . . . .	42
3.1. Microscopía de la distancia de integración . . . . .	53
3.2. Ubicación de la caché en las áreas de integración de los Pentium y Pentium III . . . . .	74
3.3. El microprocesador K7 de 0.18 micras . . . . .	82
4.1. El Pentium MMX y su disipador de calor . . . . .	142
5.1. El microprocesador Pentium II junto con su disipador de calor . . . . .	162
5.2. El microprocesador Celeron en formato Slot y Socket . . . . .	168
5.3. Aspecto de las tres versiones del Pentium III . . . . .	171
5.4. El Pentium MMX bajo Socket7 y el Pentium III bajo Socket 370 . . . . .	174
6.1. Placa base, zócalo y aspecto de un procesador Athlon de AMD . . . . .	193
6.2. Los procesadores Thunderbird, Duron y Athlon XP de AMD . . . . .	202
6.3. Dirigentes de Intel: Andy Grove, Paul Otellini y Craig R. Barret . . . . .	209
6.4. Área de integración del Pentium 4 de 0.18 micras . . . . .	212
6.5. Empaquetado y patillaje del Pentium 4 . . . . .	213

---

6.6. Las dos etapas de transporte que consumen un ciclo completo en el Pentium 4 . . .	221
6.7. Areas de integración del Pentium 4 Willamette y Northwood a escala . . . . .	226
6.8. El zócalo mPGA478 del nuevo Pentium 4 Northwood . . . . .	227
7.1. Empaquetado y patillaje de los dos modelos de K8 . . . . .	257

# Índice de tablas

---

1.1. El Silicon Valley, en cifras macroeconómicas . . . . .	9
1.2. Los diez hitos históricos más importantes durante la gestación del primer microprocesador comercial. . . . .	20
1.3. Las principales innovaciones en el campo de los microprocesadores . . . . .	22
1.4. Resumen de las cuatro generaciones de computadores . . . . .	29
2.1. Evolución de los microprocesadores de Intel en los últimos 30 años . . . . .	36
2.2. Comparativa de las familias 80x86 de Intel y 680x0 de Motorola . . . . .	38
2.3. Los microprocesadores de las seis primeras generaciones . . . . .	40
2.4. Las seis primeras generaciones de microprocesadores para PC . . . . .	45
2.5. Los hechos más relevantes en las tres décadas que lleva con nosotros el PC . . . . .	45
3.1. Evolución de la tecnología de integración . . . . .	52
3.2. Evolución en la distancia de integración . . . . .	54
3.3. Estimaciones futuras de la tecnología de integración . . . . .	55
3.4. La reducción de la distancia de integración en las sucesivas generaciones de microprocesadores . . . . .	55
3.5. Efectos directos de una reducción de la distancia de integración . . . . .	57
3.6. Frecuencia de uso de las microoperaciones de un programa en función de la naturaleza de la aplicación software . . . . .	67
3.7. La jerarquía de memoria interna de un microprocesador . . . . .	75
3.8. Analogías entre una fotocopidora y un microprocesador . . . . .	83
3.9. El tipo de caché que incorpora cada uno de los microprocesadores de quinta y sexta generación . . . . .	84
3.10. Los diferentes tipos de caché y sus buses asociados . . . . .	85
3.11. Comparativa de rendimiento de cinco configuraciones de caché . . . . .	92
3.12. Los principales parámetros relativos a la integración de las cachés L1 y L2 en los Pentium y Pentium Pro . . . . .	93
3.13. Características RISC y CISC frente a frente . . . . .	98

3.14. El conjunto de instrucciones MMX . . . . .	109
3.15. Evolución histórica de la API DirectX de Microsoft y cobertura de los conjuntos multimedia y sistemas operativos . . . . .	114
3.16. Principales aportaciones al conjunto de instrucciones de los microprocesadores para PC . . . . .	118
3.17. Resumen de las cinco principales magnitudes del procesador . . . . .	120
3.18. Principales prefijos para los múltiplos y submúltiplos de las unidades del Sistema Internacional . . . . .	121
3.19. Nuevos múltiplos binarios definidos por la ISO y la IEC . . . . .	121
4.1. Nombre comercial y código de referencia para los procesadores de quinta generación	135
4.2. El Pentium en números . . . . .	136
4.3. El Pentium MMX en números . . . . .	141
5.1. Nombre comercial y código de referencia para los procesadores de sexta generación	151
5.2. El Pentium Pro en números . . . . .	153
5.3. Valor de los contadores de la BTB del Pentium Pro . . . . .	158
5.4. El Pentium II en números . . . . .	163
5.5. Evolución de los Pentium y Pentium II respecto a la frecuencia y la memoria caché	165
5.6. El microprocesador Celeron en números . . . . .	169
5.7. La familia de microprocesadores Pentium de quinta y sexta generación resumida en sus principales parámetros . . . . .	170
5.8. El Pentium III en números . . . . .	175
5.9. El K6 en números . . . . .	178
5.10. El esqueleto de la sexta generación de Intel y AMD . . . . .	180
5.11. Ránking orientativo de los microprocesadores de Intel y AMD de quinta y sexta generación . . . . .	182
6.1. Nombre comercial y código de referencia para los procesadores de séptima generación . . . . .	191
6.2. El microprocesador K7 de AMD resumido en sus dos modelos . . . . .	200
6.3. Comparativa entre el Pentium III Coppermine y el K7 Thunderbird . . . . .	201
6.4. Frecuencia del Athlon XP equivalente a sus especificaciones comerciales . . . . .	203
6.5. Evolución tecnológica del Athlon XP y contraste con su ancestro y su rival . . . . .	205
6.6. El Pentium 4 de 1.4 GHz en números y comparativa con el Pentium III de 1 GHz . . . . .	223
6.7. Diferencias entre las distintas versiones de Pentium 4 . . . . .	227
6.8. El impacto de las dependencias de datos sobre el K7 y el Pentium 4 . . . . .	239

6.9. Comparativa de los microprocesadores K7 y Pentium 4 frente a la resolución de dependencias . . . . .	240
6.10. Resumen del marcador comparativo entre el K7 y el Pentium 4 . . . . .	243
6.11. Sensibilidad del K7 y el Pentium 4 a los conflictos software . . . . .	243
6.12. Frecuencias equivalentes para los modelos de Intel y AMD de séptima generación . . . . .	244
7.1. Nombre comercial y código de referencia para los procesadores de octava generación anunciados hasta la fecha . . . . .	256
7.2. El paralelismo a nivel de instrucción en el K8 comparado con el K7 . . . . .	258
7.3. Comparativa entre los conjuntos de instrucciones x86-64 de AMD e IA-64 de Intel. . . . .	263
7.4. Principales características de los tres modos de ejecución x86-64 . . . . .	264
7.5. Resumen de las características del Athlon 64 y el Opteron . . . . .	272
8.1. Evolución de los MIPS en las tres últimas décadas . . . . .	283
8.2. Evolución histórica de las FPU . . . . .	284
8.3. Evolución comparativa en MFLOPS entre los PCs y los supercomputadores . . . . .	285





# Índice de ejemplos

---

3.1. La amortización de las plantas de fabricación de chips . . . . .	54
3.2. Cuantificando el efecto directo de la distancia de integración sobre otras variables . . . . .	56
3.3. Cuantificando la mejora que la distancia de integración revierte sobre la frecuencia teniendo en cuenta efectos laterales . . . . .	56
3.4. El cauce de segmentación más clásico . . . . .	59
3.5. La superescalaridad en Intel . . . . .	62
3.6. El Alpha 21264, o la supersegmentación en su máxima expresión . . . . .	63
3.7. Riesgo por dependencias de datos . . . . .	65
3.8. Riesgo por dependencias de control . . . . .	65
3.9. Riesgo por dependencias estructurales . . . . .	66
3.10. La necesidad de una jerarquía de memoria . . . . .	70
3.11. La caché víctima del renovado K7 . . . . .	80
3.12. Controlador de caché L2 integrado, área de datos externa . . . . .	86
3.13. Controlador de caché L2 integrado, área de datos interna . . . . .	87
3.14. Controlador de caché L3 integrado, área de datos externa . . . . .	87
3.15. Caso 1: Búsqueda de un dato en la caché L2 externa del K6 (L2 a la mitad de la velocidad del procesador) . . . . .	89
3.16. Caso 2: Búsqueda de un dato en la caché L2 interna del K7 (L2 a 1/3 de la velocidad del procesador) . . . . .	90
3.17. Caso 3: Búsqueda de un dato en la caché L2 interna del K7 (L2 a la misma velocidad del procesador) . . . . .	90
3.18. Caso 4: Búsqueda de un dato en la caché L2 del K7 (L2 integrada) . . . . .	91
3.19. Caso 5: Búsqueda de un dato en caché L1 . . . . .	91
3.20. Uso de la ventana de relleno de una instrucción de carga retrasada por parte del compilador . . . . .	101
3.21. Uso de la ventana de relleno de una instrucción de salto retrasado por parte del compilador . . . . .	101
3.22. Optimización en el uso de registros por parte del compilador . . . . .	102

3.23. Eliminación de computación redundante desde el compilador . . . . .	102
7.1. Evolución de una arquitectura de N bits . . . . .	261
7.2. El colosal ancho de banda del K8 de 90nm . . . . .	270

# Agradecimientos

Hace unos años, jamás hubiera sospechado que mi labor docente en la Universidad terminaría satisfaciéndome de esta manera. En el camino he pasado de ser un *gran conocedor de poco*, a ser un *pequeño conocedor de mucho*. Percibo cada vez con más fuerza mi ignorancia, pero me he instalado ya en ese sentimiento. Y me considero un ser afortunado, porque lo que mejor sé hacer en esta vida es lo que más necesito: Aprender.

Los libros me han enseñado mucho, pero nada comparado con lo aprendido de la gente que me rodea. Que todos ellos me acepten este pequeño tributo.

## A nosotros

A mis compañeros del Departamento de Arquitectura de Computadores de la Universidad de Málaga, por su predisposición a la colaboración en todo momento. Muy especialmente, a Andrés Rodríguez, por su infinita paciencia en la brega cotidiana y su espíritu constructivo para mejorar muchos pasajes de mis libros.

A Salvador Salas y Daniel Moreno (Servicio Central de Informática de la Universidad de Málaga), por su amabilidad y profesionalidad en el uso del escáner de diapositivas y el lector de cuestionarios.

Al personal de Ciencia-3 (Madrid) y Megamultimedia (Málaga), por saber hacer todo tan bien y tan fácil. Especialmente, a Carlos Ortega por su pericia informática.

## A vosotros

En los últimos años, más de dos mil estudiantes quisieron elegirnos como materia optativa para completar su formación. Ellos han sido, con su espíritu crítico pero siempre constructivo, los que más han aportado para mejorar la calidad de esta obra. Agradezco profundamente a todos ellos sus enseñanzas, y espero que futuras promociones también sientan estos volúmenes como suyos, y tengan a bien hacerme llegar cuantas correcciones crean oportunas. Los que vengan detrás, a buen seguro se lo agradecerán como yo.


gracias,  
chicos

Las ventas de un libro se apoyan en tres pilares básicos: Una firma conocida, una editorial multinacional y una buena campaña publicitaria. Sin tener nada de eso, sólo el boca a boca puede lograr que el coste de una edición consiga amortizarse; así nos ha sucedido en las dos primeras ediciones de este libro, y ésa vuelve a ser nuestra única tabla de salvación en este nuevo proyecto.

el boca a boca

He detectado numerosos adictos a la serie que esperan esa futura edición mientras honran con su pausada lectura la presente, desechando la compra de publicaciones mensuales en los quioscos que no consiguen satisfacer su nivel de exigencia. A todos ellos les rogaría que en lugar de consultarme sistemáticamente por e-mail cuándo va a salir la próxima edición, tengan a bien pasarse de vez en cuando por nuestra página Web, donde daremos cumplida cuenta informativa:

a los adictos

 <http://www.ac.uma.es/~ujaldon/libro>

Web!

## A ellos

El material que tiene en sus manos ha sido creado con la sola ayuda de cuatro herramientas software de libre distribución que supieron adiestrar de forma magistral a nuestro PC: Linux como sistema operativo,  $\text{\LaTeX}$  como procesador de textos, Gimp como gestor de imágenes, y  $\text{\Xfig}$  como editor de figuras.

Habiendo contribuido al sistema únicamente con una modestísima aportación a  $\text{\LaTeX}$ , no podemos sino sentirnos en deuda con todos esos altruistas programadores que regalan sus conocimientos, su tiempo y su esfuerzo a la creación de software libre, y a quienes consideramos responsables de que esta obra haya podido ser concebida íntegramente en el entorno artesanal de su autor.

## Editorial

A los que nos gusta viajar en trenes de largo recorrido, el premio no es el destino final, sino el viaje en sí mismo. Por eso este tren no tiene llegada, y si hace escalas, es para permitir que nuevos viajeros puedan subirse en cada estación. Si usted es uno de nuestros compañeros de viaje más fieles, ya hace seis años que nos conocemos, y un par de ellos que no nos vemos. Es mucho tiempo en informática, donde ya se sabe que envejecemos con tal celeridad que...

saludos

### Cómo hemos cambiado

En nuestro último encuentro comenzaba a librarse una cruenta batalla en el campo de los procesadores: Tablas en mi análisis, y alguno se suscribió al K7 tras escucharme aquello de que por el precio de un Pentium 4 de entonces, te comprabas dos K7 y una impresora.

el K7 en 2001

Por si tocaba comprar en 2002, recuerdo haber dejado el recado de que pronto las aguas volverían a su cauce, y que el Pentium 4 sería líder en ventas a mediados de 2002. Y yo, cuando plagio las palabras de Intel, nunca me equivoco.

el Pentium 4 en 2002

En esta nueva tertulia no hablaremos de procesadores. Su mercado ha estado algo parado en todo este tiempo, y para seguirlo bastó con cubrir las versiones de 0.13 micras de las viejas arquitecturas. El lanzamiento del K8 aporta novedades, pero la octava generación está coja hasta que no responda Intel y podamos establecer la consabida comparativa. Entretanto eso ocurre, hemos aprovechado para dedicarnos en profundidad a la patata caliente que ahora es la memoria.

cambio de tercio

El otro día un vendedor me recomendaba que no se me ocurriese comprar una memoria DDRAM de 266 MHz estando ahí la RDRAM de 800 MHz por el mismo precio. Después de lo laborioso que resultó explicar en la edición anterior que la frecuencia NO siempre es caballo ganador en microprocesadores, ha sido duro descubrir que esta vez nos toca lidiar un astado de similar trapío en el ruedo de las memorias.

otra falacia más para la colección

### Renovación de contenidos

La sección dedicada a memoria principal en esta nueva edición supera el centenar de páginas (ver [capítulo 10](#)), triplicando casi la extensión de la edición anterior, con la esperanza de que los interrogantes que ahora planean sobre ella queden suficientemente despejados. Por ejemplo, lo que nuestro vendedor ignora es que RDRAM responde sólo **dos bytes** (ver [sección 10.13.6](#)), mientras que DDRAM es capaz de servir **ocho** (ver [sección 10.13.5](#)). La precipitación tampoco debería llevar a nadie a inferir que una RDRAM de 800 MHz equivale a una DDRAM de 200 MHz; la historia es bastante más compleja, y espero que le dedique la plácida lectura que merece.

• [p. 13/Vol. 2](#)  
ampliaciones  
• [p. 73/Vol. 2](#)  
• [p. 68/Vol. 2](#)

nuevos  
capítulos

Junto a los capítulos más clásicos, también incorporamos otros nuevos. Algunos, como el de diagnóstico y reparación de averías con el que finaliza el cuarto volumen, porque venían siendo ampliamente reivindicados por nuestros alumnos, y por encima de todo, nos debemos a ellos. Otros, como la trilogía en que ahora se divide nuestra cobertura de la placa base en el tercer volumen, porque este profesor lleva muy mal que en las Escuelas de Ingeniería de la Universidad Española, un ingeniero pueda titularse en computación sin haber recibido un ápice formativo acerca de la que es hoy día la pieza clave para entender conceptual y estructuralmente cualquier computador. Si es que no había libros, la excusa debe ahora ser otra para que los Planes de Estudio no tapen ¡YA! un agujero tan estrepitoso.

descartes

Y como en toda renovación, también hay descartes. Algún fiel lector se preguntará ónde están todos aquellos capítulos más técnicos de la edición anterior dedicados a las intimidades de los chips: Frecuencia, sobreaceleración, voltaje, temperatura, encapsulado y fabricación, celdas de memoria. A nosotros era la parte que más nos gustaba, y la que más necesaria creíamos para los futuros ingenieros a los que formamos. Pero no tuvimos suerte a la hora de incluirlos en una asignatura del nuevo Plan de Estudios, y con ello, perdimos parte de la motivación necesaria para completar lo que era el pretendido quinto volumen de esta obra: Chips moleculares, nanotecnología, chips biónicos, ... No obstante, trataremos de ponerlos en un estado decente de publicación para colgarlos de nuestra página Web en fechas próximas y que todo el que quiera pueda descargarlos de forma gratuita. Es una deuda que contraemos con el lector desde el momento en que decidimos colocar referencias cruzadas a ese material desde estas páginas siempre que sirva para completar la formación del alumno en los aspectos de más bajo nivel de la informática.

futuros temas

Finalmente, un par de temas que se nos han quedado en el tintero pero que tendrán cumplida cabida en nuestra próxima edición serán uno dedicado a portátiles y otro dedicado a buses inalámbricos.

## Nos vemos dentro de dos años...

recursos  
en la Web

... si ése es su deseo. Para entonces, los PC comunicarán más y procesarán menos. Y eso mismo es lo que vamos a proponernos nosotros. Si en todo este tiempo quiere mantenerse informado sobre los contenidos de la próxima edición, su fecha orientativa de publicación, los puntos de venta en librerías de ámbito nacional o por Internet para la presente edición, la lista de erratas que nuestros lectores irán identificando, e incluso si quiere bajarse los capítulos de ese quinto volumen que no será publicado en papel, todo eso y mucho más queda a su entera disposición en nuestra página Web:

🔗 <http://www.ac.uma.es/~ujaldon/libro>

Web!

el viaje

Esperamos que este nuevo vehículo de contacto no sea unidireccional, y sepa emplearlo también para trasladarnos su opinión, críticas, sugerencias, comentarios, ... En definitiva, que considere esta obra como suya para que entre todos contribuyamos a hacer de ella ese manual que yo siempre anhelé y nunca pude tener delante cuando estudiaba informática.

nuestro  
destino

Los buenos libros en inglés quedan obsoletos (y a veces destrozados) cuando son traducidos al castellano, y las revistas mensuales se publican en un idioma que parece el nuestro pero no lo es. En sus manos está, en parte, nuestra consolidación como alternativa. Apostar por una sociedad más culta tecnológicamente supone un inquietante desafío; tratar de contribuir a ello, adquirir una gratificante responsabilidad.

nuestro  
combustible

Una vez más, confío en que el alumnado, el colectivo de docentes que ya utilizan esta obra como herramienta de trabajo, y mis lectores en general, me hagan sentir que mereció la pena el esfuerzo que me ha traído hasta aquí. Por si no lo sabe, ése es el único combustible del que disponemos en nuestro largo viaje. Si esta edición se vende como la anterior, será para que la

próxima mantenga el precio ofreciendo más contenidos y mejor estética (esperemos que ya a todo color); similar metamorfosis a la que nuestra edición de 2001 ha provocado en ésta.

Sea indulgente con mis carencias, y espero que disfrute con la lectura de esta obra tanto o más de lo que yo lo hice con su escritura.

“Incluso la más larga caminata comenzó con un primer paso”

El autor

## Postdata: Cuidemos las formas

A pesar de las voces que atesoran la extinción del libro en favor de su versión electrónica, nosotros seguimos apostando por él. Leer en pantalla es un 25% más lento que hacerlo sobre un libro <sup>2</sup>, y éste además cansa menos la vista y permite aderezarlo con notas y subrayados.

Eso sí, hay libros y libros. La edición anterior de éste nos dió severas lecciones sobre lo que nunca debimos hacer. En esta nueva cita hemos buscado una evolución mayor si cabe en las formas que en el contenido, tratando de subsanar sus múltiples carencias:

aprendiendo,  
como siempre

❶ Decidimos partir la obra en cuatro volúmenes en vista de los numerosos perfiles de usuario que se han interesado en ella. Por ejemplo, hay quienes gustan sólo de lecturas sobre procesadores o memorias, y también manitas que prefieren sumergirse en el taller. Todos ellos pueden ahora adquirir su volumen temático sin necesidad de soportar la carga física y económica de la obra completa, pero disfrutando de sus propios índices de figuras, tablas, etc, glosarios de términos conceptual y comercial, cuestionarios y solucionarios. El índice global de contenidos para el conjunto de los cuatro volúmenes ya lo habrá localizado el lector en las páginas previas de este primer volumen.

modularidad

índice global

❷ Aunque la secuencia elegida para la colección sí creemos que es la óptima, la selección de los capítulos que forman parte de cada volumen puede ser algo discutible desde el punto de vista de su contenido. Por ejemplo, los capítulos dedicados al montaje del PC podrían haberse incluido en el taller (volumen 4) tanto o mejor que en la estructura del PC (volumen 3). La elección final se hizo buscando un equilibrio de páginas entre los volúmenes, pero sobre todo, porque la asignatura de la ETSI Informática de la Universidad de Málaga para la que está escrita la obra realiza cuatro exámenes parciales, cuyo contenido quisimos hacer coincidir con los tomos para comodidad de nuestros alumnos. Estos libros están hechos, ante todo, para ellos, aparte de constituir ya nuestra más preciada herramienta de trabajo.

división  
en tomos

nuestros  
alumnos

❸ Los alumnos me indicaron que a pesar de la numeración de figuras, fotos y tablas por separado, las referencias a las mismas eran abundantes desde el texto, y su localización les rompía el ritmo de lectura. Ahora tenemos una vinculación inmediata entre ambas a través del número de página, que hemos colocado como reseña en el margen lateral a la altura del texto en que aparece cada referencia.

localización  
inmediata

❹ Para los valientes que se atrean con la obra en toda su extensión hemos querido guardar una serie de incentivos. Por ejemplo, cualquier contenido de un volumen que puede enriquecerse con la consulta de otro ha sido señalizada convenientemente en el margen del texto para alentar su lectura complementaria, y poder sacar así al conjunto todo el partido que creemos encierra. En los dos primeros volúmenes, este tipo de reseñas llega casi al centenar, mientras que en los dos últimos duplica esta cantidad.

referencias  
cruzadas

<sup>2</sup>Esta y otras cifras que adjuntamos en nuestra postdata pertenecen a estudios de Jacob Nielsen, uno de los más reputados investigadores acerca de la usabilidad de sitios Web.

consulta por  
palabra clave

5.000 voces

- 5 El 79% de los lectores que abren un libro lo hojean en lugar de leerlo, normalmente buscando un contenido concreto. Nuestro libro daba miedo sólo el abrirlo, y todo el que osara transgredir semejante umbral tenía luego ante sí una desafiante búsqueda. Ahora hemos etiquetado los contenidos párrafo a párrafo, e integrado éstos junto con los glosarios de consulta que adjuntamos al final de cada volumen, donde hemos dispuesto más de 5.000 voces ordenadas alfabéticamente. El glosario del conjunto de la obra también hemos querido adjuntarlo para mayor comodidad. Lo encontrará el lector a la conclusión del último volumen, como premio a su fidelidad.

aviso a  
piratas

Hemos hecho un gran esfuerzo, incluyendo el rechazo a grandes editoriales, para lograr que esta obra tuviera un precio asequible. Fotocopiarla no le va a suponer un ahorro en el coste, ni contribuirá a abaratar sus futuras ediciones, pero sí le va a suponer una merma considerable: Las fotografías apenas se distinguen en fotocopia, los fondos oscuros de algunos entornos (tablas, ejemplos, ...) dificultan notablemente su legibilidad, y la paleta del segundo color se pierde. Le aseguro que el empleo del color encierra una didáctica tremendamente valiosa para el entendimiento de buena parte de los diagramas y figuras del texto (que he coloreado personalmente uno a uno), al permitir diferenciar señales de control y datos, piezas mecánicas y eléctricas, etc. Si prescindiera de esa ayuda, se estará poniendo a sí mismo más difícil la comprensión de unos conceptos que no se caracterizan precisamente por su sencillez.

estilo de  
redacción

En mi afán por evitar el tedio de una obra académica, he adoptado un estilo de redacción más periodístico que científico, aún siendo consciente de estar caminando en un difícil equilibrio frente a ese rigor de contenidos que se presupone a un libro universitario como éste. Mi pretensión no es otra que hacer entretenida una obra tan extensa y compleja, y de paso, porfiar en el intento de hacerla asequible a cualquier otro lector, de cualquier edad y nivel cultural, cuya inquietud le lleve a interesarse por el conocimiento del hardware de su PC.



# Génesis y evolución del PC

## Sumario

<b>1.1. La década de los 40: Desde el tubo de vacío al transistor . . . . .</b>	<b>8</b>
<b>1.2. La década de los 50: Desde el transistor al circuito integrado . . . . .</b>	<b>8</b>
<b>1.3. La década de los 60: Desde Fairchild a Intel . . . . .</b>	<b>12</b>
<b>1.4. La década de los 70: Desde el microprocesador al PC . . . . .</b>	<b>14</b>
1.4.1. La carrera por el primer microprocesador . . . . .	14
1.4.1.1. El chip ganador: El 4004 de Intel . . . . .	15
1.4.1.2. El que pudo ganar al 4004 . . . . .	19
1.4.1.3. El que pudo ganar a Intel . . . . .	19
1.4.2. Los modelos que tomaron el relevo . . . . .	20
1.4.2.1. Las innovaciones del PARC . . . . .	20
1.4.2.2. Los sucesores del 4004 en Intel . . . . .	21
1.4.2.3. Otros contendientes . . . . .	23
1.4.2.4. La pifia y el resurgimiento de Intel . . . . .	24
<b>1.5. El IBM PC . . . . .</b>	<b>25</b>
1.5.1. La fábula de Intel y el IBM-PC . . . . .	26
1.5.2. La fábula de Microsoft y el IBM-PC . . . . .	27
1.5.3. La fábula de IBM y su PC . . . . .	28
<b>1.6. El marco de las generaciones de computadores . . . . .</b>	<b>28</b>
<b>Resumen . . . . .</b>	<b>30</b>
<b>La anécdota: Intel, cadena hotelera . . . . .</b>	<b>30</b>
<b>Cuestionario de evaluación . . . . .</b>	<b>31</b>

El concepto de computador personal o PC como ahora lo conocemos tiene su origen a comienzos de los 70 bajo la idea de aplicar a los circuitos lógicos el proceso de fabricación en forma de circuitos integrados que por aquel entonces se utilizaba para los chips de memoria. La historia tiene cuatro nombres propios: Jack Kilby y Robert Noyce como co-inventores del circuito integrado en los años 50, Fairchild Semiconductor como generador de sinergias clave en las décadas de los 50 y los 60, e Intel como abanderado del PC en las tres décadas subsiguientes.

origen

## SECCIÓN 1.1

## La década de los 40: Desde el tubo de vacío al transistor

**Kelly** **1940-1945. El punto de partida.** M. J. Kelly, director de investigación y posterior presidente de los laboratorios Bell, es el primer visionario de nuestra historia. En los años 40, ya supo ver que las centrales telefónicas necesitaban de conmutación electrónica y mejores amplificadores, y ése fue el principal acicate que la ciencia utilizó para lanzarse a la búsqueda del transistor.

**Mauchly  
Eckert** **1945. El ENIAC.** John W. Mauchly y J. Presper Eckert completan la construcción del **ENIAC** (*Calculador e Integrador Numérico Electrónico*), considerado el primer computador de la historia. Cuesta trabajo imaginárselo: 90 m<sup>2</sup> de espacio, 30 toneladas de peso y 17.468 tubos de vacío que consumen más de 130 Kilovatios de electricidad. El computador realiza una multiplicación en tres milisegundos a una frecuencia de reloj de 100 KHz. La **foto 1.9.a** muestra una pequeña ilustración de esta maravilla.

pág. 19 ➔

**Brattain  
Bardeen** **1945-1950. El transistor.** En 1945 se constituyó una agrupación de física del estado sólido conocedora de los estudios teóricos sobre metales y semiconductores, que culminó en Diciembre de 1947 con unas pruebas en las que dos sondas de hilo de oro próximas entre sí eran presionadas sobre la superficie de un cristal de germanio que actuaba como base. Se pudo observar que la tensión de salida en la sonda colector era superior a la de entrada en la sonda emisor, y Brattain y Bardeen se dieron cuenta enseguida de las implicaciones del fenómeno: Había nacido el transistor. En 1956, recibirían el Premio Nobel de Física en reconocimiento a semejante hallazgo.

## SECCIÓN 1.2

## La década de los 50: Desde el transistor al circuito integrado

**Kilby** **1950-54. Los días del germanio.** Tras graduarse en ingeniería eléctrica en 1947, Jack Sinclair Kilby comenzó a trabajar en Centrallab, una empresa de Milwaukee líder en aquella época en la manufacturación de circuitos impresos y miniaturización electrónica. Se empleaban entonces transistores de germanio, allí y en las pocas empresas del ramo existentes hasta la fecha.

germanio

**silicio** **1954-57. Surge el silicio.** El año 1954, Texas Instruments anunció la fabricación de transistores de silicio. Un poco más tarde, en 1957, Sherman Mills Fairchild, fundador de Fairchild Semiconductor, reclutó un selecto grupo de ocho jóvenes científicos de California, entre los que se encontraban Gordon Moore y Robert Norton Noyce<sup>1</sup>. La **foto 1.2** muestra una foto del jefe con su grupo en aquellos tiempos, en contraste con otra de la cúpula actual de la compañía, que a fecha de hoy continúa devota al sector de semiconductores (📄 <http://www.fairchildsemi.com>).

Moore-Noyce  
pág. 10 ➔

Web!

Mountain View

Silicon Valley

pág. 9 ➔

pág. 9 ➔

El objetivo de aquel grupo era desarrollar un nuevo proceso de fabricación de transistores que fuera apropiado para la producción masiva de componentes de las más variadas especificaciones. Fairchild Semiconductor se convirtió así en la primera empresa dedicada en exclusiva la tecnología del silicio, y el enclave geográfico elegido para su ubicación, Mountain View, al sur de la bahía de San Francisco en el valle de Santa Clara, puso la primera piedra de lo que hoy es el **Silicon Valley**, el centro neurálgico de la alta tecnología a escala mundial. Silicon Valley es una expresión acuñada en 1971 por la revista Electronic News, para subrayar la entonces ya evidente concentración de empresas del sector tecnológico en el área geográfica de la parte norte del estado de California que mostramos en la **foto 1.1**. La **tabla 1.1** resume su relevancia en el contexto actual de la economía mundial.

<sup>1</sup>Los co-fundadores de Intel e inventores del microprocesador.

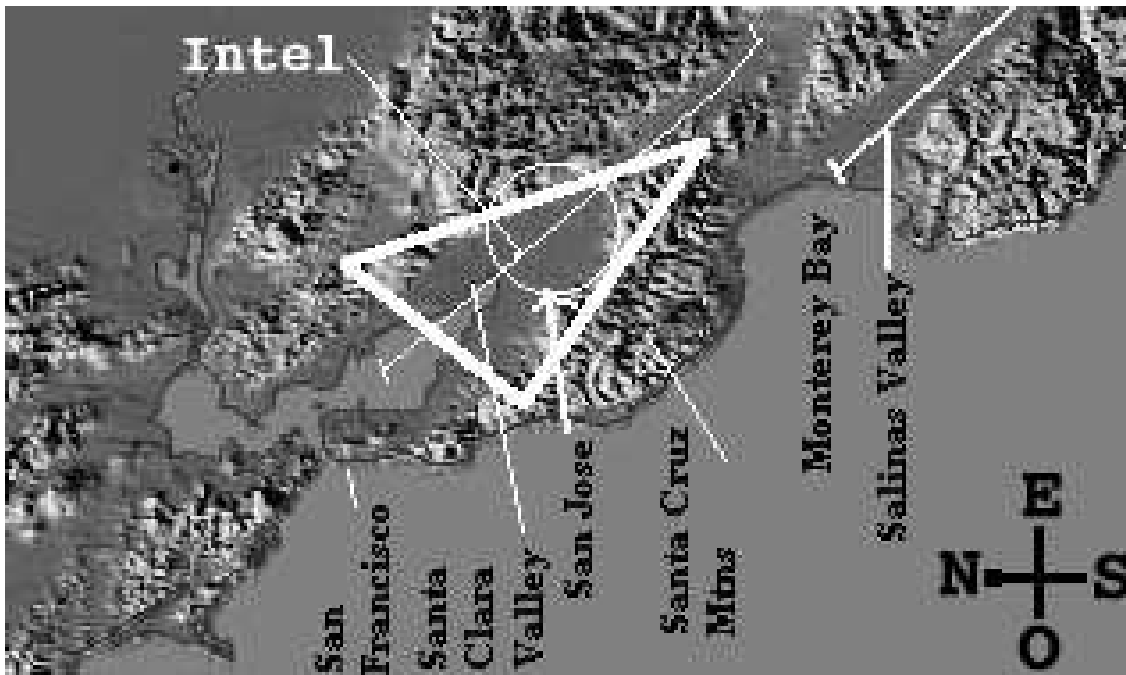


FOTO 1.1: Ubicación del Silicon Valley en el área geográfica que la Cámara de Comercio norteamericana delimita con el triángulo conformado por San Mateo (vértice inferior), el puente (arista) que lo une a Hayward (vértice superior), y el enclave al sur de Los Gatos (vértice de la derecha). La superficie total es de 3.870 km<sup>2</sup>, algo inferior a la provincia de Málaga, y en ella trabajan más de 100.000 empleados y ejercen su actividad más de 2.000 compañías. En el centro del triángulo se sitúa Palo Alto, donde tiene su sede la Universidad de Stanford, y muy cerca de allí, Intel Corporation.

Variable macroeconómica	Silicon Valley	Estados Unidos	España
Población (millones de hab.)	2.5	312	40
Productividad por empleado (dólares)	170.000	56.000	18.000
Renta per cápita (dólares)	57.400 (2000)	30.800 (2000)	18.215 (1999)
Patentes logradas	6.817 (1999)	85.000 (1999)	58 (1997)
Inversión en 2000 (millones de dólares)	21.000	100.000	2.670
Inversión en 2001 (millones de dólares)	6.000	30.000	3.100

Fuente: Datos de EEUU: Joint Venture Silicon Valley. Datos de España: OCDE.

TABLA 1.1: El Silicon Valley contrastado en cifras con la economía estadounidense y española. La cifra dada en las casillas de inversión corresponde a la partida de capital-riesgo en los casos norteamericanos, mientras que para España corresponde a la partida dedicada a I+D en los presupuestos generales del estado. El resto de cifras españolas se ha extraído de un estudio de la OCDE, donde quedamos situados un 16% por debajo de la renta europea.

**1958. Kilby y Texas Instruments.** La utilización del silicio en Texas Instruments con objeto de lograr una mayor miniaturización de los transistores despertó el interés de Kilby, que se incorporó a la plantilla de esta empresa en los inicios de 1958. Allí terminaría jubilándose, aunque aquel verano del 58 era un novato sin vacaciones que se dedicó a conocer los nuevos proyectos de la compañía.

Texas Instr.

**1958. El circuito integrado.** Kilby aportaría enseguida una nueva forma de hacer las cosas: Fabricar las resistencias y condensadores utilizando silicio en lugar de carbono (que era como se

silicio vs.  
carbono



(a)



(b)

Fotos cortesía de Fairchild Semiconductor

FOTO 1.2: El paso del tiempo en Fairchild Semiconductor. En la foto (a) vemos al fundador de la compañía, Sherman Mills Fairchild, junto al grupo de investigadores responsables de sus mayores logros, entre los que intuimos a Gordon Moore y Robert Noyce justo a su derecha. En (b), una foto de la actual cúpula directiva de la compañía, donde ya no queda nadie de aquel mítico grupo.

hacía hasta la fecha en Texas Instruments). Resultaba más caro, pero compensaba de sobra si el circuito completo se construía en un solo bloque de material semiconductor, unificando el proceso de producción y ahorrando además las conexiones de soldadura al circuito impreso.

silicio vs.  
germanio

pág. 11 ➔

Así, en Agosto de 1958, Kilby concluyó el primer circuito en el que transistores, resistencias y condensadores fueron fabricados de silicio, aunque separadamente. Un mes más tarde, integró todos estos elementos conjuntamente en un circuito oscilador, si bien en esta ocasión se usaría el germanio porque la maquinaria de Texas Instruments aún presentaba algunas limitaciones para manejarse con el silicio. La [foto 1.3.d](#) muestra el aspecto de este primer circuito integrado, invención por la que Kilby recibiría el último premio Nobel de Física del milenio (Octubre de 2000).

metodología

**1959. El proceso planar.** Una vez se conoció el primer circuito integrado de Kilby, surgieron numerosos especialistas interesados en mejorarlo. Así, Robert Noyce describía en Enero de 1959 una metodología para fabricar conjuntamente transistores, resistencias y condensadores sobre una base de silicio semiconductor en la que los transistores se disponían en el plano inferior y las conexiones metálicas formaban planos superiores paralelos al de los transistores.

aislando las  
uniones

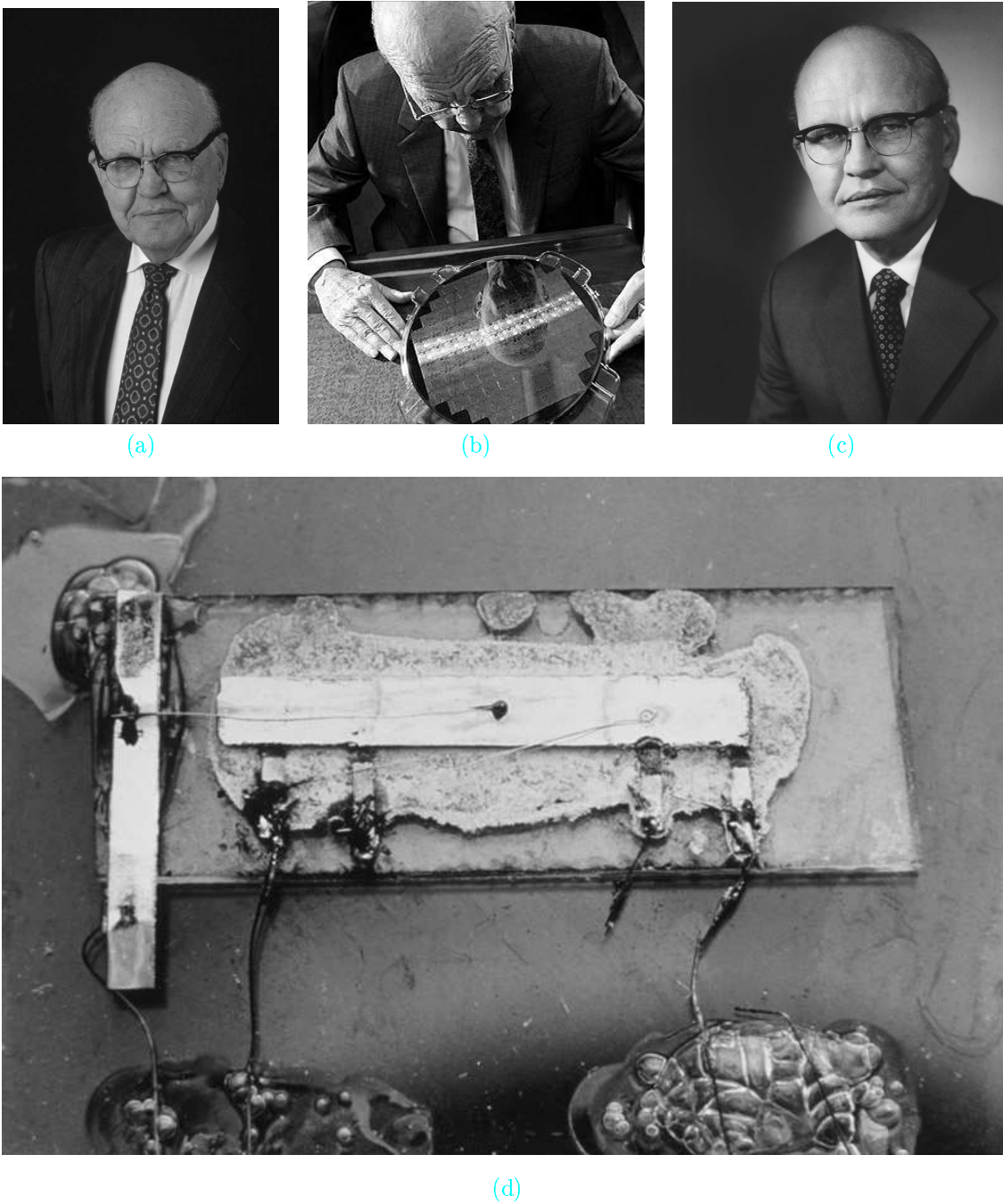
Noyce se apoyó en las técnicas de fotolitografía que un colega suyo de Fairchild, Jean Hoerni, había empleado en 1958 para fabricar el primer transistor aislando las uniones mediante una capa de dióxido de silicio. Hoerni, a su vez, vaporizó el metal a través de agujeros en el óxido siguiendo los postulados establecidos por Robert Noyce y Gordon Moore al poco de llegar a Fairchild.

soldando las  
conexiones

Toda esta secuencia de ideas bien ordenada ha sido el gran bastión sobre el que se ha apoyado la industria del chip para conseguir integrar centenares de millones de transistores en unos pocos milímetros cuadrados, pues con el proceso descrito por Kilby, estas conexiones debían ser soldadas a mano con unos diminutos cables de oro una vez la integración en silicio había concluido, de forma parecida a como ahora se procede para la vinculación del patillaje externo al área de silicio de un chip.

el primer chip

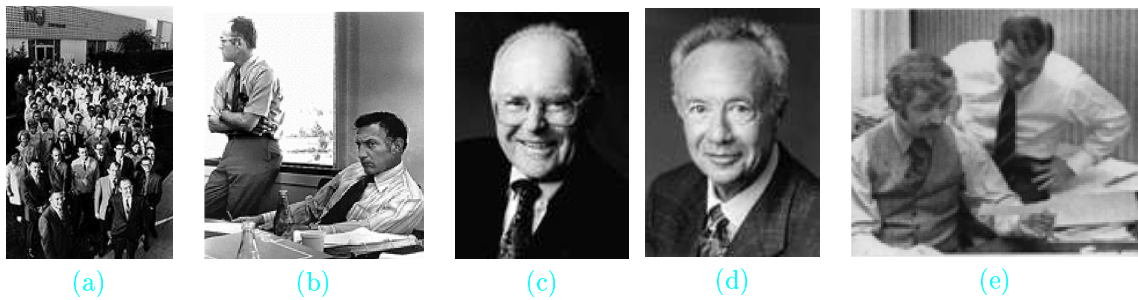
**1959. La patente.** Kilby solicitó la patente de su contribución en Febrero de 1959; Noyce la suya en Julio de ese mismo año. Años más tarde, se concedió a cada investigador su parte de innovación en lo que hoy conocemos como chip. Además, Kilby tiene el Premio Nobel por haber sido el primero; Noyce, el reconocimiento de la comunidad científica por aportar el método ingenieril



Fotos cortesía de Texas Instruments

FOTO 1.3: (a) El padre del chip, Jack Kilby, según una foto contemporánea tras recibir el Premio Nobel en Octubre del año 2000. (b) Kilby junto a una oblea actual de 300mm conteniendo decenas de chips. (c) Una foto de 1958, el año en que Kilby creó el circuito integrado, y (d) detalle de aquel primer chip, cuyas dimensiones eran de 10mm x 1.5mm, albergando un único transistor y algunos elementos pasivos sobre una base de germanio.

que desencadenó la posterior revolución, y porque, no contento con este logro, participaría en la creación del primer microprocesador doce años más tarde.



Fotos cortesía de Intel

**FOTO 1.4:** Retrospectiva de Intel y AMD. (a) La primera factoría de Intel en Santa Clara, el día en que la empresa alcanza los 100 primeros empleados. Robert Noyce y Gordon Moore son las dos personas de la parte más inferior de la fotografía. (b) Moore (izquierda) y Noyce (derecha) en una foto de 1969. (c) Foto del año 2000 de Gordon Moore. (d) Foto del año 2000 de Andy Grove, el tercer pilar de Intel, que sería presidente de la compañía durante 20 años (1979-1998). Muy a nuestro pesar, no podemos acompañar foto reciente de Robet Noyce (falleció en 1990). (e) Jerry Sanders y Rich Provite, dos de los ocho co-fundadores de AMD, en una foto también de 1969, poco después de su marcha de Fairchild Semiconductor.

### SECCIÓN 1.3

## La década de los 60: Desde Fairchild a Intel

Cursada la patente y convencidos de las posibilidades que ofrecen los nuevos descubrimientos de la electrónica, Robert Noyce y Gordon Moore orientan su actividad en Fairchild de forma exclusiva a la tecnología del silicio. Las bases para la progresiva miniaturización de los circuitos integrados estaban sentadas, y aquí fue donde la electrónica comenzó a merecer su prefijo micro-

Fairchild

**1964. La Ley de Moore.** Más técnicos que empresarios, Moore y Noyce modelan una compañía cimentada sobre las bases de la investigación básica. Fairchild comienza el desarrollo de chips de memorias utilizando la nueva tecnología de circuitos integrados para su fabricación, y observando la evolución en sus primeros cinco años de andadura, Gordon Moore extrapola su evolución futura, vaticinando que el número de transistores por chip se duplicará cada 18 meses a partir de entonces. Había nacido la archiconocida **Ley de Moore**, que no sólo se ha mantenido vigente durante 30 años, sino que ha sido respetada por otras variables del microprocesador aún más importantes, como su frecuencia de funcionamiento. La [tabla 2.1](#) muestra la trayectoria evolutiva trazada por estas variables a lo largo del tiempo, donde podremos comprobar su correlación con una representativa muestra de modelos comerciales a lo largo de todo este tiempo.

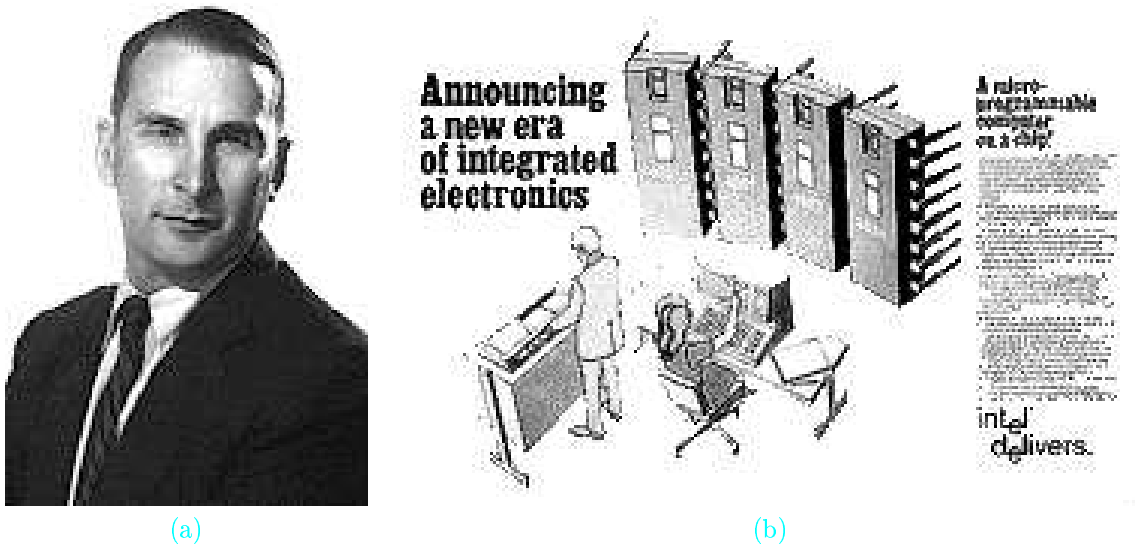
Ley de Moore

pág. 36

AMD

**1967. Entra en escena Jerry Sanders.** Otro de los ilustres personajes de esta historia se incorpora a Fairchild como manager de ventas. Hablamos de Jerry Sanders, co-fundador de AMD poco después, y único presidente de la compañía hasta 2002.

**1968. Fuga de cerebros en Fairchild.** Después de diez años juntos en Fairchild, Moore confiesa a Noyce que ya no se encuentra cómodo en la empresa. El exceso de burocracia eclipsa por momentos su actividad científica, y expresa su deseo de dejar la compañía y replantearse el negocio de nuevo. Por su parte, Noyce lleva un tiempo convencido de que la nueva tecnología de memorias basadas en semiconductores se está infrautilizando, y de que presenta un potencial suficiente como para ser capaz de integrar un computador completo en un solo chip. Así que decide asumir el riesgo de crear una nueva compañía. Noyce también manifestaría años más tarde que



Fotos cortesía de Intel

FOTO 1.5: (a) El financiero Art Rock, primer inversor del proyecto Intel. (b) El primer anuncio de la compañía tras el lanzamiento del microprocesador 4004, aparecido en la revista *Electronic News* en su número de Noviembre de 1971.

en aquellos tiempos eran tan jóvenes y arrogantes que no comulgaban con los métodos impuestos por sus colegas en Fairchild, y que necesitaban hacer las cosas a su manera. El caso es que algo extraño sucedió allí, porque se produjo una desbandada bastante notoria: Pocos meses más tarde, Jerry Sanders abandonaría también la empresa junto a siete colegas más.

fuga de cerebros

**1968. El nacimiento de Intel.** La decisión de Noyce arrastra a Moore, quien también abandona Fairchild para convertirse en el socio co-fundador de INTEgrated ELEctronics (Intel), allí mismo, en Mountain View, en Julio de 1968. Muy pronto se les une Andy Grove, componiendo un trio que Noyce denominó “el monstruo de tres cabezas” en clara referencia a un equipo que se complementa de forma extraordinaria. La [foto 1.4](#) muestra su aspecto de entonces.

NACE INTEL:  
Mountain View  
Andy Grove☛ [pág. 12](#)

**1969. El nacimiento de AMD.** La maniobra de Jerry Sanders y sus siete amigos es similar a la de Noyce y Moore: Dejar Fairchild para fundar conjuntamente su propia compañía. Esta se llamará Advanced Micro Devices (AMD), ubicándose en el que es ahora otro esplendoroso punto del Silicon Valley: Sunnyvale. AMD se dedica a la fabricación de chips, pero no tomará relevancia en el contexto del computador personal hasta la década de los noventa, momento en que volveremos a hablar de ella.

NACE AMD:  
Sunnyvale

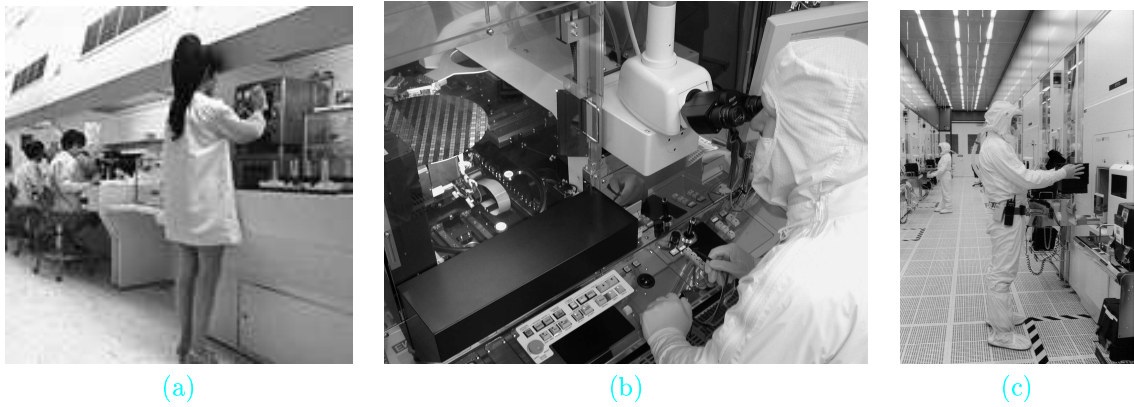
**1969. El mecenazgo.** Si ahora el dinero sale del Silicon Valley de debajo de las piedras, hubo en tiempo en que eso era pura quimera, así que los ilustres Noyce y Moore necesitaron hacerse de un aval que garantizara la viabilidad de su proyecto empresarial. Art Rock, un financiero de San Francisco cuya foto adjuntamos en la [foto 1.5.a](#), actuó como mecenas aportando el grueso del capital que Intel necesitaba para comenzar a trabajar. La firma inició su actividad diseñando memorias a escala de integración LSI, en espera de que llegara la oportunidad de poner en práctica la idea visionaria que sus fundadores tenían en mente.

Art Rock  
mecenas

El primer producto comercial de Intel fue el chip 3101, una RAM de 64 bits que en realidad dispuso de tan sólo 63 bits en su primera versión debido a un pequeño descuido en la definición manual de una de las máscaras de integración, y que se convierte así en el primer error en la historia de Intel <sup>2</sup>.

SRAM 3101

<sup>2</sup>Según fuentes de la propia compañía - ver artículo de Volk, Stoll y Metrovich en las referencias bibliográficas.



Fotos cortesía de Intel

FOTO 1.6: (a) Trabajadores en la planta de fabricación de Intel en sus inicios y (b y c) en nuestros días.

**1970. Se crea el PARC.** Aunque en el período que estamos recorriendo hubo más iniciativas empresariales brillantes, si las circunscribimos a lo que es el computador personal en sí, creemos que junto a Intel y AMD sólo nos falta una mención más que ahora puede pasarnos desapercibida: Rank Xerox. Xerox había sido creada tiempo atrás en Connecticut (costa atlántica), pero uno de sus directivos, George Pake, decidió crear un nuevo centro de investigación en Palo Alto, California, casualmente muy cerca de Sunnyvale y Mountain View, en el extrarradio de la bahía de San Francisco. Se llamó PARC (Palo Alto Research Center), componiéndose de investigadores reclutados de los principales centros patrocinados por la Agencia de Proyectos de Investigación Avanzada (ARPA) del Ministerio de Defensa Estadounidense: MIT, Stanford, Berkeley, Carnegie-Mellon, UCLA y Utah. Este grupo creó la impresora láser, el ratón y los entornos de ventanas e iconos. Incluso para algunos fue allí donde se concibió el primer PC, el Alto <sup>3</sup>, hito que visitaremos cronológicamente un poco más adelante (1973).

Rank Xerox  
Palo Alto

PARC

Alto

**1970. El Silicon Valley.** En el enclave geográfico donde Intel decidió instalarse y AMD y Xerox completaron los tres pilares básicos, estaban censadas a finales del año 2000 más de 7.000 empresas devotas al desarrollo de lo que desde entonces se acuñó como nueva economía. El Silicon Valley se ha venido consolidando desde aquella fecha como la tierra en la que se hace realidad la leyenda del tío Sam, el multimillonario hecho a sí mismo de la nada y uno de los símbolos más recalcitrantes del capitalismo de barras y estrellas. Cada año, los norteamericanos se entretienen en contarnos cuántos millonarios se producen en aquella región a cada hora que pasa, habiéndose llegado a superar la treintena en las épocas más esplendorosas (a finales del año 2000 se estimaba que vivían en Silicon Valley unos 250.000 millonarios, lo que nos da un ritmo constante de aparición de unos 20 millonarios por hora en las tres últimas décadas transcurridas).

nueva  
economía

## SECCIÓN 1.4

# La década de los 70: Desde el microprocesador al PC

## 4.1 ► La carrera por el primer microprocesador

El título que hemos puesto a esta sección encierra una paradoja en sí mismo: Una carrera sólo puede establecerse con una meta definida de antemano, y la competición que nos aprestamos a narrar desemboca en algo nuevo. Sin saberlo, la ciencia se dirige hacia el nacimiento de la piedra

nuestra piedra  
angular

<sup>3</sup>Las referencias bibliográficas contienen una simpática referencia que profundiza acerca de esta percepción.



angular de la informática: El microprocesador, reconocido por un amplio elenco de historiadores como el invento más importante del siglo XX.

El 4004 fue el primer diseño que inició su andadura con una clara vocación en este sentido, aunque a punto estuvo de dejarse tomar la delantera a mitad de camino. Afortunadamente, la historia terminó haciendo justicia, proclamando como ganador al equipo de trabajo que más méritos acumuló, formado por Masatoshi Shima, Ted Hoff, Stanley Mazor y Federico Faggin (ver [foto 1.7](#)). Así fue como todo sucedió.

4004

👉 [pág. 16](#)

#### 4.1.1 El chip ganador: El 4004 de Intel

**Abril de 1969. Busicom.** Intel acordó desarrollar un conjunto de 7 chips para la calculadora decimal Busicom (ver [foto 1.9.b](#)) de una firma japonesa <sup>4</sup>. Los chips propuestos ya habían sido implementados por Busicom utilizando lógica TTL y una ROM, y eran los siguientes: (1) control de programa, (2) unidad aritmética decimal, (3) lógica de temporización, (4) ROM, (5) registro de desplazamiento, (6) control de impresión y (7) puertos de salida. La tarea de Intel era integrarlos bajo CMOS.

7 chips

👉 [pág. 19](#)

**Junio de 1969. M. Shima.** A finales de Junio, tres ingenieros japoneses de Busicom, Masuda, Takayama y Shima, llegaron a Intel para finalizar el diseño lógico y colaborar en su plasmación en silicio. Los dos primeros volverían en Octubre, mientras que el tercero se quedaría en Intel hasta Diciembre, desarrollando los programas específicos para la calculadora.

3 ingenieros

**Julio de 1969. T. Hoff.** Marcian E. (Ted) Hoff, doctor en ingeniería eléctrica por Stanford que había construido los interfaces hardware para los computadores 1620 y 1130 de IBM, echó un vistazo al diseño inicial de Busicom, apreciando una complejidad excesiva en el conjunto. Animado por Noyce, planteó la posibilidad de considerar el diseño de un único chip de propósito general que pudiera ser programado para realizar las funciones de la calculadora, idea que agradó al trio japonés.

1 doctor

**Septiembre de 1969. S. Mazor.** T. Hoff se convierte en el duodécimo empleado de Intel, y ese mismo mes la firma contrata a Stanley Mazor para que junto al anterior adecúe el diseño al de la calculadora Busicom. Mazor trabajaba desde 1964 en Fairchild, donde había diseñado el computador Symbol (aunque siempre ejerció más labores de directivo que de ingeniero, pues sus estudios se limitaban a una diplomatura en Matemáticas).

apoyo

Ellos dos, con una pequeña contribución de Shima en las instrucciones de salto condicional, desarrollaron finalmente un conjunto de 45 instrucciones que más adelante formaron el núcleo del repertorio de instrucciones de la familia 80x86, el ensamblador que sería utilizado desde entonces por varias generaciones de programadores.

ensamblador

**Octubre de 1969. Busicom gira hacia el 4004.** Moore y Noyce se reúnen con Busicom, sugiriendo un nuevo diseño con el 4004 como único chip de propósito general. En torno a él, otros tres chips: El 4001, una memoria ROM de 2.048 bits, el 4002, una memoria RAM de 320 bits, y el 4003, un registro de desplazamiento de 10 bits. Busicom acepta el diseño, pero exige sus derechos de explotación en exclusiva, a lo que Intel accede sin vislumbrar lo que se avecina.

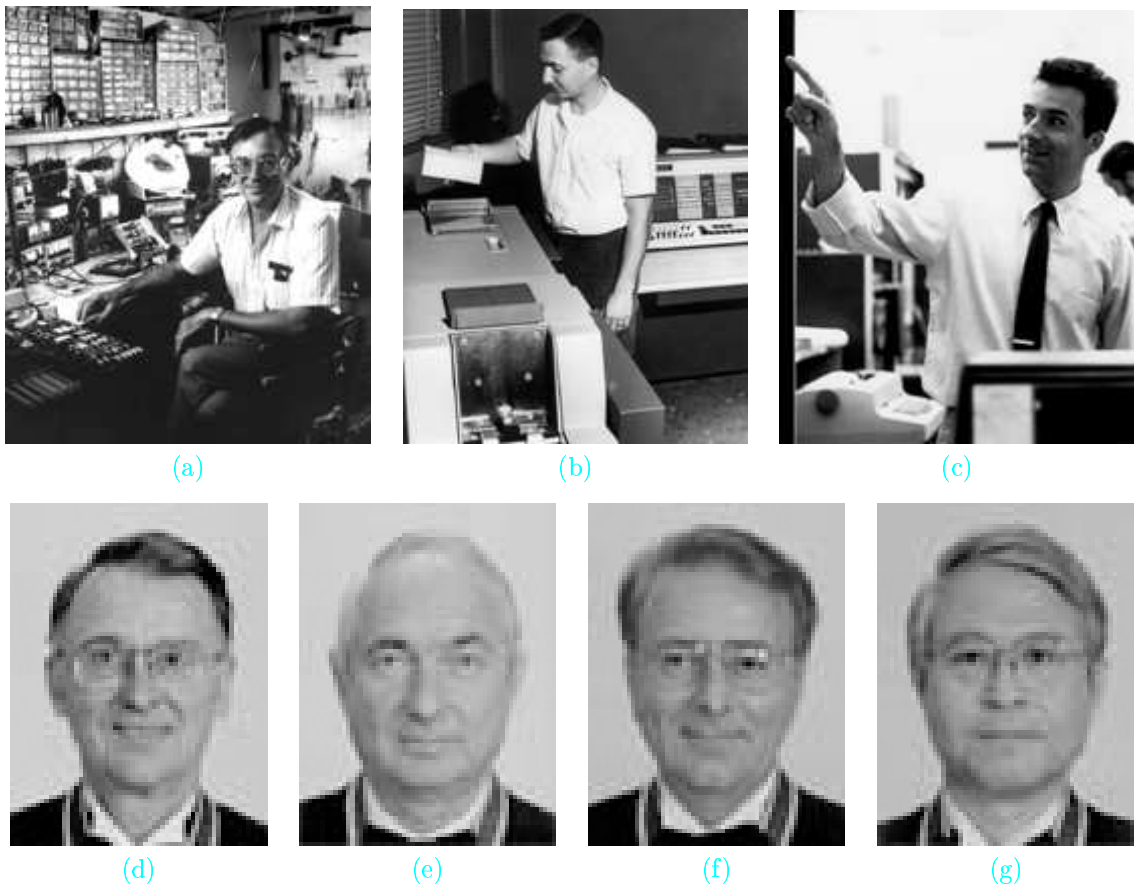
1+3 chips

**Abril de 1970. F. Faggin.** A comienzos de 1970, Intel se había comprometido a producir dos computadores en un solo chip <sup>5</sup>, pero aún no disponía del personal encargado de plasmar los diseños en silicio, ya que ni Hoff ni Mazor tenían experiencia en esa tarea. Federico Faggin, doctor en física y contratado por Intel en Abril de 1970, sería la persona elegida. Faggin tenía experiencia en el diseño de computadores, pues había creado uno para Olivetti en Italia (1961), recabando posteriormente en Fairchild, donde junto a Tom Klein concluyó en 1968 el desarrollo del proceso

2 computadores

<sup>4</sup>Estaba integrada por dos compañías: Electro-Technical Industries (desarrollo), y Nippon Calculating Machines Company (marketing). Abreviaremos denominándola directamente Busicom.

<sup>5</sup>El 4004 y el 8008, cuyo diseño se fragó en Diciembre de 1969 como luego veremos.



Fotos cortesía de Intel

FOTO 1.7: Fotos de los años 70 de (a) Ted Hoff, (b) Stan Mazor y (c) Federico Faggin. Fotos de 1997, el día en el que el cuarteto protagonista del 4004 recibió un premio de innovación tecnológica por parte de la Universidad de Kyoto: (d) Ted Hoff, (e) Stan Mazor, (f) Federico Faggin y (g) Masatoshi Shima.

de fabricación basado en tecnología CMOS, con el que diseñó poco después el primer circuito integrado comercial: El 3708, un multiplexor de 8 bits.

**Octubre-Diciembre de 1970. Las primeras obleas.** Volvió Shima de Busicom en Abril de 1970, encontrándose con el proyecto casi parado, lo que le obligó a quedarse a colaborar con Faggin a un ritmo desenfrenado para cumplir los plazos de entrega acordados. Faggin decidió diseñar los chips en el orden 4001, 4003, 4002 y finalmente, el 4004.

Hoff, Mazor y los colegas de Shima en Busicom ayudaron en las tareas de definición de la arquitectura del 4004 mientras que Faggin trasladaba al silicio el diseño de los otros tres. La [foto 1.8.a](#) muestra a Faggin afanado con el último de la serie. Las primeras obleas del 4001 se fabricaron en Octubre, testeándose sin errores. Shima volvió a Japón en Noviembre para centrarse en la programación de Busicom, mientras las primeras obleas del 4003 y el 4002 se fabricaban con pequeños errores que rápidamente fueron subsanados por Faggin. Las primeras obleas del 4004 se recibieron a final de año, pero en su primer testeo, Faggin descubrió atónito cómo sus chips no respondían ante ninguna entrada. Media hora más tarde descubrió el error, localizado en uno de los seis niveles de máscara del diseño <sup>6</sup>. Faggin confesaría posteriormente que aquella fue la media hora más larga de su vida.

<sup>6</sup> Este es el segundo error documentado que tenemos en la fabricación de chips en Intel, tras aquel bit perdido en la RAM 3101.

**Enero de 1971. El nacimiento del 4004.** Al redefinir las máscaras, hubo de reiniciarse el proceso de producción, manufacturándose una nueva remesa de chips a mediados de Enero. Faggin testeó la serie nada más recibirla, funcionando a la perfección. Aquella noche nació el primer microprocesador de la historia. Seguidamente, Faggin envió la serie 4000 a Japón, donde Shima procedió a ensamblar los diez chips de que constaba la calculadora, haciendo funcionar al conjunto correctamente. Busicom lanzó sus primeros modelos comerciales en el mes de Junio de 1971.

alumbramiento

**Mayo de 1971. La patente del 4004.** Faggin montó en cólera cuando se enteró de que la serie 4000 de Intel tenía un derecho exclusivo de explotación por parte de Busicom, pues ya intuía el esplendoroso potencial del producto. En una conversación telefónica con Shima, Faggin descubrió que Busicom atravesaba graves problemas financieros y necesitaba de una reducción de costes en la serie 4000 que se tradujera en un precio más competitivo para su calculadora. Faggin se lo chivó a Noyce, y entre ambos devolvieron la jugada a Busicom: Le bajaron el precio de la serie 4000 a condición de recuperar sus derechos de explotación, a lo que Busicom accedió siempre y cuando no suministrase el producto a sus competidores.

la jugada

**Agosto de 1971. El éxito comercial del 4004.** Tras solventar algunos problemas con la estabilidad de la RAM a altas temperaturas, Faggin se afanó en demostrar la utilidad del 4004 en otro tipo de mercados. Dov Frohman-Bentchkoswky acababa de inventar la EPROM en Intel, y Faggin y Hoff ilustraron cómo cambiando la ROM 4001 por una EPROM, el sistema podía reprogramarse para realizar multitud de tareas <sup>7</sup>.

EPROM  
versatilidad

En Agosto, la serie 4000 ya era la principal fuente de beneficios de Intel, y Moore y Noyce no tardaron en volar a Japón para comprar la totalidad de la patente a Busicom. Pagaron por ella 60.000 dólares de aquella época, cuando en el primer semestre de vida, Intel había obtenido unos beneficios de 2.600 dólares. Parecía la ruina de Intel, pero la que nunca se recuperó de aquello fue Busicom, que quebraría tan sólo unos meses después. Intel doblaría sus beneficios al siguiente semestre, y sólo con una progresión geométrica podemos explicar que treinta años más tarde, cuando la idea que está detrás del 4004 ha desarrollado ya buena parte de su potencial, la empresa obtenga unos beneficios anuales superiores a los 35.000 millones de euros <sup>8</sup>.

el negocio

**Noviembre de 1971. El lanzamiento del 4004.** Aunque había aparecido ya un anuncio en la prensa (Revista Datamation, número de Mayo de 1971), Intel era reticente a lanzar oficialmente el microprocesador para no comprometer su reputación de compañía devota al desarrollo de memorias. Pero Eduard Gelbach, que se incorporó a Intel en el verano de 1971 como vicepresidente de marketing procedente de Texas Instruments, desterró todo conservadurismo y forzó la salida oficial al mercado del 4004 el 15 de Noviembre de 1971. El número de ese mes de la revista Electronic News ya incluía publicidad del producto, que acompañamos en la [foto 1.5.b](#); posteriormente aparecería en la sección de novedades de la revista IEEE Computer en Enero y Febrero de 1972 bajo el título "Computer on a chip".

la publicidad

➔ [pág. 13](#)

Las [fotos 1.8.b](#) y [1.9.c](#) muestran, respectivamente, el aspecto interno y externo del 4004, aunque cuesta trabajo imaginárselo comparado con sus hermanos contemporáneos: Dotado de 2300 transistores en un espacio físico de 12 mm<sup>2</sup> (3.125 mm. x 4.166 mm.), disponía de tan sólo 16 patillas en formato DIP y estaba construido parcialmente de madera. El bus de datos externo era de tan sólo 4 bits y se encontraba parcialmente solapado con el de direcciones, mostrando una capacidad de direccionamiento de 640 bytes y un código de operación de 8 bits. Trabajaba a una frecuencia de 750 KHz (aunque el objetivo de diseño fue 1 MHz), pero con un peso de 20 gramos y un precio de 200 dólares de entonces ya ofrecía una funcionalidad muy similar a la del ENIAC, el primer computador de la historia (ver [foto 1.9.a](#)), cuyo peso era de 30 toneladas y su coste, superior al medio millón de euros. Hoy en día, con un proceso de fabricación de 0.13 micras, la integración en silicio del 4004 ocuparía en total la centésima parte de un milímetro cuadrado, pesaría menos

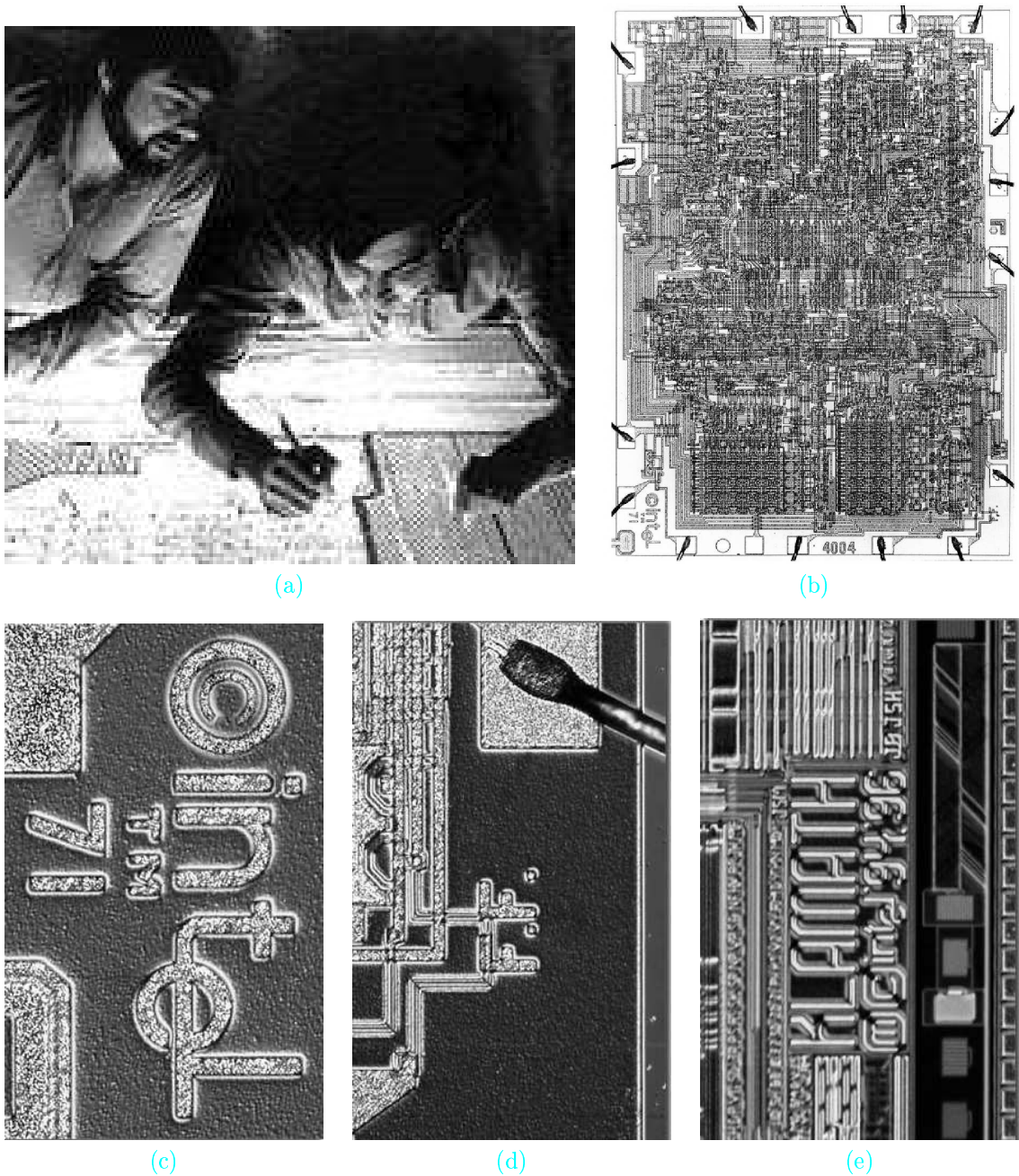
➔ [pág. 18](#)➔ [pág. 19](#)

el chip

➔ [pág. 19](#)

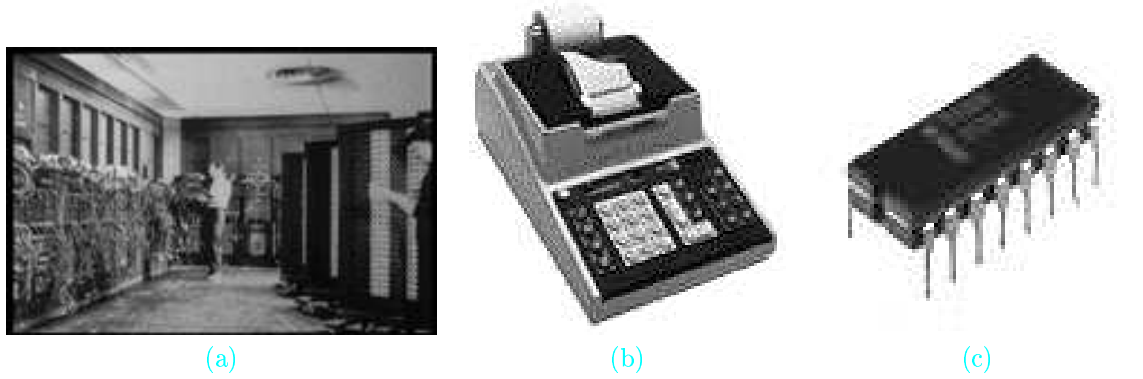
<sup>7</sup>El concepto de programa almacenado es originario del EDSAC (Electronic Delay Storage Automatic Calculator), computador que estuvo operativo en 1949, construido por Maurice Wilkes, el padre de la microprogramación.

<sup>8</sup>Datos del ejercicio correspondiente al año 2000.



Fotos cortesía de Intel

**FOTO 1.8:** (a) Federico Faggin trasladando el diseño del 4004 al *floor plan*, el plano donde se detalla la ubicación de sus 2.300 transistores. (b) Detalle de dicho plano, de 1.5 x 2 metros, 500 veces ampliado sobre su traslación a silicio (3 mm. x 4 mm.). (c) Esquina inferior izquierda del plano según la foto anterior, donde descubrimos el logo de Intel y el año de manufacturación. (d) Esquina inferior derecha de ese mismo plano, donde Federico grabó sus iniciales (en vertical), costumbre que perduró según podemos comprobar en (e) area de integración del Klamath (Pentium II de 0.35 micras), con las iniciales JSM (sobre la H de KLAMATH), y JB y JSH (a continuación del 96 - James S. Miller fue el jefe del proyecto Klamath, y Jim Bergman y Jennifer S. Hernández quienes trasladaron el diseño al silicio).



Fotos cortesía de Intel

FOTO 1.9: Algunas reliquias del pasado. (a) El ENIAC, considerado como el primer computador de propósito general de la historia, del tamaño de varios armarios roperos. El primer ENIAC se conserva en el Museo de la Ciencia de la Universidad de Pennsylvania, en Philadelphia, en honor a la procedencia investigadora de sus creadores. (b) La calculadora de sobremesa de Busicom que dió origen al proyecto del microprocesador 4004 de Intel. El sistema se compuso de un microprocesador 4004, dos chips 4002 de RAM (80 bytes en total), tres chips 4003 de ROM (1 Kbyte en total), y cuatro chips 4001 con un registro de desplazamiento cada uno (la gama avanzada de Busicom que realizaba la raíz cuadrada disponía de un 4001 adicional). (c) Aspecto externo del 4004, considerado el primer microprocesador de la historia.

de 1 miligramo y costaría aproximadamente 0.0005€.

### 4.1.2 El que pudo ganar al 4004

En apariencia, el 8008 es una extensión del 4004 a 8 bits (ver foto 1.10.a), pero la realidad fue muy diferente: Su diseño partió de algo tan diferente como una memoria RAM estática de 64 bits, el chip 3101 de Intel basado en tecnología bipolar, y bien pudo estar acabado antes que el propio 4004.

8008  
 ➔ pág. 23

El origen del proyecto 8008 estuvo en una petición que la empresa Computer Terminals Corporation (CTC) de San Antonio (Texas) hizo a Intel en Diciembre de 1969 para transformar el 3101 en una memoria tipo LIFO de 4 palabras de 16 bits que necesitaban para un terminal inteligente que CTC estaba desarrollando, el Datapoint 2200.

1a idea  
 CTC

Mazor y Hoff estudiaron la propuesta, observando que el procesador que requería aquel equipo no era mucho más complejo que la propuesta de 4004 que Moore y Noyce habían presentado un par de meses atrás a Busicom; unas semanas antes de que Intel contratara a Faggin para el 4004 llegó Hal Feeney para trabajar en el 8008. Feeney trabajó con Mazor y CTC para completar la especificación del chip, quedando pronto a las puertas de la implementación del silicio y por delante del estatus del 4004.

el demarraje

En ese momento, CTC cayó en una grave crisis financiera e Intel redujo la prioridad del proyecto, que quedó parado hasta enero de 1971, cuando, concluido el 4004, Intel lo despertó: Sería Faggin quien lo concluiría junto a Feeney, amortizando toda la metodología de diseño y testeo en silicio puesta en marcha para el 4004. En Marzo de 1972, Intel fabricaba los primeros chips comerciales del 8008, ya con 3.500 transistores.

1a pájara

1a meta

### 4.1.3 El que pudo ganar a Intel

De forma un tanto irónica, la compañía CTC mantuvo negociaciones con Texas Instruments

CTC + TI

	Hito Histórico	Microprocesador 4004 de Intel	Microprocesador 8008 de Intel	Variante 8008 de Texas Instruments
O R I G E N	Empresas implicadas	Intel y Busicom	Intel y CTC	Texas Instruments y CTC
	Firma del proyecto	Abril, 1969	Diciembre, 1969	Marzo, 1970
	Producto al que iba destinado	Calculadora Busicom	Terminal inteligente Datapoint 2200 de CTC	
D I S	Diseño del que partió	Ninguno. Todo original	Circuito Integrado SRAM I3101 de tecnología bipolar y 64 bits	
	Comienzo de desarrollo	Tercer trimestre, 1969	Primer trimestre, 1970	Segundo trimestre, 1970
E Ñ O	Finalización del testeo	Marzo, 1971	Marzo, 1972	Nunca llegó a funcionar de forma correcta según fuentes de CTC
	Ingenieros que trabajaron	Faggin, Hoff, Mazor e ins. de Busicom	Faggin, Feeney, Mazor e ins. de CTC	
V E N T A S	Producción masiva	Junio, 1971	Marzo, 1972	
	1 <sup>er</sup> producto que lo utilizó	Calculadora Busicom (04/71)	Calculadora científica programable de Seiko	
	Publicidad en prensa	Noviembre, 1971	Abril, 1972	Junio, 1971

TABLA 1.2: Los diez hitos históricos más importantes tal y como se sucedieron en apenas dos años entre los tres modelos que aspiraron a convertirse en el primer microprocesador comercial.

para que se encargara de diseñar una solución alternativa a la de Intel para su Datapoint 2200, utilizando en este caso el proceso de integración MOS y conexiones de aluminio.

El chip que derivó de aquel encargo era más del doble de grande que el 8008, y llegó a aparecer anunciado en la prensa especializada en Junio de 1971 como la primera CPU en un solo chip. Sin embargo, fuentes de CTC aseguran que aquel chip nunca llegó a funcionar en su totalidad, y su ausencia en productos comerciales de la época corroboran esta percepción.

La tabla 1.2 resume los diez aspectos más importantes que se sucedieron en la carrera que libraron los tres modelos vistos por alcanzar el hito del primer microprocesador comercial.

## 4.2 ▶ Los modelos que tomaron el relevo

### 4.2.1 Las innovaciones del PARC

**Alan Kay** **1972. El Dynabook.** Alan Kay, uno de los integrantes del PARC, comenzó en 1971 a gestar lo que consideramos como el primer computador portátil de la historia: El Dynabook. Su especificación contemplaba un diseño muy ligero y personal, carente de teclado, y que ejecutaba software basado en su innovador lenguaje de programación Smalltalk, delegando el interfaz con el usuario sobre un original modelo de trabajo tan simple como un puntero móvil. El omnipresente ratón de nuestros días daba así sus primeros pasos en 1972, y de paso Kay comenzaba a hacer realidad su frase más célebre: "La mejor manera de predecir el futuro es crearlo."

**1973. El Alto, precursor de tantas cosas.** El Dynabook parecía demasiado vanguardista, pero en 1972, otros dos ingenieros del PARC, Chuck Thacker y Butler Lampson, pidieron a Kay colaboración para hacer realidad un computador basado en la concepción de aquella pequeña máquina. Considerando el estado de la tecnología de entonces, resulta muy meritorio que tardaran sólo un

año en producir el Alto, un computador de sobremesa con una pantalla de unas 15 pulgadas y un coste de 20.000 € de aquella época que hizo que apenas tuviera calado comercial.

el Alto

Thacker y Lampson se marcharían más tarde a Digital, intentándolo de nuevo con otro portátil, el computador Lectrice, que fracasó comercialmente por su peso, superior a los 3.5 Kg. Posteriormente reclutados por Microsoft, comenzarían en el verano de 1999 el proyecto Tablet PC para crear un producto similar al Dynabook treinta años más tarde (2002), sólo que esta vez aprovechando la excelente autonomía del microprocesador Caruso de Transmeta y los adelantos en el reconocimiento de la escritura manual como innovación en el interfaz de usuario.

Lectrice  
Tablet PC

El principal problema del Alto no fue incumplir las especificaciones iniciales, pues once años más tarde (1984) vería la luz el Macintosh de Apple, claramente inspirado en el Alto, y ése arrasó en ventas hasta tal punto de que el propio Kay lo denominó desde entonces “el primer computador lo suficientemente bueno como para ser criticado”.

el Mac

El Macintosh de 1984 utilizó el 68000 de Motorola como microprocesador y un sistema operativo basado en el entorno del ratón, las ventanas y los iconos. Los mismos elementos serían también copiados ocho años después por Microsoft (1992) en su sistema operativo Windows para convertirlos en elementos tan cotidianos en nuestros días. ¿Por qué el mismo modelo que triunfó en los 80 y en los 90 no lo hizo en los 70? Por las dos mismas razones que han tumbado tantos proyectos informáticos desde entonces: Un coste elevado y una innovación excesiva. Aún en nuestros días siguen desfilando ante nuestros ojos numerosos ejemplos de productos que fracasan por ir muy por delante de las necesidades que tienen los usuarios a los que éstos van destinados. Sólo dentro de Intel tenemos un ejemplo ilustrativo en cada década: El microprocesador i432 en los años 70, el i860 en los años 80, y el Itanium en los años 90. Hacia ellos nos dirigimos.

ventanas  
e iconos

#### 4.2.2 Los sucesores del 4004 en Intel

**1972. El impacto del 8008.** La llegada al mercado del 8008 originó un doble halo: De complacencia, por incluir ya interrupciones, y de frustración, porque éstas nunca funcionaron del todo bien. Al tiempo, numerosos agoreros se mostraron escépticos sobre la viabilidad de Intel, señalando que el mercado al que estaba dirigido este computador integrado era demasiado pequeño. Después de todo, en aquellos tiempos el computador tenía un uso tan restringido que sólo existían unos pocos miles en todo el mundo, y el mercado estaba dominado por un oligopolio en el que IBM era el rey seguido a mucha distancia de lo que en el argot se conoció como la *pandilla* (significado de la palabra inglesa BUNCH y cuyas iniciales corresponden, respectivamente, a Burroughs, Univac, NCR, Control Data y Honeywell, las cinco compañías con notoria presencia en este mercado). En estas condiciones, la cuota de mercado a que podía aspirar Intel con un diseño que daba problemas era bastante cuestionable.

oligopolio

**1974. El 8080 de Intel.** Afortunadamente para Intel, el 8080 va a ver la luz en Abril de 1974. En aquel microprocesador, las interrupciones funcionan de maravilla, los buses de datos y direcciones han conseguido desacoplarse simplificando el diseño del sistema en su conjunto, y aunque el banco de registros sigue anclado en los 8 bits, las unidades de procesamiento son capaces de trabajar con 16 bits en algunas de sus instrucciones consiguiendo manejar los registros por pares. El número de transistores que contiene el chip sube hasta los 4500, y la capacidad de direccionamiento se dispara hasta los 64 Kbytes, dando a Intel el crédito que necesitaba para continuar trabajando.

16 bits

**1975. 8080 + CP/M = Altair.** Con el 8080 en la calle, el decorado cambia sustancialmente. Algunos visionarios, como Gary Kildall, fundador de Digital Research, comienzan a creer en el potencial que tienen estos dispositivos si su coste sigue bajando y se le arroja de una capa software que facilite su programación. Kildall había desarrollado el lenguaje de alto nivel PL/M para programar el 8008, que a través de un compilador podía generar código objeto para diferentes tipos de sistemas basados en el 8008 y el 8080. Guiado por su propia intuición, Kildall desarrolla el

sistema  
operativo

Firma comercial y microprocesador	Innovación que introdujo	Sistemas a los que iba dirigido	Problemas del diseño
Intel 4004 (1971)	El primer microprocesador	Calculadoras aritméticas	Recursos limitados
Intel 8008 (1972)	Bus de 8 bits. Interrupciones	Terminales y calculadoras	Deficiente gestión de interrupciones
Intel 8080 (1974)	Buses separados para datos y dirs.	Computador Altair	Programación compleja
Texas I. TMS1000 (1976)	Memoria interna al chip	Aplics específicas y de bajo coste	No permite añadir memoria externa
Intel 8086/88 (1978/79)	Bus de 16 bits.	PC y portátiles	Direccionamiento solapado
Motorola 68000 (1980)	Chip de 16 y 32 bits. Gráficos avanzados	Estaciones Unix. Videojuegos	ALU y bus datos externo de 16 bits
Intel 80286 (1982)	Memoria protegida. Direcciona 16 Mbytes	CPU estándar para PC	No gestiona faltas de página
HP FOCUS/PA (1984/86)	Primeros procesadores puros de 32 bits	Estaciones Unix de HP.	Enorme ROM de microcódigo
Intel 80386 (1985)	Bus de 32 bits. MMU interna	CPU estándar para PC	Carece de FPU y caché internas
MIPS R2000 (1986)	Primer RISC comercial. 5 etapas segmentadas	Estaciones Unix. Minicomputadores	Difícil programar. Incompatible PC
Sun SPARC (1987)	Arquitectura RISC abierta	Portátiles y estaciones	Multitud de chips. FPU externa
Intel 80486 (1989)	Incorpora caché L1 unificada y FPU	PC, aplicaciones gráficas	Diseño CISC. Baja frecuencia
Intel i860 (1989)	Primer diseño superescalar	De propósito específico	Coste elevado
MIPS R4000 (1991)	Primer micro supersegmentado	Estaciones de trabajo de SGI	Bajo rendimiento con dependencias
Digital Alpha 21064 (1992)	Frecuencia de reloj a 200 MHz	Estaciones y servidores	Temperatura y coste elevados
Motorola e IBM Power PC (1992)	Ejecución de instrs. fuera de orden	Mac, PC y servidores.	SW. no aprovecha ejec. fuera orden
Intel Pentium (1993)	Pred. dinám. salto. Bus datos 64 bits	PC y servidores	Temperatura. Legado CISC
Digital Alpha 21164 (1995)	Superescalar factor 4. Caché L2 96K. 300 MHz	Estaciones y servidores	Temperatura y coste elevados
Intel Pentium Pro (1995)	Caché L2 256/512K Sop. multiprocesador	PC de gama alta	Ineficiente con SW 16 bits. Caro
Digital Alpha 21264 (1998)	Renombre 80 regs. Ejec fuera de orden. 500MHz	Estaciones y servidores	Temperatura y coste elevados
Transmeta Crusoe (2000)	Microcódigo programable y versatilidad	Portátiles. Sist. empotrados.	Velocidad de proceso
Compaq Alpha 21364 (2000)	Multiproc. 4 canales CPU-CPU a 10 Gbps.	Supercomputadores y servidores	Coste elevado
Intel y HP Itanium (2001)	Paralelismo explícito	Servidores SMP	Dependencia de la capa software

TABLA 1.3: Las principales innovaciones en el campo de los microprocesadores.



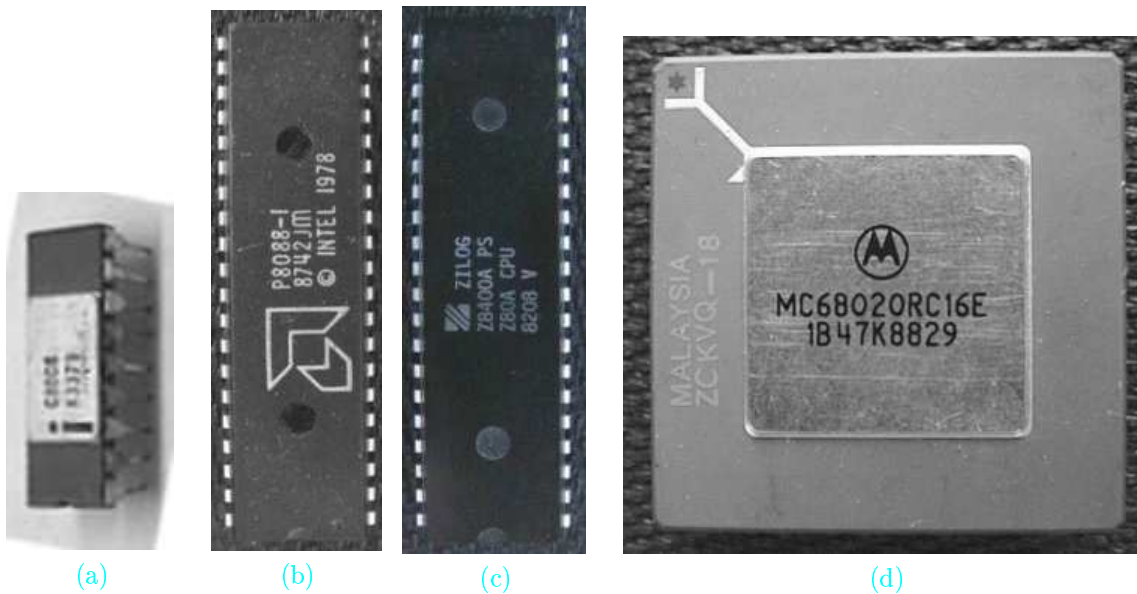


FOTO 1.10: Algunos de los primeros microprocesadores: (a) El 8008 de Intel. (b) El 8088 de Intel, donde podemos ver el sello de AMD como integrador del chip (las vueltas que da la vida). (c) El Z80 de Zilog. (d) El 68020 de Motorola.

sistema operativo CP/M sobre la plataforma del 8080, convenciendo a Intel para que comercialice el conjunto en 1975 por un precio inferior a 1.000 dólares. El computador ve la luz bajo el nombre de Altair<sup>9</sup> a un precio de 620 dólares incluyendo la CPU, el panel frontal, la fuente de alimentación y la carcasa. El volumen de ventas sobrepasa todas las estimaciones, lo que comienza a dar la razón a Kildall y, de paso, a poner nerviosos al resto de ingenieros del gremio.

#### 4.2.3 Otros contendientes

**1975. El Z80 de Zilog.** Shima y Faggin abandonaron Intel a finales de 1974 para fundar una nueva compañía: Zilog. Zilog demostró una agilidad encomiable, y en 1975 lanzó al mercado el Z80 (ver foto 1.10.c), dotado de 8.500 transistores, una frecuencia de 2.5 MHz, direccionamiento a 64 Kbytes de memoria y compatible con el 8080 (por tanto, capaz de beneficiarse de la infraestructura software del sistema operativo CP/M). El Z80 obtuvo un clamoroso éxito, llegándose a vender más de 500 millones de procesadores de 8 bits (luego vendrían el Z8 y el Z8000).

Por ejemplo, Clive Sinclair ideó los primeros computadores domésticos de bajo precio y amplia difusión que llegaron a buena parte de Europa (España incluida): El ZX80 (Febrero, 1980), dotado de 1 Kbyte de memoria, y sus sucesores, el ZX81 y el Spectrum, todos ellos basados en el microprocesador Z80 de Zilog.

Spectrum

A principios de los años 80, Zilog derivaría su actividad hacia el desarrollo de microcontroladores y chips de propósito específico, declinando todo protagonismo en la historia reciente del computador personal. Sinclair Research, por su parte, se embarcó en dos proyectos revolucionarios: El C5 (un coche eléctrico monoplaza) y el QL (*Quantum Leap - Salto Cualitativo*), un nuevo computador de prestaciones profesionales a precio doméstico. El fracaso comercial del coche fue mayúsculo, y los retrasos en la finalización del computador terminaron por llevar a Sinclair a la

<sup>9</sup>El destino de la nave Enterprise de la serie Star Trek en el capítulo correspondiente a la semana en la que Intel lanzó al mercado este producto.

la transición:  
Commodore y MSX

bancarrota, acabando reconocido como Sir y su empresa adquirida por Apple a un precio irrisorio. El Commodore-64 de la firma Commodore y el estándar japonés MSX tomarían el mercado dejado por el Spectrum, abanderando la transición hacia el IBM-PC en Europa.

sistemas  
empotrados

**1976. El TMS-1000 y el 6800.** Al calor del éxito del Z80 entran en escena nuevos fabricantes con microprocesadores de prestaciones similares, como Texas Instruments con su TMS-1000 y Motorola con su 6800 (el más madrugador de la terna, desarrollado apenas seis meses después del lanzamiento del 8080 en 1974). No obstante, ambos muestran cierta predilección por la fabricación de microcontroladores para sistemas empotrados, decisión que ha permitido a estas compañías diversificar notablemente su actividad en nuestros días.

super-  
mini-

**1976. Los minicomputadores.** Mejor suerte en el campo computacional correrían otros fabricantes como Digital, Hewlett-Packard ó Data General, que desarrollaron con éxito sistemas de tamaño más pequeño que los mainframes (los supercomputadores de IBM y compañía), los denominados minicomputadores, un nuevo mercado donde pronto alcanzarían cierta notoriedad.

micro-

**1977. Apple y el microcomputador.** En 1977 aparece en el mercado el Apple II, fabricado por Steve Jobs y Steve Wozniak en el garaje de su casa, y que con una orientación más doméstica consiguió un notable éxito en ventas, sentando así las bases del bajo coste, gran volumen y elevada fiabilidad que han caracterizado al computador personal hasta nuestros días. No obstante, la década de los 70 concluiría con Digital instalado por delante de Apple en el segundo lugar entre las empresas proveedoras de informática, y por encima de todos ellos y a gran distancia, IBM.

#### 4.2.4 La pifia y el resurgimiento de Intel

**1974-77. El 432 de Intel.** Intel comete su mayor fiasco justo después de completar el 8080 en 1974, al volcarse en un proyecto mucho más ambicioso: El i432. Este microprocesador iba muy por delante de su época, incluyendo rasgos tan sorprendentes que algunos de ellos no volvieron a ser vistos en un microprocesador hasta pasados veinte años:

- P00 **1** Incluía soporte para la programación orientada a objetos, de tal forma que cada elemento de datos tenía su tipo asociado (entero, carácter, real, ...), y el procesador siempre comprobaba que cada dato era del tipo correcto antes de utilizarlo.
- punteros **2** Alojaba los datos en memoria utilizando varios niveles de direccionamiento indirecto mediante punteros, brindando al software una gran flexibilidad en la gestión de memoria.
- robustez **3** Gestionaba los errores producidos por la memoria, lo que unido a una arquitectura tolerante a fallos le proporcionaba una extraordinaria robustez.
- SMP **4** Incluía soporte para multiprocesamiento simétrico (SMP) (más adelante veremos que este rasgo no se retomó en un microprocesador para PC hasta la sexta generación de microprocesadores, con la llegada del Pentium Pro en 1995).

el fracaso

Debido a su complejidad, la creación del i432 llevó mucho más tiempo del esperado. Intel echó en falta a Shima, y, sobre todo, a Faggin, quienes habían adquirido una visión global que les permitía diseñar la arquitectura de un procesador facilitando mucho su posterior integración. Así, cuando Intel completó la versión inicial de su i432 en 1977, comprobó atónita cómo el intrincado acceso a los datos por punteros y las extensivas comprobaciones en los tipos de datos arruinaban el rendimiento del chip, dando un rendimiento entre 5 y 10 veces inferior al del Z80 o el TMS-1000. El i432 no se lanzó al mercado hasta 1980, donde se hundió estrepitosamente convirtiéndose en uno de los fracasos más sonados de la industria del microprocesador.

el camino

Pero la diferencia entre un talante ganador y otro perdedor está en que el primero *siempre* se recupera de un fracaso, mientras que el segundo *nunca* se recupera de un éxito. Zilog acabaría desmantelándose, y sus fundadores, saltando a otras compañías. Intel, en cambio, reaccionó justo

a tiempo para reponerse a lo grande de aquel tropiezo. Digital y Hewlett-Packard habían enseñado al resto cuál era el camino bueno: Construir computadores más pequeños (los “mini”). Y Apple había rematado la faena avanzando un paso más en esta dirección con la introducción de los “micro”.

**1978. El 8086.** En cuanto Intel vió lo que se avecinaba, corrió para definir un nuevo proyecto: el 8086. Consciente de lo delicado de la situación, le concedió tan sólo tres semanas de plazo a dos de sus ingenieros, Stephen Morse y Bruce Ravenel, para desarrollar el conjunto de instrucciones y un diseño preliminar del chip, etapas que se concluyeron en 1976. Y así fue como en 1978 irrumpió en el mercado el 8086, una arquitectura de 16 bits con bus externo de 8 bits y dotada de 29000 transistores.

la urgencia

**1979. El 8088.** Un año más tarde vería la luz el 8088, la versión de 8 bits del 8086 (ver [foto 1.10.b](#)), pero a un precio mucho más atractivo. Estos dos nuevos diseños de propósito general realizaban computaciones novedosas, como el producto y la división en una arquitectura de 16 bits, consiguiendo multiplicar por diez las prestaciones del 8080 al tiempo que mantenían la compatibilidad con él.

➔ [pág. 23](#)  
velocidad a bajo coste

## SECCIÓN 1.5

## EL IBM PC

Y llegamos, por fin, a la encrucijada clave para entender la historia reciente de la informática: El momento en el que IBM, aún por entonces el rey en la escala de supercomputadores, aprendió de la batalla perdida con Digital en los minicomputadores y no quiso que el fenómeno se repitiera en el mercado de los recién emergidos microcomputadores (aquellos con coste en torno a los 3.000 dólares).

la encrucijada

Para evitarlo, la división de IBM en Boca Ratón (Florida) lanzó el proyecto Chess en 1980 para desarrollar plataformas de este tipo, lo que desembocó en el anuncio del IBM-PC tan sólo trece meses más tarde, el 12 de Agosto de 1981 (ver [foto 1.11.a](#)).

el anuncio

Por primera vez en su historia, y consciente de la premura de tiempo, IBM lanzó un producto a base de ensamblar material procedente de otros fabricantes que pudieran suministrarle componentes de forma casi inmediata y con una contrastada solvencia:

➔ [pág. 26](#)  
esos descuidos...

- ➊ A pesar de ser IBM quien había inventado la disquetera o *floppy*, no fue él quien fabricó las del IBM-PC, sino Tandon, simplemente porque tenía la capacidad de fabricar rápidamente grandes remesas a un precio asequible.
- ➋ IBM prescindió de la unidad de cinta magnética, entonces muy popular para el almacenamiento masivo, optando en su lugar por el disco duro, que él también había inventado. Pero una vez más, encargó su fabricación a otras marcas.
- ➌ Como núcleo software del sistema operativo, IBM decidió utilizar el MS-DOS de Microsoft (ver [foto 1.11.b](#)).
- ➍ Como microprocesador, se optó por el 8086 de Intel, capaz de direccionar hasta 1 Mbyte de memoria principal, 16 veces más que el Z80 de Zilog y otros competidores.

- disquetera

- disco duro

- sistema operativo

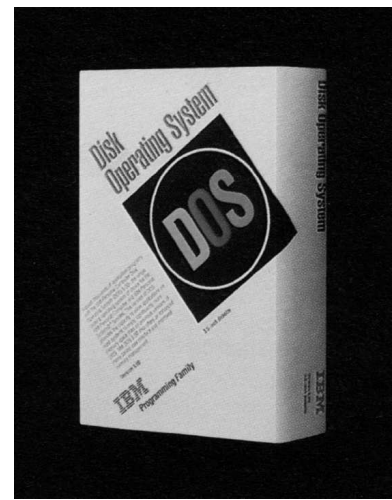
➔ [pág. 26](#)

- procesador

Para entender por qué Microsoft e Intel han ido ganando popularidad en el contexto del PC desde aquella encrucijada y a IBM le ha ocurrido lo contrario, vamos a ampliar la historia de cada compañía por separado. Su enorme trascendencia futura así lo sugiere.



(a) IBM-PC



(b) MS-DOS

Fotos cortesía de IBM

FOTO 1.11: El nacimiento del PC y su sistema operativo MS-DOS. (a) El IBM-PC original, lanzado al mercado el 12 de Agosto de 1981 con la siguiente configuración: Procesador 8088 a 4.77 MHz, 16 Kbytes de memoria ampliable a 256 Kbytes, 5 zócalos para periféricos y sistema operativo DOS 1.0. (b) El PC dobla su tamaño de memoria y disco un año más tarde, gracias a la versión 1.1 del sistema operativo MS-DOS.

## 5.1 ▶ La fábula de Intel y el IBM-PC

Al contrario que con Busicom diez años atrás, Intel no era el único candidato para desarrollar el microprocesador del IBM-PC. Sobre la mesa, había otros dos pretendientes de peso: Motorola y Zilog.

esa fecha...

Si IBM hubiese tomado aquella decisión en 1976 o 1977, muy probablemente se hubiese decantado por el Z80 de Zilog, animado por la credibilidad que le reportaron sus ventas; de haberlo hecho en 1981, quizá hubiese seleccionado el 68000 de Motorola, cuyos registros internos eran ya de 32 bits y su espacio de direcciones de 16 Mbytes. En cualquiera de esos dos supuestos, podemos estar seguros de que el mundo de los microprocesadores sería hoy muy distinto del que conocemos. Pero a IBM le tocó elegir coincidiendo en el tiempo con la novedad del 8086 y el espaldarazo dado por el 8088. La excelente fiabilidad que ambos demostraron tener y la experiencia de Intel hicieron que la balanza terminara decantándose a su favor.

la paradoja

La historia quiso que un diseño como el i432 en el que Intel trabajó durante seis años, se perdiera en el tiempo sin solución de continuidad, y otro como el 8086, en el que trabajó un año escaso, se convirtiera en el referente más importante de la historia del microprocesador.

el órdago

Pero a Intel le va a sonreír la fortuna aún más: Por aquel entonces, IBM arrastraba diversas denuncias que le acusaban de haberse apropiado de algunos productos de otras pequeñas empresas, y fue ahí donde Intel se enrocó para negarse a firmar el protocolo de confidencialidad que IBM exigía. Probablemente ni la propia Intel sabía lo mucho que arriesgaba haciéndose de rogar, pero el caso es que IBM terminó cediendo a las pretensiones de Intel.

la fiebre

La enorme popularidad que a renglón seguido adquirió el IBM PC provocó el nacimiento de toda una legión de empresas clónicas dedicadas a la fabricación de arquitecturas PC, aprovechando el hecho de que su principal pieza podía ser libremente adquirida a su creador. El gran beneficiado de todo esto fue por supuesto Intel, que rápidamente extendió su cuota de mercado a todas estas compañías.

La estimación inicial de ventas que IBM había hecho para su PC era de 240.000 unidades en sus primeros 5 años de andadura, pero ese volumen de ventas ya fue sobrepasado durante su primer mes de existencia. Entre los analistas que creyeron en el producto, los más optimistas apuntaron entonces que el número de PC podría llegar a 80 millones a finales de siglo; y el siglo concluyó, con más de 500 millones de PC instalados en todo el planeta.

el boom

Todos estos hechos sorprendieron a Intel en situación inmejorable, y así fue como se convirtió en una de las diez empresas más grandes del mundo en volumen de facturación, con un parque de microprocesadores que en 1997 sobrepasó los 300 millones de unidades, y que desde entonces sigue creciendo a un ritmo de ventas que supera los 100 millones anuales.

el gigante

## La fábula de Microsoft y el IBM-PC

◀ 5.2

Aunque estamos narrando la historia del hardware, haremos una sucinta incursión en el software. Bien merecido lo tiene Microsoft: Ninguna otra compañía ha defendido de forma tan aguerida el modelo de computador personal y a su vez debe tanto a esta firme apuesta. El matrimonio ideal, que va camino de cumplir sus bodas de plata sin una mala bronca, algo verdaderamente insólito en un mundo tan dinámico y promiscuo como el de la informática.

un fiel  
candidato

Pero el sistema operativo que la historia hubiera apadrinado encantada para el IBM-PC no era de Microsoft. Se trataba, por pura lógica y justicia, del CP/M de 16 bits: Si el CP/M se había convertido en el estándar del 8080, su nueva versión orientada al 8086 era la pieza que a IBM le faltaba para completar su puzzle una vez se había decidido por este microprocesador para su PC.

un gran  
pretendiente

El CP/M era una perita en dulce en manos de Digital, y Gary Kildall lo sabía. Así que, aún sin disponer del producto terminado, estaba pidiendo a IBM bastante más de lo que ésta estaba dispuesta a conceder. A IBM se le empezó a agotar la paciencia, y viniendo ya chamuscada de las conversaciones con Intel, decidió romper la baraja con Digital.

la ruptura

Fue entonces cuando apareció Bill Gates, co-fundador de Microsoft junto a Paul Allen. A finales de los años 70, Microsoft no era más que una pequeña compañía cuyo único mérito consistía en haber desarrollado un excelente intérprete de BASIC, el lenguaje que se había consolidado para la programación del computador Altair.

ese señor...

Gates ofreció a IBM un sistema operativo de 16 bits y enseguida las dos partes se entendieron. IBM lo comercializaría bajo el nombre de PC-DOS, mientras que Microsoft lo haría bajo MS-DOS. Trato hecho.

el compromiso

El único problema de ese sistema operativo es que no existía: Todo era un farol del señor Gates, quien únicamente conocía a alguien a quien comprárselo. En la empresa Seattle Computer Products trabajaba un tal Tim Paterson que había construido un precario sistema operativo de 16 bits para ser ejecutado en un 8086 de Intel por pura necesidad, en espera de la llegada del ansiado CP/M de 16 bits.

de farol

Paterson denominó a su sistema 86-DOS, y usaba las mismas llamadas a funciones internas que el CP/M, al tiempo que seguía las convenciones y direcciones del 8086 hechas públicas por Intel y utilizaba la idea del Microsoft BASIC de mantener una tabla de asignación de ficheros (FAT - File Allocation Table) para gestionar el espacio de datos en disco, algo que siguió vigente durante muchos años en los contemporáneos Windows, hasta que fue abolido por fin en el nuevo Windows NT.

los plagios

Pero cómo sería de malo este 86-DOS que hasta los colegas de Tim en su empresa se mofaban del producto apodándolo QDOS (Quick and Dirty Operating System - Sistema Operativo Rápido y Sucio). Por eso, cuando Bill Gates llegó por allí y les ofreció 15.000 dólares (unos 12.000€ de aquella época) para poder comercializar el producto, Paterson pensó que había hecho el negocio de su vida. Y más atónito se quedó poco después, cuando Gates volvió a Seattle Computer Products para pagar una suma bastante mayor por la totalidad de los derechos del 86-DOS: No

la apuesta

el éxito necesitó mucho tiempo para darse cuenta del filón que había encontrado. Microsoft disponía en 1981 de 85 empleados, mientras que en el año 2000 esta cifra había subido por encima de los 36.000. ¿Qué ha sido entretanto de Seattle Computer Products? Francamente: Lo desconocemos.

el maquillaje Con el tiempo, el MS-DOS, estable y útil tan sólo a partir de su versión 2.1 después de que Microsoft tapara sus muchas vías de agua, se mantuvo vigente hasta 1995, fecha en la que cedió el relevo a Windows 95. Y puesto que éste es una interfaz gráfica de usuario construida sobre los pilares de la última versión de MS-DOS, no deben extrañarnos los problemas de estabilidad que ha padecido. El y sus hermanos mayores, que purgaron los pecados de su creador.

### 5.3 ▶ La fábula de IBM y su PC

la excepción Conocemos muchos computadores con excelentes arquitecturas que han terminado siendo un fracaso por no venir convenientemente arropados en su capa software. El caso que nos ocupa es una insólita excepción. El nacimiento del IBM-PC vino acompañado de un sistema operativo como DOS y contadas aplicaciones de usuario de igual fragilidad.

el progreso Los computadores de Apple disponían entonces de una mejor infraestructura software, pero la situación cambió de forma elocuente, y muy pronto el PC pudo contar con procesadores de texto, bases de datos, hojas de cálculo, compiladores de C, y hasta juegos de sobremesa que funcionaban magníficamente sobre una máquina cuyo hardware además progresaba más que ningún otro.

PC-AT  
pág. 39

Con el nuevo microprocesador 80286 de Intel, IBM lanzó el PC-AT (ver [foto 2.1.b](#)), dotado con direccionamiento a 16 Mbytes de memoria principal, disquette de 1.2 Mbytes, disco duro de 20 Mbytes, y una tarjeta gráfica EGA de prestaciones muy superiores a la CGA original.

la competencia El florecimiento del PC no escapó a las grandes corporaciones del sector, y mientras la llegada del 80386 cogió a IBM por sorpresa, otras firmas anduvieron prestas a aprovechar el negocio: En Estados Unidos, podemos citar a Compaq como marca que comercializó PC basados en el 80386 con notable éxito, y en España, a la firma Inestrónica. Pero lo que más daño hizo a IBM fue la legión de clónicos que enseguida sembraron el mercado de PC más baratos con similares prestaciones.

p. 205/Vol. 2

IBM se dió cuenta entonces de que el mercado del PC apuntaba a una revolución que se le escapaba de las manos. En un intento por recuperar el control de la situación, introdujo la arquitectura Micro-Channel (MCA - ver [sección 15.1.5](#)), con un bus patentado muy diferente del ISA establecido en el PC-AT. El bus era una maravilla tecnológica, pero sus tarjetas tenían que diseñarse bajo fuertes limitaciones de espacio que encarecieron su precio. Para colmo, IBM hizo públicas las especificaciones MCA, exigiendo un canon por su uso, en claro contraste con la filosofía que había engrandecido al PC en sus inicios.

el desenlace La industria pensó entonces que el ISA era suficiente para dar cobertura a los dispositivos del PC, algo que se ha mantenido como una apuesta certera en los periféricos menos exigentes hasta los albores del año 2000. Tras el fiasco del MCA, IBM se limitó a verlas venir, convirtiéndose en espectador de excepción de un circo del que un día fue su fundador.

#### SECCIÓN 1.6

## El marco de las generaciones de computadores

Tras el nacimiento del microprocesador y los primeros PC, y antes de adentrarnos en sus diferentes generaciones, haremos un inciso para conocer algo más sobre los computadores que le precedieron, ya que fue en este contexto donde primeramente se acuñó el concepto generacional.

criterios Para diferenciar las generaciones de computadores se aplicaban dos criterios básicos: La tec-

Aspecto	Generación del computador			
	Primera	Segunda	Tercera	Cuarta
Período	1950-1959	1960-1968	1969-1977	1978-hoy
Innovación que la origina	Tubos de vacío	Transistor	Circuito integrado. Escala MSI	Microprocesador. Escala de integración VLSI
Modelo comercial de referencia	UNIVAC (1951)	IBM 360 (1964)	Cray-1 (1976)	Pentium Pro 200 MHz (1996)
Memoria	48 Kbytes	64 Kbytes	32 Mbytes	16 Mbytes
Potencia	124 KW.	10 KW	60 KW	0.3 KW
Tamaño	28 m <sup>3</sup>	1.68 m <sup>3</sup>	1.62 m <sup>3</sup>	0.05 m <sup>3</sup>
Coste en € (*)	5.500.000	4.570.000	9.340.000	4.850
Rendimiento	1	263 × UNIVAC	332 × IBM360	2.5 × Cray-1

TABLA 1.4: Las cuatro generaciones de computadores: Marco temporal en el que transcurren, innovación que las origina, modelo comercial más representativo y coste (\* = trasladando su valor en el momento del lanzamiento al mercado a euros del año 2001 según los coeficientes de inflación anual de la economía norteamericana y asumiendo la paridad entre el dólar y el euro). Se incluyen además unas breves reseñas acerca de su configuración, como el tamaño de su memoria principal, la potencia disipada, el espacio físico que ocupan y el rendimiento que proporcionan en relación a su predecesor. Los números delatan que un equipo basado en el Pentium Pro es 1134 veces más barato y 220000 veces más rápido que el UNIVAC de los años 50.

nología utilizada en la fabricación de sus componentes y los métodos utilizados para su manejo y programación. Las cuatro generaciones de computadores vienen así marcadas por la aparición de innovaciones revolucionarias a razón de una por década, y que se encadenan de forma sucesiva durante los años 40, 50, 60 y 70, tal y como hemos resumido en la [tabla 1.4](#).

pág. 29 ➔

Ninguna otra innovación de importancia relevante a las cuatro mostradas en esa tabla ha ocurrido desde entonces, por lo que siguiendo ese criterio aún continuamos inmersos en la cuarta generación o “generación del microprocesador”. No obstante, observando esta evolución histórica, no resulta difícil imaginar que, cuando uno se sitúa en los años 80, se encuentre claramente sugestionado por los acontecimientos y tenga fundadas expectativas para pensar que la década que le toca vivir va a regalarle igualmente una innovación que marcará un punto y aparte en la evolución de la informática. Ocurre de esta manera que se busca con una cierta vehemencia ese nuevo invento que convulsione el mercado, lo que provoca la aparición de falsas alarmas.

la generación que nunca llegó

En la primera mitad de la década de los 80, se decía que la quinta generación vendría marcada por los computadores pensantes, esto es, dotados de una estructura que más que realizar cálculos a gran velocidad permitiese al computador *razonar* acerca de lo que procesaba. La ventaja que nos da observar aquellos vaticinios veinte años después nos permite esbozar una irónica sonrisa: la realidad dictamina que un robot sigue siendo algo torpe para percibir los objetos que tiene delante, que los sistemas expertos no gozan del uso masivo que de ellos se esperaba, y que hay que invertir bastantes millones de euros para conseguir desarrollar un computador capaz de tumbar a Garry Kasparov jugando al ajedrez.

computadores pensantes

Semejantes descalabros permitieron mitigar esa euforia desmedida que los visionarios de la tecnología arrastraban tras cuatro décadas de éxito ininterrumpido, lo que permitió entrar en la década de los años 90 con actitud menos pretenciosa: Los investigadores ya se habían dado cuenta de que era necesario contemplar la posibilidad de que en esa década tampoco tuvieran cabida inventos revolucionarios, como así terminaría sucediendo.

cautela

Los últimos 30 años pertenecen por tanto a la generación del procesador, y ahí continuaremos inmersos hasta que la tecnología diga lo contrario. Muchos pueden juzgar este hecho como un

la generación del procesador

fracaso de la capacidad de innovación de la humanidad, pero nuestro talante es bastante más positivo: La ausencia de grandes innovaciones ha permitido a la tecnología del microprocesador madurar y mostrarnos todo su potencial. Y si después de treinta años seguimos usando microprocesadores y nadie reniega de ellos, hay que empezar a pensar que a lo mejor no nos hace falta una nueva revolución, y que la tecnología existente puede ser suficiente para cubrir las necesidades de la sociedad a la que sirve.

maduración

suficiencia

Más aún, demos por supuesto que pronto aparece ese invento que pone de nuevo todo patas arriba. Cuando menos, ahora se enfrenta a un escenario bastante más adverso que las otras generaciones, al tener que vencer la resistencia de un mercado ya consolidado. Porque cuando las válvulas de vacío tienen un uso mínimamente extendido, ocupan un espacio enorme y se calientan de lo lindo, es fácil dilapidarlas con la llegada del transistor. Pero cuando el microprocesador está tan enquistado en nuestra sociedad y además no demuestra fisura alguna, provocar su desaparición en un corto espacio de tiempo parece hartamente complicado; al menos, mientras siga inmerso en su meteórica evolución y continúe dando respuesta al incremento del rendimiento sobre el que se ha venido sustentando toda la industria de la informática.

enquistación

pág. 35

El capítulo 2 realiza un recorrido histórico por la evolución del microprocesador, mostrándonos cómo ha madurado durante estos treinta años en los que ha abanderado el progreso tecnológico. Después de eso nos espera toda una singladura por sus generaciones más recientes.



## Resumen



Aunque resulta imposible resumir la historia del computador personal o PC sin ser injusto, vamos a intentarlo.

En relación a los inventos que lo han hecho posible, nos quedaríamos con el transistor, el circuito integrado y el microprocesador. La siguiente tabla sintetiza la ubicación temporal y los protagonistas asociados a cada uno de ellos.

Año	Invencción	Protagonistas
1947	El transistor	Brattain y Bardeen
1958	El circuito integrado	Jack Kilby y Robert Noyce
1971	El microprocesador	Intel Corporation

Respecto a los productos que mejor contribuyeron a la concepción del computador personal, nos hemos quedado también con tres, que hemos resumido en la siguiente tabla:

Año	Antecedente del computador personal	Empresa fabricante
1973	El Alto	PARC
1977	El Apple II	Apple
1981	El PC	IBM

Las tres décadas subsiguientes (los ochenta, los noventa, y la presente), son bastante más tranquilas, porque a pesar de los extraordinarios logros conseguidos, muchos de ellos son continuación de tendencia de conceptos que tienen su origen en la historia que acabamos de narrar. Esperemos que muy pronto, con la saturación de los avances en la integración del transistor de silicio, la comunidad científica viva un nuevo Renacimiento y podamos presenciar algún invento que revolucione el progreso de la Humanidad como lo hizo hasta la fecha nuestro adorado chip.





## La anécdota: Intel, cadena hotelera



Pintoresca cuando menos resulta la historia de cómo Gordon Moore y Robert Noyce eligieron el nombre para su compañía Intel.

El creador de algo siempre tiene querencia a considerarlo parte de su ser, y así suele inclinarse inicialmente por un nombre que derive del suyo propio. El caso que ahora nos ocupa no fue una excepción, derivando la primera intentona en *Moore-Noyce*, curiosa composición del apellido de la pareja pionera que no cristalizó por su sorprendente semejanza fonética con *more noise*, que en inglés significa *más ruido*, un término de connotaciones particularmente desaconsejables en el entorno de la microelectrónica en el que pretendía abrirse camino la empresa.

Descartada esa elección, y centrados ya en reminiscencias más de tipo tecnológico, desarrollaron la lista de 13 nombres siguientes (recordar que tanto Moore como Noyce residían en Mountain View, parte norte del estado norteamericano de California):

CALCOMP	CALifornia COMPuter
CALECOM	CALifornia Electronic COMPuter
CALEX	CALifornia ElectroniCS (*)
COMPTEK	COMPuter TECHnology (**)
COMPUTEK	COMPuter TECHnology (**)
DIGICOM (***)	DIGItal COMPuter
DISTEK	Digital Solid State TECHnology (**)
ECTEK	Electronic Computer TECHnology (**)
ELCAL	ELEctronics of CALifornia
ELCOM	ELEctronic COMPuter
ESSCOTEK	Electronic Solid State Computer TECHnology (**)
INTEL	<b>IN</b> Tegrated <b>E</b> LEctronics
TRONICOM	ElecTRONic COMPuter

(\*) La CS se pronuncia en inglés como una X.

(\*\*) La CH se pronuncia en inglés como una K.

(\*\*\*) El nombre de DIGICOM sería utilizado veinticinco años más tarde por Michael Crichton en una afamada novela cuya trama se desarrollaba en el entorno de una empresa dedicada a la fabricación de chips para computadores personales. La historia sirvió además de base para el guión cinematográfico de la película "Acoso" (Disclosure), dirigida por Barry Levinson y protagonizada por Michael Douglas y Demi Moore (por cierto, que la Moore no es hija ni descendiente directo del co-fundador de Intel: son, una vez más, caprichos del azar).

Elegido el nombre de Intel, Moore y Noyce procedieron a registrarlo, encontrándose con la sorpresa de que ya lo estaba: Una cadena de moteles de pequeña escala lo llevaba utilizando desde hacía un tiempo (hipótesis personal: reminiscencias sonoras con *Inn-Bed*, traducible por algo así como "cama y posada"). El caso es que nuestra pareja protagonista se aferró a ese nombre hasta el punto de buscar al propietario de la cadena hotelera y hacerle una jugosa oferta por sus derechos de explotación para poder utilizarlo comercialmente de forma legal a partir de entonces.

Y fue así como este acrónimo entró a formar parte de nuestras vidas. Con la perspectiva que dan cuarenta años de historia, si uno repasa los candidatos de la lista anterior, encuentra el nombre de Intel como uno de los más acertados: Estéticamente, a mí sólo me parecen más bonitos los que llevan a California como parte del mismo, pero entonces habría que explicar ahora cómo un nombre de genuino sello local como "Electronics of California" (por ejemplo) es una de las empresas más universales del panorama empresarial mundial. En fin, que Moore y Noyce tuvieron intuición hasta para esto...

## 📄 Cuestionario de evaluación 📄

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

**1** 🗣️ ¿Qué investigador jugó un papel relevante en los dos puntales clave del desarrollo tecnológico que culminaron con el nacimiento del PC en Agosto de 1981?

- a** Jack Kilby: El transistor y el circuito integrado.
- b** Robert Noyce: El circuito integrado y el microprocesador.
- c** Gordon Moore: El microprocesador y las sucesivas escalas de integración.
- d** Federico Faggin: El 4004 y el Z80.

**2** 🗣️ ¿Quién está considerado desde los años 70 como el precursor de los entornos de ventanas e iconos en el entorno PC tal y como se conocen en Windows?

- a** Alan Kay, director científico del Palo Alto Research Center.
- b** Bill Gates, fundador de Microsoft.
- c** Gary Kildall, fundador de Digital Research.
- d** Steve Wozniak, fundador de Apple Computer.

**3** 🗣️ ¿Cuál es el primer computador de propósito general de la historia?

- a** El IBM PC.
- b** El Apple Macintosh.
- c** El Busicom.
- d** El ENIAC.

**4** 🗣️ ¿Qué dispositivo desencadenó la aparición del microprocesador?

- a** Un mainframe.
- b** Una calculadora.
- c** Una memoria estática.
- d** Un lector de tarjetas perforadas.

**5** 🗣️ ¿En qué marco temporal se suceden las distintas generaciones de computadores?

- a** En la década de los 40 la primera, en los 50 la segunda, en los 60 la tercera, en los 70 la cuarta, en los 80 la quinta, en los 90 la sexta y en la década actual la séptima.
- b** Lo anterior sólo es correcto hasta la cuarta generación.
- c** Lo anterior sólo es correcto hasta la cuarta generación, y además, en las dos décadas siguientes se suceden seis generaciones de microprocesadores.
- d** Todo es incorrecto, ya que nuestro punto de partida, el primer PC, aparece en 1980.

**6** 🗣️ ¿Qué generación de computadores se origina con la llegada del circuito integrado o chip?

- a** La segunda, ya que aparece en esa década.
- b** La tercera, ya que no se aplica de forma extendida hasta la década siguiente.
- c** La cuarta, ya que su desarrollo no tiene lugar en la industria informática hasta la década de los 70 en que entra en escena el microprocesador.
- d** Todas a partir de la cuarta, lo que ha provocado la existencia de subgeneraciones.

# PRIMERA PARTE

# MICROPROCESADORES



# El legado generacional

## Sumario

<b>2.1. La década de los 80: Intel frente a Motorola</b> . . . . .	<b>37</b>
2.1.1. Primera generación (1978-1982) . . . . .	38
2.1.2. Segunda generación (1982-1985) . . . . .	38
2.1.3. Tercera generación (1985-1989) . . . . .	39
<b>2.2. La década de los 90: Omnipresente Intel</b> . . . . .	<b>40</b>
2.2.1. Cuarta generación (1989-1993) . . . . .	41
2.2.1.1. Despedimos al coprocesador matemático . . . . .	41
2.2.1.2. Y recibimos a la caché . . . . .	41
2.2.2. Quinta generación (1993-1997) . . . . .	41
2.2.3. Sexta generación (1997-2000) . . . . .	42
<b>Resumen</b> . . . . .	<b>43</b>
<b>La anécdota: IBM, ese celestino entre el chip y el procesador</b> . . . . .	<b>45</b>
<b>Cuestionario de evaluación</b> . . . . .	<b>46</b>

El hecho de que las generaciones de computadores acaben en la cuarta con la llegada del microprocesador en absoluto indica que la tecnología no haya evolucionado desde entonces. Un 4004 de los años 70 tiene muy poco que ver con un Pentium 4 de los tiempos actuales, y hasta llegar a él se han sucedido multitud de pasos intermedios a una velocidad de vértigo (ver [tabla 2.1](#)). Esta evolución no ha sido uniforme; está jalonada de hitos en los que se entremezclan éxitos rotundos con otros más discretos, e incluso algún que otro resbalón de cierta enjundia.

Con objeto de poder discriminar los grandes cambios que se han producido en la arquitectura de un microprocesador frente a aquellos que sólo suponen un refinamiento y/o perfeccionamiento de modelos ya existentes, los fabricantes de hardware han querido seguir utilizando el concepto de generación, sólo que esta vez lo han aplicado al microprocesador, dándonos la oportunidad de clasificar en generaciones de microprocesadores los sucesivos modelos que han ido desfilando en los veinte años de existencia del PC. Dado que Intel es la compañía de referencia durante todos estos años, son sus modelos quienes llevan la manija del cambio generacional.

Ahora bien, la historia de todas estas generaciones acontece en el seno de un marco empresarial que tiene una notable incidencia en el desarrollo de los diferentes productos, siendo en no pocas ocasiones factores exógenos al puramente científico los encargados de encumbrar o hundir un determinado modelo de microprocesador y/o sus prestaciones como promotor del cambio generacional. Este trasfondo mercantil tiene dos partes bien diferenciadas que pueden hacerse coincidir en el tiempo, de forma aproximada, con la década de los 80 y la década de los 90.

Microprocesador	Fecha de lanzamiento	Frecuencia en MHz	Ancho bus	T. l. Micras	Número de transistores	Rendimiento aproximado
4004	15/11/71	Entre 0.1 y 0.75	4	12	2.300	
8008	1/4/72	Entre 0.1 y 0.75	8	10	3.500	2 x 4004
8080	1/4/74	2	8	6	6.000	10 x 8008
8086	8/6/78	4,77 8 10	16	3	29.000	10 x 8080 (0.8 MIPS a 8 MHz)
8088	1/6/79	4.77 8	8	3	29.000	8 x 8080
80286	1/2/82	8 10 12	16	1.5	134.000	4 x 8086 (2.7 MIPS a 12 MHz)
80386DX	17/10/85	16, 20, 25, 33	32	1	275.000	2.5 x 80286 (6.0 MIPS a 20 MHz)
80386SX	16/6/88	16, 20	16	1	275.000	2 x 80286
80486DX	10/4/89	25, 33, 50	32	1 0.8	1.2M	4 x 386DX (20 MIPS a 25 MHz)
80486SX	22/4/91	16, 20, 25, 33	32	0.8	1.2M	3 x 386SX
Pentium	22/3/93	60-200	64	0.8 0.6 0.35	3.1M	5 x 486DX (100 MIPS a 60 MHz)
Pentium Pro	27/3/95	133-200	64	0.6 0.5 0.35	5.5 M	2.5 x Pent (440 MIPS a 200 MHz)
Pentium II	7/5/97	233-450	64	0.35 0.25	7.5 M	4 x Pent (466 MIPS a 266 MHz)
Pentium III	28/2/99	450-1000	64	0.25 0.18 0.13	8.2 M (28.1 M)	2 x P.II (1000 MIPS a 500 MHz)
Pentium 4	20/11/00	1300-4000	64	0.18 0.13	(42 M) (55 M)	3 x P.III
Estimac. Intel	2011	10000	?	0.07	1000M	?

**TABLA 2.1:** Evolución de los microprocesadores de Intel en los últimos 30 años. En la columna del número de transistores, los datos entre paréntesis incluyen a la caché L2 integrada. En la columna del rendimiento, el factor multiplicador es una estimación del autor basada en la frecuencia inicial de cada modelo, mientras que los datos entre paréntesis proceden de Intel, siendo Dhrystone MIPS para el modelo a la frecuencia que se indica. Aunque los MIPS apenas se utilizan ya, son la única medida de rendimiento que abarca el extenso marco temporal considerado.

## La década de los 80: Intel frente a Motorola

Durante la década de los 80, las dos compañías fabricantes de microprocesadores que mayor atención prestaron al sector de la informática doméstica fueron Intel y Motorola, protagonizando una dura rivalidad por la hegemonía de este mercado.

El abanderado de Intel durante esta década fue la familia 80x86, de la que se sucedieron hermanos mayores del 8086: El 80286, el 80386 y el 80486. Motorola, con una estrategia similar, fue replicando a las novedades de Intel con modelos de su familia 68000: El 68020, el 68030 y el 68040.

familia 80x86

familia 68000

Cada compañía tenía su propia idiosincrasia y acumulaba su legión de adeptos, pero al margen de valoraciones subjetivas, creemos que pueden repartirse halagos para ambas partes.

De Intel alabaríamos dos cosas:

INTEL:

- ❶ Supo encontrar los aliados empresariales más oportunos en cada momento.
- ❷ Supo fidelizar a sus clientes con nuevos productos sin renunciar a la compatibilidad con modelos pasados. Estas ventajas se funden al nivel de la capa software del equipo dando extraordinarias sinergias: El usuario no sólo tiene la garantía de que los programas más conocidos del mercado siempre tendrán la versión que le permitirá disfrutarlos en su equipo, sino que además, el día que quiera adquirir un PC más potente, podrá seguir utilizando toda la infraestructura software de que disponía en su viejo equipo.

- aliados

- fidelización

De Motorola también vamos a destacar dos cosas:

MOTOROLA:

- ❶ Su diseño arquitectural, siempre completo y vanguardista aunque para ello tuviese que sacrificar algunos MHz de frecuencia respecto a los modelos de referencia de la competencia.
- ❷ La elegancia de su lenguaje ensamblador, hecho que se desprende del anterior y que contrasta con el de su rival, el ensamblador del 80x86, en el que uno se preguntaba más a menudo de lo deseable cómo podían ser las cosas tan complicadas. La razón hay que buscarla precisamente en la política de Intel de hacer diseños compatibles con arquitecturas anteriores: No hay nada malo en tratar de amortizar un diseño reutilizándolo en generaciones sucesivas, pero para eso el punto de partida debe estar a la altura. Y ya hemos visto que el origen de la familia 80x86 de Intel es el 4004 y el 8008, dos arquitecturas inmaduras y precipitadas que han hipotecado la sencillez de programación de siete generaciones de microprocesadores. Afortunadamente para Intel, los sistemas operativos y los lenguajes de alto nivel estaban al quite y desmotivaron lo suficiente el uso del ensamblador como para que pronto entrara en claro desuso y no afectase seriamente el liderazgo que ya ostentaba la compañía.

- arquitectura

- ensamblador

Y así es como llegamos a la conclusión de que aquella batalla que se libró en los años 80 por el mercado de microprocesadores para PC tuvo un ganador en cada parcela: Intel, en ventas, y Motorola, en el diseño del chip y su ensamblador asociado. La [tabla 2.2](#) compara los diferentes modelos de cada fabricante durante esta década y arroja los números suficientes para que cada cual saque conclusiones propias.

veredicto

 [pág. 38](#)

Precisamente por estar los modelos de Intel más extendidos y por haber tenido una mayor continuidad en el tiempo, pensamos que son más adecuados para una perspectiva histórica como la que aquí tratamos de ilustrar, siendo ésa la principal razón para haberlos escogido como referencia en la breve semblanza que a continuación exponemos para las tres generaciones de microprocesadores que acontecen en el espacio temporal de la década que nos ocupa.

Procesador (Año)	Frecuencia (MHz)	Transistores / Dist. integr.	Cachés instr./datos	Buses datos/dirs.	Segmentado / FPU (segm.)
8086 (1978)	4.77, 8, 10	29.000 3 micras	No No	16 16	No No
68000 (1979)	8, 10, 12, 16, 20	68.000	No No	16 24	No No
8088 (1979)	4.77, 8	29.000 3 micras	No No	8 16	No No
68008 (1982)	8, 10, 12	- -	No No	8 20	No No
80286 (1982)	8, 10, 12	134.000 1.5 micras	No No	16 24	No No
68020 (1984)	12, 16 20, 25, 33	103.000 1.5 micras	256 bytes No	8,16,32 32	3 etapas No
80386 (1985)	16, 20, 25, 33	275.000 1 micra	No No	32 32	4 etapas No
68030 (1987)	16, 20, 25, 33, 40, 50	270.000 1.2 micras	256 bytes 256 bytes	32 32	3 etapas No
80486 (1989)	16, 20, 25 33, 50	1.185.000 1-0.8 micras	8 Kbytes Unificada	32 32	5 etapas Sí (5 etapas)
68040 (1989)	25, 33, 40, 50, 66	1.200.000 0.8 micras	4 Kbytes 4 Kbytes	32 32	6 etapas Sí (6 etapas)
Pentium (1993)	60-200	3.100.000 0.8-0.35 micras	8 Kbytes 8 Kbytes	64 32	5 etapas Sí (8 etapas)
68060 (1994)	50, 66, 75	2.400.000 0.5 micras	8 Kbytes 8 Kbytes	32 32	6 Sí (8 etapas)

TABLA 2.2: Comparativa de las familias 80x86 de Intel y 680x0 de Motorola modelo a modelo.

## 1.1 ▶ Primera generación (1978-1982)

8086 El pistoletazo de salida en las generaciones de microprocesadores para PC viene marcado de forma natural por su primer representante: El 8086 de Intel. Esta primera generación se encuentra caracterizada por la potencia bruta de cálculo: Las mejoras en la integración de circuitos permiten que el diseñador no encuentre las limitaciones tan fuertes que condicionaron el diseño del 4004 y el 8008, lo que posibilita la adopción de un ancho de 16 bits en las unidades de procesamiento y una complejidad inusual en las ALU, que ahora habilitan instrucciones máquina para efectuar el producto y la división de números enteros.

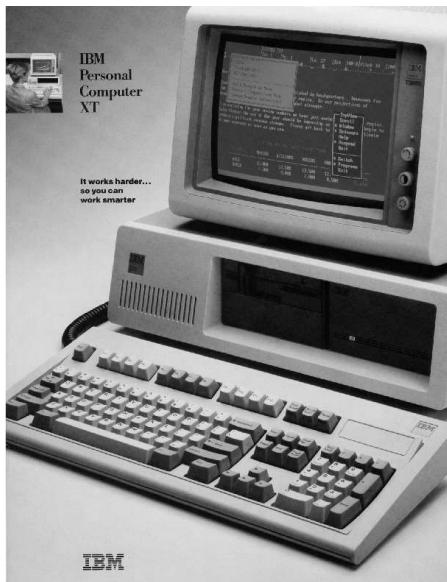
68000 El 68000, lanzado al mercado por Motorola a finales de 1979, era un diseño microprogramado de 16 bits (aunque con registros internos de 32 bits) que rompió con la tendencia de Unidad de Control cableada que imperaba en aquella época. Su éxito comercial fue indudable, siendo adoptado no sólo por el Macintosh de Apple, sino también por las estaciones de trabajo de Sun y Hewlett-Packard hasta que éstos decidieron crear sus propios procesadores RISC.

## 1.2 ▶ Segunda generación (1982-1985)

bus de direcciones

En la segunda generación, el protagonismo se traslada a la memoria. La capa software comienza a madurar y demanda un espacio de direcciones mayor. La respuesta del hardware es, por un





(a) PC-XT



(b) PC-AT

Fotos cortesía de IBM

FOTO 2.1: El PC en sus dos primeras generaciones. (a) El PC-XT, lanzado al mercado el 8 de Marzo de 1983 con un disco duro de 10 Mbytes, 256 Kbytes de memoria ampliable a 640 Kbytes, 8 zócalos de expansión y DOS 2.0 que incluye la gestión del disco duro. (d) El PC-AT (1984), con el nuevo procesador 80286 de 6 MHz capaz de direccionar hasta 16 Mbytes de memoria principal y buses ISA de 16 bits para periféricos. Viene equipado con un disco duro de 20 Mbytes y DOS 3.0, que a su capacidad multiusuario y multitarea agrega la gestión de discos flexibles de alta densidad (1.2 Mbytes).

lado, ensanchar el bus de direcciones rebasando ya la frontera del Megabyte direccionable, y por el otro, habilitar el concepto de memoria virtual que permite al procesador ejecutar programas mayores que el espacio físico de memoria disponible, siempre que las líneas del bus de direcciones permitan direccionarlo. Así, al final de esta generación se alcanza ya la cota del Gigabyte direccionable, tanto en los modelos de Intel como en los de Motorola.

memoria virtual

### Tercera generación (1985-1989)

### ◀ 1.3

La tercera generación es la continuación natural de las mejoras introducidas en la segunda. La memoria virtual supone tener que realizar una traducción de dirección virtual a física por cada acceso a memoria realizado, tarea de la que se encarga la MMU (Memory Management Unit).

MMU interna

Esto es algo ciertamente costoso en un contexto en el que el espacio de direcciones comienza a ser grande y aún no existen memorias cachés internas al microprocesador que alivien el caudal de peticiones a memoria principal. Por todo ello, se decide optimizar la implementación de estas traducciones introduciendo en el interior del microprocesador la circuitería capaz de llevarlas a cabo. Esto ahorra todo el tiempo necesario para dialogar con la MMU (Memory Management Unit) externa, al tiempo que alivia el tráfico que soporta el bus local del microprocesador.

Adicionalmente, Intel retoma el diseño dual que ya había utilizado con el tándem 8086/8088, desarrollándose una versión recortada de su procesador insignia 80386 para el mercado de bajo coste, a la que se coloca el sufijo SX. Esta idea se consagrará definitivamente en la sexta generación con los modelos Celeron que ahora conocemos, extendiéndose incluso a una tercera gama, más alta, con los modelos Xeon. En esta estrategia, Intel cabalgará en solitario hasta la séptima generación, donde los modelos de AMD Duron en la gama baja y con sufijo MP en la gama alta copian su proceder. Un ejemplo más de la vista comercial que siempre ha caracterizado a Intel.

gama ampliada

Generación de microprocesadores	Modelo más representativo	Variante de bajo coste y rendimiento	Otros modelos de esa misma generación
Primera (1978-82)	Intel 8086	Intel 8088 (8 bits datos)	Motorola 68000
Segunda (1982-85)	Intel 80286	No hubo	Motorola 68010 AMD 80286
Tercera (1985-89)	Intel 80386(DX)	Intel 80386SX (16 bits datos)	Motorola 68020 AMD 80386 Cyrix 80386
Cuarta (1989-93)	Intel 80486(DX)	Intel 80486SX (sin FPU)	Motorola 68030 AMD 80486 Cyrix 80486
Quinta (1993-97)	Intel Pentium	Intel Pentium Overdrive	Pentium MMX AMD K5 Cyrix M1
Sexta (1997-00)	Intel Pentium II	Intel Celeron (sin caché L2)	Pentium Pro y III AMD K6, 2 y III Cyrix M2 y MXi

TABLA 2.3: Los modelos comerciales más representativos de las seis primeras generaciones de microprocesadores.

## SECCIÓN 2.2

### La década de los 90: Omnipresente Intel

Power PC

Motorola se desmarca

En los años 90 la situación cambia ostensiblemente. Motorola se lanza al diseño del Power PC en alianza con IBM y Apple, y lo que en un primer momento se ve como una vuelta de tuerca más en su guerra contra Intel por la hegemonía del mercado de microprocesadores para PC, supone en realidad una completa redefinición de las estrategias de mercado de la compañía.

servidores...

❶ Motorola apunta un poco más arriba de la informática doméstica, allá donde se encuentran ubicados los servidores y estaciones de trabajo, por lo que renuncia voluntariamente al protagonismo en el entorno PC.

... y otras muchas cosas

❷ La compañía desarrolla una marcada política de diversificación, comercializando multitud de productos en el ámbito de las comunicaciones y la telefonía móvil, hecho por el que hoy es más conocida si cabe.

El desmarque de Motorola supone un pequeño cataclismo, al quedar Intel como líder hegemónico del mercado. Sucede así que las tres generaciones de esta década no sólo destilan personalidad propia a nivel arquitectural, sino también a nivel empresarial.

Intel-Motorola

Intel sola  
Intel-AMD

Recopilando: En la cuarta, se vive el último episodio de la rivalidad Intel-Motorola. En la quinta, Intel se queda sola y arrasa con su Pentium, que se consolidará rápidamente como la arquitectura estándar de referencia en lo concerniente a microprocesadores domésticos de 32 bits. Y en la sexta, asistimos al primer capítulo de la rivalidad Intel-AMD.

## Cuarta generación (1989-1993)

### ◀ 2.1

La potencia de cálculo vuelve a centrar la atención de los diseñadores de microprocesadores, y si en la primera generación es la computación de números enteros la protagonista, ahora lo será la computación de números reales o de punto flotante.

cálculo

### 2.1.1 Despedimos al coprocesador matemático

El microprocesador comienza a mostrar una cierta dependencia de su coprocesador matemático o FPU (Floating-Point Unit), la unidad funcional que realiza estas operaciones. Se decide así introducir ésta dentro del procesador para ganar velocidad y reducir el precio del conjunto. Desde este momento, el concepto de microprocesador cambia y presupone ya la existencia de esta unidad funcional como un elemento más de su ser.

promociona  
la FPU

Los dos exponentes más importantes de esta generación son el 80486 y el 68040. Lanzados casi al unísono en 1989, disponen de FPU propia, que segmentan además en cinco y seis etapas. Este diseño se quedará muy corto conforme se vayan incorporando operaciones de mayor complejidad en generaciones posteriores.

segmentación

### 2.1.2 Y recibimos a la caché

Ya hemos bajado de la micra en la distancia de integración de los transistores, y eso nos permite alguna licencia más.

Junto a la FPU, también se introducen las primeras memorias cachés en el interior del microprocesador. Desde el punto de vista del rendimiento global, la memoria ya no es un problema en espacio direccionable, pero sí lo es en lo referente a su velocidad.

cachés

Respecto a su diseño, Motorola viene algo más rodada al haber incluido porciones testimoniales en sus modelos de segunda y tercera generación, y probablemente por eso acierta más separando el flujo de instrucciones y el de datos, otorgando 4 Kbytes a cada uno de ellos.

separadas

Intel, por el contrario, se decanta por un diseño unificado de 8 Kbytes, que se muestra menos efectivo en los números de rendimiento. A partir de aquí, nunca más se volverá a ver una caché de primer nivel (L1) que no se encuentre escindida en las dos partes mencionadas.

unificadas

## Quinta generación (1993-1997)

### ◀ 2.2

La frecuencia del microprocesador se desliga de la asociada al resto de la circuitería, provocando la aparición de un multiplicador de reloj que mide este desfase en velocidad y que pronto alcanza un valor superior a 3. Este hecho favorece también la modularidad del sistema tal y como ahora la conocemos: Cada elemento funciona a su velocidad, dispone de zócalo propio, y puede adquirirse y montarse por separado.

multiplicador  
modularidad

Las mejoras en la tecnología de integración, que físicamente se han aprovechado para subir la frecuencia, permiten también contar con un par de millones de transistores más, que se van a emplear preferentemente en dotar al microprocesador de un agresivo paralelismo a nivel de instrucción, en el que destacan una segmentación en 5 o más etapas (ver [sección 3.3.1](#) para una explicación de este concepto) y la llegada de los primeros diseños superescalares (así mismo, ver [sección 3.3.2](#)).

frecuencia  
segmentación  
▶ [pág. 59](#)  
superescalar  
▶ [pág. 60](#)  
Pentium  
▶ [pág. 134](#)

Con esas premisas va a nacer el Pentium (ver [sección 4.1.1](#)), un procesador al que la historia

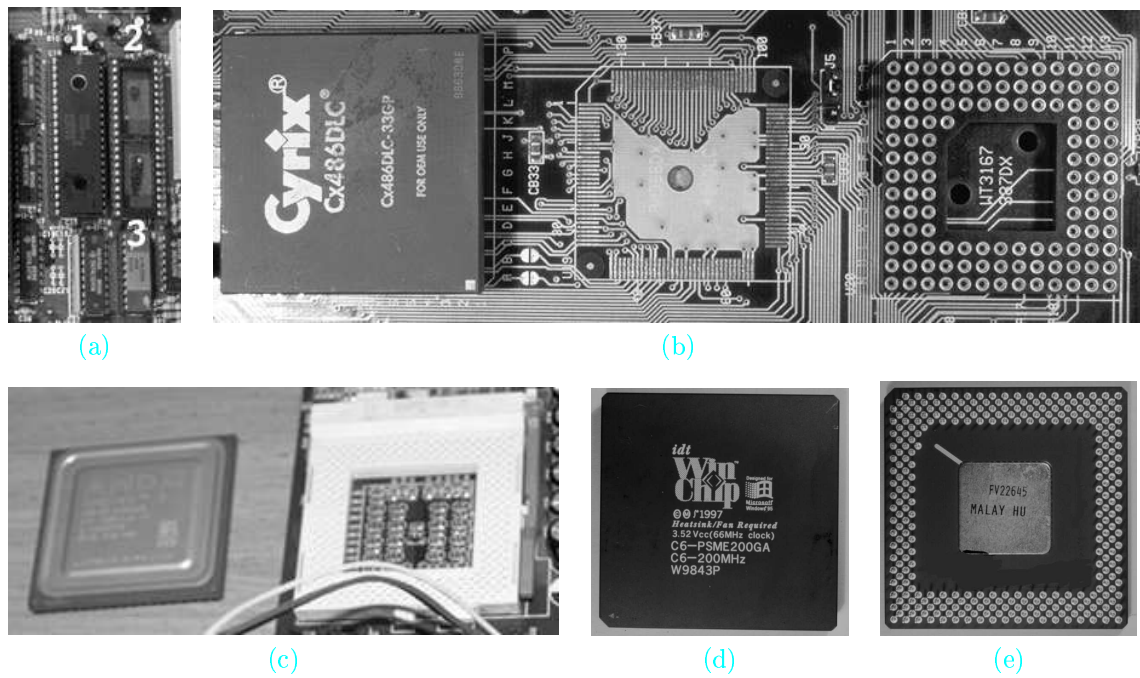


FOTO 2.2: Microprocesadores de las cuatro firmas más conocidas en sus seis primeras generaciones: (a) El mítico 8086 de Intel (1), y junto a él, el zócalo para el coprocesador matemático 8087 (2) y el controlador de bus local, 8288 (3). (b) El procesador 80486 de Cyrix (izquierda), espacio para acoplar un 80386 (centro), y zócalo para el coprocesador matemático 80387 (derecha). (c) Un K6 de AMD junto a su zócalo Super 7. (d y e) Envés y revés de un procesador Winchip C6 de IDT.

le va a hacer un inmenso favor: Tiene todo el mercado para él, y a pesar de ser una arquitectura terminal para Intel, será recordado por la firma como su éxito comercial más contundente.

K5  
pág. 144

La intención de AMD era lanzar el K5 (ver [sección 4.2.1](#)) para competir con el Pentium en esta generación, pero el retraso en su lanzamiento le obligó a asumir el rol de perdedor desde el primer momento (y a nosotros, a que casi tengamos que enviarlo a la sexta generación a tenor de la fecha en que pisó por fin el mercado).

### 2.3 ▶ Sexta generación (1997-2000)

multiprocesador

Se reduce la inventiva de los diseñadores. La única idea fresca de este período consiste en dotar al microprocesador para PC de un carácter multiprocesador, incorporándole la circuitería necesaria para sincronizarse y colaborar con otros microprocesadores.

SMP

En términos de rendimiento, esto queda un peldaño más arriba que el paralelismo interno a nivel de instrucción de la generación anterior, aunque uno más abajo que un multiprocesador puro, dado que se comparten todos los componentes de la placa base (buses, memoria principal y periféricos). Para este concepto se acuña la terminología *multiprocesamiento simétrico* ó **SMP (Symmetric MultiProcessing)**.

Pentium Pro

La idea del SMP no va a cuajar en el mercado doméstico como se esperaba. Primero, por el elevado coste que supone para el usuario medio en el que sus aplicaciones no justifican una búsqueda de rendimiento tan exagerada. Además, la capa software muestra una desmesurada inmadurez para sacar partido de las posibilidades que el SMP abre. Así que esta novedad, cuyo abanderado es el Pentium Pro de Intel, repliega velas para situarse desde entonces en mercados

más pretenciosos como las estaciones de trabajo y los servidores.

En el segmento doméstico, tras el lanzamiento del Pentium, el mercado se reordena. Aparecen nuevos competidores para Intel, surgidos al reclamo de captar la cuota de mercado dejada por Motorola, que contraatacan con nuevos diseños que simplemente dan una vuelta de tuerca más a ideas ya utilizadas en la cuarta y quinta generación.

reclamo

De esta manera, aparece una nueva caché interna, la L2, ubicada en un nivel posterior a la ya existente, y también se reutilizan las técnicas de paralelismo a nivel de instrucción: la segmentación en la FPU, que ahora consta de más de diez etapas (14 en el Pentium II, por ejemplo), y la superescalaridad, combinando esta unidad con las tradicionales ALU para alcanzar un factor de 3 en adelante.

caché L2

Intel aprovecha el diseño del Pentium Pro para derivar de él su Pentium II, montándolo sobre un nuevo zócalo (formato Slot) para tratar de disuadir a sus competidores y dar cabida a todas las novedades citadas.

Pentium II

AMD aprende la lección del K5 y lanza el K6 (ver [foto 2.2.c](#)) para competir con el Pentium II en su mismo marco temporal. Esta vez las fuerzas quedan bastante más igualadas: Intel, con una infraestructura mucho más fuerte, y AMD, mejor posicionada en los aspectos de diseño y con la variable temporal a su favor para el lanzamiento del K7 con un año y medio de antelación al Pentium 4.

K6

☛ [pág. 42](#)

Otras compañías de menor calado también trataron de plantar cara a Intel y AMD. Cyrix fue la tercera en discordia, pero a pesar de ser auspiciada por IBM, nunca demostró la agilidad necesaria en un mercado tan competitivo. Su decadencia quedó patente en el tramo final de la década, donde fue adquirida por National Semiconductor, entrando en quiebra técnica en Mayo de 1999. Unos meses más tarde, VIA, el popular fabricante de juegos de chips para placas base, reflató de nuevo la compañía poniendo en marcha su línea de productos existentes en la actualidad.

Cyrix

Completamos el ramillete de compañías menos conocidas citando a Rise e IDT, dos firmas cuyos productos quedaron restringidos al ámbito norteamericano. La [foto 2.2](#) muestra un Winchip C6 de IDT, uno de sus últimos productos dentro de la sexta generación.

Rise

IDT

☛ [pág. 42](#)

Winchip C6

En definitiva, aunque en sexta generación las alternativas a Intel distan mucho de ser el rival que tenía en Motorola, la segunda mitad de la década de los 90 marca un meteórico crecimiento de AMD, llegándose de nuevo a una bicefalia en la séptima generación, donde se encuentran Intel y AMD seguidos del resto a una clara distancia.

estatus final

En capítulos posteriores realizaremos un completo recorrido por los modelos más importantes de microprocesadores de la última década, entre los que a buen seguro se encuentra el que hay ubicado en el corazón de su PC. En concreto, tras conocer en el [capítulo 3](#) los secretos que encierra internamente un microprocesador, desmenuzaremos en el [capítulo 4](#) la quinta generación, en el [capítulo 5](#) la sexta generación, en el [capítulo 6](#) la séptima generación, y por fin, en el [capítulo 7](#) la octava generación para llegar a nuestros días. Incluso sobrepasaremos el momento presente para hacer una breve incursión en el futuro a lo largo del [capítulo 8](#).

☛ [pág. 47](#)☛ [pág. 133](#)☛ [pág. 149](#)☛ [pág. 189](#)☛ [pág. 255](#)☛ [pág. 275](#)

## Resumen



La [figura 2.1](#) desglosa las principales novedades que abanderan las seis generaciones de microprocesadores que hemos visitado, incorporándole progresivamente nuevos elementos conforme las mejoras en la tecnología de integración pone a nuestra disposición una mayor cantidad de transistores.

☛ [pág. 44](#)

las piezas

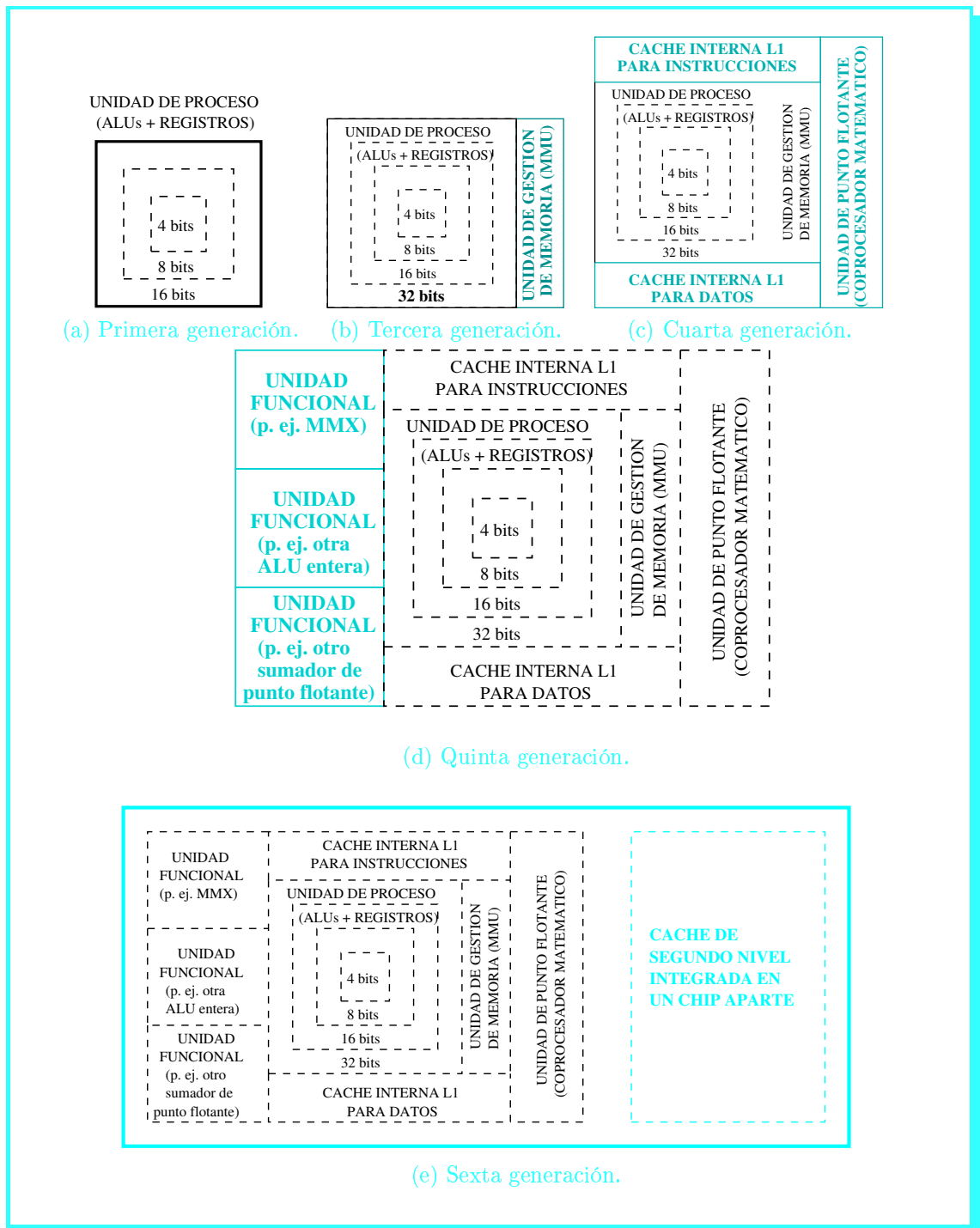


FIGURA 2.1: Diagrama de bloques de un microprocesador a lo largo de sus diferentes generaciones.

☛ pág. 45  
los hitos

La [tabla 2.4](#) resume los hitos más importantes acaecidos durante el período de vigencia de cada una de las seis generaciones que hemos repasado a lo largo del presente capítulo.

el mercado  
pág. 45 ☛

Y como ya advertimos que los avances en el procesador están jalonados de factores exógenos, hemos sintetizado en la [tabla 2.5](#) los hechos que nos han parecido más significativos durante este período.

Generación de microprocesadores para PC	Período de vigencia aproximado	Principales novedades que la caracterizan
Década de los 80		
Primera	1978-1982	Potencia bruta de cálculo: Operaciones de producto y división sobre operandos de 16 bits
Segunda	1982-1985	Gestión de la memoria: Memoria virtual, memoria protegida, mayor memoria direccionable
Tercera	1985-1989	La Unidad de gestión de memoria (MMU) se integra dentro del chip microprocesador
Década de los 90		
Cuarta	1989-1993	- El coprocesador matemático (FPU) y la caché de primer nivel (L1) se incluyen en el microprocesador
Quinta	1993-1997	- Se dispara la frecuencia de reloj del microprocesador - Primeros diseños superescalares
Sexta	1997-2000	- Paralelismo a nivel de instrucción - Gran caché L2 interna al chip - Soporte multiprocesador

TABLA 2.4: Las seis generaciones de microprocesadores para PC que se suceden en las dos últimas décadas.

Década	Hechos más relevantes	Protagonistas
1981-1990	Las familias de microprocesadores	Rivalidad Intel-Motorola
	El boom de la informática de consumo	IBM y Microsoft
1991-2000	Metamorfosis arquitectural del equipo	Rivalidad Intel-AMD
	Auge de aplicaciones software	Multimedia e Internet
2001-2010	Agotamiento de la tecnología de silicio Desarrollo de tecnologías alternativas	

TABLA 2.5: Los hechos más relevantes en las tres décadas que lleva con nosotros el PC.

La casualidad ha querido que las generaciones **pares** hayan concedido protagonismo a la memoria (memoria virtual en la segunda, caché de primer nivel en la cuarta y caché de segundo nivel en la sexta), mientras que las generaciones **impares** han centrado sus avances sobre el núcleo arquitectural más interno del procesador (aritmética, frecuencia, segmentación y superescalaridad). Esto puede servirnos como nemotécnico para recordar la serie cronológica completa.

par, memoria

impar, núcleo



### La anécdota: IBM, ese celestino entre el chip y el procesador



Ahora que conocemos la importancia del chip y el microprocesador, cuesta trabajo imaginar que el primero estuviera entretenido durante 13 años fabricando memorias y minicomputadores sin fusionar su extraordinario potencial con el segundo. Pero si nos trasladamos a la década de los 60, encontramos

hechos que lo que hacen no es sino plantearnos por qué esos trece años no fueron bastantes más.

En un informe interno de IBM fechado en septiembre de 1963 pudo leerse la siguiente sentencia: “Los circuitos integrados no constituirán una amenaza competitiva ni ahora ni en los próximos cinco años”. La amenaza era para el System/360, el minicomputador más emblemático de IBM en aquellas fechas, pionero de buena parte de los conceptos que conforman los Sistemas Operativos actuales, y construido con una tecnología de estado sólido en la que diversos componentes se depositaban en un substrato cerámico con contactos metálicos impresos: Algo así como una placa base actual donde sus múltiples chips serían condensadores y resistencias bastante más apiñados.

Menos mal que otras empresas como SDS y RCA adoptaron mucho antes que IBM los circuitos integrados para fabricar sus primeros minicomputadores (en 1965 y 1966, respectivamente), y de paso, obligaron a éste a rectificar poco más tarde en el diseño de su serie siguiente, la System/370. De no haber sido así, con la supremacía que IBM ostentaba en el mercado por aquel entonces, hubiéramos visto al chip flirtear más de la cuenta con las memorias mientras el microprocesador esperaba para conformar con él una unión que engendraría a la informática de consumo, y en su regazo, todas las generaciones de procesadores que han desfilado por el presente capítulo.

Y es que con un largo matrimonio monógamo entre el chip y la memoria, la historia de la informática sería mucho menos espectacular que la que ahora conocemos.

## 📖 Cuestionario de evaluación 📖

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

1 📖 ¿Qué procesadores de Intel para PC incluyeron por primera vez la caché L1 y L2 dentro del propio producto comercial?

- a El 80486 la L1 y el Pentium la L2.
- b El 80486 la L1 y el Pentium Pro la L2.
- c El Pentium la L1 y el Pentium Pro la L2.
- d El Pentium la L1 y el Pentium II la L2.

2 📖 Entre un procesador de primera y quinta generación, el número de transistores utilizado en el diseño se ha multiplicado de forma aproximada por un factor...

- a Diez.
- b Cien.
- c Mil.
- d Diez mil.

3 📖 ¿En qué generación se encuentra Intel más desprovisto de competencia en el diseño de procesadores?

- a En la primera.
- b En la cuarta.
- c En la quinta.
- d En la sexta.

4 📖 ¿Qué compañía de microprocesadores cambió tres veces de propietario durante el lustro 1995-2000?

- a Intel.
- b AMD.
- c Cyrix.
- d Transmeta.



# Las principales magnitudes

## Sumario

<b>3.1. Frecuencia de reloj</b>	<b>49</b>
<b>3.2. Tecnología de integración</b>	<b>50</b>
3.2.1. Evolución y significado	50
3.2.2. Efectos directos sobre otras variables	54
3.2.3. Efectos laterales entre las variables afectadas	56
3.2.4. Cómo dar empleo a un ejército de transistores	58
<b>3.3. Paralelismo a nivel de instrucción</b>	<b>58</b>
3.3.1. Segmentación (pipelining)	59
3.3.2. Superescalaridad	60
3.3.3. Combinación de segmentación y superescalaridad	62
3.3.4. Supersegmentación	63
3.3.5. Dependencias: Las enemigas del paralelismo	64
3.3.5.1. Ejecución fuera de orden	67
3.3.5.2. Predicción de salto	68
<b>3.4. Memoria caché integrada</b>	<b>69</b>
3.4.1. Breve sinopsis histórica	72
3.4.2. Jerarquía	73
3.4.3. Optimizaciones	77
3.4.3.1. Buses desacoplados	77
3.4.3.2. Caché no bloqueante	77
3.4.3.3. Caché segmentada	78
3.4.3.4. Caché con lectura anticipada	79
3.4.3.5. Caché víctima	79
3.4.3.6. Caché de tercer nivel (L3)	80
3.4.4. Proximidad al núcleo del procesador	81
3.4.5. Ubicación del controlador de caché	85
3.4.6. Velocidad	88
3.4.7. Análisis del rendimiento de caché en relación al procesador	88
3.4.8. Análisis del coste asociado a una caché	93
3.4.8.1. Caché interna	93
3.4.8.2. Caché integrada	94
<b>3.5. Conjunto de instrucciones</b>	<b>94</b>
3.5.1. CISC versus RISC	95
3.5.2. Diseño RISC	97

3.5.2.1. Selección del conjunto de instrucciones . . . . .	98
3.5.2.2. Soporte software para una arquitectura RISC . . . . .	100
3.5.3. Diseños VLIW . . . . .	104
3.5.4. Instrucciones multimedia . . . . .	106
3.5.4.1. El concepto: SIMD . . . . .	108
3.5.4.2. El embrión: MMX . . . . .	109
3.5.4.3. Criterios para la selección de instrucciones . . . . .	110
3.5.4.4. Compatibilidad . . . . .	112
3.5.4.5. Ampliaciones al conjunto MMX . . . . .	114
3.5.4.6. Otras extensiones multimedia . . . . .	118
<b>Resumen . . . . .</b>	<b>119</b>
<b>La anécdota: Magnitudes oficiales y oficiosas . . . . .</b>	<b>121</b>
<b>Cuestionario de evaluación . . . . .</b>	<b>122</b>

definición	<b>E</b> l microprocesador es el cerebro del computador y el centro neurálgico de sus actividades. Se trata de un circuito integrado o chip cuya función consiste en interpretar y ejecutar instrucciones máquina, para lo cual se divide en dos grandes unidades funcionales: La Unidad de Proceso, que ejecuta las instrucciones siguiendo una secuencia de pasos, y la Unidad de Control, encargada de interpretarlas y generar las señales eléctricas que controlan dicha secuencia. El conjunto se denomina también CPU (del inglés, <i>Central Processing Unit - Unidad Central de Proceso</i> ).
U. Proceso	
U. Control	
CPU	
conjunto de instrucciones	Las instrucciones que entiende un microprocesador conforman su repertorio o conjunto de instrucciones máquina, el alfabeto con el que se articulan todos los programas que en él se ejecutan. Estos programas se encuentran almacenados en la memoria del equipo, junto con los datos con los que operan. El microprocesador accede a la memoria a través de los buses del sistema, al igual que al resto de componentes del equipo, incluidos los periféricos. Todos estos componentes se conectan o vienen ya integrados en un soporte rígido poco mayor que las dimensiones de un folio de papel, que recibe el nombre de placa base. El sistema informático se completa con los periféricos, como los discos, las distintas tarjetas, la impresora, el ratón y el teclado.
sistema informático	
secuencia de estudio	La memoria, los buses y la placa base serán analizados, por este orden, más adelante, pero antes de eso, comenzaremos nuestro estudio con el microprocesador. En el presente capítulo, conoceremos las principales variables que inciden en su funcionamiento y rendimiento, y cómo éstas se encuentran ligadas entre sí. Los capítulos posteriores nos descubrirán las características de todos sus modelos, clasificados por generaciones.
cinco magnitudes	El microprocesador es un sistema extremadamente complejo, por lo que una de las primeras cosas que debemos aprender en su estudio es acotar los parámetros que realmente influyen en su rendimiento y saberlos distinguir de aquellos que son meros elementos decorativos y/o reclamos publicitarios. Sus cinco magnitudes más importantes son las siguientes: <ul style="list-style-type: none"> <li>❶ Frecuencia de reloj.</li> <li>❷ Tecnología de integración.</li> <li>❸ Paralelismo a nivel de instrucción.</li> <li>❹ Memoria caché integrada.</li> <li>❺ Conjunto de instrucciones.</li> </ul>
secuencia e interrelación	El orden elegido para su tratamiento no ha sido caprichoso. Comenzamos por las variables de bajo nivel, más ligadas a su constitución eléctrica, proseguiremos en un nivel intermedio con las que se encuentran más asociadas a su diseño, y finalizaremos a más alto nivel, con la frontera de diálogo con el software. La <a href="#">figura 3.1</a> muestra la secuencia que seguiremos en nuestro tratamiento, y la interrelación entre las distintas magnitudes atendiendo a criterios de vecindad dentro de la pirámide que adjuntamos.

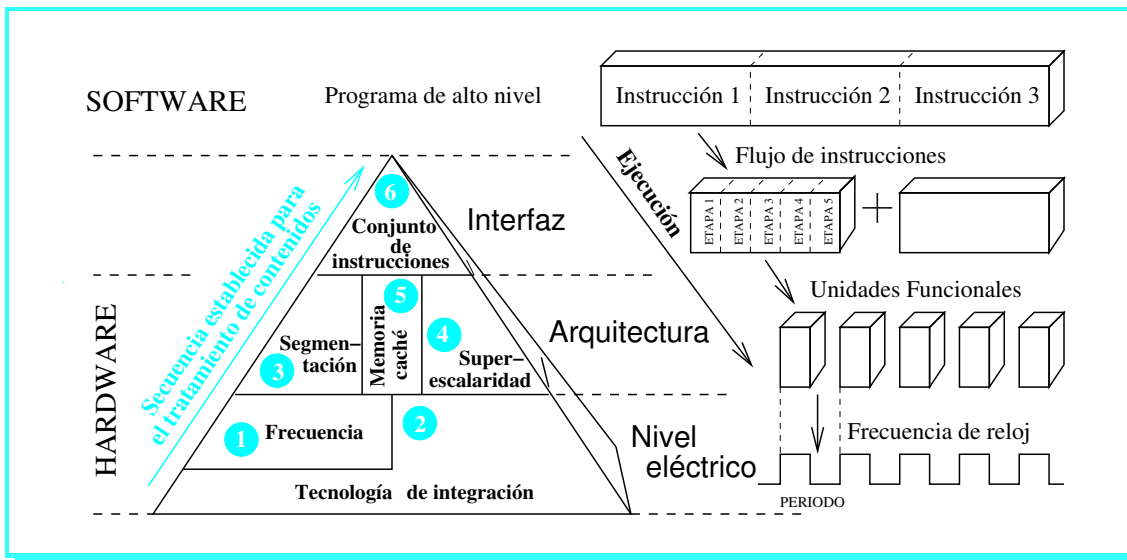


FIGURA 3.1: La secuencia en el tratamiento de las principales magnitudes del procesador y su interrelación y dependencia, atendiendo a donde se encuentra cada una apoyada en nuestra pirámide. Fuera de ella, podemos también observar la relación con la vertiente software del procesador.

## SECCIÓN 3.1

## Frecuencia de reloj

La frecuencia de reloj es un factor cuantitativo que indica la velocidad del microprocesador. Tiene su origen en un cristal de cuarzo, que ante la aplicación de un voltaje comienza a vibrar (oscilar) a una frecuencia armónica determinada por la forma y el tamaño del cristal.

Las oscilaciones emanan en forma de una corriente que sigue la función senoidal correspondiente a su frecuencia armónica, y que una vez filtrada en un circuito PLL se convertirá en la secuencia de pulsos digitales, cuadrados, periódicos y síncronos cuya cadencia marcará el ritmo de trabajo de los distintos chips del computador. Para mayor información sobre la distribución de esta señal por toda la placa base del PC, consultar la [sección 17.2](#).

El oscilador suele integrarse ya dentro del juego de chips de la placa base, por lo que cada vez es menos frecuente advertir su presencia en la circuitería mediante una exploración visual. No obstante, siguen existiendo muchos modelos de tarjetas y placas base en las que se monta como un elemento aparte, en forma de una pequeña pastilla de estaño que reviste su delgada lámina de cuarzo. Las [fotos 17.3](#), [30.1](#) y [30.2](#) muestran el aspecto de varios de estos osciladores y su circuito PLL asociado.

La magnitud inversa de la frecuencia es el **período de reloj**. De esta manera, si la frecuencia se expresa en Megahercios, o millones de pulsos por segundo, el período lo hará en microsegundos, y si la primera lo hace en Gigahercios, el segundo se regirá en nanosegundos. La [figura 3.2](#) muestra el aspecto de esta señal en contraste con otra de tipo analógico. Para tener conciencia de las magnitudes que manejamos, diremos que un reloj de 1 GHz realiza dos millones de ciclos en el mismo tiempo que dura un leve parpadeo de nuestros ojos.

origen

forma

☛ [p. 8/Vol.3](#)

oscilador

☛ [p. 9/Vol.3](#)

☛ [Vol.5 en Web](#)

período de reloj

magnitudes

☛ [pág. 50](#)

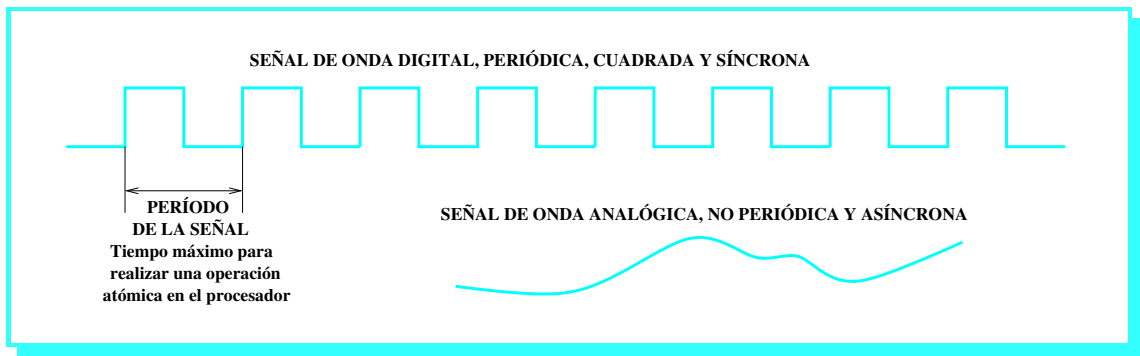


FIGURA 3.2: Aspecto de la señal de reloj que sincroniza la velocidad de un microprocesador y comparativa con otra señal de naturaleza analógica.

Las instrucciones máquina que acepta el microprocesador se descomponen en una sucesión de operaciones atómicas o etapas, cada una de ellas consumiendo un período o ciclo de reloj, tal y como hemos ilustrado en la [figura 3.1](#). Así, cuanto mayor sea la frecuencia del chip, menor será el tiempo que tardará en ejecutar las instrucciones de un programa, con el consiguiente aumento en el rendimiento de nuestro PC.

pág. 49  
descomposición

Mucha gente resume la potencia de un microprocesador en el valor de su frecuencia de trabajo, y los fabricantes y distribuidores de hardware siguen aprovechando esta práctica instaurada entre sus clientes para tratar de encandilarnos con ella. Realizar tal simplificación en un sistema de la sofisticación de un procesador con decenas de millones de transistores es una temeridad. Ni siquiera con el estudio de las cinco magnitudes que aquí proponemos puede uno pretender conocerlo, aunque sí creemos que puede forjarse ya una impresión válida del conjunto.

estereotipos

Remitimos al lector al [capítulo 29](#), dedicado en exclusiva a la frecuencia del procesador, para profundizar en su caracterización física y la incidencia que tiene sobre el resto de variables eléctricas del chip. Si quiere aprender más sobre la farsa que este parámetro representa desde hace ya un tiempo, le recomendamos la lectura del [capítulo 30](#). En los próximos capítulos encontrará ejemplos esclarecedores del peso relativo que tiene la frecuencia en el rendimiento de un microprocesador, siendo el Pentium 4 (ver [sección 6.4](#)) un buen exponente de lo que decimos.

Vol. 5 en Web

Vol. 5 en Web  
información  
adicional  
pág. 208

## SECCIÓN 3.2

### Tecnología de integración

**definición** La tecnología de integración es un indicador más cualitativo que cuantitativo. En general, puede definirse como *la mínima resolución de la maquinaria responsable de integrar los circuitos mediante técnicas de litografía*.

**importancia** Estamos frente a la magnitud de la que más dependen las otras cuatro, ya que disminuir esta resolución supone reducir el coste por cada chip integrado y su voltaje de alimentación, y aumentar la frecuencia y el número de transistores disponibles. Profundizaremos más sobre todas estas implicaciones, pero después de acercarnos a su significado intrínseco.

#### 2.1 ► Evolución y significado

transistor MOS

Dentro del microprocesador existen millones de minúsculos conmutadores, los **transistores**, cuyo funcionamiento comporta dos estados que son interpretados como '0' o '1' para componer el sistema lógico binario con el que se procesa toda la información dentro del computador. En su

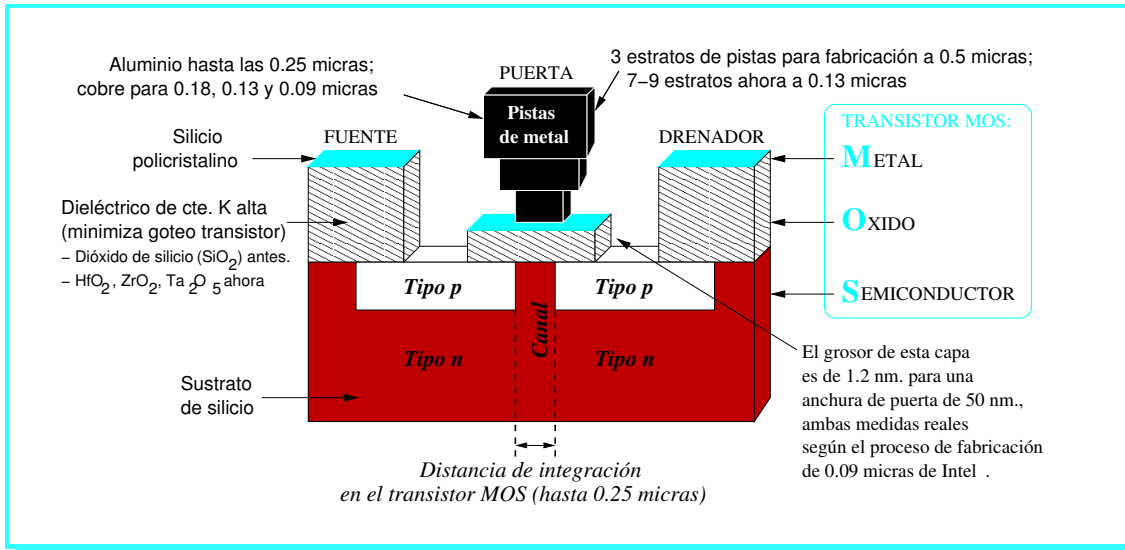


FIGURA 3.3: Un típico transistor MOS, donde indicamos los materiales utilizados para su fabricación y su distancia de integración. El silicio dopado con impurezas actúa como semiconductor cuya dualidad conductor/aislante permite la codificación de un bit al nivel más interno.

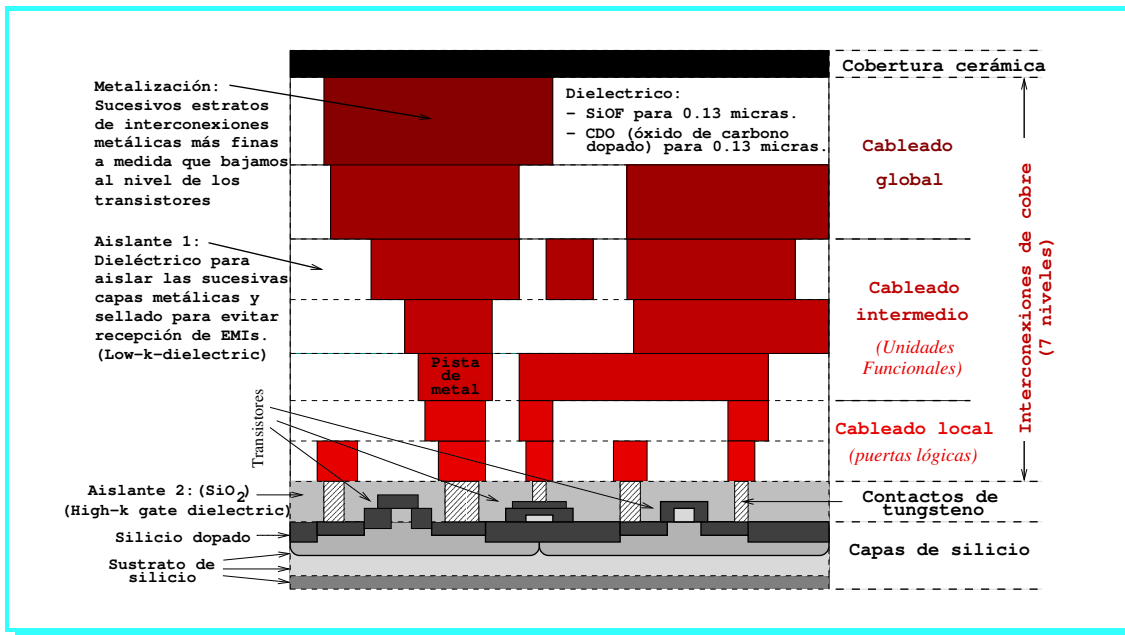


FIGURA 3.4: Sección transversal de un chip, donde se aprecian los niveles de interconexión de sus transistores: Entre 6 u 8 estratos de pistas metálicas de aluminio o cobre, responsables de definir la funcionalidad del chip.

versión MOS (*Metal Oxide Semiconductor*), los transistores se fabrican con un material semiconductor, y se conectan entre sí por medio de un metal (ver figura 3.3).

La integración de estos dos componentes en el área del chip tiene lugar mediante sofisticadas técnicas de encapsulado de materiales, donde los transistores, que no son apilables, se disponen en una estructura bidimensional que ocupa la capa más inferior, y se interconectan mediante un enrejado compuesto por entre seis y ocho capas de aluminio o cobre como metal situado por

	Tecnología de integración (micras)									
	1.00	0.80	0.50	0.35	0.25	0.18	0.13	0.09	0.065	0.045
Anchura (micras) para la puerta del transistor	1.00	0.80	0.50	0.35	0.20	0.13	0.07	0.05	0.035	0.025
Area (micra <sup>2</sup> ) para la celda de caché (6 transistores)	220	111	44	21	10.6	5.6	2.09	1.00	?	?
Voltaje de alimentación (voltios)	5.0	5.0	3.3	2.5	1.8	1.5	1.3	1.2	?	?
Maquinaria y año de fabricación	P648 1989	P650 1991	P852 1993	P854 1995	P856 1997	P858 1999	P860 2001	P1262 2003	P1264 2005	P1266 2007

TABLA 3.1: Evolución de la tecnología de integración y su relación con la anchura de la puerta de sus transistores y otras variables eléctricas en las plantas de fabricación de Intel. La maquinaria de fabricación se renueva aproximadamente cada dos años.

encima del silicio (ver [figura 3.4](#) - para más información puede consultarse el [capítulo 34](#)).

Vol. 5 en Web

CMOS

La tecnología de integración más ampliamente utilizada durante el proceso de fabricación es la CMOS (Complementary Metal Oxide Semiconductor), responsable del 75 % del volumen total de chips manufacturados a escala mundial según la Semiconductor Industry Association.

bipolar

De todos los microprocesadores que analizaremos, tan sólo el Pentium y el Pentium Pro no utilizan íntegramente esta tecnología: Son encapsulados CMOS en su parte más interna, pero la zona perimetral donde se sitúa su patillaje (287 y 386 pines resp.), están integrados con tecnología bipolar. Esta alternativa se utilizaba para las zonas del chip que necesitaran una mayor intensidad de corriente, pero ha entrado en claro desuso debido a la fuerte disipación de potencia que esto conlleva y los problemas de temperatura que padecen los microprocesadores actuales.

distancia de integración

pág. 51

pág. 55

El parámetro clave de una tecnología de integración concreta es la **distancia de integración**. En tecnología CMOS, por ejemplo, esta distancia coincidió durante las dos últimas décadas con la anchura del canal con que se fabrican sus transistores de silicio (ver [figura 3.3](#)), aunque dicha coincidencia dejó de cumplirse con la llegada de las 0.25 micras (los datos de las [tablas 3.1](#) y [3.3](#) son bastante clarificadores a este respecto); en otras tecnologías de fabricación de transistores, la distancia de integración se asocia con la anchura de la pista de metal que une los transistores.

de la micra al nanómetro

Tradicionalmente, la distancia de integración ha venido midiéndose en **micras**, diminutivo del *micrómetro*, que representa la millonésima parte del metro. Sin embargo, la evolución tan espectacular que ha seguido la miniaturización de los transistores ha dejado grande a esta escala, y cada vez es más usual emplear el nanómetro (abreviado nm. - mil millonésima parte del metro). Por ejemplo, los últimos modelos de K7 y Pentium 4 se fabrican a 0.13 micras o 130 nm., y la próxima mejora en este sentido nos lleva en 2004/2005 a las 0.09 micras o 90 nm.

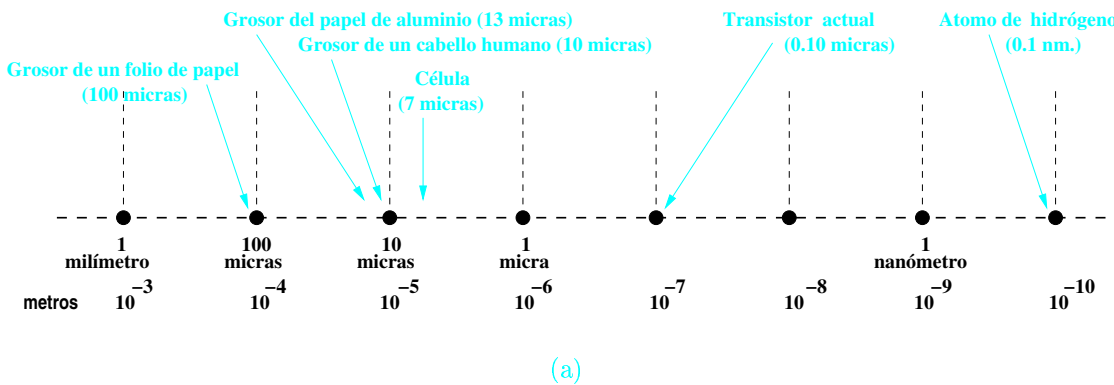
escala

pág. 53

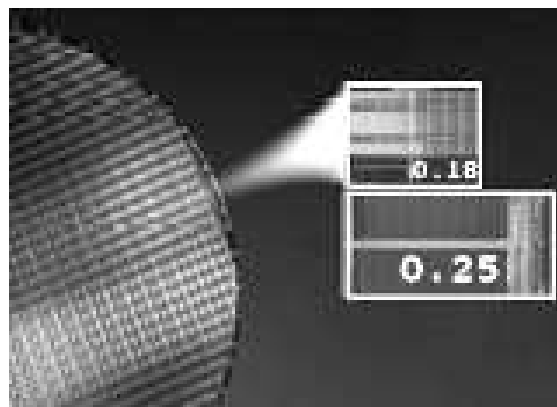
Para tener una referencia visual de lo que todo esto significa, diremos que tendríamos que apilar la anchura de más de 750 transistores de 0.13 micras para alcanzar el grosor de una hoja de este libro. La [foto 3.1](#) muestra una escala comparativa donde ubicamos una serie de elementos de dimensiones microscópicas.

nomencIatura

Puesto que la tecnología de integración va a ser siempre CMOS (salvo en las dos excepciones obsoletas ya reseñadas), tenderemos a omitirla a partir de ahora, y para simplificar las cosas englobaremos también en este término a la distancia de integración en micras. Es decir, diremos que un microprocesador se fabricó, por ejemplo, con *tecnología de integración de 0.13 micras* cuando formalmente tendríamos que haber dicho que se fabricó con *tecnología de integración CMOS a una distancia de integración de 0.25 micras para las puertas de sus transistores*.



(b)



(c)

Fotos cortesía de Intel

FOTO 3.1: Microscopía de la distancia de integración. (a) Escala de distancias (logarítmica) que permite relativizar el tamaño de un transistor actual respecto a otros elementos. Si la distancia de 0.25 micras fuese este libro abierto, las dimensiones del átomo de hidrógeno en que acaba la escala se corresponderían con el punto en que acaba esta frase. (b) Un grano de sal fotografiado junto a la orografía de un circuito integrado. (c) Comparativa espacial (horizontal) entre las 0.18 micras (arriba) y las 0.25 micras (abajo).

De forma más reciente en la que esa correspondencia ya no existe (por ejemplo, para 130 nm., la anchura de la puerta del transistor es de apenas 70 nm. - ver [tabla 3.1](#)), estaríamos ya obligados a utilizar la definición más general que hace referencia a la resolución mínima de la maquinaria de litografía con la que se fabricó el microprocesador, pero seguiremos utilizando la jerga a la que estamos acostumbrados aún reconociendo que científicamente no es lo más purista.

La tecnología de integración no avanza de forma continua, ya que una vez instaurado un proceso de fabricación debe transcurrir un tiempo hasta la amortización de sus plantas de fabricación. Ciertamente es que una empresa como Intel o AMD produce cientos de millones de chips cada año, pero el coste de sus plantas de fabricación es también enorme.

plantas de  
fabricación

Los plazos establecidos por la industria del chip para la renovación de sus plantas de fabricación están en torno a los dos años, tal y como se refleja en la [tabla 3.2](#), donde resumimos los valores utilizados por Intel y AMD. Ese período ha demostrado ser lo suficientemente extenso como para amortizar el enorme coste de la planta, y al mismo tiempo, lo suficientemente efímero como para mantener a estas empresas a la vanguardia del tren de la alta tecnología.

renovación  
cada 2 años  
pág. 54

Distancia de integración (micras)	Año de implantación	Modelo más representativo	
		Intel	AMD
1	1989	80486	80486
0.8	1991	80486DX	-
0.8 (0.5, 0.35)	1993	Pentium	-
0.6 (0.35)	1994	Pentium Pro	-
0.35 (0.25)	1995	Pentium MMX	K5
0.35 (0.25)	1997	Pentium II	K6
0.25 (0.18, 0.13)	1999	Pentium III	K7
0.18 (0.13)	2001	Pentium 4	Athlon XP
0.13	2003	Pentium 4	K8

TABLA 3.2: La evolución en la distancia de integración ha transcurre siempre a pasos discretos. En esta tabla recogemos los valores que ha tomado esta variable en los últimos diez años. Los números entre paréntesis corresponden a evoluciones posteriores.



### Ejemplo 3.1: LA AMORTIZACIÓN DE LAS PLANTAS DE FABRICACIÓN DE CHIPS

La planta de fabricación de microprocesadores más avanzada instalada en suelo europeo se encuentra emplazada en Dresden (Alemania), y pertenece a la empresa AMD.

Inició su actividad en Noviembre de 1998, fabricando desde Enero de 1999 los K6 de 0.25 micras y aluminio, desde Enero de 2000, los K7 de 0.18 micras y cobre, y desde Noviembre de 2002 los Athlon XP de 0.13 micras y 9 niveles de metal. Para mediados de 2003 fabricará los futuros K8, y ya en 2004, integrará éstos a 0.09 micras.

Su ritmo de producción es de más de un millón de chips por semana (5000 obleas de 20 centímetros de diámetro cada una, para ser exactos), y si dicho ritmo disminuye sólo un 10% durante una jornada laboral de 8 horas, las pérdidas para AMD sobrepasan los 10 millones de euros.

A esa velocidad de manufacturación, lo primero que uno piensa es que la amortización de costes es casi inmediata. No tanto: La inversión inicial superó los 2.000 millones de euros.

ASML Por cierto, que la empresa líder en suministrar a los fabricantes la maquinaria de litografía más puntera con la que hacer realidad sus chips es precisamente europea: La holandesa ASML.

## 2.2 ▶ Efectos directos sobre otras variables

El potencial que esconde una reducción de la distancia de integración es tan grande que consigue hasta cuatro efectos benignos sobre las variables físicas ligadas a la constitución interna de un microprocesador:

número de transistores

- 1 Consigue aumentar de forma cuadrática el **número de transistores** que se pueden integrar en un mismo espacio físico. Es decir, una reducción desde 1 micra hasta las 0.5 micras para un mismo diseño permitiría albergar cuatro veces más transistores, y una reducción hasta las 0.25 micras, dieciséis veces más. Esto hace crecer de forma considerable el patrimonio de que dispone el diseñador del microprocesador para aumentar sus prestaciones: incorporando funcionalidad adicional, aumentando el tamaño de las cachés integradas, y un sinfín de posibilidades más que iremos desvelando un poco más adelante.



	Tecnología de integración (micras)				
	0.13	0.10	0.07	0.05	0.035
Año de su puesta en marcha	2002	2005	2008	2011	2014
Anchura de puerta del transistor	0.085	0.065	0.045	0.030	0.020

Fuente: International Technology Roadmap for Semiconductors. Edición 1999

**TABLA 3.3:** Estimaciones para los sucesivos valores de la tecnología de integración y la anchura de puerta del transistor, junto a su puesta en funcionamiento en un marco temporal de quince años en adelante. Las predicciones han sido realizadas por la Semiconductor Industry Association. A finales de 2002, IBM y AMD firmaron un acuerdo para la fabricación de chips que confirma la llegada de las 0.065 micras para 2005 y las 0.045 incluso para 2007 en una nueva planta de fabricación con obleas de 30 cm. de diámetro a construir en Singapur.

Generación y modelo de microprocesador	Distancia de integración en micras							
	1	0.8	0.6	0.5	0.35	0.25	0.18	0.13
Cuarta - 80486	33-100MHz							
Quinta - Pentium	66MHz — 233MHz							
Sexta - Pentium Pro	133 — 200MHz							
Quinta - Pentium MMX					133-233MHz			
Sexta - K6					166-333MHz			
Sexta - Pentium II					233-450MHz			
Sexta - K6-2					266-500MHz			
Sexta - Pentium III					0.45 — 1.3 GHz			
Séptima - K7					0.5 — 2.5 GHz			
Séptima - Pentium 4					1.4 — 4 GHz			

**TABLA 3.4:** La reducción de la distancia de integración en las sucesivas generaciones de microprocesadores es una de las claves sobre las que se sustenta la consecución de frecuencias más elevadas.

- 2 Aumenta la velocidad de operación del transistor, y con ello, la **frecuencia** del chip. La agilidad para conmutar entre los estados lógicos 0 y 1 viene dada por el tiempo que debe transcurrir para que la corriente que provoca ese cambio fluya entre la fuente y el drenador del transistor (ver [figura 3.3](#)). Como la distancia de integración es precisamente la que separa estas dos zonas, cuanto menor sea ésta, menor será el tiempo de paso y mayor la velocidad de conmutación. La física del transistor nos dice en este sentido que una reducción de su distancia de integración conlleva un aumento en su frecuencia de similares proporciones, es decir, que como cada nuevo proceso de fabricación contempla una reducción de distancia de factor 0.7x respecto a su predecesor, la frecuencia aumentará en un factor 1.5x.
- 3 Disminuye el **voltaje de alimentación** que requiere el chip. La principal beneficiaria de esta disminución es la menor **potencia disipada** en forma de calor, lo que repercutirá en un descenso de la **temperatura** del chip. En el [capítulo 29](#) descubrimos que la temperatura es uno de los parámetros que más limitan la frecuencia de reloj de un microprocesador, por lo que una reducción del voltaje complementa la consecución de altas frecuencias.
- 4 Permite disminuir el **coste de fabricación** del microprocesador, ya que al acortarse las distancias entre transistores, se reduce el área de silicio que se necesita para la integración de un diseño dado un número de transistores fijo. En definitiva, aumenta la densidad de integración y el número de chips que caben en cada oblea de silicio. Como la distancia de integración es una magnitud lineal y el precio del chip se determina en función del coste por oblea (área circular, y por tanto, bidimensional), resulta que una reducción en la distancia de integración abarata el coste de forma cuadrática.

frecuencia

→ [pág. 51](#)

voltaje  
potencia  
disipada  
temperatura

→ [Vol.5 en Web](#)

coste de  
fabricación

 **Ejemplo 3.2:** CUANTIFICANDO EL EFECTO DIRECTO DE LA DISTANCIA DE INTEGRACIÓN SOBRE OTRAS VARIABLES

Supongamos que los últimos Pentium (P55C) y primeros Pentium II (Klamath), todos de 0.35 micras, hubiesen sido fabricados con las distancias de integración de 0.18 micras de los últimos Pentium III (Coppermine) y primeros Pentium 4 (Willamette).

En ese caso, el Pentium 200 MHz hubiese sido un Pentium 400 MHz, y el Pentium II 300 MHz hubiese sido un Pentium II 600 MHz. Asimismo, el Pentium, de 3.1 millones de transistores habría podido tener 12.4 millones, y el Pentium II, de 7.5 millones, hasta 30 millones. Bastantes más de los que realmente dispone el Coppermine, y en un área de silicio muy similar. Con estos transistores hubiera sido posible: (a) Dotar al Pentium de una caché L2 de 128 Kbytes sincronizada a su misma velocidad utilizando la tecnología de aquella época, pues se necesitaban entonces unos 8 millones de transistores para ello, y (b) Incorporar al Pentium II una caché L2 interna de 384 Kbytes a su misma velocidad, algo que ni siquiera tuvo a su alcance el Pentium Pro en sus inicios con un coste superior a los 1.200 €. (El ejemplo cuantifica sólo la incidencia directa entre variables, porque dado que en la práctica existen multitud de efectos laterales, sería extraordinariamente difícil precisar un valor real exacto).

### 2.3 ▶ Efectos laterales entre las variables afectadas

Desgraciadamente, las cosas no son tan sencillas como acaban de ser expuestas. Ya avisamos que estamos ante un sistema extraordinariamente complejo en el que se producen multitud de efectos laterales, y al menos debemos hacer referencia a los más importantes:

- 1 Si contrastamos la evolución de la distancia de integración con la de la frecuencia, vemos que si los diseños de una micra rondaban los 33 MHz y los de 0.13 micras se encuentran en torno a los 3 GHz, en el mismo espacio temporal en el que un parámetro se ha reducido en un factor de 7.5, el otro ha aumentado en un factor de 100. Esto representa 13 veces más de lo esperado según la incidencia lineal de uno sobre otro que acabamos de postular, pero lo que ha ocurrido aquí es que otros aspectos que también inciden positivamente sobre la frecuencia han sido responsables del rango de mejora restante. Estos otros aspectos se encuentran documentados en el [capítulo 29](#).

velocidad

Vol.5 en Web 

 **Ejemplo 3.3:** CUANTIFICANDO LA MEJORA QUE LA DISTANCIA DE INTEGRACIÓN REVIERTE SOBRE LA FRECUENCIA TENIENDO EN CUENTA EFECTOS LATERALES

Si un viejo Pentium Pro se fabricara ahora utilizando distancias de integración de 0.18 micras, alcanzaría su techo de frecuencia en torno a los 1.2 GHz. Esto significa tres veces más del valor esperado según el [ejemplo 3.2](#), lo que evidentemente pone de manifiesto la incidencia de los efectos laterales comentados.

Parámetro eléctrico	Efecto de una reducción de la distancia de integración	Efecto lateral sobre los demás			
		Núm.	Frec.	Coste	Volt.
Núm transistores	Aumento cuadrático	=	Baja	Sube	Sube
Frecuencia	Aumento lineal	Baja	=	=	Sube
Coste	Reducción cuadrática	Baja	=	=	=
Voltaje	Reducción lineal	Baja	Baja	=	=

TABLA 3.5: Efectos directos que produce una reducción de la distancia de integración y la repercusión que éstos tienen a su vez en forma de efectos laterales sobre los demás parámetros eléctricos de un chip.

- ② Tampoco es correcto considerar que el precio del chip se reduzca de forma cuadrática, pues hemos cuantificado su coste en función de la materia prima utilizada, pero no en la dificultad de integración y testeo, que obviamente son tareas más complejas al realizarse sobre más transistores y más diminutos, lo que exige mayores inversiones en infraestructura de litografía.

coste

Además, no todos los transistores de un chip son iguales en tamaño. Los de la caché son muy pequeños en silicio pero muy complejos en sus capas de metalización (interconexiones), mientras que los que contiene una ALU ocupan un área de integración mayor. A efectos prácticos, y de forma implícita, estamos considerando un tamaño medio común a todas las unidades funcionales con objeto de poder establecer una proporcionalidad entre el número de transistores utilizados y el área de integración del chip.

espacio

- ③ El voltaje y la frecuencia distan mucho de ser independientes entre sí. Si subimos la frecuencia para aprovechar las mejoras que nos brinda la tecnología, ésta tirará a su vez para arriba del voltaje, contrarrestando los niveles de tensión inferiores que la tecnología ponía a nuestro alcance. El resultado puede ser que incluso se haga necesario aumentar el voltaje a distancias de integración más pequeñas.

voltaje

- ④ El número de transistores y la frecuencia tampoco son independientes, porque si nos decidimos a aprovechar el mayor número de éstos, a buen seguro que aumentaremos la cantidad de unidades funcionales del chip y el número de elementos que debe atravesar su **camino crítico** (aquel que atraviesan las señales eléctricas cuando el microprocesador ejecuta su operación atómica más lenta), y la longitud de éste condiciona fuertemente la máxima frecuencia de funcionamiento del conjunto. El resultado contrasta de nuevo con las cuatro premisas establecidas en el apartado anterior, ya que una reducción de la distancia de integración ha beneficiado a la funcionalidad del diseño, pero ha perjudicado a la frecuencia.

camino crítico

En consecuencia, podemos concluir que no existe de antemano una estrategia ganadora en la construcción de un microprocesador. Y el mercado es, una vez más, quien nos proporciona la lección más soberana: En él coexisten diseños que priman descaradamente la frecuencia de reloj (como el Alpha 21264 de Digital que en 1993 superó los 500 MHz), frente a otros orientados claramente a aumentar el número de unidades funcionales (como la familia del Power PC de Motorola, que para ese mismo año disponía de diseños con un factor superescalar de seis). Y entre ambos extremos tenemos toda una gama de soluciones intermedias que ponderan de diferente manera uno y otro aspecto. Un buen ejemplo que recorre todo este territorio intermedio sería la familia de los Pentium, que basculó progresivamente desde el lado de las unidades funcionales con el Pentium Pro hacia el lado de la frecuencia con los Pentium II, III y 4<sup>1</sup>.

Alpha

Power PC

Pentium

<sup>1</sup> Escribiremos los ordinales de los microprocesadores en números romanos o arábigos según se haya contemplado en la marca registrada de su fabricante.

## 2.4 ► Cómo dar empleo a un ejército de transistores

Llegado este punto, conocemos que las sucesivas reducciones en la distancia de integración traen consigo la posibilidad de aumentar la frecuencia o el número de transistores del diseño, pero no ambas, dados los conflictos apuntados como efectos laterales.

bajo nivel La opción de aumentar la frecuencia exige ciertos conocimientos sobre el funcionamiento interno del transistor a bajo nivel, de ahí que hayamos decidido abordarla más adelante dentro del contexto microelectrónico que predomina en el volumen 5, disponible en nuestra Web. En concreto, en el [capítulo 29](#) contamos los fundamentos teóricos, mientras que en el [capítulo 30](#) pasamos a la acción para su manipulación.

nivel arquitectural Ahora, lo que toca es centrarse en el nivel arquitectural del chip, o lo que es lo mismo, conocer qué infraestructura habilitar para aprovechar el creciente número de transistores que la tecnología de integración va a ir poniendo a nuestro servicio de forma sucesiva. Mostraremos cómo esa circuitería adicional puede colaborar en el aumento del rendimiento de un microprocesador a pesar de que la frecuencia del chip pueda verse en ocasiones resentida. De esta manera, destruiremos ese falso mito que muchos profanos de la informática tienen en la cabeza: “el mejor procesador es aquel que funciona a más MHz”.

Las vías que se han utilizado para estas mejoras del rendimiento giran en torno a tres ideas principales que ocuparán los tres tramos siguientes de este capítulo:

paralelismo ❶ El paralelismo a nivel de instrucción ([sección 3.3](#)), o la facultad de poder procesar varias instrucciones de forma simultánea, todas ellas procedentes de un único programa en ejecución.

caché ❷ La incorporación de memorias cachés como un elemento más del diagrama de bloques del microprocesador, o la virtud para hacerle llegar grandes volúmenes de datos con la presteza que los necesita ([sección 3.4](#)).  
pág. 69 ➡

instrucciones ❸ Las ampliaciones y/o modificaciones del conjunto de instrucciones máquina, o la potestad para especializar al microprocesador en las nuevas operaciones multimedia que van requiriendo los programas, al tiempo que se logra también optimizar el procesamiento de las más tradicionales ([sección 3.5](#)).  
pág. 94 ➡

### SECCIÓN 3.3

## Paralelismo a nivel de instrucción

1a idea En su descomposición funcional más sencilla, un microprocesador se compone de una Unidad de Control y una Unidad de Proceso. En esta última convivían inicialmente la ALU y el banco de registros en solitario, pero enseguida se les unieron nuevas unidades funcionales en la búsqueda de un aumento del rendimiento. La mayoría de estos aditivos llevan a la práctica alguna forma de **paralelismo a nivel de instrucción**, idea que de forma genérica consiste en romper con la ejecución secuencial de instrucciones (una detrás de otra en el tiempo) para simultanear su ejecución (varias a la vez). Distintas formas de explotar este paralelismo son la segmentación, la superescalabilidad y la supersegmentación.

## Segmentación (pipelining)

### ◀ 3.1

Un procesador **segmentado** es aquel que divide el proceso de ejecución de una instrucción en  $N$  etapas de similar duración, con el objetivo último de procesar  $N$  instrucciones simultáneamente, encontrándose cada una en una etapa diferente de su ejecución.

definición

Conseguimos así que  $N$  unidades funcionales del procesador estén trabajando a la vez, lo que redundará en un factor  $N$  de mejora en el rendimiento del chip en circunstancias ideales.

rendimiento

Las instrucciones fluyen secuencialmente por las distintas unidades funcionales del procesador de igual forma que el agua fluye por el cauce de una tubería, de ahí que en la jerga de la calle se les conozca también como procesadores *pipeline*<sup>2</sup>. A nosotros nos ha parecido más elegante emplear el término *cauce segmentado* o simplemente *cauce*, denominación que seguiremos a partir de ahora.

etimología



#### Ejemplo 3.4: EL CAUCE DE SEGMENTACIÓN MÁS CLÁSICO

El modelo de segmentación más repetido en la primera mitad de los años 90 es el compuesto por las siguientes cinco etapas: Búsqueda de la instrucción, decodificación, lectura de operandos, ejecución de la operación asociada y escritura de su resultado.

De esta manera, mientras el procesador escribe una instrucción, ejecuta simultáneamente la siguiente, busca los operandos de una tercera, decodifica una cuarta y busca de memoria una quinta instrucción, lográndose en el caso ideal una aceleración de cinco para el rendimiento del chip.

Tanto el MIPS, el procesador modelo de esta técnica desarrollado en Stanford, como el Pentium de Intel, presentan una división en cinco etapas muy similar a la descrita.

Todos los microprocesadores actuales se encuentran segmentados, habiendo aumentado el número de etapas con el paso de las generaciones. Así, lo normal en séptima generación es encontrarnos con cauces de ejecución entera compuestos de hasta 20 etapas de segmentación, como ocurre, por ejemplo, en el procesador que lidera este rasgo en la arquitectura PC de 2003, el Pentium 4 de Intel.

Junto al cauce de ejecución entero se sitúan otros cauces por donde circulan los otros tipos de instrucciones, principalmente multimedia y de punto flotante. Como hasta que no concluye la fase de decodificación de instrucción no se puede realizar la pertinente ramificación, las primeras etapas de segmentación son siempre comunes a todos los cauces del procesador.

variedad de cauces

El empleo tan superlativo del concepto de segmentación es algo que no sorprende desde el momento en que se conoce su gran activo subyacente: El aumento del rendimiento se consigue con sólo reorganizar las unidades funcionales existentes, es decir, no supone incremento de coste para la Unidad de Proceso. El diseño de la Unidad de Control sí se complica un poco a medida que aumenta el número de etapas y el grado de concurrencia en la ejecución de instrucciones, pero en cualquier caso, con carácter marginal frente al espectacular incremento logrado en el lado del rendimiento.

coste

<sup>2</sup>El término anglosajón que significa precisamente *tubería*.

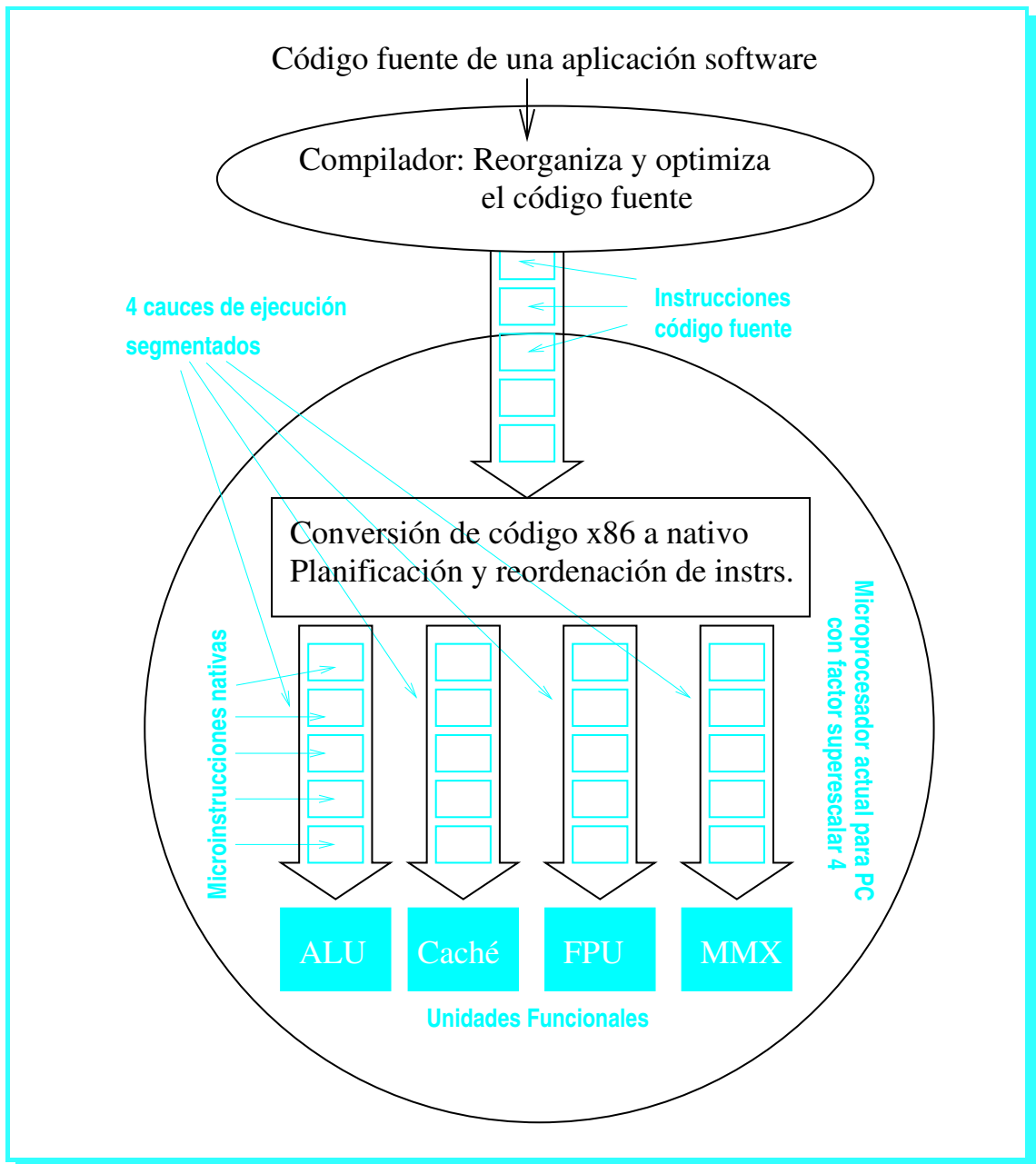


FIGURA 3.5: La forma tradicional de ejecutar instrucciones en un microprocesador de quinta y sexta generación utiliza los conceptos de segmentación y superescalaridad combinados.

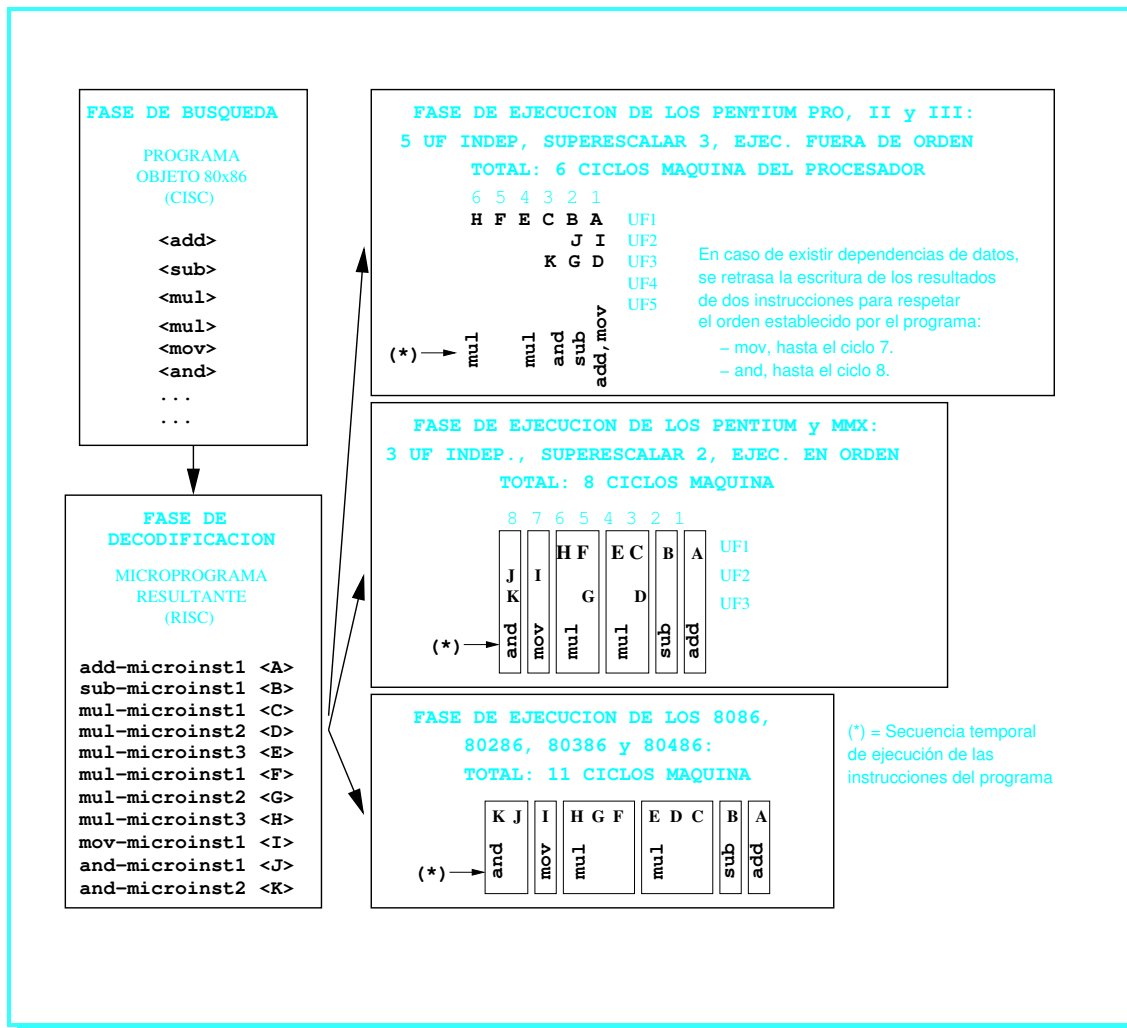
### 3.2 ▶ Superescalaridad

concepto

Un procesador **superescalar** de factor  $N$  es aquel que replica  $N$  veces la circuitería de alguna de sus unidades funcionales con el fin de poder ejecutar  $N$  instrucciones en sus respectivas etapas de computación.

rendimiento

Al igual que la segmentación, a medida que las mejoras en la integración de circuitos han permitido concentrar más y más componentes dentro de un único chip, los diseñadores de microprocesadores han aumentado el factor de superescalaridad, logrando un incremento del rendimiento que al igual que en la segmentación alcanza el factor  $N$  en circunstancias ideales.



MAGNITUDES

FIGURA 3.6: Evolución de los procesadores de Intel de los últimos veinte años en relación a su ejecución superescalar tomando como ejemplo un breve programa secuencial. El proceso de decodificación que tiene lugar en la parte izquierda del diagrama es una conversión de código x86 al código nativo de cada uno de los procesadores indicados con objeto de mantener la compatibilidad hacia atrás que ha estado presente en todos los modelos de Intel y AMD hasta la fecha. El esquema se repite en diseños tan contemporáneos como el K7 y el Pentium 4.

Ahora bien, el coste de crecer por esta vía es superior al de la segmentación, puesto que aquí tenemos un incremento lineal de complejidad en la unidad de proceso. El hecho de que el coste de integración por transistor haya seguido siempre una evolución descendente explica que estemos ante otro negocio altamente rentable para los diseñadores del procesador.

coste

El primer procesador superescalar, el i960, fue diseñado por Intel en 1989, y podía ejecutar dos instrucciones por ciclo de reloj. Ya en 1995, lo normal era encontrar diseños de 4 instrucciones por ciclo, y aunque aparecieron algunos como el Power PC capaces de ejecutar hasta seis, entre la complejidad y el choque frontal que supone con la forma en que están escritos los programas, el mercado volvió rápidamente sobre sus pasos y se encuentra cómodamente instalado en factores de superescalaridad de entre 3 y 4. La figura 3.5 muestra cómo se complementan la superescalaridad y la segmentación para ejecutar los programas de forma más eficiente.

1989  
1995

2003



### Ejemplo 3.5: LA SUPERESCALARIDAD EN INTEL

El Pentium de Intel y su versión MMX son procesadores superescalares de factor 2. Disponen de dos ALU para operar con números enteros y una tercera para operar con números reales, permitiendo ejecutar de forma simultánea dos instrucciones aritméticas de tipo entero, o bien una de tipo entero y otra de tipo real (con ciertas limitaciones). Las versiones posteriores de Intel, como el Pentium Pro, II y III, adoptan todos un factor tres de superescalaridad, que puede aplicarse sobre un total de cinco unidades funcionales de ejecución independiente.

Aunque en las secciones dedicadas a cada microprocesador desglosaremos ampliamente todas estas estrategias, la [figura 3.6](#) resume en un sencillo ejemplo las diferencias básicas existentes entre el esquema de ejecución del 80486, el Pentium, y sus hermanos mayores. Los nuevos diseños como el Pentium 4 siguen respetando este mismo factor tres de superescalaridad.

## 3.3 ► Combinación de segmentación y superescalaridad

pág. 60 ➔  
compatibles...

Como ha quedado de manifiesto en la [figura 3.5](#), las dos filosofías de diseño anteriores son perfectamente compatibles. De hecho, puesto que la segmentación llega antes al diseño del procesador y es más barata de implementar, no conocemos de modelos comerciales que sean superescalares sin estar segmentados.

...pero  
antagónicas  
pág. 49 ➔

No obstante, existen ciertos conflictos a la hora de poner en práctica las dos ideas simultáneamente, y es que, tal y como ilustramos en nuestra pirámide de la [figura 3.1](#), cada técnica exige unos requisitos diferentes a la capa de bajo nivel del procesador:

- La segmentación descansa fundamentalmente sobre la base de una elevada frecuencia, en el sentido de que sólo un período de reloj muy corto permitirá descomponer cada instrucción en un número elevado de etapas.
- La superescalaridad, por el contrario, necesita de un ingente número de transistores para poder ser llevada a la práctica, y esto sólo se consigue con mejoras en la tecnología de integración.

conflictos

En definitiva, a la segmentación le estorba la superescalaridad porque ésta acarrea un desdoble de la circuitería existente, y por el famoso dicho de la microelectrónica “cuanto más grande, más lento”, tenemos un perjuicio sobre la frecuencia de reloj, perdiendo esperanzas de lograr un elevado número de etapas de segmentación.

De forma similar, a la superescalaridad le estorba la segmentación, porque cuando las etapas son tan minúsculas, se hace difícil incrementar su complejidad replicando circuitería.

dos escuelas

En la práctica, ocurre que los diseños fuertemente segmentados no utilizan un factor de superescalaridad elevado, y que los que apuestan por la superescalaridad reducen el número de etapas de segmentación del diseño. Por ejemplo, el Pentium 4 llega a las 20 etapas de segmentación pero sólo tiene un factor tres de superescalaridad, mientras que el K7 consigue un factor cinco de superescalaridad a costa de reducir a catorce el número de etapas en su cauce de ejecución entero. Al final, las dos estrategias alcanzan un grado de paralelismo inherente en torno a las



60-70 instrucciones simultáneas compatibilizando ambos conceptos, pero dando prioridad sólo a uno de ellos.

## Supersegmentación

### 3.4

La palabra supersegmentación apunta en primera instancia una cosa que no es, puesto que nos lleva a la tentación de aplicar la ecuación *superescalar + segmentado = supersegmentado*, cuando en realidad la ecuación correcta es *segmentado + segmentado = supersegmentado*.

equivoco

En efecto, un procesador **supersegmentado** es aquel que aplica dos veces el concepto de segmentación, la primera al nivel del diseño global, y la segunda al nivel interno de sus unidades funcionales.

concepto

Por ejemplo, una descomposición en cinco etapas de segmentación como la del Pentium, compuesta de etapas de búsqueda, decodificación, lectura, ejecución y escritura, deja abierta la puerta a aplicar una nueva segmentación sobre cada una de las unidades funcionales que intervienen en cada ciclo: la caché de instrucciones, el decodificador de instrucción, la caché de datos, la ALU o el banco de registros, respectivamente.

Tomando como referencia la caché de instrucciones referenciada en la primera etapa, ésta puede ser una caché segmentada como las que invadieron el mercado en la segunda mitad de los años 90 (ver [sección 11.1.4](#)), donde una subdivisión en dos etapas permite simultanear la búsqueda de una instrucción y la localización de su celda de caché asociada, con la lectura en sí del código de instrucción de la anterior y su volcado al bus de datos camino del procesador. Llegamos así a los dos niveles de segmentación mostrados en la [figura 3.7.d](#).

caché  
segmentada  
p. 118/Vol. 2

pág. 64

Notar que como buena segmentación, ha provocado un desdoble de la señal de reloj, que discurre por el eje temporal de abscisas al doble de frecuencia que en los otros tres diagramas mostrados. En general, la supersegmentación lleva asociada la subdivisión del ciclo de reloj en tantos ciclos como etapas se hayan establecido para el segundo nivel de segmentación, lo que desemboca diseños de muy elevada frecuencia.

frecuencias  
elevadas

Aunque en teoría la supersegmentación sea un concepto independiente de la **superescalaridad**, en la práctica, necesita de ella. Esto es así porque una buena implementación supersegmentada siempre incluye la ejecución de instrucciones *fuera de orden* para minimizar el efecto negativo que las dependencias de datos tienen sobre su rendimiento. Y como veremos enseguida, la ejecución fuera de orden sólo tiene cabida en procesadores que sean superescalares.

superescalar



### Ejemplo 3.6: EL ALPHA 21264, O LA SUPERSEGMENTACIÓN EN SU MÁXIMA EXPRESIÓN

Un clásico ejemplo de diseño supersegmentado es el procesador Alpha 21264 de Digital, procesador RISC de principios de la década de los 90 y hermano menor del Alpha 21364 de Compaq, el que decodificó el mapa del genoma humano en la empresa Celera Genomics. El segundo nivel de segmentación de esta arquitectura incluye una descomposición en **nueve** etapas para el caso de la caché (ver [figura 3.11](#)), lo que dió lugar a una frecuencia de reloj de 600 MHz para la versión comercializada a finales de 1997, el marco temporal en el que los Pentium II y demás modelos para PC estaban a menos de la mitad de esa frecuencia.

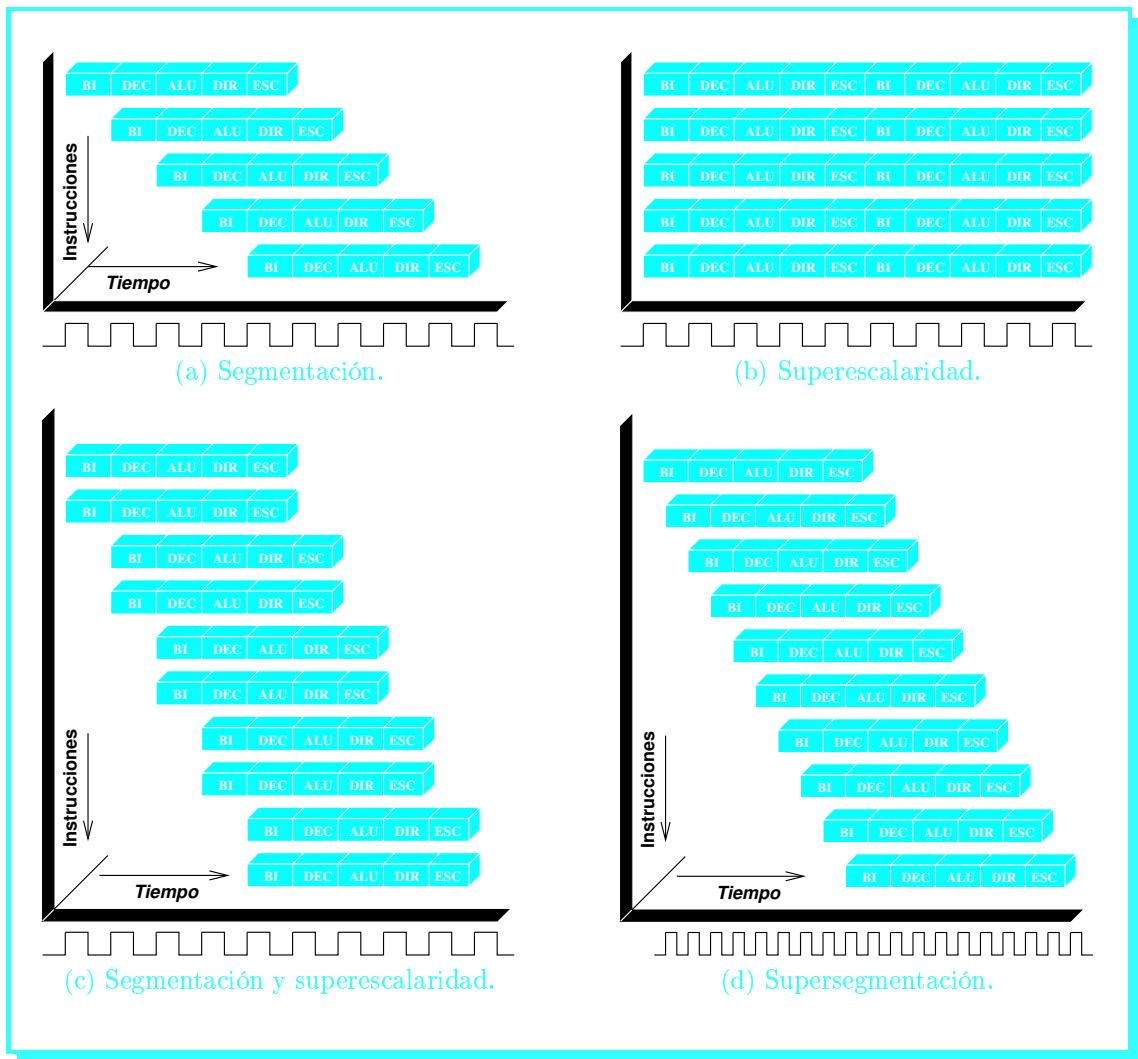


FIGURA 3.7: Comparación de las cuatro técnicas de paralelismo a nivel de instrucción que podemos encontrar en los microprocesadores actuales: (a) Segmentación. (b) Superscalaridad. (c) Segmentación y superscalaridad combinadas. (d) Supersegmentación. En los cuatro casos hemos supuesto que todas las instrucciones pueden descomponerse en cinco etapas: Búsqueda del código de operación (BI), decodificación de instrucción y búsqueda de operandos (DEC), ejecución de operación (ALU), generación de la dirección de destino (DIR) y escritura del resultado (ESC).

### 3.5 ► Dependencias: Las enemigas del paralelismo

dos mundos  
diferentes

Cualquier forma de paralelismo a nivel de instrucción ve mermado su potencial de mejora de manera considerable debido a la estructura secuencial que guardan los programas en la capa software, que ignoran cualquier tipo de ejecución simultánea de instrucciones. Es decir, estamos diseñando un procesador que no se corresponde con la forma en que va a ser utilizado por la capa software.

un mundo  
virtual

Podría pensarse que la multiprogramación, o más recientemente, la ejecución *multithread*, ayuda a sacar provecho de estos recursos hardware, pero no es así. Estos conceptos tan sólo reparten el tiempo del procesador entre un número de procesos o *threads*, pero en cada momento lo que se ejecuta en su interior es un único flujo de instrucciones.

La percepción de concurrencia que tenemos en nuestro PC no es más que una ilusión pro-

vocada por la enorme diferencia existente entre las dos escalas temporales, la nuestra y la del procesador. Dicho de otra manera, si echamos una foto al procesador en cualquier instante, lo pillaremos ejecutando siempre un único código, escrito para que sus instrucciones se ejecuten una detrás de otra. El compilador y el propio hardware colaboran para aprovechar los recursos disponibles redefiniendo la ejecución del código para habilitar algún tipo de concurrencia, pero siempre tienen la obligación de respetar la secuencialidad definida por el programador. Estos condicionantes introducen riesgos como los siguientes:

el mundo real

- 1 **Dependencias de datos.** Si alguno de los operandos fuente (o de lectura) de una instrucción B es el operando destino (o de escritura) de una instrucción anterior A, B no puede comenzar su ejecución hasta que A no haya finalizado.



### Ejemplo 3.7: RIESGO POR DEPENDENCIAS DE DATOS

Sea la siguiente pareja de instrucciones máquina consecutivas dentro de un programa cualquiera.

```
.....  
Instr. A:   add R1, R2, R3      # R1 = R2 + R3  
Instr. B:   sub R7, R1, R8      # R7 = R1 - R8  
.....
```

Podemos ver que la instrucción B toma uno de sus operandos de lectura de R1, por lo que necesita leer el valor de este registro para comenzar a trabajar. Sin embargo, puesto que este mismo valor es escrito por A justo antes de finalizar su ejecución, esto nos obliga a ejecutar B después de A en lugar de simultáneamente, deshabilitando cualquier forma de paralelismo a nivel de instrucción que pudiera tener implementado el microprocesador.

- 2 **Dependencias de control.** Una instrucción de salto condicional impide conocer la siguiente instrucción a ejecutar hasta que no se evalúe su condición de salto, y durante todo ese tiempo deberá detenerse la ejecución concurrente de instrucciones.



### Ejemplo 3.8: RIESGO POR DEPENDENCIAS DE CONTROL

Sea la siguiente terna de instrucciones máquina consecutivas dentro de un programa cualquiera.

```
Instr. A:   beq R1, R2, C       # Salta a la instr. C si R1=R2  
Instr. B:   sub R10, R11, R12  # R10 = R11 - R12  
Instr. C:   add R20, R10, R20  # R20 = R10 + R20  
.....
```

En este caso, la instrucción B no puede simultanear su ejecución con la de A: Debe esperar al menos a que ésta evalúe su condición de salto, ya que si resulta que los registros R1 y R2 contienen el mismo valor, la instrucción B no deberá ser ejecutada. Tampoco podemos simultanear la ejecución de las instrucciones A y C, pues C tomará el valor calculado en B si el programa finalmente no salta.

- ③ **Dependencias estructurales.** Una instrucción necesita en uno de sus ciclos de ejecución una unidad funcional que está siendo utilizada por otra instrucción en ese mismo instante.



### Ejemplo 3.9: RIESGO POR DEPENDENCIAS ESTRUCTURALES

Sea la siguiente pareja de instrucciones máquina consecutivas dentro de un programa cualquiera.

```
.....
Instr. A:  lw R1, Memoria(1000)           # Carga en R1 el contenido de
                                           # la posic. de memoria 1000)
Instr. B:  lw R2, Memoria(2000)         # Carga en R2 el contenido de
                                           # la posic. de memoria 2000)
.....
```

Si el dato solicitado por la instrucción A a memoria se encuentra en la memoria caché, es obtenido de forma casi inmediata, y enseguida se podrá proceder a ejecutar la instrucción B. Pero si el dato que A necesita no se encuentra en caché, se deberá solicitar de memoria principal, lo que consumirá al menos un centenar de ciclos del procesador, tiempo durante el cual la instrucción B deberá esperar en el caso de que necesite también utilizar la memoria principal.

penalización

pág. 67

Los conflictos anteriores reducen notablemente las oportunidades que pueden presentarse en un código para la ejecución simultánea de instrucciones, y el impacto que cada una de estas dependencias tiene sobre un código dependerá de la naturaleza de la aplicación software. La [tabla 3.6](#) resume el porcentaje de uso de cada unidad funcional del procesador con objeto de darnos una idea de la penalización que supone cada tipo de dependencia en una aplicación entera y de punto flotante. Una regla heurística que se viene cumpliendo tradicionalmente es que una de cada seis instrucciones de un programa es una instrucción de salto que provoca un riesgo por dependencias de control.

soluciones

Las dependencias estructurales se producen por las limitaciones del hardware, y están ligadas a su disponibilidad. En general, su efecto puede mitigarse replicando unidades funcionales o incorporando un banco de registros o una memoria multipuerto.

Las dependencias de datos y control, por el contrario, pueden reducirse mediante técnicas software. A continuación describiremos las dos más importantes: La *ejecución fuera de orden*, para solventar los riesgos por dependencias de datos, y la *predicción de salto*, que hace lo propio con las dependencias de control.

Tipo de microoperación	Tipo de datos de la aplicación	
	Números enteros	Números reales
Ops ALU con datos enteros	50 %	25 %
Ops ALU con datos reales	0 %	30 %
Lectura de operandos	17 %	25 %
Escritura de operandos	8 %	15 %
Salto	25 %	5 %

**TABLA 3.6:** Frecuencia de uso de las microoperaciones de un programa en función de la naturaleza de la aplicación software. Esto nos da una idea de la influencia de cada tipo de dependencia en el rendimiento del código sobre un procesador con paralelismo a nivel de instrucción, ya que el riesgo por dependencias de datos aparece cada vez que se escribe un operando (8 % ó 15 % para códigos que utilizan exclusivamente números enteros y de punto flotante, respectivamente), y el riesgo por dependencias de control, cada vez que se produce un salto (25 % y 5 %, respectivamente).

### 3.5.1 Ejecución fuera de orden

La **ejecución fuera de orden** (del inglés, *out-of-order execution*) es una estrategia consistente en alterar, en tiempo de ejecución de las instrucciones, su orden de finalización preestablecido en el programa.

El riesgo que se asume esta vez es mucho más elevado: Cuando una instrucción detiene su ejecución ante una dependencia de datos, el procesador comenzará a ejecutar la siguiente en lugar de quedarse parado. Esta segunda instrucción puede así finalizar antes que la primera, por lo que hay que asegurarse de que no viole la semántica del código establecida por el programador, en particular, que no escriba en algún registro cuyo valor vayan a necesitar las instrucciones precedentes que se encuentren detenidas.

riesgos

El número de verificaciones a realizar para garantizar la consistencia de una ejecución fuera de orden es innumerable, y la complejidad de la unidad de control responsable, desbordante. Piénsese que en un procesador actual, el número de instrucciones que pueden estar activas en un momento dado puede superar perfectamente la cincuentena, y que cada una de ellas puede quedarse atascada en el cauce por razones muy variopintas.

complejidad

Es importante hacer notar que la ejecución fuera de orden sólo puede implementarse sobre un procesador que ya sea segmentado y superescalar. Esto es así porque la idea de la ejecución fuera de orden supone que unas instrucciones *adelanten* a otras durante su paso por el cauce, y ésto sólo es posible si alguna de sus etapas dispone de varias unidades funcionales que permitan a otras instrucciones progresar cuando la anterior se encuentra detenida en esa etapa.

sólo si es superescalar



#### Analogía 3.1: LAS INSTRUCCIONES Y EL TRÁFICO EN SU FORMA DE FLUIR

Para entender la necesidad de contar con un procesador superescalar si pretendemos incorporarle ejecución fuera de orden, resulta útil ver el cauce de ejecución como una carretera, sus unidades funcionales como los carriles de la misma, y las instrucciones como los vehículos que por ella circulan, todos en la misma dirección: Un coche sólo puede adelantar a otro cuando dispone de carriles alternativos por los que circular que no le hagan depender de la velocidad del vehículo que le precede.

Al igual que en carretera hay vehículos lentos y otros más veloces, en el cauce de ejecución también coexisten instrucciones rápidas (las enteras) con otras que no lo son tanto (multimedia) y un tercer grupo bastante más lento (las de punto flotante), sin contar con aquellas que mientras no se ejecuten no se sabrá su velocidad (las de acceso a memoria, que para no complicar el ejemplo vamos a descartarlas).

Desde la llegada del Pentium MMX, el procesador dispone de vías de circulación separadas para cada una de estas tres clases de instrucciones, y no por ello debemos pensar que ya es superescalar. La superescalaridad es un aspecto más ligado a la replicación de unidades funcionales de ejecución, esto es, a la facultad de ejecutar simultáneamente una serie de instrucciones *de la misma clase*. La ejecución fuera de orden actúa de forma separada sobre cada una de estas clases, ya que el hecho de que cada clase disponga de su banco de registros propio, le impide entrar en conflicto con las demás.

Así, el tráfico es siempre común en los dos primeros tramos de carretera correspondientes a la búsqueda y decodificación de instrucción, y a partir de ahí, la carretera se bifurca en derivaciones secundarias específicas para cada tipo de vehículo, donde cada uno de ellos sólo podrá adelantar a los de su misma clase, y únicamente en aquellos puntos en los que la arquitectura haya puesto un desdoble de carriles de circulación, esto es, múltiples unidades de ejecución. Esta multiplicidad será el aspecto que determinará el grado de superescalaridad del chip en cada una de sus etapas de ejecución.

La siguiente tabla trata de sintetizar la analogía comentada:

Flujo de vehículos	Flujo de instrucciones	Comportamiento
Coches deportivos	Instrucciones enteras	Muy rápido
Utilitarios	Instrucciones multimedia	Rápido
Tráfico pesado	Instrs. de punto flotante	Lento
Carretera principal durante búsqueda y decodificación	Tramo común del cauce segmentado	Tráfico sincronizado
Derivaciones secundarias durante lectura, ejecución y escritura de operandos	Subcauces especializados	Tráfico desacoplado
	Superescalaridad $\Rightarrow$	Adelantamiento

En cualquier caso, los cambios que se producen durante la ejecución de las instrucciones del programa no deben nunca alterar la tarea global a realizar por el mismo. Por lo tanto, la unidad de control del procesador monitorizará todos y cada uno de los posibles adelantamientos para examinar su licitud, impidiendo aquellos que puedan desembocar en resultados erróneos. La circuitería que se hace necesaria para ello provoca un incremento del área de integración del procesador, pudiendo incluso alterar su camino crítico, con el consiguiente impacto sobre la frecuencia de reloj.

verificaciones

### 3.5.2 Predicción de salto

concepto

La técnica de **predicción de salto** trata de eliminar las dependencias de control de un programa a través de una predicción en la que el microprocesador intenta adivinar lo que hará una instrucción de salto condicional *antes* de que se evalúe su condición de salto. Puesto que sólo puede hacer dos cosas, saltar o no saltar, la probabilidad de acertar en dicha predicción es bastante elevada, pero como en cualquier caso nuestra apuesta es una conjetura, a esta técnica también se le denomina **ejecución especulativa**.

La predicción puede ser de dos tipos:

- **Estática.** El procesador apuesta siempre por la misma premisa, a saber, “el programa salta

siempre que ejecuta una instrucción de salto” o “el programa no salta nunca”.

- **Dinámica.** El procesador es capaz de hacer una suposición u otra dependiendo del comportamiento del programa en su pasado más reciente.

Para implementar la predicción dinámica es necesario registrar el comportamiento histórico de las instrucciones de salto del programa. De esto se encarga la **BTB (Branch Target Buffer)**, o *búfer para los destinos de los saltos*, una nueva unidad funcional del procesador que guarda en cada una de sus entradas los siguientes **campos de información**:

- 1 El código de una instrucción de salto del programa.
- 2 Su dirección de salto, esto es, por donde prosigue su ejecución el código en caso de que la instrucción salte realmente.
- 3 Un grupo de bits que conforman el *historial*, siendo este grupo más numeroso cuanto mayor sea el pasado temporal que quiera archivarse para dicha instrucción. Con el historial crece el coste de la BTB, pero también su probabilidad de acertar en la predicción.

La BTB **funciona** de forma similar a como enseguida veremos que trabaja una memoria caché:

- Si la instrucción de salto no se ha ejecutado nunca, entrará en ella sustituyendo a otra que sea de poco interés para el procesador. En ese caso, al no disponerse de historial para la instrucción, el procesador puede activar mecanismos de predicción alternativos como la propia predicción estática. En la práctica, casi todos los procesadores encadenan estas dos estrategias.
- Si la instrucción ya se ha ejecutado antes, se encontrará en ella con mayor probabilidad cuanto mayor sea el tamaño de la BTB y menor sea el número de instrucciones de salto que se hayan ejecutado entre medias. En cada ejecución anterior, el procesador ha registrado en la BTB si esta instrucción saltó o no, y ahora utiliza dicha información para mostrar un comportamiento adaptativo: Si la instrucción presenta una marcada tendencia al salto, la consigna de la unidad de predicción de saltos será “suponer salto realizado”, y si su historial muestra un perfil secuencial, la consigna será “proseguir ejecutando secuencialmente”.

Por otra parte, aunque hasta ahora sólo hemos hablado de los saltos condicionales, la BTB también soluciona las dependencias de control debidas a **saltos incondicionales**. Aunque pueda pensarse que éstos no introducen conflicto alguno debido a que su predecible comportamiento (siempre saltan), hay que resaltar que un salto incondicional lleva asociado el cálculo de una dirección de salto. Este cálculo se efectúa normalmente en una etapa bastante tardía del cauce segmentado, lo que detiene la entrada de nuevas instrucciones en él hasta tanto no se sepa por dónde prosigue el programa. Dado que un acierto en la BTB ya proporciona la dirección del salto, el conflicto desaparece, no siendo necesario esperar a la conclusión de dicha etapa.

En este sentido, podemos considerar a la BTB responsable de llevar a cabo una estrategia de anticipación en el tratamiento de las dependencias de control, de igual manera que también existen mecanismos más sofisticados para llevar a cabo la anticipación de valores en la resolución de las dependencias de datos.

BTB

información en la BTB:

- código

- dirección

- historial

funcionamiento

fallo en la BTB:

predicción estática

acierto en la BTB: comportamiento adaptativo

saltos incondicionales

anticipación

## Memoria caché integrada

Tradicionalmente se le presta mucha atención a la rapidez de cómputo, cuando en el interior del procesador tienen lugar muchos más accesos a datos que operaciones en sí.

Veámoslo con un ejemplo. La [figura 3.8](#) muestra la instrucción más típica de un procesador. En ella se requieren cuatro accesos a distintas fuentes de información para realizar una sola operación. El procesador se encuentra así mucho más limitado por la obtención de datos que por la celeridad de su computación.

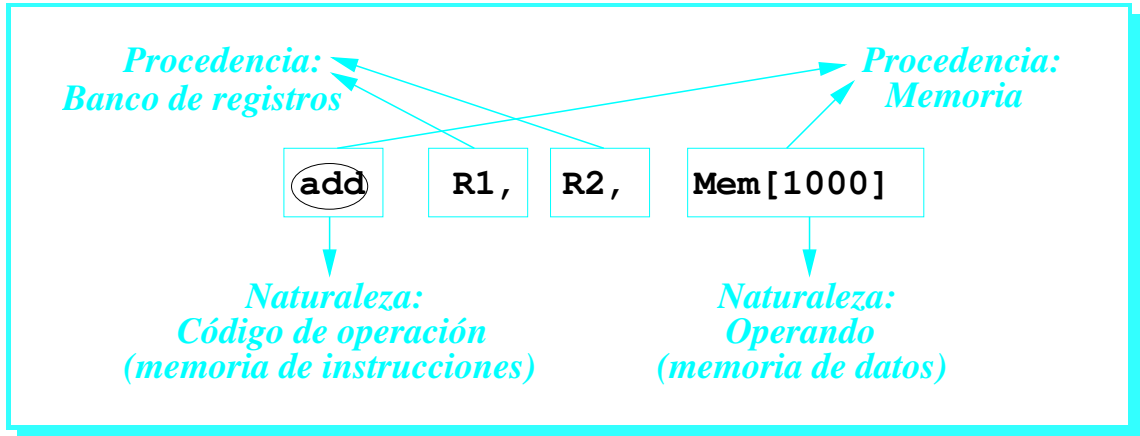


FIGURA 3.8: Las tareas involucradas en la computación de una instrucción típica por parte del procesador. Se han señalado con una elipse las operaciones a realizar, y con un rectángulo los accesos a datos. La proporción es de cuatro a uno en favor de estos últimos.

De los datos de la instrucción `add` en la [figura 3.8](#), dos proceden de memoria y otros dos del banco de registros del procesador. Pero como ya sabemos que un procesador actual dispone de un elevado grado de paralelismo a nivel de instrucción, la pregunta que surge es: ¿Cómo puede el procesador alimentarse del enorme volumen de información que esa operativa demanda, y todo ello a una velocidad superior a 1 GHz?

el dilema

la solución

pág. 73

Diremos en primer lugar que como la memoria no es tan rápida, el número de etapas de segmentación se verá ampliado cuando se procese información de este tipo. Y en segundo lugar, que se habilitarán hasta cinco unidades funcionales distintas dedicadas a la exclusiva tarea de proporcionar datos al procesador. Estas unidades funcionales serán presentadas en la [sección 3.4.2](#). Lo que ahora pretendemos ilustrar es su necesidad.



### Ejemplo 3.10: LA NECESIDAD DE UNA JERARQUÍA DE MEMORIA

Según nuestras propias estimaciones, hoy en día la velocidad real de acceso a memoria principal es unas 3.000 veces más lenta que el acceso al banco de registros interno del procesador.

Esto significa que de no contar con una jerarquía de memoria que haga de intermediario entre estos dos niveles, un Pentium 4 de 2 GHz trabajaría en la actualidad a 666 KHz, puesto que al menos tendría que acceder a memoria una vez por cada instrucción que ejecuta para recoger su código de operación. Esa frecuencia de trabajo está por debajo de la frecuencia original del ENIAC, el primer computador de la historia, que data del año 1945.





### Analogía 3.2: EL PAPEL DE LA FOTOCOPIADORA, O LOS RECURSOS QUE HACEN FALTA PARA SUMINISTRAR DATOS CON CELERIDAD AL PROCESADOR

Imagínese que compramos una fotocopidora en Málaga, nos preocupamos por adquirir un modelo ultrarrápido que realiza una copia en 3 sg., y luego tenemos que ir a Almería a por el folio de papel cada vez que queramos efectuar una copia. Este ejemplo coincide con un procesador sin caché, pues las 2 horas y media del trayecto encajan con el ratio 1:3000 que acabamos de apuntar en el ejemplo anterior.

Habilitar el cajón alimentador de papel tiene un propósito similar al del banco de registros del procesador: Garantizar el suministro inmediato, y es tan necesario, que el sistema no se concibe sin él. Pero se trata de un recurso limitado, así que cuando la bandeja agota sus folios, tenemos de nuevo visita pendiente a Almería.

Apilar unos paquetes de folios junto a la fotocopidora tiene el mismo efecto que habilitar una caché de primer nivel en el procesador: Cuando el sistema agote la bandeja, tomará el papel de otro sitio al que se tarda tres o cuatro veces más en acceder, pero que permite el funcionamiento autónomo por un tiempo, evitando tener que salir al exterior.

Con esta solución, el sistema funciona muy bien, pero si ahora decidimos adquirir una docena de fotocopadoras más, la solución se queda pequeña, como también se le quedó pequeña una sola caché al procesador cuando optó por replicar unidades funcionales en sus primeros diseños segmentados y superescalares.

Esto motivó la aparición de un segundo nivel de caché interno al procesador, dotado de una mayor capacidad, pero algo más lento. Lo mismo que si decidimos adquirir para nuestra fotocopidora un almacén contiguo en el que dar cabida a varios millones de folios. El proveedor principal de papel sigue estando en Almería, pero ahora nuestro sistema ha conseguido una autonomía propia de varios millones de copias. Las mismas operaciones que hoy en día puede realizar un microprocesador sin requerir el concurso de la memoria principal.

La memoria caché es objeto de nuestro estudio en el [capítulo 11](#), por lo que el lector puede realizar una incursión allí si no se encuentra suficientemente familiarizado con este concepto. Respecto a nuestras estimaciones sobre la lentitud de acceso a memoria principal comparada con caché, provienen de un análisis realizado en la [sección 13.2.2](#) en el contexto del futuro que dibujamos para la memoria caché.

← p.115/Vol.2

← p.160/Vol.2

Y es que cuando se trata de hablar de la memoria caché, ya no sabemos donde encaja mejor: Por el nombre y por su pasado, el [capítulo 11](#) es su ubicación natural; por su presente y futuro, su sitio está aquí, dentro de la cobertura dedicada al procesador, pues para eso hace ya un tiempo que al menos dos niveles de caché van enquistados dentro de él. Así que aquí la trataremos como una extensión del banco de registros del procesador, ilustrando las diferencias en su funcionamiento, su interrelación con las unidades funcionales del procesador, y los diferentes buses por los que establece su estrecha vinculación con él. Y dejaremos para el [capítulo 11](#) su cobertura como ente propio, donde visitaremos su tecnología distintiva y la serie de implementaciones comerciales de que dispone como chip independiente.

nuestro  
tratamiento

Cuando uno afronta la compra de un procesador, su caché es uno de los puntos clave en los que debe fijarse. Y sin embargo, la documentación que los distribuidores proporcionan al respecto es pírrica, indicando únicamente su tamaño, y gracias, porque en más ocasiones de las que sería

pieza clave

deseable ni siquiera indican a qué caché se están refiriendo, y hay al menos tres dentro de un procesador actual.

**metamorfosis** A nivel técnico, la laguna tampoco es despreciable. Conocemos multitud de libros que explican de maravilla los principios de localidad espacial y temporal y los conceptos de línea o conjunto de caché, parámetros que irán eternamente ligados a su estructura interna. Pero la caché ha sufrido tal grado de metamorfosis en los últimos años, que podemos indicarle que en los libros anteriores a 1995 encontrará buenos fundamentos sobre ella, pero una contribución más bien escasa acerca de las claves de su rendimiento actual.

**desafío** La cantidad de variantes que han surgido desde esa fecha ha supuesto para nosotros todo un desafío a la hora de poner un poco de orden en esta sección. Trataremos de apoyarnos en continuas referencias a modelos comerciales con objeto de situar cada variante en su contexto dentro del mercado.

**acelerador** La complejidad anunciada no deja de ser, además, una paradoja. Conserve en mente una verdad de perogrullo: *Toda caché tiene como objetivo exclusivo acelerar el acceso a memoria; ninguna variante incorpora funcionalidad adicional al sistema, o espacio de almacenamiento suplementario. Por tanto, por lo que al diseño de cachés respecta, las cosas, o se hacen muy rápido, o no sirven absolutamente para nada.* Como esta premisa se encuentra bastante reñida con las ideas sofisticadas, veremos que en no pocas ocasiones el camino más simple será también el más acertado.

#### 4.1 ▶ Breve sinopsis histórica

**banco de registros** En sus orígenes, los microprocesadores apenas disponían de unas pocas celdas de memoria dentro del propio chip para almacenar los operandos fuente y destino de sus instrucciones. Estas celdas se estructuraban en un banco de registros para almacenar los operandos fuente y destino de las instrucciones, que se referenciaban de forma directa y explícita en el propio código de instrucción

**en los años 80** Ya durante los años 80, el concepto de memoria interna al chip se extendió a una memoria caché de unos pocos kilobytes, hito que situamos en la cuarta generación de microprocesadores para PC (1989) con la llegada del 80486 de Intel y el 68040 de Motorola. Esta memoria respondía mucho más rápido que la memoria principal externa ubicada en la placa base, por lo que si el cuello de botella del sistema se situaba en el acceso a memoria, se lograba una sustancial mejora en el rendimiento del microprocesador en general.

**en los años 90** En la década de los 90, los diseñadores de microprocesadores fueron aumentando progresivamente el tamaño de estas memorias y estructurándolas en niveles. El buque insignia de estas transformaciones fue el procesador Alpha de Digital, que en su versión 21264 (1995) disponía ya de 112 Kbytes de caché interna organizada en tres memorias a dos niveles, resultando una jerarquía de memoria interna que es fiel reflejo de lo que hoy encontramos en los procesadores contemporáneos.

**p. 8/Vol. 2** La [figura 9.2](#) ilustra la ubicación de todos estos niveles jerárquicos de memoria en el contexto de la cuarta, quinta, sexta y séptima generación de microprocesadores.

Dentro ya de los modelos comerciales, tenemos ejemplos concretos en abundancia:

**pág. 74** ❶ El Pentium. La [foto 3.2](#) muestra su aspecto externo (arriba) y su área de integración (abajo), donde las dos cachés L1 para datos e instrucciones se han delimitado en su parte izquierda.

**pág. 76** ❷ El Pentium Pro. La [figura 3.9](#) nos muestra una radiografía de este procesador, donde la caché L2 se suministra ya junto al microprocesador en lugar de venir integrada en la placa base.

- ③ El Pentium II. En la [figura 3.10](#) podemos observar un diagrama de cómo tiene dispuestos sus dos niveles de memoria caché. ➡ [pág. 77](#)
- ④ El Pentium III. En su versión de 0.18 micras, la caché L2 se incluye ya dentro del propio chip del procesador, según se aprecia en la parte derecha de la [foto 3.2](#). Este estatus permanece vigente durante toda la séptima generación, alcanzando ya el momento presente. ➡ [pág. 74](#)

La [sección 3.4.4](#) formalizará todas las posibles ubicaciones. Hasta llegar allí, la [sección 3.4.2](#) se ocupará de profundizar más en los diferentes niveles de memoria de forma separada, y la [sección 3.4.3](#) dará cuenta de algunas optimizaciones interesantes. ➡ [pág. 81](#)  
➡ [pág. 77](#)

## Jerarquía

### ◀ 4.2

Los niveles de memoria internos al procesador dan lugar a una jerarquía en función de la proximidad a su núcleo de ejecución. Comenzando por la capa más interior, podemos distinguir:

- ① **El banco de registros.** Su **tamaño** suele estar comprendido entre 32 y 512 registros, siendo mayor en microprocesadores de tipo RISC (ver [sección 3.5.2](#)). Con respecto a la **anchura** de cada registro, coincide con el tamaño de palabra que puede procesar la ALU. Por ejemplo, en un Pentium III, el banco de registros para datos enteros es de 32 bits, mientras que el banco de registros que alberga a los datos reales o de punto flotante tiene una anchura de 80 bits, pues ésa es la anchura de las unidades funcionales sumadora, multiplicadora y divisora que conforman su FPU. tamaño  
➡ [pág. 97](#)  
anchura
- ② **El búfer de prebúsqueda de instrucciones/datos.** Relacionado con el paralelismo a nivel de instrucción en general, y dentro de él más con el carácter superescalar del procesador, ya que la ejecución simultánea de varias instrucciones en cada ciclo de reloj asume implícitamente la existencia de algún mecanismo que suministre las instrucciones al ritmo que el procesador las ejecuta. superescalar  

Este búfer no es más que un almacén intermedio, normalmente implementado mediante una cola <sup>3</sup>, donde se han traído las instrucciones candidatas a ser ejecutadas en un futuro inmediato, y desde donde pueden dirigirse al procesador con una agilidad extrema en cuanto sean requeridas por éste. De esta manera, se cubre una doble función:

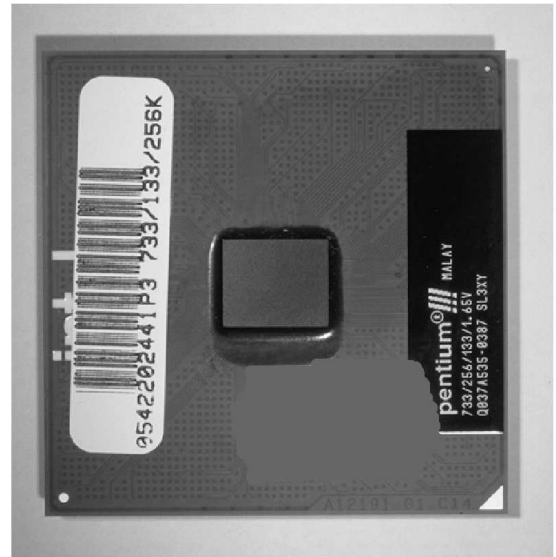
  - ① Desacoplar la entrada de información a las unidades funcionales del proceso de ejecución de instrucciones en sí. desacoplo
  - ② Actuar de acelerador en la captura de dicha información. acelerador
- ③ **La caché de primer nivel (L1).** Suele estar separada en dos: Una para datos y otra para instrucciones. Esta separación viene como consecuencia del carácter segmentado del procesador. En efecto, dado que una de las etapas en que se divide la ejecución de instrucciones accede a la caché de instrucciones para traerse el código de la instrucción a ejecutar, y otra etapa accede a la caché de datos para traerse los operandos fuente o almacenar el operando destino, la ejecución simultánea de varias instrucciones en distintas etapas del cauce segmentado pasa por implementar estas cachés como componentes independientes. segmentación  

Obsérvese que, vista esta dualidad entre instrucciones y datos en el primer nivel de caché, sólo existe un nivel que le precede en dirección al procesador, ya que el banco de registros se interpone entre éste y la caché de datos, mientras que el búfer de prebúsqueda de instrucciones se ubica por delante de la caché de instrucciones. dualidad
- ④ **La caché de segundo nivel (L2).** Una caché más grande pero más lenta que la anterior, albergando conjuntamente datos e instrucciones, que se coloca en la placa base hasta el ubicación

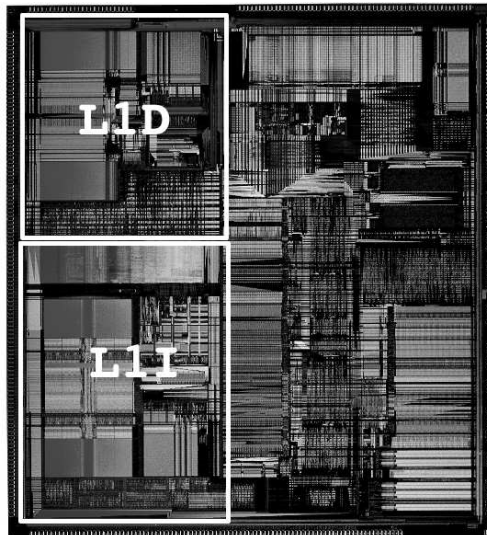
<sup>3</sup>Una estructura de datos FIFO - First In First Out - en la que los datos salen en el orden en el que llegan.



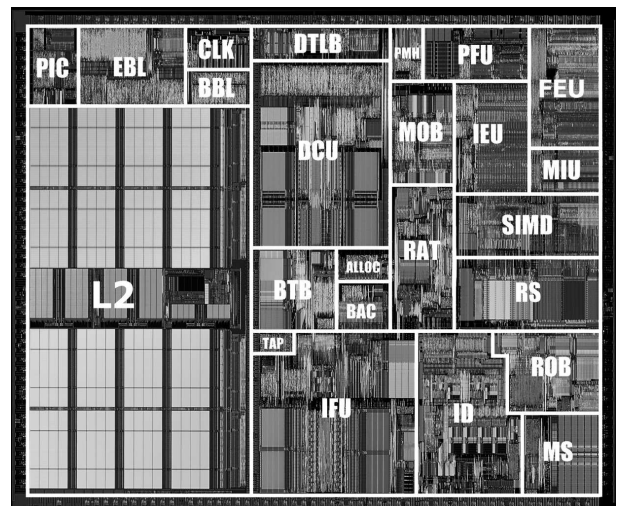
(a)



(b)



(c)



(d)

Fotos de las áreas de integración: Cortesía de Intel

FOTO 3.2: Las memorias cachés se van integrando en el chip procesador conforme se avanza en su integración. En el primer Pentium de 0.35 micras (foto a) sólo se disponía de dos cachés L1 para datos e instrucciones, ambas de 8 Kbytes, y marcadas en el área de integración (foto c). En el Pentium III de 0.18 micras (foto b) estas cachés son ya de 16 Kbytes cada una, marcadas en la foto (d) como DCU - centro arriba - e IFU - centro abajo. Además se incluye una caché L2 de 256 Kbytes - parte izquierda. Las fotos (a) y (b) están tomadas a escala real, por lo que puede apreciarse que el área de integración de los 28 millones de transistores del Pentium III (rectángulo de la parte central) es muy inferior a la del Pentium de sólo 3.1 millones.

Pentium (accesible a través del bus local), y que ha pasado a formar parte integrante del propio chip del procesador en casi todos los diseños contemporáneos.

Esta mejora ha venido obligada por la evolución de las aplicaciones software, cuya mayor dimensión ha hecho que la caché de primer nivel sea insuficiente para contener el uso masivo de datos externos al microprocesador. La caché L2 se hace entonces necesaria pa-

necesidad

Tipo de memoria interna al microprocesador	Hecho con el que puede relacionarse	Generación del microprocesador en la que aparece
Banco de registros	Indispensable en la arquitectura de un microprocesador	Primera
Búffer de prebúsqueda de instrucciones	Paralelismo a nivel de instrucción: Superescalaridad	Quinta
Memoria caché de primer nivel (L1)	4 Gbytes de memoria direccionable Paralelismo a nivel de instrucción: Segmentación Mejoras en la integración	Tercera Cuarta Quinta
Memoria caché de segundo nivel (L2)	Nuevo formato PCB multichip con zócalo tipo Slot para el microprocesador	Sexta

TABLA 3.7. Los diferentes niveles que componen la jerarquía de memoria interna de un microprocesador junto al hecho más relevante que provoca su aparición y la generación de microprocesadores en que esto sucede.

ra aprovechar la propiedad de *localidad* de los programas en el acceso a memoria, que se encuentra un tanto diluida por el enorme área de memoria sobre el que éstos se extienden.

La [figura 3.9](#) muestra la ubicación física de todos estos niveles de memoria para el primer microprocesador de Intel que los incorporó conjuntamente: El Pentium Pro. En la parte superior de la figura puede apreciarse el enorme espacio ocupado por el patillaje externo y la composición de su núcleo multichip: la caché L2 se integra separadamente en el chip de la derecha, mientras que el chip de la izquierda contiene al resto de unidades funcionales de la CPU (incluyendo las cachés de primer nivel), cuya distribución geográfica presentamos en la parte inferior de la figura.

El chip de la caché aglutina 15.5 millones de transistores, mientras que el de la CPU sólo contiene 5.5 millones. Y es que ya avisamos en la [sección 3.2](#) de que los transistores de la caché ocupan un espacio de silicio inferior al de otras unidades funcionales de la CPU. En general, los diseños actuales tienen una marcada tendencia a incluir cachés cada vez más grandes, que en los casos más extremos han llegado a copar hasta el 80 % del total de transistores.

A modo de recopilación para los niveles de memoria que componen la jerarquía interna del microprocesador, mostramos en la [tabla 3.7](#) la secuencia cronológica de aparición de todos ellos y su relación con las distintas generaciones de microprocesadores.

La arquitectura de un microprocesador con todos estos niveles de memoria internos contempla dos vías de comunicación separadas (ver [figura 3.10](#)):

- ❶ El **bus backside**, que conecta los dos chips que componen el microprocesador, y que se utiliza para el transporte de datos desde la caché L2 a la caché L1 y el microprocesador.
- ❷ El **bus frontside**, que conecta el patillaje externo del microprocesador con el juego de chips de la placa base y que se encarga de transferir la información entre la memoria principal y esta caché de segundo nivel.

A partir de ahora, utilizaremos las denominaciones de *bus trasero* para el primero y *bus frontal* para el segundo, sinónimo para nosotros del sempiterno bus local de la placa base que desemboca en el patillaje del procesador cuando no existen cachés de por medio.

➡ [pág. 76](#)  
visión de conjunto

➡ [pág. 50](#)  
espacio en silicio

recopilación

➡ [pág. 77](#)

bus trasero

bus frontal

sinónimos

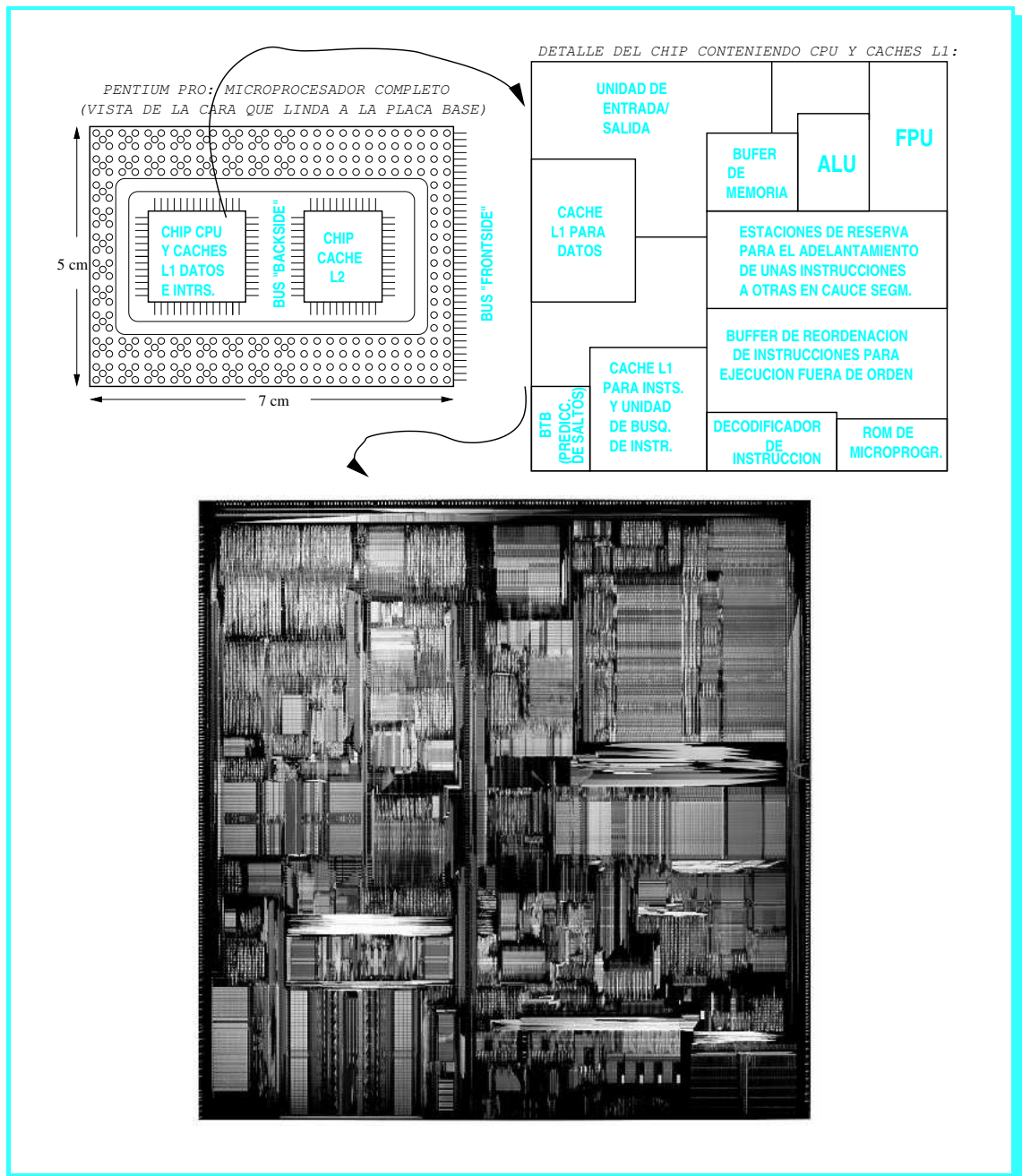


Foto interior: Cortesía de Intel

FIGURA 3.9: Radiografía del Pentium Pro para ilustrar la ubicación de cada uno de los niveles de su jerarquía de memoria. Los diagramas están realizados a escala real para mostrar el enorme espacio físico dedicado a los 387 pines del patillaje externo. Rodeados del mismo, arriba a la izquierda podemos distinguir dos chips: Uno para la caché L2, y otro para la CPU en sí y las cachés de primer nivel. A la derecha: Desglose de las unidades funcionales. Abajo: Detalle del área de integración en una foto real que detalla la constitución interna del chip.

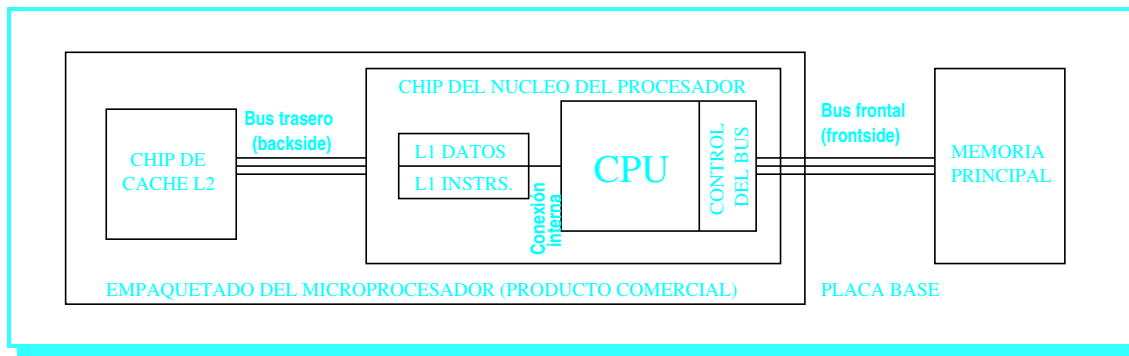


FIGURA 3.10: Las memorias cachés y los buses asociados a la arquitectura de un microprocesador de sexta generación.

## Optimizaciones

### 4.3

Tomando como partida un microprocesador dotado de los niveles de memoria caché L1 y L2 y de los buses frontal y trasero ya descritos, el abanico de optimizaciones que pueden llevarse a cabo en su interior es muy amplio. A continuación recogemos las que a nuestro juicio son más representativas de la situación actual del mercado.

#### 4.3.1 Buses desacoplados

La lógica de control y sincronización interna del microprocesador permite el funcionamiento independiente y simultáneo de los buses frontal y trasero. Esta característica ha sido desarrollada por muchos fabricantes. Intel la incorpora en sus procesadores domésticos a partir del Pentium Pro con el nombre de **Dual Independent Bus (DIB)** (ver [sección 14.4.3](#)).

DIB

➔ p.197/Vol.2

Esta optimización entra en juego cuando el dato que el procesador busca no se encuentra en ninguna de las cachés internas, requiriéndose un acceso a memoria principal. En ese caso, se debe traer una línea de datos íntegra de memoria principal, pero el procesador podrá trabajar con el dato que solicitó mientras se procede de forma simultánea a la carga de la línea en la caché L2.

#### 4.3.2 Caché no bloqueante

Cuando una caché es **no bloqueante**, pueden realizarse nuevos accesos a la misma mientras uno o más fallos están siendo cursados.

La conjunción de esta característica con el uso de buses desacoplados nos permitirá que cuando un dato no se encuentre en la caché L2, puedan cursarse nuevas peticiones a L2 por el bus trasero mientras se resuelve el fallo anterior por el bus frontal. Si es en la caché L1 donde no hemos encontrado el dato, el funcionamiento es idéntico: El bus trasero se encarga de traer este dato de L2 al tiempo que nuevas peticiones del procesador pueden acceder internamente a L1.

funcionamiento

Las cachés no bloqueantes se incorporan a los microprocesadores para PC coincidiendo en el tiempo con la llegada del segundo nivel L2. En el caso que nos sirve de referencia, la L2 es la aliada natural de la arquitectura DIB en el Pentium Pro, perdurando ambas en todos los diseños posteriores de Intel (Pentium II, Celeron, Pentium III y Pentium 4).

origen

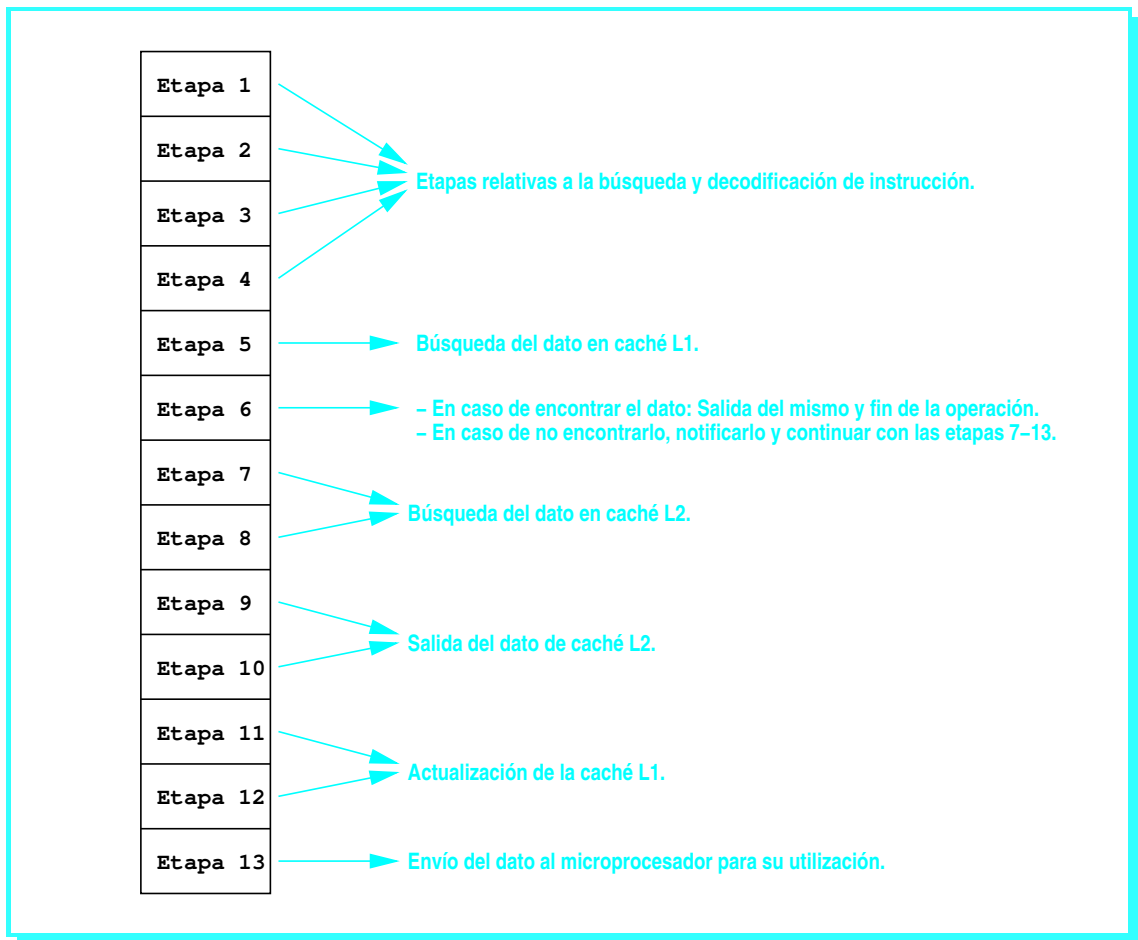


FIGURA 3.11: El diseño supersegmentado de los procesadores 21164 y 21264 de Digital contempla hasta un máximo de nueve etapas de segmentación dedicadas a la operativa de funcionamiento de sus múltiples cachés segmentadas.

### 4.3.3 Caché segmentada

Se trata de aplicar al funcionamiento interno de la memoria caché la idea de la **segmentación** que ya explicamos sobre los procesadores en la [sección 3.3.1](#).

← pág. 59

2 o 4 etapas

Por la velocidad intrínseca de la caché, lo más común es encontrarla segmentada en dos etapas si es interna o en cuatro si es externa y/o funciona a la mitad de frecuencia que el núcleo del procesador: En la primera etapa se realiza la consulta del dato a localizar en L1, y en la segunda se extrae dicho dato en caso de ser encontrado en L1 o se lanza la petición a L2 en caso contrario; suponiendo que el mayor tamaño de ésta última le obliga a funcionar a la mitad de velocidad que el procesador (y su caché L1), su funcionamiento sería similar pero consumiendo dos etapas para la salida del dato correspondiente.

supersegmentados

Alpha

La segmentación de la caché interna encuentra su ubicación más natural en procesadores supersegmentados, en los que una subdivisión del ciclo de reloj asociado a la respuesta de las cachés posibilita la distinción de estas subetapas de búsqueda y salida de datos. Por ejemplo, en los diseños 21164 y 21264 del procesador Alpha de Digital, esta subdivisión permitió la consecución de una frecuencia de reloj que marcó el techo de referencia durante toda la década de los 90, y donde un acceso a caché puede gastar hasta nueve ciclos, segmentados según se detalla en el diagrama de la [figura 3.11](#).



Desde el punto de vista del rendimiento de una jerarquía de memoria dotada de dos niveles de caché L1 y L2, podría resultar un tanto paradójico el hecho de que existan casos como el del Pentium Pro o los Pentium II y III Xeon en los que la caché L2 funciona a la misma velocidad que la caché L1. Sin embargo, son las tres optimizaciones vistas hasta ahora las que dan sentido a semejante configuración:

- ❶ La presencia de vías de comunicación desacopladas para el acceso a estas cachés permite reducir la probabilidad en la aparición de riesgos estructurales en procesadores con un elevado grado de paralelismo a nivel de instrucción, al permitirse que una instrucción acceda a la caché L1 de instrucciones, otra a la L1 de datos, y una tercera a la caché L2. acceso múltiple
- ❷ La utilización de cachés no bloqueantes y segmentadas permite al nivel L1 independizar el funcionamiento del procesador y la caché L2, y así, cuando no se encuentre un dato en L2, L1 puede estar sirviendo otros datos al procesador al tiempo que L2 busca el suyo en memoria principal. funcionamiento  
desacoplado
- ❸ Al nivel de una petición individual a memoria, la diferencia entre acceder a caché L1 ó L2 la detallaremos en la [sección 3.4.7](#), dedicada al análisis de su rendimiento. acceso  
individual  
☛ pág. 88

#### 4.3.4 Caché con lectura anticipada

Esta caché supone un refinamiento en el modo de proceder de una caché no bloqueante. Tiene por objeto acelerar la lectura de un dato por parte del procesador cuando éste no se encuentra en caché, simultaneando allí la recepción del bloque que lo contiene con el envío del dato concreto al procesador. refinamiento

En este sentido, disponemos de dos estrategias básicas: dos estrategias

- **Early restart** (reinicio prematuro). Aguarda a que llegue la palabra del bloque requerida por el procesador para enviarla al mismo, evitando por tanto la espera restante hasta que se complete la transferencia del bloque.
- **Critical word first** (primero la palabra crítica). Fuerza al siguiente nivel de memoria a enviar primero el dato requerido por el procesador seguido del resto del bloque.

Aunque la primera técnica pueda parecer peor a simple vista, en la práctica suele ser mucho más frecuente encontrarla en los diseños comerciales por el hecho de ser la única que tiene en cuenta las múltiples estrategias que las memorias habilitan en la actualidad para enviar datos consecutivos por grupos o ráfagas. Las [secciones 10.13.4.2](#) y [10.13.5.1](#) nos descubren este tipo de estrategias en el contexto de las memorias SDRAM y DDRAM, respectivamente. uso comercial  
☛ p. 60/Vol. 2  
☛ p. 69/Vol. 2

#### 4.3.5 Caché víctima

Se trata de incorporar una diminuta caché L2 de unas pocas líneas, normalmente entre 8 y 16, que actúa de repositorio donde se le va a dar una oportunidad de permanecer a los **descartes** más recientes que tenga que realizar la caché L1, en lugar de enviarlos directamente al destierro que supone su expeditivo alojamiento en memoria principal. segunda  
oportunidad



### Ejemplo 3.11: LA CACHÉ VÍCTIMA DEL RENOVADO K7

Con la llegada de las 0.18 micras al microprocesador K7, AMD decidió mejorar la eficiencia de su sistema de cachés integradas incorporando un búfer o caché víctima con capacidad para albergar 8 líneas de caché (cuyo tamaño por cierto aumentó de 32 bytes a 64 bytes respecto a los primeros K7).

Se habilitó así un espacio intermedio de 512 bytes que restaba una media de 3 ciclos de penalización en el caso de no encontrar el dato solicitado en el primer nivel de caché.

#### 4.3.6 Caché de tercer nivel (L3)

Adicionalmente a los niveles de memoria descritos hasta la caché L2, podríamos plantearnos la incorporación de una **caché de tercer nivel (L3)**, en los casos en los que viniese integrada en una placa base que a su vez acoplara un microprocesador con memoria interna hasta el nivel L2.

Sin embargo, no parece que este tercer nivel de caché vaya a introducirse en el interior del microprocesador a corto plazo: Fijándonos en la forma en que emergen los sucesivos niveles de memoria a lo largo de la historia, podemos inferir que la aparición de un nivel nuevo se encuentra asociado al paulatino distanciamiento de sus niveles anterior y posterior, hasta crear un vacío intermedio lo suficientemente amplio como para que sea ocupado por este nuevo nivel. Y en el contexto actual de la evolución de las memorias, no está sucediendo esto. Repasemos los puntos clave:

L2 crece

- 1 La L2 interna está creciendo bastante en tamaño, y aunque la memoria principal es previsible que también lo haga, la distancia se sigue manteniendo en el mismo orden de magnitud. Esto reduce la dependencia que el procesador tenía del exterior.

MP y bus local  
más veloces  
p.73/Vol.2  
p.165/Vol.2

- 2 La memoria principal está ganando en velocidad, con sucesivas mejoras a 166 y 200 MHz y diseños que tienen a su alcance frecuencias muy superiores (ver [secciones 10.13.6](#) y [13.3.3](#)). Con ella, la frecuencia del bus local va también en progresivo aumento. Bajo esta situación, una hipotética caché L3 externa se encontraría bastante más próxima en velocidad a la memoria principal que a la caché L2 interna, justo lo contrario de lo que sería deseable para justificar su presencia.

L1 y L2  
progresan

- 3 La brecha en velocidad que abre el procesador respecto a la memoria no es ahora como lo fue antaño, puesto que arrastra con él a sus cachés L1 y L2. Esto retrasa la creación del hueco que la caché L3 necesita para instalarse cómodamente en la arquitectura PC.

esquema similar

Sea como fuere, el funcionamiento de un microprocesador con tres niveles de caché es muy similar al descrito para dos niveles, prevaleciendo todas las ideas referentes a mayor tamaño y menor velocidad conforme nos alejamos del núcleo del procesador, y habilitando una nueva vía que comunicaría la caché L2 con la L3, dejando el bus frontal (frontside) para las comunicaciones externas con memoria principal.

Las posibilidades para adoptar operaciones concurrentes serían ahora aún mayores, pero si hemos visto un cauce segmentado de trece etapas cuando abordábamos un diseño concurrente

con dos niveles, debemos presuponer que para tres niveles sobrepasaríamos fácilmente las veinte etapas, y no resulta nada fácil controlar un cauce de semejante profundidad sin que la complejidad de la circuitería resultante desborde nuestro presupuesto.

## Proximidad al núcleo del procesador

## ◀ 4.4

Antes de nada, conviene tener claro a qué nos referimos con “caché interna (o externa) al microprocesador”, ya que existen diferentes acepciones y formas de entender las cosas que se confunden más a menudo de lo que sería deseable.

- ❶ Desde un punto de vista comercial, la frontera entre lo externo e interno se encuentra delimitada por lo que nos suministra el vendedor cuando lo adquirimos (esto es, lo que viene *de serie* con él). comercial

Bajo esta perspectiva, todas las cachés L1 desde el 80486 serían internas, mientras que para las L2 serían internas las del Pentium Pro Pentium II, CeleronA, Pentium III y Pentium 4, y todos los Xeon e Itanium por parte de Intel, y las del K6-III y K7 por parte de AMD, mientras que serían externas las del Pentium, MMX, K5, K6 y K6-2. ejemplos
- ❷ Desde una vertiente funcional, podemos entender como *interno* todo aquello que fuese capaz de transmitir los datos al procesador siguiendo el mismo ritmo de su frecuencia de reloj. funcional

Según este criterio, sólo una pequeña parte de las cachés internas desde el punto de vista comercial lo serían atendiendo a criterios de funcionalidad. Concretamente, seguirían siéndolo todas las cachés L1, mientras que para las L2, permanecerían en este grupo todas las del Pentium Pro, Celeron y Pentium III a 0.18 micras, Xeon, K6-III y sólo algunas configuraciones selectas del Itanium y K7. ejemplos
- ❸ Desde el punto de vista de la integración, *interno* sería todo aquello que cohabita en el mismo chip de silicio de la CPU. integración

Esto mantendría como internas todas las cachés de primer nivel, mientras que descartaría un gran número de cachés de segundo nivel, dejando tan sólo la de 128 Kbytes del CeleronA y la de 256 Kbytes de los Pentium III y 4 (todos en su versión de 0.18 micras) por parte de Intel, y la de 256 Kbytes del K6-III y los K6-2+ y K7 de 0.18 micras por parte de AMD. ejemplos

Todos los Pentium Pro, II y las versiones Katmai del Pentium III integran la L2 en un segundo chip, y de igual forma proceden los K7 Athlon de 0.25 micras de AMD; además, a excepción del Pentium Pro, que integra conjuntamente el área de datos y el controlador, los Pentium II, III y K7 aprovechan su formato de cartucho cerámico para escindir el controlador de caché en un tercer circuito integrado que se coloca entre el chip del procesador y el chip del área de datos de caché.

La razón por la que se integran estos elementos en chips aparte en cuanto el conjunto alcanza un tamaño de cierta consideración tiene una clara justificación en el coste asociado a su fabricación: Uno de los principios básicos del diseño de circuitos integrados cuantifica el coste de integración de un circuito como una función de la cuarta potencia del parámetro “área de integración”. función de coste

Por lo tanto, sobrepasado un umbral en el que aún tiene cabida la L1, la fabricación en un sólo área es ocho veces más cara que la variante consistente en integrar dos áreas con la mitad de superficie, una para L2 y otra para el microprocesador. No debe extrañarnos por ello que los modelos de microprocesador que evolucionan fusionando la caché L2 en el mismo chip del procesador suelen recortar el tamaño de ésta, e incluso esperar para hacer el cambio con la transición hacia distancias de integración más pequeñas. ocho veces más cara



FOTO 3.3: (a) El microprocesador K7 de 0.18 micras, que cuenta ya con una caché L2 integrada de 256 Kbytes. (b) Detalle del área de integración donde se encuentra esta caché.

Aunque desde una perspectiva rigurosa la mejor definición de caché interna a un microprocesador es la que se encuentra integrada de forma indivisible en su interior, somos conscientes de que el usuario percibe como interno a su microprocesador todo aquello que le suministran dentro del producto que adquiere en la tienda, así que para evitar inconsistencias, hemos adoptado la siguiente **nomenclatura**:

nomenclatura

- externa**      ❶ Será **externa** toda caché ubicada en la placa base y por tanto ajena a lo que es la adquisición del procesador. Cuando lleguemos a la descripción de las características de cada procesador, ésta en concreto vendrá simbolizada mediante el símbolo ✕.
- interna**      ❷ Será **interna** toda caché que venga con el procesador cuando lo compramos pero que se encuentre colocada en un chip aparte. Distinguiremos esta característica en lo sucesivo con el símbolo ✓.
- integrada**    ❸ Será **integrada** toda caché incluida con el procesador e integrada en el mismo chip de la propia CPU. Este rasgo lo denotaremos mediante el símbolo ☆. La [foto 3.3](#) muestra el aspecto de un K7 de 0.18 micras y un detalle de su área de silicio en la que se encuentra integrada una caché L2 de 256 Kbytes.

🏠🏠 **Analogía 3.3: LA UBICACIÓN DEL PAPEL DE LA FOTOCOPIADORA COMO MEDIO PARA RECONOCER CADA TIPO DE CACHÉ**

Continuemos con esa visión de la caché L2 como un almacén contiguo del que proveer con millones de folios a esa fotocopidora que era nuestro procesador. Si ese almacén fuese una ampliación del habitáculo donde está la máquina, estaríamos ante lo que es una caché integrada: Compartiendo el mismo espacio físico que el procesador y la caché L1, y sólo ligeramente más lenta que ésta, por no encontrarse tan a mano.

Proveedor de papel en una máquina fotocopidora	Proveedor de información en un procesador actual
Bandeja de carga de folios	Banco de registros
Torre de folios apilados siempre en el habitáculo de la máquina	Caché L1 siempre integrada junto al procesador
Almacén junto a las dependencias de la máquina	Caché L2 integrada en el procesador
Almacén en el mismo edificio de nuestra empresa	Caché L2 interna adquirida con el procesador
Papelería ajena de ámbito local	Caché L2 externa
Proveedor en Almería	Memoria principal

TABLA 3.8: Resumen del conjunto de analogías existentes entre una fotocopidora como procesador y el papel para copias como los datos necesarios para la ejecución de sus instrucciones.

La ubicación del almacén en el mismo edificio de nuestra empresa es el símil de la caché interna: Hay que salir del habitáculo, pero a otra dependencia de nuestra propiedad, y aunque el viaje es ya un poco molesto, aún podemos regresar con cierta presteza. Finalmente, el no disponer de almacén y buscar el proveedor en una papelería local sería el caso de contar con una caché externa: Ya hay que utilizar la vía urbana compartida con otros componentes (placa base) y acudir a un ente ajeno (chip adquirido separadamente), con lo que el retraso comienza a ser bastante grande, pero en cualquier caso, siempre será mejor que desplazarse a Almería a por el papel (tomar el dato de memoria principal).

En la [tabla 3.8](#) resumimos todas las similitudes encontradas entre el tándem fotocopidora-papel y procesador-caché. Por otro lado, la [tabla 3.9](#) utiliza la nomenclatura propuesta para clasificar todos los modelos comerciales de quinta y sexta generación con objeto de que vayamos familiarizándonos con ella al tiempo que afianzamos los conocimientos sobre las cachés de nuestros procesadores. Esta tabla constituye además un magnífico compendio que resume la tendencia del mercado en los últimos cinco años:

- ❶ **Caché L1: Siempre integrada.** Desde hace ya algún tiempo, asume el rol de banco de registros extendido para el procesador, y nadie cuestiona ya su ubicación.
- ❷ **Caché L2 externa.** Difícilmente vamos a verla ya. El acceso a memoria principal de una DDRAM o RDRAM actual se encuentra más penalizado por la conexión a través del bus que por la latencia del dispositivo, y una caché externa sólo consigue enjugar ésta última. En otras palabras: Flaco aliado es un bus lento para una caché en la que la velocidad es su razón de ser. Ni siquiera los nuevos buses de 400 y 533 MHz hacen que este sombrío panorama cambie, ya que en ellos el procesador es también mucho más rápido, y lo que cuenta aquí es la diferencia entre ambos. Además, la L2 ya no tiene razón para quedarse fuera del procesador, pues la evolución de la tecnología de integración juega a su favor.
- ❸ **Caché L2 interna.** Los procesadores en formato Slot supusieron su consolidación, al llevar un zócalo cuya razón de ser era precisamente la colocación de una L2 interna sin incurrir en un coste excesivo. Ha quedado demostrado que aquella solución sólo era una estación temporal en el camino de la L2 hacia la integración conjunta con el procesador, y no esperamos una vuelta atrás: El camino ha venido marcado por la reducción de las micras, y ésta es calle de una sola dirección.
- ❹ **Caché L2 integrada.** Aunque alguna marca ya se atrevió con ella en la época de las 0.25 micras, la L2 se integra en el procesador con la llegada de las 0.18 micras. No esperamos

← pág. 84

incuestionable

en extinción

coyuntural

consolidada

Caché de primer nivel ó L1		Caché de tercer nivel ó L3	
<b>Integrada</b> en todos los procesadores de Intel y AMD a partir del Pentium		<b>Interna</b> en aquellos en los que está presente, que son únicamente el Cascades y Foster de Intel, esto es, el Pentium III Xeon y el Pentium 4 Xeon de 0.18 micras	

Tipo de caché de segundo nivel ó L2 y su símbolo asociado			
	Externa - X (todos en formato Socket)	Interna - ✓ (todos en formato Slot)	Integrada - ☆ (todos en Socket)
I N T E L	Pentium Pentium MMX Celeron (0.35)	<b>CeleronA (0.25)</b> Pentium II <b>Pentium II Xeon</b> Pentium III (0.25) <b>P III Xeon (0.25)</b>	<b>CeleronA (0.18)</b> <b>Pentium III (0.18)</b> P III Xeon(*) (0.18) <b>Pentium 4</b> <b>Pentium 4 Xeon(*)</b>
	En general, todos los de 0.35 micras	En general, todos los de 0.25 micras	En general, todos los de 0.18 micras
A M D	K6 K6-2	K7 (0.25)	K6-2+, K6-III, K6-III+ K7 (0.18), Duron

**TABLA 3.9:** Clasificación del tipo de caché que incorporan todos los modelos comerciales de microprocesadores de quinta y sexta generación. Entre paréntesis, el valor de la distancia de integración cuando actúa como factor discriminante al respecto. Vemos que Intel lo ha tenido bastante claro en relación a cómo fabricar las L2 de sus procesadores en función de su tecnología de fabricación. También existe una correlación clara con el tipo de zócalo utilizado, salvo en los dos casos señalados con(\*), donde los Xeon se desmarcan en formato Slot debido a la presencia de la L3 interna. En negrita, los procesadores cuya L2 funciona a la misma velocidad del procesador.

que se mueva de aquí, puesto que en una arquitectura actual es su ubicación natural. Con la llegada de las 0.13 micras, aprovecha además para consolidar su posición y aumentar de tamaño.

en extinción

⑤ **Caché L3 externa.** En vías de extinción por razones muy similares a las esgrimidas para la L2 externa.

sólo para servidores

⑥ **Caché L3 interna.** Apenas sí hemos visto esta modalidad hasta la fecha. Tiene sus opciones para consolidarse en aquellas arquitecturas en las que esquivar el bus local sea casi una obligación. Un claro ejemplo son las versiones Xeon de los Pentium III y 4, sistemas concebidos para que múltiples procesadores se acoplen en una única placa base que comparte la vía de acceso a memoria. En todos los casos, se utilizará para ella una solución basada en zócalo Slot como ya ocurrió con la L2.

inédita

⑦ **Caché L3 integrada.** Aún inédita. Conforme la tecnología de integración avance y veamos chips con más de cien millones de transistores en la segunda mitad de la década, probablemente algún fabricante se anime a gastar parte de este ingente patrimonio en integrar una L3. En este caso, la L3 repetiría el mismo peregrinar ya realizado por la L2. No obstante, volvemos a recalcar que nos parece mejor opción decantarse por aumentar los tamaños de la L1 y la L2.

relación con los buses  
pág. 85 ➡

Establecidas las tres posibilidades para la ubicación de la caché, debemos también clarificar la conexión al procesador para cada una de ellas. La [tabla 3.10](#) sintetiza la correspondencia entre las tres modalidades de caché vistas y los diferentes buses relacionados con el procesador que ya conocemos.

Modalidad de caché en relación al procesador	Ubicación en el sistema	Denominación(es) de su conexión al procesador
Externa	En placa base	Bus local o bus frontal (frontside bus)
Interna	En un segundo chip junto al de la CPU	Bus trasero (backside bus)
Integrada	Como parte del propio chip de la CPU	Conexión interna

TABLA 3.10: Interrelación entre los diferentes tipos de caché, su ubicación en la arquitectura del sistema y el bus que las liga al microprocesador.

## Ubicación del controlador de caché

## 4.5

Muchas veces, cuando describimos el procesador, solemos fijarnos en su Unidad de Proceso, que es donde realmente se efectúan las operaciones, y pocas veces recaemos en la Unidad de Control. Con la memoria caché ocurre algo similar: Describimos su tamaño, su velocidad, sus líneas, ... en definitiva, todo lo que conforma su área de datos, sin reparar en que también existe un área de control desde donde se gobiernan todas las operaciones.

área de control

La parte de control más importante de una caché es su **directorio caché**, que es donde se consultan las etiquetas de memoria principal para determinar si un dato buscado se encuentra allí o no, y a partir de ahí, obtener su dirección de acceso.

directorio  
caché

Las posibilidades de ubicación de una memoria caché dentro del conjunto del sistema se completan con una eventual separación de las áreas de datos y control, pudiéndose situar el controlador en tres emplazamientos diferentes:

- ❶ Fuera del procesador, en un chip aparte junto con su área de datos. VARIANTE 1
- ❷ Dentro del procesador, pero gestionando los datos de una caché ubicada en el exterior. VARIANTE 2
- ❸ Dentro del procesador, integrado junto con los datos que controla. VARIANTE 3

A continuación estudiaremos por separado estas tres posibilidades, que al haber sido enumeradas en orden cronológico, irán apareciendo en el recorrido temporal que vamos a efectuar.

En los primeros PC, cuando la caché era considerada un artículo de lujo, algunas placas base habilitaban un zócalo donde opcionalmente se podía incorporar una pequeña placa de circuito impreso con la caché y su controlador si se disponía del dinero suficiente para hacer frente a su coste. El aspecto de esta placa es muy parecido al de los módulos de memoria principal actuales, tal y como se aprecia en la [foto 10.2](#).

zócalo opcional

➔ p. 27/Vol. 2

Posteriormente, la caché se hizo imprescindible y las ventas masivas y la competencia en el sector encargado de su fabricación le hicieron perder ese elitismo. Cuando la caché comenzó a montarse de serie en la placa base, el controlador de caché se incluía de forma independiente en la geografía de la placa base, junto al circuito integrado que albergaba las celdas de datos.

popularización

VARIANTE 1

Con posterioridad se instaló en la arquitectura de las placas base una corriente que trató de reducir espacio y minimizar el número de chips. Apareció así el concepto de juego de chips, o serie de circuitos integrados que aglutinan multitud de controladores que antes se encontraban diseminados por la placa base en chips independientes (DMA, interrupciones, temporización, ...).

juego de chips

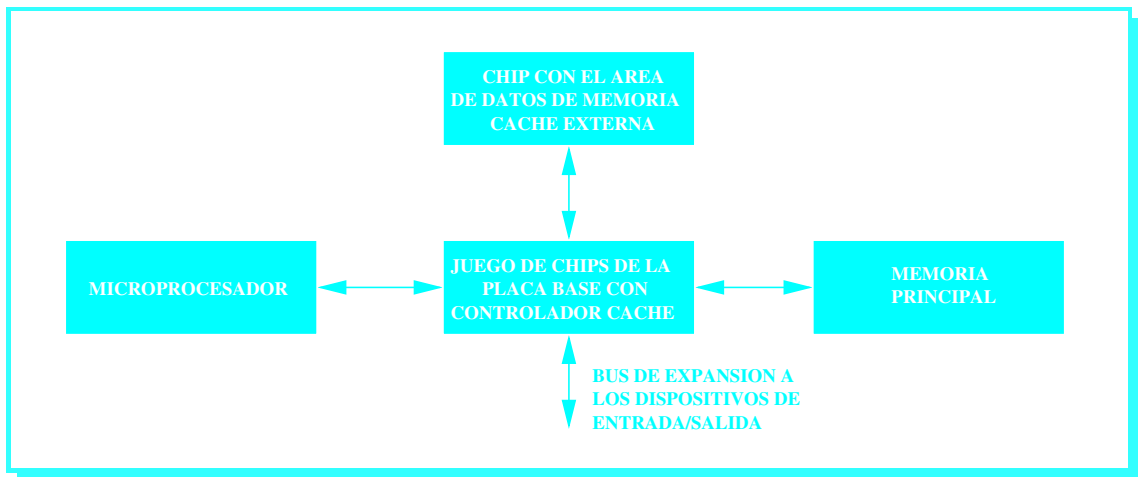


FIGURA 3.12: Ubicación del área de datos de caché y su controlador en el contexto de una placa base de quinta y sexta generación.

absorción

El controlador para la caché L2 no fue una excepción, y también fue absorbido por este juego de chips central. El área de datos se mantuvo en sus chips dedicados independientes, y como el fabricante de la placa base es el responsable de montar ambos componentes, no existe conflicto de compatibilidad alguno. El resultado es el esquema que aparece en la [figura 3.12](#).

VARIANTE 2

La situación anterior conduce a una importante conclusión: No es el microprocesador el que dialoga con la caché externa de forma directa, sino el juego de chips donde se ubica el controlador, y que hace de puente entre ambos. Esto provocó más adelante que cuando el microprocesador comenzara a absorber a la caché externa, en ocasiones no pudiese abarcar tanto el controlador como el área de datos, decantándose por integrar el primero, que es con quien realmente dialoga, y dejando fuera al segundo. Esto permite al procesador averiguar de forma local si el dato que se busca se encuentra o no en la caché externa, con lo que saldremos al bus únicamente cuando estemos seguros de encontrar el dato buscado, minimizando su uso.

descoordinación

Los microprocesadores que implementan esta aproximación son escasos, ya que dependen en exceso de la bondad del fabricante de la placa base: Si éste decide incorporar una caché externa, debe saber que el controlador está ubicado en el microprocesador, y por tanto limitar el área de datos al rango que éste puede gestionar. Por otro lado, si el fabricante de la placa base decide sacar productos de gama más baja que carecen de caché, el microprocesador contendrá un controlador de caché al que no vamos a sacar provecho pero que hemos pagado religiosamente.



### Ejemplo 3.12: CONTROLADOR DE CACHÉ L2 INTEGRADO, ÁREA DE DATOS EXTERNA

En el ámbito de los RISC de gama alta, encontramos algunos ejemplos de este tipo de configuraciones en la primera mitad de la década de los noventa. Quizá el más significativo sea el de la familia del microprocesador R10000 de Silicon Graphics, que evoluciona pasando por el R4000 con sendas cachés L1 internas de 16 Kbytes y ausencia de caché L2, el R8000, con idénticas cachés L1 pero que ya incorpora el mencionado controlador interno para la L2, y los R5000 y R10000, en los que se duplica la capacidad del primer nivel de caché manteniendo internamente el controlador para la caché L2 externa.



**Ejemplo 3.13: CONTROLADOR DE CACHÉ L2 INTEGRADO, ÁREA DE DATOS INTERNA**

El mejor ejemplo que tenemos de esta configuración es el K7 de 0.25 micras, que incorporó una caché L2 de 512 Kbytes interna en un chip separado de la CPU pero dentro del cartucho cerámico del procesador. El directorio caché L2 estaba integrado en el chip de la CPU, y daba cobertura completa a este área de datos. Si posteriormente se optaba por aumentar el tamaño del chip de la caché L2, entonces este directorio caché proporcionaba sólo una parte de las etiquetas de dirección necesarias, teniéndose que colocar la parte sobrante en el chip de datos de la caché.

Resultaba así una implementación flexible y eficiente al mismo tiempo. En la práctica, lo de distribuir el directorio caché L2 en dos porciones hubiera dado lugar a un algoritmo de consulta de dos niveles, actuando el primero en el chip de la CPU y eventualmente el segundo en el chip de la caché, pero no llegó a utilizarse porque el K7 nunca llegó a superar el tamaño de 512 Kbytes de caché L2 en su versión de 0.25 micras.

En los diseños en los que se opta por incorporar la L2 conjuntamente con su área de datos de forma integrada en el propio chip procesador, lo que se plantea es si además se incorpora el controlador de la caché L3. Aquí también disponemos de algunos ejemplos.

VARIANTE 3

**Ejemplo 3.14: CONTROLADOR DE CACHÉ L3 INTEGRADO, ÁREA DE DATOS EXTERNA**

Los ejemplos comerciales de esta variante se encuentran, en la primera mitad de los años 90, con alguna versión puntual de la familia de procesadores Alpha de Digital, y en la segunda mitad de los noventa, con el K6-III de AMD.

Este último caso representó de lo poco malo que tenía aquel procesador, ya que se apostó por encontrar una L3 en la placa base que muy pocos fabricantes de ésta realmente incorporaron (por entonces, la idea de la L3 ya se encontraba en franco declive). Paradójicamente, fueron las placas base para K6-2, que sí llevaban caché externa (hacia las veces de L2 para ese procesador), las que, al compartir el zócalo Super 7 con el del K6-III permitieron a éste utilizar la L2 como L3.

Finalmente, llegamos a la situación actual, en la que todo este caos intermedio ha quedado despejado: La L2 está toda integrada dentro del chip del microprocesador, y la L3, de asomarse, lo hace en su versión interna donde el chip dedicado a ella también integra conjuntamente datos y controlador.

situación  
actual

## 4.6 ▶ Velocidad

Estamos ante otro de los parámetros que más confusión genera en una caché por la ligereza con que se trata. La publicidad suele proclamar “Microprocesador XXX con caché a la misma velocidad del procesador”. Inocente frase, pero difícilmente más ambigua:

- tipo
  - ❶ Primera omisión: El tipo de caché de que se trata. Si ésta es integrada, la información acerca de la velocidad de caché es superflua (toda caché integrada en el mismo chip procesador funciona a su misma velocidad). Del contexto de la frase hemos de suponer que se trata de una caché interna, pero no siempre es así, y cada vez lo va a ser menos.
- nivel
  - ❷ Segunda omisión: El nivel de caché de que se trata. Con buena voluntad, pensamos que la frase debe estar refiriéndose a la L2, y esta vez sí tenemos una elevada probabilidad de acertar.

otra falacia más  
 No cuesta tanto incorporar una letrita y un numerito a la susodicha frase para evitar confusiones, pero es que todo esto esconde una negligencia aún mayor: La de desconocer el significado real de la frase. La publicidad señala la velocidad de la caché como único responsable de su rendimiento, pero es otra falacia más del mercado.

pregunta 1  
 pág. 91 ➔  
 Un usuario de Pentium 4 ó de K6-III podría preguntarse: “Mi caché L1 funciona a la misma velocidad del procesador, y mi L2 también, ¿Qué diferencia existe entre que mi procesador recoja el dato de la primera o que lo haga de la segunda?”. La respuesta lógica es *ninguna*; la respuesta válida, *otra bien distinta*, que será desvelada tras el [ejemplo 3.19](#).

pregunta 2  
 pág. 90 ➔  
 Un usuario de Pentium III ó K7 de 0.25 micras podría también haberse formulado la siguiente pregunta cuando salió la versión de 0.18 micras de su procesador: “Si mi L2 funciona a la mitad de la velocidad del procesador, ¿Significa esto que mi procesador tarda el doble en obtener el dato de ella con respecto a los nuevos modelos de 0.18 micras en los que la caché L2 funciona a la misma velocidad del procesador?”. La respuesta lógica es *afirmativa*; la respuesta válida, *otra bien distinta*, que será desvelada tras el [ejemplo 3.17](#).

información oculta  
 La explicación a todo esto comienza por desvelar que sólo se está proporcionando al usuario una parte marginal de la información que necesita para conocer el rendimiento real de su memoria caché. Se dice que (a) la caché responde a cierta velocidad, pero se omite que (b) antes de que la caché responda, debe llegarle la petición del procesador, y que (c) una vez devuelto el dato solicitado, éste, y todos los que le acompañan en su misma línea de caché, deben emprender todo el camino de regreso hacia el procesador.

cinco configuraciones  
 Para aprender a relativizar la importancia que tiene la velocidad de una caché frente al resto de parámetros aquí estudiados, analizaremos a continuación cinco configuraciones comerciales distintas en las que hemos cubierto un amplio espectro de variantes. Eso nos permitirá seguir un criterio certero en la identificación de los modelos de microprocesador más ventajosos cuando presentemos sus características en nuestra cobertura generacional de capítulos posteriores.

## 4.7 ▶ Análisis del rendimiento de caché en relación al procesador

pág. 91 ➔  
 Con objeto de realizar un análisis del rendimiento que sea mínimamente comprensible y didáctico, no queda más remedio que simplificar suponiendo peticiones aisladas a memoria. No vamos a tener en cuenta el alto grado de concurrencia que tiene lugar entre los buses, las distintas cachés y el procesador, y que han sido el recurso utilizado por algunas de las optimizaciones vistas, pero aún con estas limitaciones podremos ordenar muchas de las estrategias en relación al potencial de mejora que revierten sobre el sistema de una forma bastante realista. Cuando lleguemos al [ejemplo 3.18](#), comprobaremos que efectivamente nuestro análisis apenas difiere de los

resultados hechos públicos por AMD para el rendimiento de sus cachés comerciales actuales.



**Ejemplo 3.15: CASO 1: BÚSQUEDA DE UN DATO EN LA CACHÉ L2 EXTERNA DEL K6 (L2 A LA MITAD DE LA VELOCIDAD DEL PROCESADOR)**

Comenzaremos con el primer K6 que vió la luz, aunque el ejemplo es también válido para su versión más reciente de 0.25 micras y para el K6-2 (no ya para el K6-2 de 0.18 micras, que goza de una pequeña L2 integrada). Los parámetros que completan su configuración de memoria son los siguientes:

- Tamaño de línea L2: 32 bytes.
- Tamaño de línea L1: 32 bytes.
- Anchura del bus local que en este caso conecta L1 y L2: 64 bits (8 bytes).

Si en la obtención del dato de una instrucción se tarda un ciclo en el acceso directo al banco de registros, en el acceso a memoria L2 se tardarían los siguientes:

- Ciclo 1: Conversión de la dirección relativa a dirección virtual (el direccionamiento relativo es el más utilizado en las instrucciones de acceso a memoria).
- Ciclo 2: Traducción virtual a física en L1 de datos (todos los procesadores emiten direcciones virtuales que hay que mapear sobre el espacio de direcciones físico de cada usuario en función de la cantidad de memoria de que disponga).

La **TLB (Translation Look-Aside Buffer)** es la unidad funcional responsable de retener las traducciones más recientes (funcionando como una caché cuyos datos son direcciones físicas, y si se acierta en el acceso a la misma, la traducción tiene lugar en un solo ciclo; con objeto de simplificar nuestro estudio, supondremos que siempre va a ser así).

- Ciclo 3: Selección de la línea de caché L1 a sustituir y emisión de la petición al chip L2 por el controlador del bus local.
- Ciclo 4: El controlador del bus local arbitra la obtención de este recurso compartido. Supongamos que gana el bus en el ciclo 10. Mientras tanto, ya se ha actualizado la etiqueta en la línea de caché L1 donde se alojará el dato cuando se obtenga.
- Ciclos 10, 11, 12 y 13: Envío de la dirección al controlador de la caché L2 ubicado en el juego de chips. Suponemos que el bus local funciona con un divisor de 4 respecto al procesador (esto es, 100 MHz para el bus local y 400 MHz para el K6-2 según la configuración más vendida de este procesador).
- Ciclo 14: Traducción virtual a física en L2 (TLB).
- Ciclo 15: Consulta en directorio caché de L2, obteniéndose la dirección de caché en la que se encuentra el dato de memoria principal.
- Ciclos 16 y 17: Tiempo de respuesta de la caché L2 (dos ciclos por el hecho de que la caché responda a la mitad de velocidad del procesador; éste es el único paso que se ve afectado por la velocidad de la caché).
- Ciclo 18: Comienza la arbitración del bus local para el viaje de regreso. Suponemos que se conserva la apropiación de este recurso.
- Ciclos 19 al 34: Transporte de la línea de 32 bytes hacia la L1 (4 viajes de 8 bytes por el bus local, gastándose 4 ciclos para cada viaje).
- Ciclo 35: Almacenamiento en L1 de la línea procedente de L2 (parte de esta operación se solapa con los últimos viajes por el bus).
- Ciclo 36: Envío al procesador de la palabra que solicitó.

**Conclusión:** Cuando el bus local entra en juego, es el responsable de la mayor parte de la penalización en el acceso a memoria caché dada su lentitud y arbitración. No nos extrañe pues la paulatina extinción de las cachés externas del mercado.

 **Ejemplo 3.16: CASO 2: BÚSQUEDA DE UN DATO EN LA CACHÉ L2 INTERNA DEL K7 (L2 A 1/3 DE LA VELOCIDAD DEL PROCESADOR)**

Estamos ante el primer Athlon que salió al mercado, el de 0.25 más orientado al segmento doméstico. Estos son los parámetros que completan su configuración de memoria:

- Tamaño de línea L2: 64 bytes.
- Tamaño de línea L1: 32 bytes.
- Anchura del bus trasero que conecta L1 y L2: 64 bits (8 bytes).

Y ésta, su operativa de funcionamiento:

- ▶ Ciclo 1: Cálculo de la dirección virtual procedente del direccionamiento relativo.
- ▶ Ciclo 2: Traducción virtual a física en L1 de datos (TLB).
- ▶ Ciclo 3: Consulta en directorio caché de L1. El dato no está.
- ▶ Ciclo 4: Selección de línea víctima en L1 y emisión de la petición a L2 por el bus trasero.
- ▶ Ciclo 5: En L1, actualización de la etiqueta en la línea de caché donde se alojará el dato una vez obtenido. En L2, traducción virtual a física (TLB).
- ▶ Ciclo 6: Consulta en directorio caché de L2. Obtención de la dirección para el dato.
- ▶ Ciclos 7, 8 y 9: Tiempo de respuesta de la caché L2.
- ▶ Ciclos 10, 11, 12 y 13: Transporte de la línea de 32 bytes hacia la L1 (4 viajes de 8 bytes por el bus trasero).
- ▶ Ciclo 14: Almacenamiento en L1 de la línea procedente de L2.

- ▶ Ciclo 15: Envío al procesador de la palabra solicitada (el K7 dispone de la estrategia *critical word first* para enviar primero la palabra de la línea que ha solicitado el procesador).

**Conclusión:** En el caso 1, con una caché externa más rápida se gastan más del doble de ciclos que en el caso 2 con una caché interna más lenta. Se pone de manifiesto la importancia del tipo de caché frente a lo anecdótico de su velocidad.

 **Ejemplo 3.17: CASO 3: BÚSQUEDA DE UN DATO EN LA CACHÉ L2 INTERNA DEL K7 (L2 A LA MISMA VELOCIDAD DEL PROCESADOR)**

Se trata de la versión del Athlon de 0.25 micras que fue candidata en su tiempo al segmento de servidores, incorporándole una caché L2 DDR SRAM que se sincronizaba a la velocidad del microprocesador, entonces por los 600 MHz. Los modelos de Athlon que se vendieron bajo este aditivo fueron muy escasos debido a su elevado coste.

En relación a su caché L2, el comportamiento de este procesador es mimético al de todos los modelos de Pentium II/III Xeon de Intel: La caché es más grande y esta vez funciona a la misma velocidad del procesador, aunque sigue siendo interna.

Esta variante ahorra los ciclos 8 y 9 anteriores, reduciendo el montante total a sólo 13 ciclos.

**Conclusión:** Parecía que el acceso a L2 en el caso 3 se aceleraría en un factor de 3 en relación al caso 2 (esta era la respuesta lógica a nuestra pregunta 2 de la sección anterior). Sin embargo, la respuesta válida es una aceleración de apenas un 13%. Dejando a un lado otro tipo de mejoras, la velocidad no es un parámetro tan decisivo como sugiere a primera vista.

**Ejemplo 3.18: CASO 4: BÚSQUEDA DE UN DATO EN LA CACHÉ L2 DEL K7 (L2 INTEGRADA)**

Este procesador es el Thunderbird, el K7 de 0.18 micras comercializado por AMD bajo formato SocketA. Al integrarse la L2 dentro del chip CPU, se prescindió del bus trasero y se instaló un puerto de conexión L1-L2 de 256 bits.

Como resultado, la operativa del caso 2 conlleva ahora sólo 10 ciclos, pues se siguen ahorrando los ciclos 8 y 9 del caso 3 debido a la menor latencia del dispositivo, y además, se ahorran los ciclos 11, 12 y 13 puesto que la comunicación de la línea completa puede ahora completarse en un solo ciclo.

Según AMD, bajo la ejecución de aplicaciones comerciales, el tiempo medio que su procesador K7 de 0.18 micras tarda en tomar un dato en el caso de no encontrarlo en la caché L1 y sí en la L2 puede oscilar entre 11 y 20 ciclos dependiendo de la actividad del procesador, esto es, considerando todas las variantes de flujos concurrentes que pueden darse. En la gran mayoría de los casos se sitúa en el valor mínimo de 11 ciclos, al que se le restarían tres ciclos más por la actuación de una caché víctima compuesta de 8 líneas de caché que se sitúa entre la L1 y la L2 como aditivo que la compañía incorporó en la transición a 0.18 micras. Nosotros no hemos considerado la intervención de esta nueva caché aquí porque restaría limpieza a la comparativa que tratamos de ilustrar, en la que sólo intervienen L1 y L2.

**Conclusión:** Cuando aparentemente parecía que no lograríamos mejoras en relación a la configuración del caso 3 (pues su velocidad era la misma), la ganancia ha resultado ser del 23%. De nuevo tenemos ante nosotros una prueba inequívoca de la importancia del tipo de caché, aunque eso sí, el salto en rendimiento desde caché externa a interna es muy superior al que se produce desde caché interna a integrada.

**Ejemplo 3.19: CASO 5: BÚSQUEDA DE UN DATO EN CACHÉ L1**

Este caso es similar en todos los procesadores, y su operativa de funcionamiento, muy sencilla:

- ▶ Ciclo 1: Conversión de la dirección relativa de memoria a dirección virtual.
- ▶ Ciclo 2: Traducción de la dirección virtual a física en L1 de datos (se utiliza su TLB).
- ▶ Ciclo 3: Consulta en directorio caché de L1. Encontramos el dato y lo enviamos al procesador.

**Conclusión:** Creíamos que daba igual traer el dato de la L2 que de la L1 simplemente porque ambas iban igual de rápido (esta era la respuesta lógica a la que fué nuestra pregunta 1 en la sección anterior). Acabamos de ver la importancia del nivel de caché: La aceleración del caso 5 respecto al caso 4 es superior a un factor 3.

Aprovecharemos para recalcar que si el dato es servido por el banco de registros en lugar de

Modelo de microprocesador	Memoria caché analizada			Bus que entra en juego	Número ciclos
	Nivel	Tipo	Velocidad		
K6 ó K6-2	L2	Externa	La mitad	Bus local	36
K7 0.25 micras	L2	Interna	Un tercio	Bus trasero	15
K7 0.25 micras	L2	Interna	La misma	Bus trasero	13
K7 0.18 micras	L2	Integrada	La misma	Ninguno	10
K7 0.18 micras	L1	Integrada	La misma	Ninguno	3

TABLA 3.11: Comparativa de rendimiento de cinco configuraciones distintas de memoria caché escogidas según su distinta relación con el microprocesador.

por la caché L1, la aceleración también es superior a un factor 3, pues aunque ambos respondan igual de rápido, la dirección de acceso al banco de registros se obtiene de forma directa a partir del código de operación de la instrucción, ahorrándose los dos primeros ciclos de traducción de la operativa anterior (y un eventual tercer ciclo si fallamos en el acceso a la TLB y hemos de realizar la traducción de forma manual).

Curiosa circunstancia: La respuesta lógica que dimos a la pregunta 1 de la sección anterior estuvo muy cerca de ser la respuesta válida a la pregunta 2, y la respuesta lógica de la pregunta 2, muy cerca también de convertirse en la respuesta válida para la pregunta 1. Menudo trabalenguas hemos compuesto. Es el signo más evidente de lo traicioneras que pueden resultar las cosas de la caché si no son escudriñadas con esmero. Pasemos a limpiar las conclusiones obtenidas, pues condensan todo un recital didáctico.

La [tabla 3.11](#) resume los resultados de los cinco casos analizados, y el balance que éstos arrojan en la interacción del procesador con la caché nos sirve para ordenar los parámetros estudiados según su influencia en el rendimiento. El orden de mayor a menor importancia es el siguiente:

respuestas con  
trabalenguas

RÁNKING FINAL:

- nivel ❶ El nivel de caché: Su posición en la jerarquía.
- tipo ❷ El tipo de caché dentro de un mismo nivel: Externa frente a interna incide más que interna frente a integrada.
- bus ❸ El bus en el tipo de caché externa ó interna: Incide más su velocidad en el primer caso, y su anchura en el segundo, pero en ambos casos, tanto en el camino de ida como en el de vuelta.
- velocidad ❹ La velocidad de la caché respecto a la del procesador (influye sólo en el acceso).
- tamaño línea ❺ El tamaño de la línea de caché (influye sólo en el transporte de vuelta).
- tamaño TLB ❻ El tamaño de la TLB (influye en la traducción de ida).

labores ocultas

El directorio caché, la TLB y el bus trasero son elementos que casi nadie suele tomar en consideración. Sin embargo, aunque no tengan en su mano mejoras porcentuales de tres dígitos, sí realizan una labor fundamental en el conjunto del sistema caché que conviene destacar:

- dir. caché TLB
  - El directorio caché y la TLB, porque son los responsables de localizar un dato en caché a partir de una dirección que no es la suya, sino la de memoria principal. A la caché se le obliga a ser ultrarrápida al tiempo que transparente al resto del sistema, así que nadie puede facilitarle información más útil que aquella destinada a otras partes del sistema.
- bus
  - La velocidad del bus que conecta la caché y el procesador, que puede oscilar entre 1, 2 ó 3 en las diferentes implementaciones de caché L2 interna incluso dentro de un mismo modelo de microprocesador, originando así diferentes alternativas de coste y rendimiento de cara al usuario sin modificar un ápice la arquitectura interna del procesador.

		Pentium clásico	Pentium Pro L2 256 Kbytes	Pentium Pro L2 512 Kbytes
Fecha lanzamiento		Marzo, 1993	Marzo, 1995	Junio, 1995
Patillaje		296 pines	387 pines	387 pines
Número de chips		1 (CPU+L1)	2 (Uno con CPU+L1, otro con L2)	
Zócalo		Socket 5 y 7	Socket 8	Socket 9
Número de transistores	Bloque CPU	2.1M	4.5M	4.5M
	Bloque L1	1M	1M	1M
	Chip L2	Externo	15.5M	31M
Espacio físico ocupado	Chip CPU+L1	91 mm <sup>2</sup>	306 mm <sup>2</sup>	195 mm <sup>2</sup>
	Chip L2	Externo	202 mm <sup>2</sup>	242 mm <sup>2</sup>
	Encapsulado	25 cm <sup>2</sup> (**)	35 cm <sup>2</sup>	35 cm <sup>2</sup>
Integración L2 (micras)		SRAM (0.8) (*)	CSRAM (0.6)	CSRAM (0.35)
Velocidad caché L2		Asíncrona (*)	Síncrona con el procesador	
Coste inicial del conjunto		935 dólares (**)	1200 dólares	1600 dólares

**TABLA 3.12:** Los principales parámetros relativos a la integración de las cachés L1 y L2 en los dos modelos de Intel más representativos a este respecto durante la pasada década de los 90. Hemos incluido numerosos parámetros del chip CPU en sí para poder comparar sus valores con los de las cachés. (\*) = Estos datos hacen referencia a una caché L2 externa de 256 Kbytes incorporada en la placa base. (\*\*) = Datos sin contar la caché L2 de la placa base, para cuyo tamaño de 256 Kbytes estimamos unas dimensiones de 250 mm<sup>2</sup> y un coste de unos 100 dólares de aquellas fechas.

## Análisis del coste asociado a una caché

### 4.8

#### 4.8.1 Caché interna

La gama más alta, compuesta por los modelos con caché sincronizada a la velocidad del microprocesador, es la más cara. Por ello, si son modelos orientados al segmento doméstico, tendrán un tamaño reducido con objeto de que el coste no se dispare (128 Kbytes en la L2 del primer Celeron y 512 Kbytes en el primer Athlon, por ejemplo).

tamaños  
domésticos

Para tamaños superiores, hay que buscar en la L3 y en procesadores del segmento servidor. Por ejemplo, los Xeon de Intel disponen de configuraciones de este tipo hasta los 8 Mbytes, pero el coste de la configuración básica, que comienza en 1 Mbyte, es ya superior a los 3000 €.

tamaños para  
servidor

La [tabla 3.12](#) compara el coste de dos configuraciones de tipo servidor frente a otra de corte doméstico para los primeros sistemas servidores de 1995. El coste que se muestra se expresa en dólares de aquella época, es decir, sin actualizar por el efecto de la inflación. Tratamos de comparar los tres tipos de caché (externa, interna e integrada) y los dos primeros niveles (L1 y L2), y para encontrar un sistema real con tantas variantes, no queda más remedio que remontarse un poco atrás en el tiempo. En nuestra defensa, diremos que la tecnología de caché avanza a un ritmo más lento que el procesador, y los precios y prestaciones de la tabla no están tan lejos de la situación actual del mercado como podría pensarse.

comparativa

El directorio caché es también un ingrediente de particular relevancia en la formación de precios. La presteza con la que éste tiene que llevar a cabo la búsqueda del dato solicitado le obliga a utilizar una memoria asociativa para alojar las etiquetas, de forma que todas ellas puedan ser consultadas simultáneamente en el mismo ciclo de reloj. El problema de esta memoria asociativa es que su coste es exponencial con el tamaño, por lo que un directorio caché el doble de grande resulta cuatro veces más caro.

memoria  
asociativa

### 4.8.2 Caché integrada

tamaños  
domésticos

El tamaño de las cachés integradas está entre los 32 y los 64 Kbytes para la L1, y entre los 256 y los 512 Kbytes para la L2 a 0.13 micras (Northwood, Barton), y ya en 1 Mbyte para la L2 a 0.09 micras (K8).

En cualquiera de estos casos, la caché L2 se llevará, ella sola, más de la mitad de los transistores del chip microprocesador, de los que buena parte de ellos se encontrarán en su controlador ó directorio caché.

Ahora bien, en las áreas de integración de los procesadores, algunas que ya hemos mostrado y otras que irán desfilando más adelante, puede observarse que en ningún caso una L2 interna ocupa más de la mitad del área de silicio. Esto es así por dos razones básicas:

silicio

- 1 La celda básica del área de datos de caché que almacena un bit está optimizada para ser integrada con seis transistores proporcionalmente más pequeños que los de otras unidades funcionales del procesador, y por lo tanto, ocupa bastante menos espacio en silicio.

metal

- 2 En segundo lugar, el directorio caché contiene fundamentalmente conexiones de metal. En realidad, la densidad de este retículo metálico es tan grande, que no pocos fabricantes habilitan sabiamente la superficie de silicio que queda despejada por debajo de él para colocar circuitería de otras unidades funcionales del procesador. En estos casos, resulta injusto atribuir al cableado el espacio ocupado en lugar de a estas unidades extra allí ubicadas.

dos frenos

Pero tampoco podemos aspirar a cachés integradas muy grandes aunque estemos decididos a asumir el elevado coste que supone en transistores y silicio. Porque el área de silicio influye exponencialmente en el coste de integración, pero también incide en la velocidad del conjunto: Los retardos de las señales eléctricas en el interior de un chip grande suponen en la práctica uno de los frenos más claros para la frecuencia del procesador.

un remedio

Estas dos razones explican que el máximo tamaño de caché integrada haya estado históricamente condicionado por la tecnología, pues una distancia de integración más corta pone remedio a esos dos obstáculos: Primero, hace que el transistor resultante ocupe menor área de silicio, y segundo, permite disfrutar de mayor velocidad de conmutación.

varios ejemplos

Así, durante la época de las 0.35 micras no vimos a la caché L2 integrada en el procesador, y ya durante las 0.25 micras empezaron a asomar los primeros modelos, como el K6-III. Con la llegada de las 0.18 micras, fue ya una práctica generalizada a todos los modelos existentes.

## SECCIÓN 3.5

### Conjunto de instrucciones

personalidad

El conjunto de instrucciones máquina que es capaz de entender un procesador es un parámetro clave para entender su diseño, y condiciona lo que podríamos catalogar como su personalidad. Esta sintetiza cuatro aspectos básicos:

abstracción

- 1 El nivel de abstracción con que se le proporcionan las instrucciones. En un mayor nivel de abstracción, sólo diríamos qué queremos hacer, encontrándose las instrucciones más cercanas a nuestro lenguaje natural. Un menor nivel de abstracción aboga en cambio por un mayor nivel de detalle, una visión más cercana a la circuitería en la que ya se dice cómo se ejecutan las instrucciones en su arquitectura interna, trascendiendo los aspectos de su diseño a la capa software de más bajo nivel.



Hace unos años, esta capa podía ser el propio usuario si éste era capaz de fajarse al nivel del lenguaje ensamblador de la máquina. En la actualidad, los programadores que están dispuestos a hacer así las cosas son una especie en vías de extinción, pues se busca cada vez más simplificar el desarrollo de los programas frente a la consecución de unos puntos porcentuales de rendimiento extra.

- ② La amigabilidad del interfaz. Un mayor nivel de abstracción debería facilitar el diseño de un interfaz más amigable, más cómodo al usuario, aunque en la práctica no ha sucedido así: El conjunto de instrucciones 80x86 es uno de los que mayor nivel de abstracción presenta, y sin embargo, parece un lenguaje diseñado por el peor de nuestros enemigos. En cambio, ciertos diseños RISC, cuyo nivel de abstracción es siempre bajo, presentan un lenguaje tremendamente sencillo de manejar aprovechándose de su simpleza.

amigabilidad

- ③ La rapidez de asimilación, o cómo de rápido decodifica e interpreta las instrucciones que le llegan. Si todo el software se escribiera para el procesador sobre el que va a ser ejecutado, estaríamos hablando siempre de una única operación de decodificación que gastaría un solo ciclo, pero muchos programas ejecutables son compilados para una plataforma anterior a otra que los reutiliza para garantizar la compatibilidad con las aplicaciones software ya existentes en el mercado.

asimilación

En el mundo del PC, esta historia nos es muy familiar, pues llevamos veinte años ejecutando en nuestros procesadores código escrito para el procesador 8086. El cómo lleve internamente a cabo el procesador esta labor de conversión al que es su código nativo es un bastión nada despreciable en su rendimiento.

- ④ La riqueza del lenguaje, o el arte de diseñar un conjunto de instrucciones que responda a lo que los programadores desean ejecutar en la máquina. En este sentido, el lenguaje máquina evoluciona de la mano de las aplicaciones, y ahí está la retahíla de conjuntos de instrucciones multimedia que han emergido en los últimos cinco años al calor de la fiebre por las aplicaciones gráficas, de sonido e Internet.

riqueza

Remontándonos atrás en el tiempo, la historia ha sido **pendular** respecto al comportamiento del conjunto de instrucciones. Comienza con una primera fase que data de finales de los años 70 y principios de los 80 en la que el conjunto de instrucciones va engordando paulatinamente, y a finales de los 80 invierte su tendencia y evoluciona en sentido opuesto hacia conjuntos de instrucciones cada vez más simples. Ultimamente, la tendencia parece invertirse de nuevo, con la llegada de las instrucciones multimedia que amplían el conjunto de instrucciones del procesador y aumentan su complejidad, como las populares MMX y 3DNow! en quinta generación, sus sucesoras las SSE y Enhanced 3DNow! en sexta generación, o iniciativas más complejas como la VLIW que se enmarca ya dentro de la séptima generación de microprocesadores.

historia  
pendular

## CISC versus RISC

### ◀ 5.1

Volviendo a los orígenes, diremos que los microprocesadores comienzan su andadura con un repertorio de instrucciones simples. A finales de los años 70 se origina la primera corriente evolutiva hacia diseños de mayor complejidad, la cual vino respaldada por cuatro aspectos básicos:

- ① **El auge de los lenguajes de programación alto nivel.** El programador escribe sus programas en un lenguaje cada vez más potente y alejado del lenguaje ensamblador. La responsabilidad de generar código eficiente ya no es del programador: Se ha trasladado hacia el compilador.
- ② **La aparición de familias de microprocesadores.** Los fabricantes utilizan una estrategia de marketing en la que cada nuevo microprocesador es compatible con el anterior, pero mejorado con nuevas características, lo que supone ampliar su conjunto de instrucciones y complicar su circuitería.

compilador

compatibilidad

velocidad ③ **La migración de funciones desde el software hacia el hardware**, motivada por la ganancia en velocidad que la implementación hardware de una instrucción proporciona frente a su homóloga software.

empaquetado ④ **La lentitud de la memoria respecto al procesador**. Esto ralentiza la fase de búsqueda de una instrucción, por lo que se trata de empaquetar muchas instrucciones en una sola con el fin de minimizar el número de operaciones de búsqueda necesarias para completar un programa.

CISC De esta manera, el procesador va incorporando cada vez más modos de direccionamiento de operandos, más funciones potentes y especializadas en tareas concretas, y más registros de propósito general, surgiendo el diseño CISC, o de *conjunto de intrucciones complejo* (del inglés, *Complex Instruction Set Computer*). Este diseño se caracteriza por una extensa circuitería, sobre la que la capa software ha delegado parte de sus funciones. Un buen ejemplo es la saga de microprocesadores 80x86 de Intel.

Pero la tendencia CISC se rompe a finales de los 80, con la llegada de nuevos personajes que cambian el trasfondo de la situación:

- ① Aparecen las memorias caché, provocando una drástica disminución del tiempo de búsqueda de una instrucción y posibilitando así una eventual descomposición de las instrucciones en otras más sencillas.
- ② Se alcanza un punto en el cual la incorporación de nuevas instrucciones proporciona una funcionalidad cada vez más rebuscada, que apenas puede ser aprovechada por el compilador, y que por el contrario complica el diseño del microprocesador, haciéndolo cada vez más lento y costoso.
- ③ Atendiendo a las necesidades de los programas más populares, se demuestra que el código máquina de éstas contiene un aplastante predominio de instrucciones sencillas.

RISC Estos tres motivos van a provocar, en primer lugar, un freno a la ampliación del conjunto de instrucciones de un procesador, y, posteriormente, su paulatina disminución. Se eliminan así aquellas instrucciones más complejas que puedan implementarse mediante otras más simples, lo que poco a poco desemboca en una filosofía de diseño tipo RISC, o de conjunto de instrucciones reducido (del inglés, *Reduced Instruction Set Computer*). Así, la responsabilidad de obtener una ejecución rápida se traslada de nuevo a las capas software del sistema, como el compilador y el sistema operativo.

diseños híbridos Los microprocesadores que aquí estudiaremos se encuentran en una extraña confluencia entre las corrientes CISC y RISC. Disponen de ciertos rasgos RISC que cada vez tratan de acentuarse más, pero no pueden considerarse como tales debido a su obligada compatibilidad con diseños CISC de la familia Pentium a la que tributan vasallaje.

dos escuelas Sea como fuere, hemos de admitir CISC y RISC como una dualidad más en el diseño de computadores, cada una de ellas con sus ventajas y con sus carencias, y serán siempre factores exógenos los que sobreponderen las ventajas de uno frente a las de otro, provocando un desplazamiento del mercado en esa dirección.

un solo objetivo A continuación vamos a comparar estas dos grandes escuelas, y de paso demostraremos cómo, desde perspectivas enfrentadas, ambas persiguen un mismo objetivo: Minimizar el tiempo que un microprocesador invierte en la ejecución de un programa.

tres factores: Este tiempo puede obtenerse como el producto de tres factores:

- NI ① NI = Número de instrucciones máquina en que se descompone el programa fuente.

- ② CPI = Número medio de ciclos de reloj que se necesitan para ejecutar cada una de las instrucciones máquina anteriores. CPI
- ③ T = Tiempo del ciclo de reloj anterior (o su frecuencia F como magnitud inversa). T

Una filosofía de diseño CISC trata de reducir el primero de esos factores, proporcionando para ello instrucciones de muy alto nivel capaces de llevar a cabo operaciones complejas. Por el contrario, un diseño RISC está orientado a minimizar el segundo de los factores. Las dos alternativas tratan de aprovecharse de las mejoras en la tecnología de integración de chips (velocidad de conmutación de los transistores) para reducir al máximo la duración del ciclo de reloj del procesador, el tercero de los factores. CISC: NI y T  
RISC: CPI y T

La Unidad de Proceso de un RISC es del tipo carga/almacenamiento, esto es, las operaciones de lectura y escritura a memoria se aíslan del resto y el compilador las trata de forma separada para conseguir un alto grado de concurrencia en la ejecución de instrucciones. En cambio, una arquitectura CISC no puede aislar estas operaciones, al estar presentes en un mayor número de instrucciones. U. Proceso

La Unidad de Control, que en un procesador RISC es cableada, en uno de tipo CISC se implementa de forma microprogramada. Cada instrucción de un CISC tarda una serie de ciclos (entre 4 y 20 aproximadamente), y para cada uno de ellos la Unidad de Control tiene que activar unas señales de control que gobiernen el funcionamiento de la Unidad de Proceso. U. Control

La palabra de control de cada ciclo se almacena en una memoria de microprograma, donde la secuencia de palabras de control perteneciente a cada instrucción se agrupa formando microrutinas. Este diseño facilita la posterior modificación del procesador con un simple cambio en su memoria de microprograma, cualidad que han sabido aprovechar muy bien diseños contemporáneos como el reciente Crusoe de Transmeta. También ha permitido a otros fabricantes como Intel corregir sobre la marcha errores descubiertos en sus modelos con posterioridad a su lanzamiento al mercado, como el archiconocido de la unidad de punto flotante del Pentium. La cara negativa de la memoria de microprograma es que, puesto que la Unidad de Control tiene que esperar a que ésta responda para cada ciclo de ejecución del procesador, su funcionamiento se ralentiza bastante, y además, ocupa bastante área de integración en silicio. memoria de microprograma  
versatilidad  
ralentización

La necesidad de ejecutar una instrucción por ciclo obliga al RISC a cablear la Unidad de Control primando la velocidad por encima de la versatilidad. Como además disminuye su espacio de integración, tiene en su mano la consecución de frecuencias más elevadas. ventajas RISC

La mayoría de los aspectos negativos de un RISC aparecen precisamente como consecuencia de sus ventajas: La simplicidad de las instrucciones, por ejemplo, provoca que el rendimiento de una máquina RISC dependa mucho de la eficiencia del compilador. Por ello, el tiempo de desarrollo del software para una máquina RISC es potencialmente más elevado que para una CISC. El mayor número de instrucciones máquina que un programa RISC posee también repercute negativamente en el espacio que el programa ejecutable ocupa en memoria. inconv. RISC

La [tabla 3.13](#) muestra una comparativa que resume las principales diferencias entre ambas alternativas de diseño. ← pág. 98

## Diseño RISC

El denominador común de un microprocesador avanzado de los años 90 se asienta sobre los principios básicos de la filosofía de diseño RISC. Son los procesadores que se montan en la amplia gama de computadores que existen por encima de los PC: Estaciones de trabajo, servidores, computadores paralelos, ...

En este punto del capítulo, estamos en condiciones de dar un paso adelante para ilustrar cómo trabajan al nivel más ligado al conjunto de instrucciones.

	RISC	CISC
<b>Hardware</b>		
Tiempo de desarrollo	Bajo	Alto
Unidad de Control	Cableada	Microprogramada
Banco de registros	Extenso (256 o más)	Reducido (16, 32)
Espacio de integración	Reducido	Grande
<b>Software</b>		
Tiempo de desarrollo	Alto	Bajo
Compilador	Complejo y fundamental para la eficiencia	Más sencillo y menos crítico
Programa objeto	Muy largo	Compacto (20-30 % menor)
Grado de abstracción del HW	Bajo	Alto
<b>Diseño</b>		
Formato de las instrucciones	Fijo	Variable
Conjunto de Instrucciones	Pequeño (128 o menos)	Grande (200-500)
Modos de direccionamiento	Pocos y simples (hasta 4)	Muchos y Complejos
Características de las instrucciones	Sencillas y rápidas (1 ciclo de duración)	Potentes y lentas (de 4 a 20 ciclos)
<b>Ejemplos comerciales</b>	PowerPC(Motorola/IBM) R10000(SGI), Alpha(DEC)	80x86(Intel) 68000(Motorola)

TABLA 3.13: Características RISC y CISC frente a frente.

☛ pág. 58  
segmentación y  
superescalaridad

En primer lugar, diremos que un diseño tipo RISC favorece la implementación de las estrategias de paralelismo a nivel de instrucción vistas en la [sección 3.3](#): Su reducido conjunto de instrucciones simples hace que todas ellas tengan una duración similar, lo que permite una mejor segmentación y superescalaridad al estar sus etapas de ejecución más compensadas entre sí. Por otro lado, la propia simplicidad del procesador deja espacio de silicio libre para incluir cachés internas, ejecución fuera de orden, extensiones multimedia, y alguna que otra maravilla más.

proceso  
iterativo

El diseño de un procesador RISC transcurre como un proceso iterativo compuesto de dos fases que se realimentan entre sí con el fin de optimizar al máximo el resultado final: La selección del conjunto de instrucciones del procesador, y el diseño de la circuitería sobre la que éstas se ejecutan.

### 5.2.1 Selección del conjunto de instrucciones

base

Paso 1. Elección del núcleo básico. El proceso de obtención del conjunto de instrucciones del procesador parte de la selección de un núcleo de instrucciones básico, compuesto por instrucciones imprescindibles en cualquier procesador. Para ello se utiliza la experiencia previa que proporcionan los diseños de procesadores anteriores, y por intersección de los conjuntos de instrucciones más populares, llegamos al que será nuestro punto de partida.

ampliación

Paso 2. Selección de candidatos. A partir de ahí, se considera la extensión de este conjunto de instrucciones mínimo con operaciones y modos de direccionamiento candidatos a formar parte de la funcionalidad del procesador. La selección de candidatos se realiza en base al carácter que se le quiera dar al procesador y a parámetros de afinidad y coste.

purga

Paso 3. Criba de candidatos. Los candidatos que hayan pasado todos los filtros anteriores son entonces sometidos a pruebas de rendimiento sobre aplicaciones reales para cuantificar el beneficio que producen cuando el compilador las utiliza para la generación de código. Si la instrucción candidata produce una mejora significativa en la mayoría de códigos testeados, la instrucción

es finalmente incluida en el conjunto de instrucciones. Tanto el porcentaje de mejora como el de aplicaciones sobre las que produce el efecto deseado son parámetros que determinan el grosor del conjunto de instrucciones. A mayores porcentajes, mayor es la criba de candidatos y menor el conjunto de instrucciones, la funcionalidad, y el coste del microprocesador resultante. (Ejemplo: En el diseño del procesador MIPS, una instrucción se admitió si mejoraba en un 1 % el código de al menos el 90 % de los programas que se escogieron para las pruebas).

**Paso 4. Completitud.** El resultado de todo este proceso es un conjunto de instrucciones en buena sintonía con las necesidades reales de uso de un lenguaje de alto nivel. Cada instrucción es, o bien estructuralmente necesaria (esto es, no puede obtenerse en función de otras ya existentes), o bien ampliamente demandada durante el proceso de compilación de un programa.

**Paso 5. Eficiencia.** Una vez seleccionado un conjunto de instrucciones simple, debemos ocuparnos del segundo de los objetivos inherentes al diseño RISC: La ejecución de una instrucción por ciclo de reloj. De entre las instrucciones que dificultan este logro, sobresalen las de acceso a memoria y las de salto. A continuación comentaremos las optimizaciones más sobresalientes que un RISC realiza sobre ellas para lograr salirse con la suya.

#### □ Acceso a memoria: Carga retrasada

Para realizar operaciones con valores almacenados en memoria, tan sólo necesitamos estructuralmente una operación de carga del valor de una posición de memoria en un registro y su operación inversa de almacenamiento (escritura en memoria desde el banco de registros). El resto de operaciones necesitan referirse únicamente al banco de registros para obtener operandos y/o guardar resultados. Por eso se dice que una máquina RISC implementa una arquitectura de carga/almacenamiento. Las principales ventajas de este tratamiento en el acceso a memoria son básicamente tres:

- ① La reducción del número de accesos a memoria. Puesto que se dispone de un gran banco de registros, muchos de los valores requeridos por las instrucciones pueden encontrarse allí, ahorrando un eventual acceso a memoria. Esto permite relajar los requerimientos de ancho de banda entre el procesador y la memoria.
- ② El hecho de que todas las operaciones se realicen con los registros simplifica el conjunto de instrucciones y los modos de direccionamiento necesarios.
- ③ La eliminación de operaciones con memoria posibilita una mejor estrategia de alojamiento de valores en el banco de registros por parte del compilador. Esto termina de optimizar el número de accesos a memoria a la vez que reduce el ratio del número de instrucciones necesarias para llevar a cabo una tarea.

Estos tres factores ponen al alcance la ejecución de una instrucción por ciclo de reloj del procesador. Para los accesos a memoria que sean inevitables (incluido el fallo en caché), el procesador se ralentiza en principio el número de ciclos que la memoria tarde en responder.

Una forma de aprovechar estos ciclos de espera del procesador consiste en utilizar instrucciones de carga retrasada, esto es, redefinir la semántica de la instrucción de carga para que lleve asociada la ejecución de una serie de  $k$  instrucciones de relleno de forma inmediatamente consecutiva. Una instrucción de relleno puede ser cualquiera del conjunto de instrucciones siempre que reúna las siguientes dos condiciones: (a) No utilizar como operando el valor que se está trayendo de memoria en la operación de carga anterior, y (b) respetar la secuencia de ejecución de todas aquellas dependencias de datos y control que contenga el programa.

El valor de  $k$  o tamaño de la ventana de relleno para una instrucción de carga vendrá determinado por el tiempo de respuesta de la memoria en ciclos del procesador. Los buenos compiladores conocen este valor y se encargan de buscar instrucciones máquina que cumplan las condiciones

sintonía

presteza

arquitectura de carga/almacen.

tres ventajas:

menos accesos

menos modos

menos instrucciones

carga retrasada

instrucciones de relleno

ventana de relleno

compiladores

dependencias

de relleno, así como de reestructurar el código objeto para llenar las ventanas de relleno de las instrucciones de carga en la medida de lo posible. Los compiladores son bastante eficientes realizando este tipo de tareas, aunque su porcentaje de éxito será menor cuanto mayor sea el número de dependencias del programa y/o el tamaño de la ventana de relleno. En el peor de los casos, la ventana de relleno se completa con instrucciones NOP (de no operación) que simplemente dejan al procesador inactivo hasta que llegue de memoria el dato con el que ponerse a trabajar.

#### □ Instrucciones de salto: Salto retrasado

dirección de salto

El principal problema que introducen las instrucciones de salto proviene de su negativo impacto en el cauce segmentado que todo procesador RISC implementa para la ejecución de instrucciones: La dirección de destino del salto normalmente no se conoce hasta la última etapa de segmentación, es decir, una vez la instrucción de salto ha sido buscada y decodificada, se han obtenido sus operandos, y se ha evaluado la condición de salto en la etapa de ejecución.

Por tanto, el procesador comienza la etapa de búsqueda de la instrucción que sigue a la del salto cuando ésta se encuentra en su fase terminal de ejecución. Esto produce ciclos en los que tenemos varias unidades funcionales paradas (por ejemplo, las correspondientes a las fases de decodificación, búsqueda de operandos y ejecución).

salto retrasado

Es posible aprovechar estos ciclos ociosos utilizando para las instrucciones de salto la misma técnica de ventana de relleno ya utilizada para las instrucciones de carga: Redefiniendo la semántica de las instrucciones de salto con saltos retrasados para que contengan una ventana de relleno de tres instrucciones. Notificando esto al compilador, éste puede buscar instrucciones del programa que puedan ser insertadas en las posiciones de relleno de los saltos y reestructurar el código máquina de forma apropiada para que se aprovechen los ciclos de penalización asociados a la instrucción de salto.

### 5.2.2 Soporte software para una arquitectura RISC

Uno de los aspectos más criticados en los diseños RISC es el elevado número de instrucciones máquina en el que tiene que transformarse un programa para ser ejecutado. Dado que esta transformación es responsabilidad del compilador, resulta inevitable ligar la popularidad de los procesadores RISC con las mejoras en las técnicas de compilación.

#### □ Compilador

Puede decirse que no hay una técnica de compilación específica para un procesador RISC. Los métodos que se presentan a continuación también se aplican con arquitecturas CISC. Sin embargo, la simplicidad de una máquina RISC hace que el compilador encuentre en ella muchas más oportunidades de optimización que en una CISC.

Los compiladores más actuales son el resultado de una evolución en el proceso de traducción de lenguaje de alto nivel a lenguaje máquina. La eficiencia de un compilador se mide básicamente por el tamaño y la velocidad del código objeto que genera. Las técnicas avanzadas de compilación que mejor rendimiento producen en una arquitectura RISC son las siguientes:

- **Planificación de instrucciones:** La primera tarea que se exige a un compilador es la de aprovechar la presencia de instrucciones de carga y salto retrasado en el conjunto de instrucciones del procesador. Para ello, el compilador debe identificar las instrucciones máquina del código que puedan ser utilizadas como instrucciones de relleno y, posteriormente, reorganizar la ejecución del programa para que estas instrucciones cubran las ventanas de relleno de las instrucciones retrasadas.



### Ejemplo 3.20: USO DE LA VENTANA DE RELLENO DE UNA INSTRUCCIÓN DE CARGA RETRASADA POR PARTE DEL COMPILADOR

Considerar el programa fuente  $C := A + B$ ;  $F := 10$ ; transformado en el siguiente programa objeto:

```
Load R1, A
Load R2, B
Add R3, R1, R2      ← Espera la llegada de A y B de memoria
Load R4, 10         ← Instrucción de relleno
```

Un compilador más optimizado generaría la siguiente secuencia de instrucciones:

```
Load R1, A
Load R2, B
Load R4, 10         ← Se ejecuta la instrucción...
Add R3, R1, R2     ← ...mientras se esperan datos de memoria
```



### Ejemplo 3.21: USO DE LA VENTANA DE RELLENO DE UNA INSTRUCCIÓN DE SALTO RETRASADO POR PARTE DEL COMPILADOR

Considerar el siguiente programa objeto:

```
Move R1, R2
Move R3, R4         ← Instrucción de relleno
Add R1, R1, 1
Jump R1, 0, A      ← Instrucción de salto retrasado
...
A: Sub R5, R5, 1
```

Un compilador optimizado aprovecharía la ventana de relleno de la instrucción de salto para ejecutar la segunda de las instrucciones justo a continuación de aquella:

```
Move R1, R2
Add R1, R1, 1
Jump R1, 0, A
Move R3, R4        ← Instrucción cambiada de lugar
...
A: Sub R5, R5, 1
```

- **Uso optimizado de los registros:** El compilador emplea registros para almacenar los datos más frecuentemente utilizados, con el fin de minimizar el número de accesos a memoria.



### Ejemplo 3.22: OPTIMIZACIÓN EN EL USO DE REGISTROS POR PARTE DEL COMPILADOR

El siguiente fragmento de código objeto:

```
Load R1, B
Load R2, C
Add R3, R1, R2
Store R3, A
```

que ejecuta la sentencia  $A := B + C$  podría reducirse a una sola instrucción máquina si los valores de  $A$ ,  $B$  y  $C$  residen en los registros del procesador  $Ra$ ,  $Rb$  y  $Rc$  respectivamente. Esto es:

```
Add Ra, Rb, Rc
```

- **Eliminación de redundancias:** El compilador busca oportunidades para reutilizar resultados parciales y eliminar así computaciones redundantes.



### Ejemplo 3.23: ELIMINACIÓN DE COMPUTACIÓN REDUNDANTE DESDE EL COMPILADOR

Sea el siguiente código objeto, que ejecuta las sentencias  $A := B + X * Y$  y  $D = C + X * Y$ :

```
Mul R1, Rx, Ry
Add Ra, Rb, R1
Mul R2, Rx, Ry
Add Rd, Rc, R2
```

Para computar el valor a almacenar en  $D$ , el compilador puede aprovecharse de que el producto ya se computó anteriormente y está aún alojado en  $R1$  para tomarlo directamente de allí en lugar de volverlo a computar. Como resultado ahorramos una instrucción. El programa ahora quedaría:

```
Mul R1, Rx, Ry
Add Ra, Rb, R1
Add Rd, Rc, R1
```



- **Optimización de bucles:** El compilador optimiza también las operaciones que aparecen dentro de los bucles de un programa, con el fin de identificar expresiones invariantes y sacarlas fuera de éstos. invariantes
- **Optimización de operaciones:** En ocasiones una misma operación de alto nivel puede llevarse a cabo con distintas instrucciones máquina. En este caso, el compilador debe seleccionar aquella que sea más rápida. Por ejemplo,  $A := A + 1$  puede efectuarse a nivel máquina mediante `ADD Ra, Ra, 1` como suma en la ALU, o mediante `INC A`, que es mucho más rápida al tratarse únicamente de un incremento.
- **Eliminación de llamadas a subrutinas:** La ingeniería del software ha propugnado siempre la escritura de programas en estilo procedural, esto es, descomponer el programa en una serie de rutinas y/o procedimientos que encapsulan funciones a las que se llama desde el programa principal. Sin embargo, una instrucción máquina `CALL` (o de llamada a subrutina) resulta muy costosa de ejecutar para un RISC, principalmente por la necesidad de salvar el contexto del programa (el contador de programa, los registros a utilizar por la subrutina, ...) previamente al salto que realiza.

Podemos ahorrarnos estas operaciones efectuando desde el compilador lo que se conoce como **code inlining**, esto es, suprimir la instrucción `CALL` a costa de duplicar literalmente el código de la subrutina en el programa principal cada vez que se llama a ésta. Como resultado, el programa se ejecuta más rápidamente a costa de ocupar un mayor espacio en memoria. code inlining

Esta estrategia de compilación está teniendo una popularidad creciente que se sustenta en el hecho de que hoy día el tiempo de ejecución de un programa es un parámetro más prioritario que el espacio que ocupa en memoria. No obstante, aunque el *inlining* lo soportan muchos compiladores, la mayoría de ellos no lo realiza a no ser que el usuario así se lo indique de forma explícita mediante alguna de las opciones o niveles de compilación disponibles.

#### □ Sistema Operativo

Mientras los sistemas operativos orientados a máquinas CISC suelen proporcionar un conjunto de servicios muy elaborados, los principios de diseño RISC apuestan más por la calidad que por la cantidad de los servicios prestados, tratando en todo momento de evitar la complejidad salvo en casos plenamente justificados. Se favorece así a las operaciones que son más utilizadas, proporcionando una buena velocidad de operación a través de controles mínimos y simples. minimalista

Algunos de los mecanismos que los diseños RISC utilizan a nivel de sistema operativo para aumentar el rendimiento de una máquina sin añadir una complejidad excesiva a su hardware son los siguientes: recursos

- **Búfer para la traducción de direcciones virtuales a físicas (TLB):** Agilizar esta traducción que tiene lugar por cada operación de acceso a memoria resulta esencial para la implementación de un potente sistema operativo. Aunque la TLB no es un mecanismo exclusivo de los procesadores RISC, sí es cierto que su reducido espacio de integración permite que la TLB pueda ser integrada dentro del propio chip o extenderse a lo largo de un espacio mayor de silicio, ahorrando el tiempo que se pierde para transferir la dirección virtual a una TLB externa en el primer caso, y reduciendo el riesgo de no encontrar la traducción en la TLB en el segundo. TLB
- **Mecanismos de protección:** Los sistemas operativos utilizan modos de funcionamiento que restringen el acceso del usuario a ciertas partes delicadas del sistema que son gestionadas en exclusiva por parte del sistema operativo. Frente a los múltiples modos y mecanismos de

protección que se suministran en una arquitectura CISC, los RISC proporcionan un control que normalmente se limita a la simple distinción entre modo usuario y supervisor.

- **Gestión de interrupciones:** La gran mayoría de los eventos externos al procesador son gestionados por éste a través de mecanismos de interrupción. Muchos procesadores CISC proporcionan controladores hardware dedicados a la gestión de interrupciones (como el PIC 8259 usado en la familia de los Intel 80x86) para salvar una gran cantidad de información de estado del procesador y generar la dirección del vector de interrupción al que transferir el control en respuesta a la interrupción. Esto añade complejidad hardware, pero no necesariamente simplifica la tarea del sistema operativo. Por ejemplo, muchos sistemas operativos no usan los diferentes vectores de interrupción, sino que en su lugar ejecutan un manejador de interrupciones común a todas ellas que determina de forma precisa las necesidades de procesamiento de la interrupción y la información de estado del procesador que se necesita salvar.

vectores de  
interrupción

### 5.3 ▶ Diseños VLIW

En el diseño de los nuevos conjuntos de instrucciones como el IA-64 de Intel para su Itanium y el x86-64 de AMD para su K8, se ha retomado un concepto que data de comienzos de los años 80: El **VLIW (Very Long Instruction Word)**.

origen  
compilador

Este concepto emerge en un contexto histórico muy particular, al calor de los primeros resultados que arrojan un balance favorable en la capacidad de un compilador para identificar las oportunidades de ejecución simultánea que esconde un programa secuencial. Se trata así de que el compilador conecte directamente con alguna(s) de las formas de paralelismo a nivel de instrucción descritas en la [sección 3.3](#).

pág. 58

información  
privilegiada

El compilador puede ser capaz de generar un código en el que se da por hecha la presencia de, por ejemplo, tres unidades funcionales de suma (superescalaridad), desgranando operaciones para cada una de ellas de forma explícita en el código de la instrucción. También puede conocer la presencia de, por ejemplo, diez etapas de ejecución segmentadas, así como la incidencia que tienen las dependencias en el código que pasa por sus manos, siendo (supuestamente) capaz de analizarlo y transmitir esta información en el propio formato de instrucción máquina.

pág. 105

idea

pág. 60

La [figura 3.13](#) muestra la idea que hay detrás de una filosofía de ejecución de instrucciones en una arquitectura VLIW. Podemos contrastarla con la [figura 3.5](#), en la que mostrábamos la ejecución segmentada y superescalar utilizada en todos los diseños de la quinta y sexta generación de microprocesadores.

el pionero

El primer diseñador que apostó por esta idea fue Josh Fisher en su proyecto ELI (Univ. Yale - 1981), lo que derivó en la iniciativa empresarial Multiflow, fundada por él en 1984. Según unas fuentes, Multiflow vendió más de 100 multiprocesadores, algunos hasta con 28 microprocesadores trabajando en paralelo. Según otras, sólo se vendió una máquina que tampoco terminó de funcionar del todo bien. El caso es que económicamente la idea resultó un fiasco, y la empresa cerró en 1990 tras serios problemas financieros.

ambición

Pero una cosa es el mundo mercantil, y otra muy distinta el ingenieril. La historia de la computación está llena de proyectos sabiamente concebidos que no tuvieron calado en el mercado y otros a los que éste apadrinó aún no se sabe cómo. Desconocemos si la implementación física de aquella máquina estaba o no a la altura de la idea, pero hay algo que sí cargaríamos en el debe de Fisher: Su excesiva pretenciosidad. El concepto VLIW vale para aplicarlo a cuatro o seis caminos de ejecución independientes (tres ha sido el número escogido por Intel y HP en el IA-64 de su Itanium, y cuatro el seleccionado por Transmeta en su Crusoe), pero colocar un formato de instrucción de muchos cientos de bits y tratar de coordinar con él hasta 28 caminos de ejecución de forma simultánea cuando el software continúa conceptualmente varado en una ejecución secuencial, parece un sueño demasiado bonito como para hacerse realidad sin contratiempo alguno.

utopía

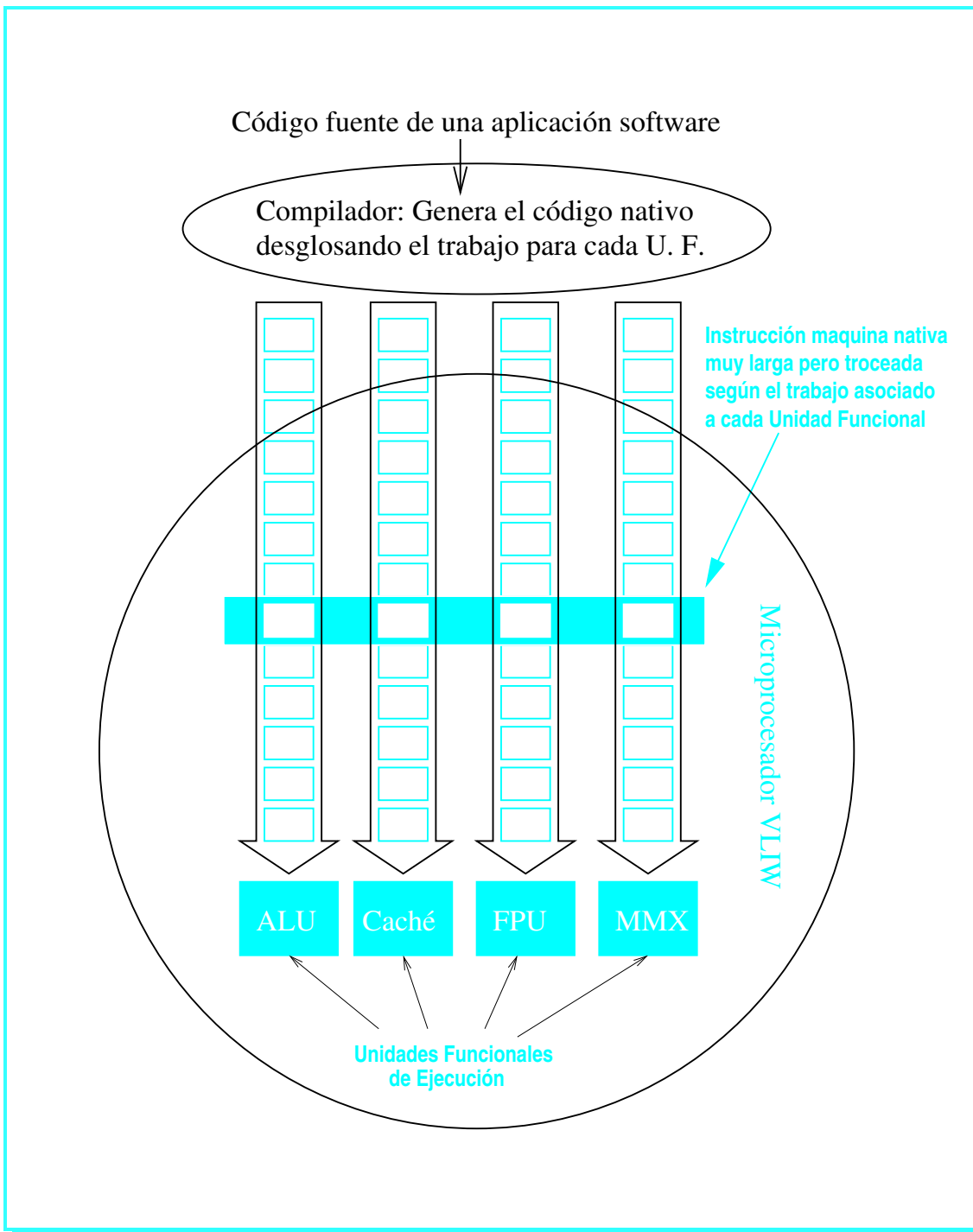


FIGURA 3.13: La filosofía de diseño VLIW (Very Long Instruction Word). Se asume un código máquina muy largo compuesto de compartimentos estancos dedicados a cada unidad funcional existente en el hardware, y el responsable de este desglose es el compilador en la capa software del sistema.

En general, la aspiración de una arquitectura VLIW consiste en crear una máquina con muchas unidades funcionales que puedan trabajar de forma simultánea, cada una según le indica un segmento del formato de instrucción. De este hecho podemos deducir que la instrucción tiene

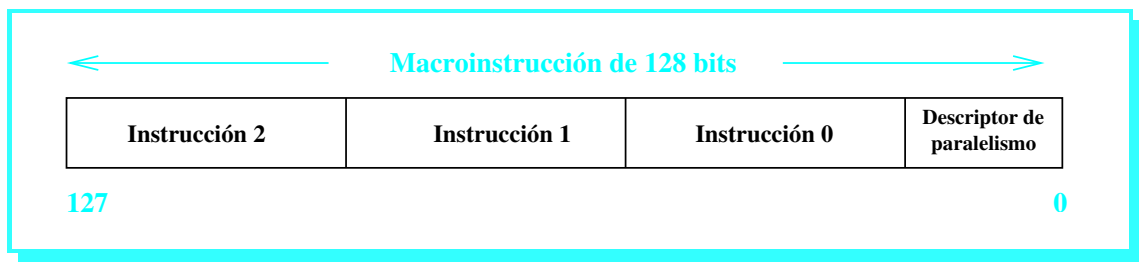


FIGURA 3.14: Formato de instrucción IA-64 de 128 bits correspondiente al microprocesador Itanium de 64 bits. Consta de tres instrucciones máquina agrupadas, junto con un campo adicional que describe el grado de paralelismo existente entre ellas.

**calidades** una anchura muy generosa, **calidad** de la que precisamente deriva el nombre VLIW. Se retoma así la máquina microprogramada ancha (claro concepto CISC), en la que las microinstrucciones que controlan la ejecución de una instrucción no son generadas por la unidad de control (como ocurre en los CISC), sino directamente desde el compilador (hecho más ligado a los RISC).

**híbrido RISC/CISC** En realidad, el mejor resumen que podemos hacer de VLIW es que trata de quedarse con lo mejor de los RISC y de los CISC; desgraciadamente, también arrastra de forma compartida algunas de sus carencias. La principal es una mayor dependencia del compilador que las máquinas RISC, lo que ya nos parece excesivo. Digamos que RISC sabe quedarse en un punto de equilibrio: Aquel en el que se le puede sacar más partido al compilador de lo que lo hace el CISC, porque se aprovechan tareas en las que el compilador se desenvuelve con maestría, pero sin llegar a responsabilizarle de otras tareas en las que se encuentra desamparado porque además de no realizarlas tan bien, el programa fuente y la circuitería se encuentran mirando para otro lado.

En el mundo de la informática se han sucedido recientemente dos iniciativas VLIW de notable repercusión, aunque ninguna de ellas tiene como epicentro la arquitectura PC:

**Itanium** ■ El Itanium, antaño conocido como Merced, en el mercado de grandes estaciones de trabajo y servidores. El formato de instrucción para su conjunto de instrucciones IA-64 se adjunta en la [figura 3.14](#).

**Crusoe** ■ El Crusoe de Transmeta orientado al segmento de los portátiles de muy bajo consumo. La [figura 3.15](#) muestra su formato de instrucción, mientras que en la [figura 3.16](#) incluimos un bosquejo de su arquitectura.

[pág. 107](#) ➔  
[pág. 107](#) ➔

## 5.4 ▶ Instrucciones multimedia

**revolución** El fenómeno de las instrucciones multimedia ha constituido toda una revolución en lo concerniente al conjunto de instrucciones de un procesador, provocando fuertes repercusiones sobre su arquitectura hardware.

**CISC operandos** La senda descrita por las instrucciones multimedia supone un nuevo acercamiento hacia una filosofía CISC, en tanto en cuanto se apuesta por instrucciones de compleja decodificación que llevan encapsulado en su formato una serie de operandos variable en número y longitud.

**propósito específico** Además, nos encontramos frente a una discontinuidad en el concepto de conjunto de instrucciones para un microprocesador de propósito general, ya que se incluyen instrucciones cuya finalidad está claramente sesgada por las aplicaciones específicas a las que se encuentra orientado.

**exigencias del mercado** La decisión de incorporar instrucciones multimedia al repertorio de instrucciones de un procesador hay que buscarla en las exigencias del mercado. Si la tecnología siempre se ha construido en función de las demandas establecidas por la sociedad a la que sirve, el giro dado por el mer-

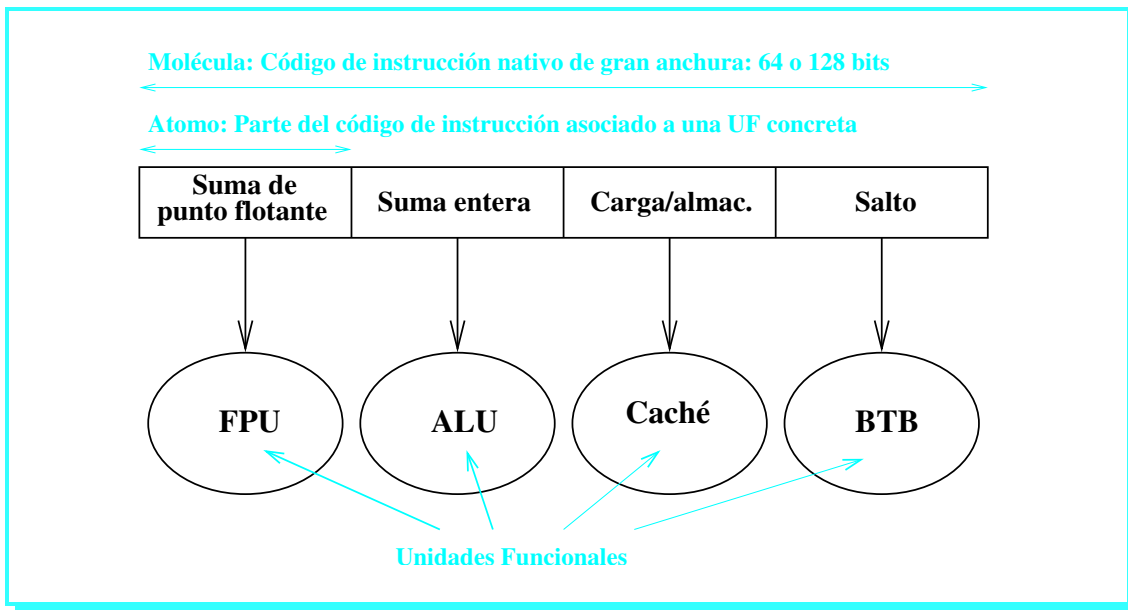


FIGURA 3.15: La filosofía de diseño VLIW (Very Long Instruction Word) aplicada sobre el código de instrucción del microprocesador Crusoe de Transmeta.

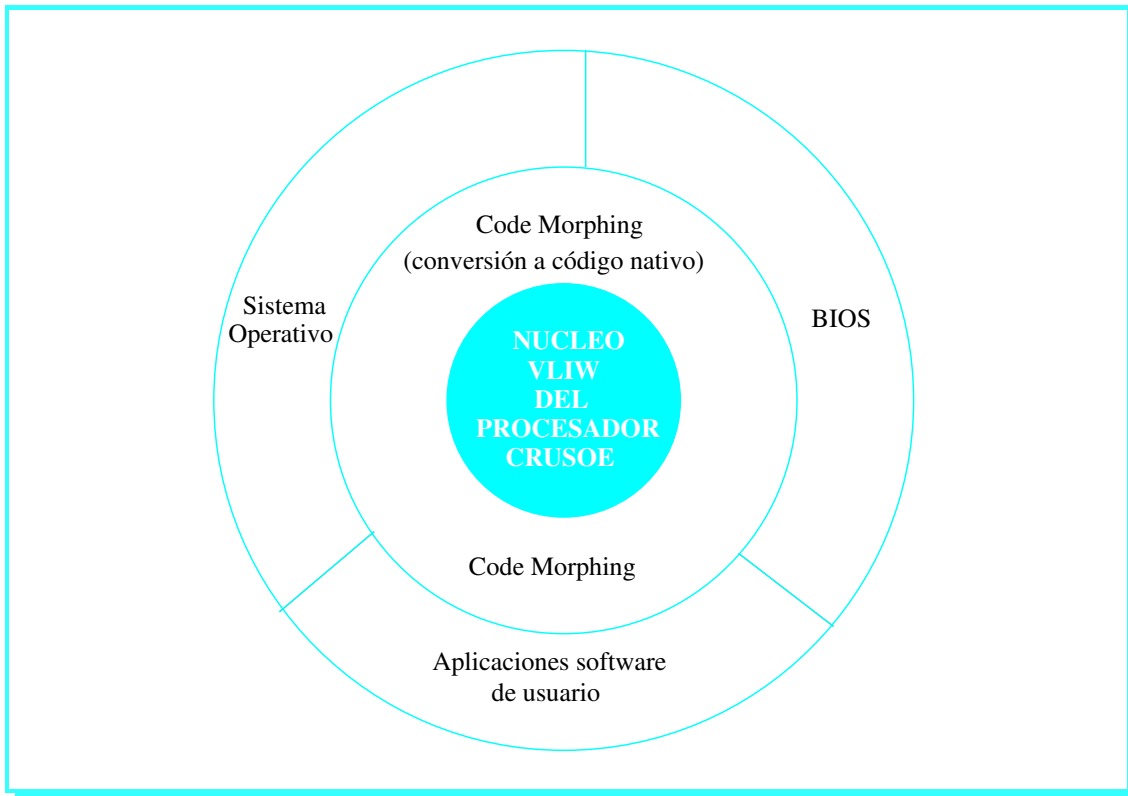


FIGURA 3.16: Composición del procesador Crusoe de Transmeta y relación con la capa software con la que habilita una frontera muy difusa.

cado hacia aplicaciones multimedia (vídeo interactivo, gráficos 3D, animación, sonido, realidad virtual, ...) exigía a los fabricantes de procesadores estar a la altura de las circunstancias.

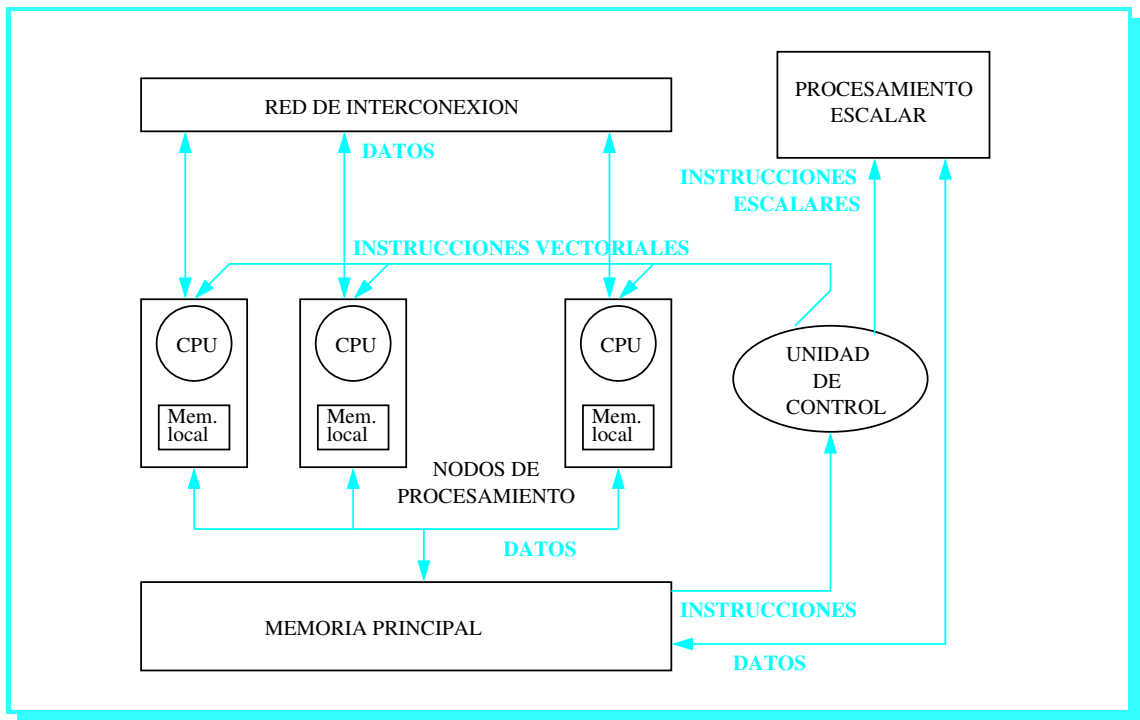


FIGURA 3.17: Diagrama de bloques de un computador SIMD (Simple flujo de Instrucciones, Múltiple flujo de Datos).

#### 5.4.1 El concepto: SIMD

La idea básica sobre la que subyacen todas las instrucciones multimedia parte del concepto SIMD, **originado** en los años 70 en el ámbito de los supercomputadores, esto es, arquitecturas compuestas de múltiples procesadores.

Un computador **SIMD (Simple Instruction Multiple Data)** se compone de un conjunto de nodos de procesamiento y un procesador escalar, todos operando bajo las órdenes de una Unidad de Control común que centraliza el funcionamiento de toda la máquina (ver figura 3.17).

La Unidad de Control busca y decodifica instrucciones de la memoria principal y, dependiendo de su tipo, envía las correspondientes señales de control al procesador escalar o a los nodos de procesamiento para su ejecución. Así, si se trata de una instrucción escalar, sólo funcionará el procesador escalar; en caso contrario, funcionarán todos los nodos de procesamiento en paralelo, los cuales ejecutarán la misma instrucción pero sobre datos diferentes.

La aplicación del concepto SIMD a un solo microprocesador es análoga a la ya comentada. Una única Unidad de Control busca y decodifica las instrucciones convencionales y las SIMD. Cuando llega una instrucción normal, el procesador actúa como siempre, en semejanza con el procesador escalar anterior. En cambio, cuando se trata de una instrucción SIMD, la Unidad de Control envía señales de control a cada una de las unidades en punto flotante, las cuales ejecutan la misma operación pero sobre distintos datos almacenados en sus bancos de registros.

Conceptualmente, SIMD trata de explotar el paralelismo que presenta el conjunto de datos de una aplicación, en contraposición con el paralelismo a nivel de instrucción, donde se paraleliza la secuencia de ejecución de las instrucciones de la aplicación. Por tanto, el rendimiento de un procesador con extensiones SIMD será mucho mayor en programas con abundante cálculo sobre vectores o *arrays* de datos de grandes dimensiones. Por otra parte, el sincronismo de instrucción inherente a la sección SIMD del procesador deja poca flexibilidad para su diseño.

Instrucción	Operandos	Variantes	Duración	Descripción de la operación
Aritméticas				
PADD	B, W, D	3	1	Suma con redondeo
PADDSS	B, W	2	1	Suma con acarreo
PADDUS	B, W	2	1	Suma sin acarreo
PSUB	B, W, D	3	1	Resta con redondeo
PSUBS	B, W	2	1	Resta con acarreo
PSUBUS	B, W	2	1	Resta sin acarreo
PMULHW	HW	1	3	Producto del byte alto
PMULLW	LW	1	3	Producto del byte bajo
PMADDWD	WD	1	3	Explicado en el pié de figura
Comparativas				
PCMPEQ	B, W, D	3	1	Comparación <i>igual-que</i>
PCMPGT	B, W, D	3	1	Comparación <i>mayor-que</i>
Lógicas				
PAND	Q	1	1	Aplican los operandos AND, NAND, OR y XOR resp. sobre el operando fuente y destino, guardando el resultado en este último.
PNAND	Q	1	1	
POR	Q	1	1	
PXOR	Q	1	1	
De conversión				
PACKUSWB	WB	1	1	Empaqueta sin acarreo
PACKSS	WB, DW	2	1	Empaqueta con acarreo
PUNPCKH	BW, WD, DQ	3	1	Desempaqueta datos de mayor peso
PUNPCKL	BW, WD, DQ	3	1	Desempaqueta datos de menor peso
Para el desplazamiento de bits en un registro				
PSLL	W, D, Q	6	1	Desplazamiento lógico a izquierda
PSRL	W, D, Q	6	1	Desplazamiento lógico a derecha
PSRA	W, D	4	1	Desplazamiento aritmético a derecha
Para el movimiento de datos entre registros				
MOV	D, Q	4	1	Transferencia entre registros MMX
Para el cambio de estado de los registros				
EMMS	E	1	1	Pone a cero el byte de estado MMX

TABLA 3.14: El conjunto de instrucciones MMX. Los operandos que acepta cada instrucción van en función de las variantes que admite y su duración. La instrucción PMADDWD multiplica por separado las 4 palabras empaquetadas del operando destino por las cuatro del origen. Los dos resultados de multiplicar las dos parejas de palabras de menor peso se suman, y el resultado se redondea para ser almacenado en la palabra doble de menor peso del operando destino. Con las dos palabras de mayor peso se procede exactamente igual.

#### 5.4.2 El embrión: MMX

El punto de partida en la inclusión de instrucciones multimedia en los microprocesadores para PC fue el conjunto MMX (MultiMedia eXtensions en primera instancia, y luego redefinido por Intel como Matrix Math eXtensions), desarrollado al alimón por Intel y AMD para sus procesadores Pentium MMX y K6, respectivamente.

Con anterioridad a la llegada de las instrucciones MMX, el procesamiento de tipos de datos de 8 o 16 bits en los microprocesadores infrautilizaba los recursos hardware, ya que el ancho de banda para datos en las unidades funcionales de cálculo de los microprocesadores era de 32 o 64 bits, de los que sólo se empleaban los 8 o 16 bits menos significativos.

antecedentes

SIMD

MMX agrupa estos datos en grupos de 64 bits que luego son procesados individual pero concurrentemente mediante la aplicación del concepto SIMD ya comentado, con lo que se aprovechan mejor los recursos de que dispone el microprocesador. Esto, unido a la explotación del paralelismo inherente a la mayoría de algoritmos multimedia, pone al alcance de estas aplicaciones mejoras en velocidad de entre el 50% y el 100%.

mejoras en  
velocidadindependencia  
de la  
arquitectura  
secuelas

Otra importante característica con que se dota la implementación MMX es la de conservar su independencia de la arquitectura microprogramada, tanto del Pentium como del K6, con el fin de que el juego de instrucciones MMX resultara fácilmente escalable con futuros diseños arquitecturales o frecuencias de reloj más elevadas. Esta virtud se vería ampliamente refrendada con el paso del tiempo, al potenciar la fácil aparición de secuelas que fueron paulatinamente extendiendo el conjunto de instrucciones multimedia original.

57  
instrucciones  
pág. 109

Este primer conjunto de instrucciones estuvo formado por un total de 57 instrucciones que se resumen en la [tabla 3.14](#). Como podemos apreciar, muchas de ellas son variantes de una misma operación, que puede ser definida sobre diferentes subconjuntos de datos, todos ellos de tipo entero. Las distintas variantes han sido abreviadas de la siguiente forma:

- ❖ B: Byte. 8 operandos de entrada de 8 bits cada uno.
- ❖ W: Word. 4 operandos de entrada de 16 bits cada uno.
- ❖ D: Double word. 2 operandos de 32 bits.
- ❖ Q: Quad word. Un único operando de entrada de 64 bits.
- ❖ WB: Word - Byte. 4 operandos de entrada de 16 bits cada uno y 8 operandos de salida de 8 bits cada uno.
- ❖ DW: Double word - Word. 2 operandos de entrada de 32 bits cada uno y 4 operandos de salida de 16 bits cada uno.
- ❖ QD: Quad Word - Double Word. Un operando de entrada de 64 bits; 2 de salida de 32 bits.
- ❖ BW: Byte - Word. 8 operandos de entrada de 8 bits y 4 de salida de 16 bits.
- ❖ WD: Word - Double Word. 4 operandos de entrada de 16 bits y 2 de salida de 32 bits.
- ❖ DQ: Double Word - Quad Word. 2 operandos de entrada de 32 bits y 1 de salida de 64 bits.
- ❖ HW: High Word. Los 8 bits más significativos de un dato de 16 bits.
- ❖ LW: Low Word. Los 8 bits menos significativos de un dato de 16 bits.
- ❖ E: Etiqueta. El byte de etiqueta que señala para cada registro de punto flotante su uso como tal o como registro MMX. Se utiliza un bit de la etiqueta por cada 8 bits de datos MMX.

### 5.4.3 Criterios para la selección de instrucciones

Seleccionar una instrucción para que forme parte del repertorio que acepta un procesador es una tarea bastante más peliaguda de lo que a simple vista nos parece. Para que la finalidad última de mejorar el rendimiento se cumpla, las instrucciones elegidas deben satisfacer las tres premisas siguientes:

sencillas

- ❶ Ser sencillas. Se trata de que no desentonen con las ya existentes. Si estamos en un procesador CISC, tendremos algo más de margen, pero si es un RISC, una excesiva complejidad afectará a la velocidad que alcanzaba el procesador sobre las instrucciones antiguas.



- ② Ser utilizadas por una amplia mayoría de aplicaciones multimedia. Se trata de identificar las primitivas más representativas de sus cualidades intrínsecas. Sólo así se amortizará el coste de la circuitería responsable de su ejecución. Una nueva instrucción tampoco es gratuita para el tiempo de decodificación de instrucción, ni para el tiempo que dedica la Unidad de Control a secuenciar todos los eventos en el corazón del microprocesador. Recordemos esa máxima del hardware: *Más grande, más lento*. útiles
- ③ Poder aprovechar las mejoras tecnológicas por venir, principalmente un número creciente de transistores y una mayor frecuencia de reloj. optimizables

El último criterio quedó satisfecho con el diseño escalable que ya comentamos. Para cumplir el primero, lo primordial es conocer el conjunto de instrucciones existentes. Y para cumplir el segundo, lo esencial es estudiar el comportamiento de una aplicación multimedia, cuya caracterización pasamos a desglosar a continuación.

#### ❑ Caracterización software de una aplicación multimedia

Las aplicaciones multimedia tienen todas unos rasgos muy similares. Desde la perspectiva software más ligada a su constitución, destacaríamos los tres siguientes:

rasgos  
software

- ① **Tipos de datos de tamaño reducido organizados en estructuras grandes.** Por ejemplo, una imagen en la pantalla se compone de infinidad de píxeles de 8 bits (12 ó 16 en media y alta resolución), y una partitura musical, de muestras de sonido de 16 bits (24 bits si el sonido es alta fidelidad). datos pequeños
- ② **Operaciones repetitivas simples y regulares.** Por ejemplo, el suavizado de una imagen (actualizar cada píxel con la media de una pequeña región de la imagen centrada en él con la finalidad de reducir su contraste). operaciones repetitivas
- ③ **Alto grado de paralelismo inherente.** Por ejemplo, el suavizado anterior puede realizarse concurrentemente sobre distintas partes de la imagen. paralelismo

#### ❑ Caracterización hardware de una aplicación multimedia

Desde la perspectiva hardware más ligada al microprocesador, señalaríamos cuatro rasgos como los más sobresalientes de una aplicación multimedia:

rasgos  
hardware

- ① **Respuesta en tiempo real.** Por ejemplo, a la hora de visualizar una secuencia de vídeo, resulta más adecuado prescindir de algunos fotogramas que mostrar todos y ralentizar la imagen. Se hace necesario reservar recursos y anticipar el tiempo necesario para realizar una tarea. respuesta
- ② **Pobre localidad temporal y alta localidad espacial en el acceso a los datos.** En general, el volumen de datos referenciado más recientemente por el procesador (concepto de **conjunto de trabajo** para una aplicación) es superior al primer nivel de memoria caché, siendo necesaria la intervención del segundo nivel. localidad
- ③ **Control de múltiples flujos simultáneos.** Por ejemplo, los datos de una secuencia de imágenes y sus efectos especiales de sonido asociados. control
- ④ **Elevado ancho de banda.** Principalmente, entre los datos y las unidades funcionales del procesador en que son procesados. ancho de banda

Para cumplir con los dos primeros requisitos, se dobló la capacidad de la caché L1 en la versión MMX del propio Pentium, y posteriormente la velocidad de la L2 en los procesadores que

recursos  
hardware

albergaron las futuras extensiones, como las SSE de Intel y las Enhanced 3DNow! de AMD. Estas dos también habilitaron instrucciones especiales para un control más exhaustivo de los recursos de la memoria caché, así como bancos de registros dedicados.

Respecto a las dos últimas cualidades, se trataron de cubrir mediante la ampliación del carácter superescalar del procesador en lo que respecta a la interrelación de las unidades funcionales con las demás (replicación de unidades multimedia) y la habilitación de puertos de conexión independientes para ellas.

#### 5.4.4 Compatibilidad

##### ❑ A nivel hardware: Registros extensos

la clave Un aspecto fundamental para el éxito de la iniciativa MMX era garantizar la compatibilidad con los mismos modelos de microprocesadores sin extensiones MMX. Se sabía que si se suministraba rendimiento adicional pero no se podían ejecutar las aplicaciones software ya existentes, el mercado daría la espalda a la idea por las muchas limitaciones que tendría que soportar el usuario final.

coexistencia Pero el problema era más complejo aún, pues se necesitaba asegurar la coexistencia de las viejas aplicaciones con las nuevas MMX en una ejecución multiproceso. Esto se consiguió conmutando el procesador a modo de ejecución en punto flotante cuando los procesos multimedia solicitaban sus servicios, reutilizándose el banco de 8 registros en punto flotante de 80 bits de que disponían tanto el Pentium como el K6 para almacenar los operandos multimedia de 64 bits.

cómo se ejecutan De esta manera, desde el punto de vista de las instrucciones, la compatibilidad estaba asegurada al definir las instrucciones MMX como enteras normales, y desde el punto de vista de los datos, los tipos de datos MMX de 64 bits se mapeaban sobre registros en punto flotante de 80 bits. Con esta operativa en marcha, cuando una aplicación se ejecuta, comprueba antes la presencia de hardware MMX: Si el procesador se encuentra dotado de él, utiliza las nuevas instrucciones y se ejecuta disfrazándose de aplicación en punto flotante; en caso contrario, se ejecutará como una aplicación entera normal.

##### ❑ A nivel software: Compiladores, ensambladores y API

varios enfoques La tecnología MMX se aprovechó de la capa software del sistema para transformar las aplicaciones existentes en otras que pudieran beneficiarse del nuevo repertorio de instrucciones. La conexión puede aprovecharse a diferentes niveles, obteniéndose en cada caso un aumento de rendimiento distinto respecto a una misma aplicación:

reescribir código **1 Vía 1: Priorizar el rendimiento.** El método de programación más eficaz consiste en transformar la aplicación en un programa nativo para instrucciones MMX. Para ello, reescribiremos las funciones que consumen más tiempo para que puedan llamar a las nuevas instrucciones, lo cual puede realizarse de forma automática o manual dependiendo del software que nos ayude:

- compilador
- Si disponemos de un compilador que soporte las nuevas instrucciones, bastaría con recompilar el programa fuente (escrito en C, por ejemplo) y utilizar el nuevo fichero ejecutable.
  - La alternativa manual consiste en instalar un parche para el Macro Assembler de Microsoft que haga a éste generar los códigos de operación de las nuevas instrucciones. Tanto Intel como AMD han proporcionado estos parches a los clientes de sus microprocesadores con extensiones multimedia a través de sus páginas Web. De optar por esta vía, deberemos codificar directamente en lenguaje ensamblador las rutinas cuyo rendimiento se desee mejorar, y hacer uso de los mnemotécnicos de las nuevas instrucciones.
- ensamblador

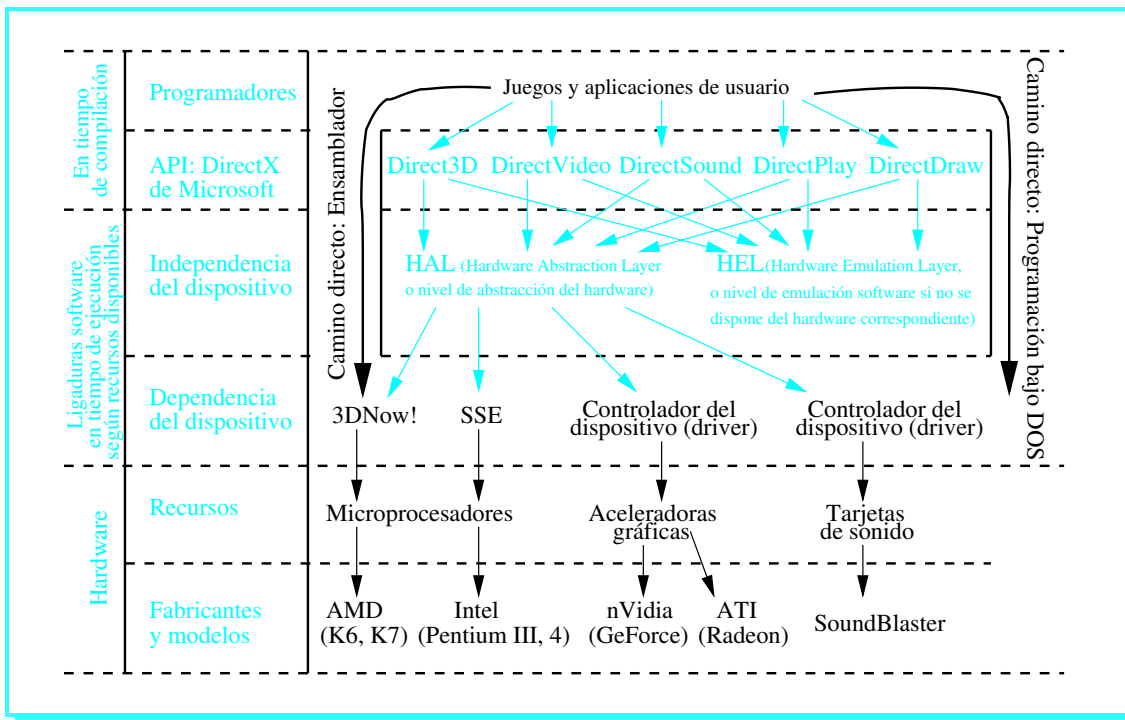


FIGURA 3.18: Las dos vías para la programación de aplicaciones utilizando los conjuntos de instrucciones multimedia. El empleo de un mayor número de capas sacrifica el rendimiento del hardware en favor de la compatibilidad y la facilidad de programación software.

🔗 **Vía 2: Priorizar la facilidad de programación.** También es posible conseguir un aumento de rendimiento si se utiliza una API que esté escrita a bajo nivel utilizando las extensiones MMX.

Una **API (Application Program Interface)** es un conjunto de funciones y librerías especializadas que define un interfaz a bajo nivel para los programas de aplicación de los usuarios. Estas funciones se ejecutan desde los programas como las tradicionales llamadas al sistema, sólo que en vez de suministrarlas el sistema operativo, lo hace una capa superior que se suministra en un paquete software instalable de forma similar a un *driver*.

De esta manera, llamando a las funciones de esa API especial para MMX, nos aprovecharíamos indirectamente de las nuevas instrucciones, aunque la ganancia esperada sería inferior a la obtenida en el caso anterior.

La **figura 3.18** sintetiza las dos alternativas descritas. Ambas ofrecen una dicotomía software similar a la que existe entre programar en ensamblador y en un lenguaje de alto nivel, apostando las principales compañías del sector por la segunda alternativa con objeto de reducir al máximo el tiempo de desarrollo de sus productos.

En esta línea, los principales ejemplos son firmas como 3Dfx Interactive con su API Glide para MMX, Silicon Graphics con el estándar OpenGL y Microsoft con DirectX, que aglutina diferentes API para gráficos, sonido y todo tipo de dispositivos en el contexto de los juegos para PC.

DirectX ha proporcionado sucesivamente versiones más ampliadas para dar cobertura a las posteriores extensiones multimedia, como DirectX 6.0, API que da cobertura al conjunto de instrucciones 3DNow!, DirectX 7.0, API para SSE y Enhanced 3DNow!, y finalmente DirectX 8.0, API para SSE2 y 3DNow! Professional. Microsoft también incluye paulatinamente las nuevas versiones de DirectX en sus viejos sistemas operativos. La cobertura de las API de Microsoft para las distintas extensiones multimedia se resume en la **tabla 3.15**.

MAGNITUDES

usar una API

qué es una API

dicotomía

principales API para MMX

versiones de DirectX

Versión de DirectX	API que incorpora y características más reseñables	Cobertura multimedia	Primer S.O. que lo incluye
1.0 (Abr'95)	DirectDraw, DirectInput, DirectPlay, DirectSound		Windows'95
2.0	Direct3D		
3.0 (primera vers estable)	_Sound3D suplanta a _Sound y se amplía DirectInput	MMX básico	
5.0	Setup (autoconfiguración), multimonitor, media layer, vertex buffers	MMX completo	Windows'98
5.2	Amplía DirectPlay		
6.0 (Oct'98)	Compresión de texturas, stencil buffers bump mapping, nuevo panel de control	MMX óptimo, 3DNow!	
6.1 (Feb'99)	DirectMusic		
7.0 (Oct'99)	Interfaz con VisualBasic, vertex blending, mejores gráficos y sonido 3D	SSE,Enhanced 3DNow!	Windows'2000, Windows Me
8.0 (Nov'00)	_Draw y _3D se funden en _Graphics, _Music y _Sound se funden en _Audio, DirectShow	SSE2, 3DNow! Professional	
8.1 (Ago'01)	Actualización de imagen más rápida, soporte para amplio número de jugadores		Windows'XP

TABLA 3.15: Evolución histórica de la API DirectX de Microsoft y cobertura de los sucesivos conjuntos de instrucciones multimedia y sistemas operativos en cada una de sus versiones.

La clara orientación de DirectX hacia la industria de los juegos es lo que explica su floja penetración en Windows'NT y la reciente ampliación de DirectX en el ámbito de la video-consola X-Box. DirectX actúa así como capa software independiente de la plataforma hardware, con dos claras ventajas:

- 1 Para la industria de los juegos para PC, permite que las novedades de su ingente mercado puedan salir al mismo tiempo para X-Box sin coste de desarrollo adicional.
- 2 Para el usuario de PC, permite que cada vez que éste incorpore nuevos recursos hardware, pueda aprovechar sus prestaciones multimedia sin más que actualizar la versión de DirectX, disponible a través de la World Wide Web en <http://www.microsoft.com/directx>. Siempre que la aplicación del usuario esté programada utilizando el API DirectX, todo quedará transparente al usuario.

Web!

Fuera del mundo Microsoft, tanto en arquitecturas más potentes y de grandes recursos gráficos (como las estaciones de trabajo) como en otros sistemas operativos (léase Linux), la mejor elección es utilizar el estándar OpenGL.

#### 5.4.5 Ampliaciones al conjunto MMX

Después de una guerra de pleitos entre Intel y AMD por atribuirse la autoría de la idea MMX (guerra que terminó en los juzgados con decisión salomónica), se sucedió una cruenta batalla por desarrollar versiones mejoradas de la original.

AMD respondió antes, a mediados del año 1998 con su 3DNow! para el K6-2, mientras que Intel lo hizo de forma más contundente un año más tarde con su SSE para el Pentium III (70 nuevas instrucciones frente a sólo 21 en 3DNow!). La historia volvió a repetirse poco más tarde: AMD extendió el conjunto con 24 nuevas instrucciones en el Enhanced 3DNow! para su K7 (verano de

la batalla por la sucesión

3DNow!  
SSE

Enhanced 3DNow!

PC y X-Box

estaciones  
Linux

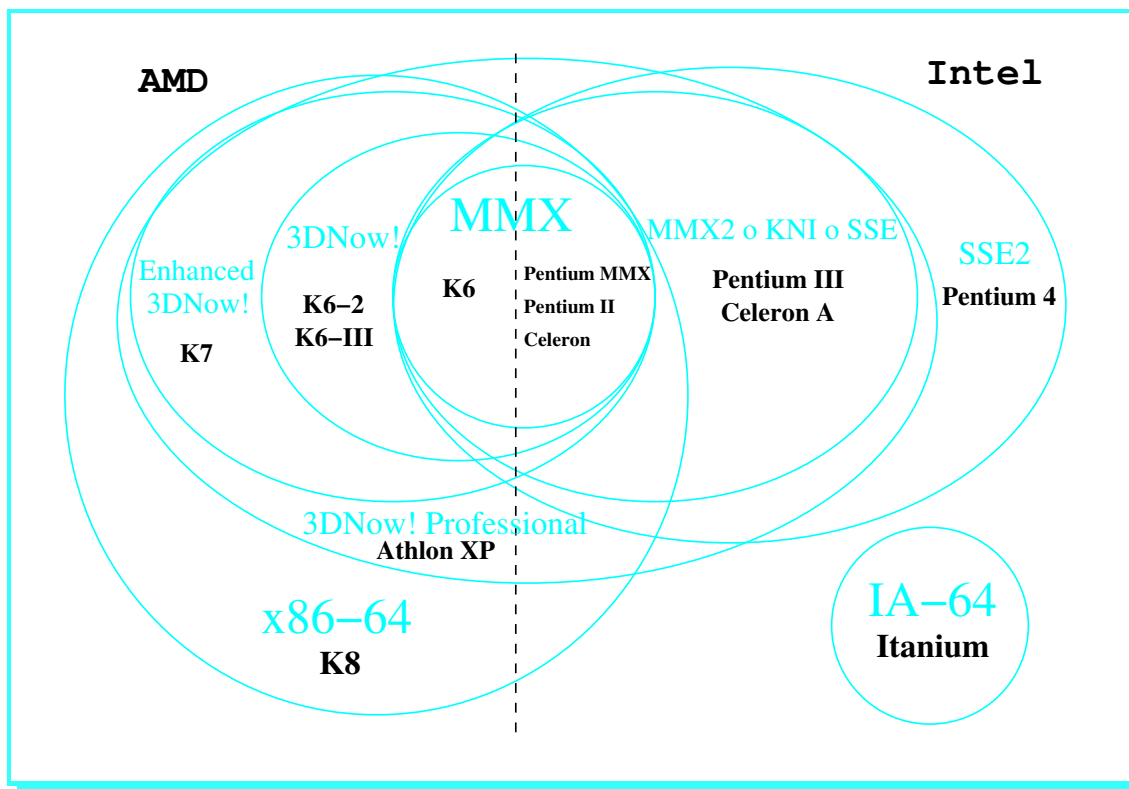


FIGURA 3.19: Las sucesivas ampliaciones del conjunto de instrucciones multimedia desarrolladas por Intel y AMD en años pretéritos frente a las inminentes iniciativas que nos aguardan con la llegada de nuevos conjuntos de instrucciones de 64 bits.

1999), e Intel le respondió un año más tarde con 144 nuevas instrucciones en el conjunto SSE para Pentium 4. La figura 3.19 resume toda esta evolución.

SSE

pág. 115

Todas estas iniciativas no son más que un refinamiento de la idea original, así que para no aburrir con su descripción, vamos a detallar únicamente sus aspectos más sobresalientes:

□ 3DNow! de AMD

Fue la primera ampliación del conjunto de instrucciones MMX. Comandada por AMD y finalizada en el verano de 1998, se incorporó al mercado como una de las principales novedades del procesador K6-2. Además, algunos fabricantes de microprocesadores compatibles de aquella época, como Cyrix y Centaur, decidieron no esperar más y pronto anunciaron la utilización de la tecnología 3DNow! en sus productos, lo que supuso un importante respaldo por parte del mercado.

Al igual que en la tecnología MMX, la base de 3DNow! es la técnica SIMD para el procesamiento simultáneo de operadores asociativos sobre muestras de datos de un mismo tipo. En el repertorio de operaciones disponibles tampoco aparecen variaciones con respecto a aquéllas. La novedad estuvo en el rango de operandos contemplados, incorporándose la manipulación de datos en formato real de simple precisión. Para ello, se habilitó un nuevo tipo de datos *Paired Simple* de 64 bits que empaquetaba dos de estos datos (32 bits cada uno) y que era precisamente el tipo de dato utilizado por el API Direct3D de Windows 98 para la representación gráfica de polígonos 3D, o la reproducción de vídeo MPEG-2 y sonido AC-3.

base

rango de operandos

La forma de implementar las instrucciones 3DNow! es también similar a la MMX, renombrando los registros de punto flotante como MM0..MM7 y utilizando su gran longitud para empaque-

implementación

tar una serie de datos más cortos. La novedad de la circuitería subyacente estuvo en la incorporación de una unidad de procesamiento adicional a la MMX, dedicada en exclusiva al cálculo de las nuevas instrucciones y en la que además se podía disponer de superescalaridad de factor 2 por encontrarse replicada.

21 instrs.  
2 grupos

Las nuevas instrucciones son un total de 21, y se dividen en dos grupos distintos: vectoriales y escalares. Las primeras operan simultáneamente sobre dos operandos de 32 bits colocados en las mitades inferior y superior de un registro MM o de una palabra de memoria, mientras que las segundas lo hacen sobre un único operando de 32 bits colocado en la parte baja de dicha palabra.

funcionalidad:

Respecto a la funcionalidad, cubre las siguientes:

- enteros
  - 3 de aritmética entera: Media de 8 datos de 8 bits, permutación de datos de 16 bits, y multiplicación con redondeo.
- reales
  - 10 de aritmética de punto flotante: Suma y resta, suma y resta acumuladas, multiplicación, tres tipos de comparaciones, y máximo y mínimo.
- conversión
  - 4 de conversión: De entero de 16 y 32 bits a punto flotante y viceversa.
- raíz cuadrada
  - 2 para computar la raíz cuadrada de números de punto flotante, que se desglosan en un total de 5 subinstrucciones.
- conmutación
  - Salida del modo MMX del procesador.
- prebúsqueda
  - Prebúsqueda de una línea de caché en L1D.

#### □ MMX2 ó KNI ó SSE de Intel

etimología

Intel anduvo bastante vacilante con el bautismo de su primera secuela al conjunto MMX. Primero, pensó eternizar esas tres letras, optando por MMX2; después eligió KNI para reflejar en él el nombre del primer procesador suyo que lo incorporaba <sup>4</sup>; finalmente optó por enfatizar el carácter SIMD de la idea original, acuñando como nombre comercial SSE (*Streamind SIMD Extensions*), que es el que ha quedado para la posteridad.

70 instrs.

El conjunto SSE es bastante más numeroso que el 3DNow!: Cuenta con 70 instrucciones en total, aunque buena parte de ellas son desdobles de una misma operación sobre operandos cuyo abanico de posibilidades se expandió a la computación de punto flotante, incorporándose los formatos estándar IEEE utilizados por Intel: 32 bits para simple precisión, y 64 bits para doble precisión. La otra contribución novedosa de SSE estuvo en las instrucciones dedicadas al control explícito de los contenidos de memoria caché.

El desglose de las 70 instrucciones puede resumirse de la siguiente forma:

- reales
  - 50 nuevas instrucciones para computación de punto flotante siguiendo el mismo desdoble de operandos que posibilita la idea SIMD.
- caché
  - 8 nuevas instrucciones para el control de caché que no tienen nada que ver con la filosofía SIMD desde el punto de vista de su implementación. Controlan de forma individual sobre cada línea de caché: La política de *write-through* o actualización simultánea en memoria principal de las líneas de caché, la compartición de líneas entre varios procesadores acoplados a una misma placa base, y los riesgos de inconsistencias en que puede incurrirse al actualizar la información.
- NMI
  - 12 instrucciones denominadas NMI, de corte muy similar a las ya existentes en el conjunto MMX original.

<sup>4</sup>KNI significa *Katmai New Instructions*, y Katmai era el código de referencia del primer Pentium III.

Dados los requisitos, se hizo necesaria la utilización de un nuevo banco de registros cuya anchura fuese al menos de 128 bits. Así, más que en la concepción, la principal novedad estuvo en su implementación. Este nuevo banco contaba con 8 registros exactamente de 128 bits, y se denominó XMM, incorporándose a los microprocesadores de Intel a partir del Pentium III. La arquitectura del procesador se amplió con una nueva unidad de procesamiento conectada al resto del sistema por un puerto separado en el que no interfería el tráfico de las instrucciones MMX originales, ni tampoco el de las de punto flotante con las que éstas se solapaban.

recursos  
hardware

#### ❑ Enhanced 3DNow! de AMD

La nueva réplica de AMD tuvo lugar casi de forma inmediata al lanzamiento de las SSE por parte de Intel. Son un total de 24 instrucciones, que podemos desglosar en tres grupos:

24 instrs.

- 12 típicas de cálculo de números enteros desdoblados. enteros
- 7 dedicadas a optimizar el uso de las cachés, controlando explícitamente la permanencia de datos en ellas e incluso la habilitación de un bypass en el acceso a caché. caché
- 5 aceleran funciones relacionadas con MP3 (MPEG2), sonido Dolby Digital (AC3) y módem ADSL. MP3, AC3, ADSL

Las 19 primeras son sospechosamente parecidas a otras tantas ya existentes en SSE, siendo las últimas 5 realmente novedosas.

Las instrucciones Enhanced 3DNow! se incorporaron en primer lugar al procesador K7 en Junio de 1999, donde ofrecieron una distinción entre tratamiento escalar y vectorial muy similar al que ya comentamos para sus antecesoras sobre el K6-2.

#### ❑ SSE2 de Intel

Aprovechando el lanzamiento de la nueva arquitectura Pentium 4 (Noviembre de 2000), Intel incorporó a este procesador una extensión del conjunto multimedia SSE, denominado SSE2. Consta de un total de 144 nuevas instrucciones, donde vuelve a utilizarse el banco de registros XMM de 128 bits y el puerto de conexión separado de las MMX y de punto flotante.

recursos  
hardware  
144 instrs.

Respecto a su funcionalidad, el desglose por grupos afines es el siguiente:

- 7 controlan el uso de la caché de forma explícita. caché
- 16 se dedican a efectuar todas las conversiones posibles entre datos enteros empaquetados de 32 bits y punto flotante de simple y doble precisión. conversión
- 8 se dedican al movimiento de datos entre registros MMX y XMM. transporte
- 68 son operaciones enteras similares a las MMX, pero aplicadas sobre los registros XMM, que al ser más anchos, admiten una mayor cantidad de operandos. enteros
- El resto son operaciones de punto flotante. reales

#### ❑ 3DNow! Professional de AMD

La última extensión multimedia de AMD, denominada 3DNow! Professional no introduce novedad alguna en su implementación, pero sí en su concepción.

Son 72 instrucciones adicionales basadas de nuevo en el paradigma SIMD y orientadas tanto a computación entera como a computación de punto flotante. Pero tienen una peculiaridad: 52 de ellas son compatibles con las SSE de Intel. El primer procesador en el que se incorporan es el Athlon XP, lanzado a finales de 2001.

72 instrs.

Tipo	Conjunto de instrucciones		Nº bits de los registros	Conjunto de instrs. antecesor	Primer procesador que lo incorpora		Fecha de salida
	Nombre	Nuevas instrs. incorp.			Intel	AMD	
MUL	MMX	57	64	IA-32	Pent MMX	K6	May'97
	KNI ó SSE	70	128	MMX	Pent III	-	Mar'99
TI	SSE2	144	128	SSE	Pentium 4	-	Nov'00
ME	3DNow!	21	128	MMX	-	K6-2	Ago'98
DIA	Enh. 3DNow!	24	128	3DNow!	-	K7	Jun'99
	3DNow! Prof.	52	128	E. 3DNow!	-	Athlon XP	Nov'01
NUE	IA-64	Todas	64	Ninguno	Itanium	-	May'01
VO	x86-64	N/D	64	3DNow! Prof.	-	K8	Abr'03

TABLA 3.16: Resumen de las principales aportaciones al conjunto de instrucciones de los microprocesadores para PC en los últimos cinco años.

estandarización

Con esta maniobra por parte de AMD se da un certero paso en la estandarización que el software demandaba y en el rendimiento que el hardware anhelaba tras la concepción de las instrucciones multimedia a mediados de los años 90. Unificando este interfaz de bajo nivel, las API definidas en DirectX (que constituyen el siguiente estrato software según ilustramos en la [figura 3.18](#)) ya no necesitan preguntar el hardware que se tiene disponible y ramificarse por un flujo de ejecución diferente en función de él, permitiendo concentrar las optimizaciones del código por una única senda.

pág. 113

DirectX

Caminar de espaldas a Intel y Microsoft durante tantos años ha enseñado a AMD que los grandes fabricantes de software se deben siempre al líder en ventas, y por ello, la mejor manera de garantizar a los clientes del Athlon XP que aprovecharán todas sus prestaciones es comulgar con Microsoft en un soporte plenamente consolidado como el API DirectX.

resumen

La [tabla 3.16](#) recopila a modo de resumen de toda esta sección las principales novedades que han aparecido en los microprocesadores comerciales de Intel y AMD en el periodo 1997-20002 al respecto del conjunto de instrucciones.

pág. 119

relación con  
DirectX

La información se completa con la [figura 3.20](#), donde se ilustra su relación con las diferentes versiones del API DirectX de Microsoft así como su secuencia evolutiva, primero siguiendo caminos divergentes desde Intel y AMD, y poco a poco convergiendo junto con Microsoft.

#### 5.4.6 Otras extensiones multimedia

Hay que decir como colofón que prácticamente cada fabricante de microprocesadores tiene ahora en el mercado su propio conjunto de instrucciones multimedia. Los enumeramos a continuación:

Hewlett-Packard

- 1 MAX (*Multimedia Acceleration eXtensions*). Desarrollado por Hewlett-Packard para su familia de microprocesadores RISC PA7x00-PA8x00.

Sun

- 2 VIS (*Visual Instruction Set*). Desarrollado por Sun Microsystems para su familia de microprocesadores UltraSparc..

Sil. Graphics

- 3 MDMX (*Mips Digital Media eXtensions*). Elaborado por Silicon Graphics para su familia de microprocesadores MIPS-V y R2000-R10000.

Compaq

- 4 MVI (*Motion Video Instructions*). Diseñado por Digital (ahora Compaq) para su familia de microprocesadores Alpha.

Motorola

- 5 AltiVec. Desarrollado por Motorola para su familia de microprocesadores Power PC.



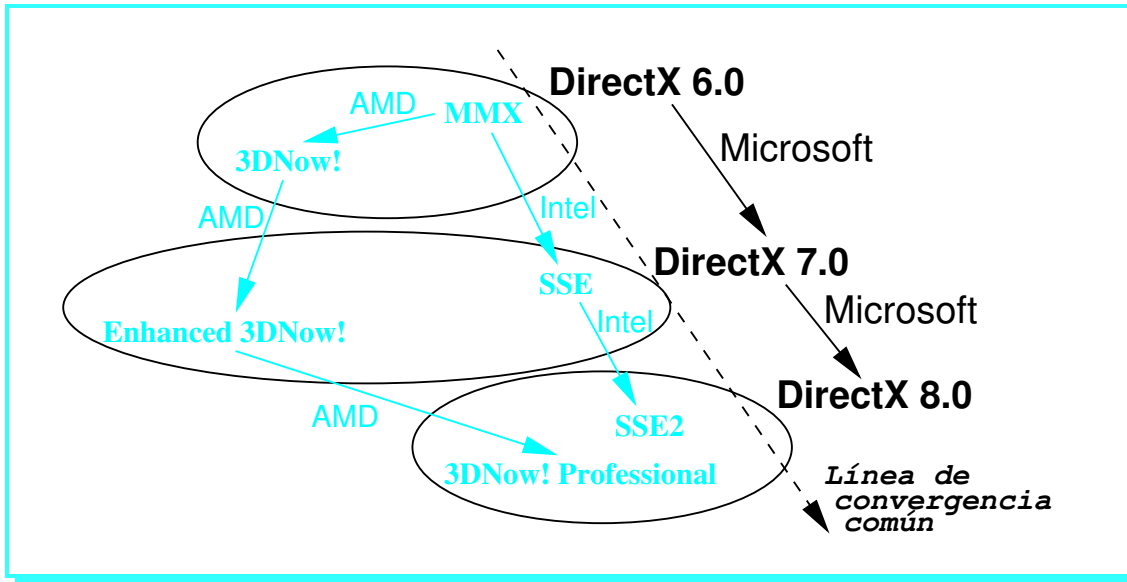


FIGURA 3.20: Secuencia evolutiva de los conjuntos de instrucciones multimedia y su relación con las sucesivas versiones del API DirectX de Microsoft, donde destacamos la divergencia en los primeros pasos y la convergencia hacia un estándar común en los últimos.

La proliferación de conjuntos de instrucciones multimedia no es sino el síntoma más inequívoco de que están siendo utilizados como verdaderos reclamos publicitarios por los fabricantes de microprocesadores. En la práctica, son las compañías de software las responsables de que el usuario saque partido de ellas, y lo cierto es que su potencial está muy infrautilizado en la actualidad. Cuando su utilidad es clara incluso en un contexto de mero cálculo científico, ni siquiera las firmas de juegos por ordenador que son las más beneficiadas apuestan de forma unánime por ellas, habiéndonos encontrado demasiados casos en los que sus posibilidades son sencillamente ignoradas.

infra-utilizadas

🕒 Resumen 🕒

El microprocesador es un componente extraordinariamente complejo. Entre la multitud de parámetros que influyen en su rendimiento, hemos seleccionado los cinco que consideramos más importantes. La [tabla 3.17](#) sintetiza los aspectos del procesador en los que cobra mayor protagonismo cada uno de ellos.

complejidad

➔ [pág. 120](#)

- ❶ La frecuencia mide la velocidad con que se suceden los ciclos del procesador, y no necesariamente revierte sobre su rendimiento. Puede ser un reflejo del mayor número de etapas por instrucción o incluso de una extrema simplicidad en el hardware de la arquitectura.
- ❷ La distancia de integración del transistor es la magnitud menos conocida, pero es, con diferencia, la que mayor influencia ejerce sobre las otras cuatro. En el pasado ha sido la principal responsable del progreso simultáneo de todas ellas, pero dado que los límites del transistor de silicio ya acechan por el horizonte, el relevo debería ser tomado en magnitudes de más alto nivel como las que prosiguen.
- ❸ El paralelismo a nivel de instrucción rompe con la ejecución del código tal y como la planteó el programador, introduciendo una enorme complejidad en su procesamiento lógico.

frecuencia:  
reclamo

integración:  
la clave

Nivel de abstracción	Parámetro	Incidencia sobre el microprocesador
Bajo: Circuitería	Frecuencia	❖ Velocidad. ❖ Precio final del producto.
	Tecnología de integración	❖ Potencial futuro de mejora. ❖ Comportamiento de variables eléctricas. ❖ Coste de fabricación.
Intermedio: Arquitectural	Paralelismo a nivel de instrucción	❖ Carácter: Rendimiento frente al software a través de la ejecución concurrente de instrucciones.
	Memoria caché interna	❖ Independencia: Protección frente a ralentizaciones procedentes de componentes externos.
Alto: Funcional	Conjunto de instrucciones	❖ Compatibilidad con la capa software. ❖ Favoritismo ante ciertas aplicaciones de moda (RISC, multimedia, ...).

TABLA 3.17: Resumen de las cinco magnitudes más sobresalientes del procesador y principales aportaciones de cada una de ellas.

paralelismo:  
tedioso

Los riesgos en los que incurre se solventan introduciendo circuitería dedicada: En este nivel residen las estrategias de diseño más imaginativas, pero la carencia de ideas ha sido alarmante a lo largo de las generaciones que serán objeto de nuestro seguimiento: En no pocas ocasiones hemos visto al mercado tomar una dirección y posteriormente salir en la dirección opuesta buscando sus orígenes. Por lo tanto, no podemos trazar una tendencia aquí, sino tan sólo hablar de movimientos cíclicos.

caché:  
complemento

- ④ La memoria caché tiene un nombre (L1D, L1I, L2) que denota su posición en la jerarquía y revela su tamaño aproximado, y un apellido (externa, interna o integrada) que delata su velocidad. Se encarga de alimentar al procesador con instrucciones y datos a gran velocidad, erigiéndose en su complemento ideal.

instrs:  
publicidad

- ⑤ El conjunto de instrucciones x86 es el responsable de la compatibilidad software desde hace veinte años, proporcionando un estándar para comunicarse con el procesador. Su enrevesada concepción ha dificultado la innovación y desaprovechado grandes recursos hardware, surgiendo entonces las extensiones multimedia como parches para ayudar a la escritura de nuevos programas. Estas nuevas instrucciones gozan de mejor diseño, pero la ausencia de un estándar ha dificultado su aprovechamiento por parte de la capa software, convirtiéndose más en un reclamo publicitario.

efectos  
laterales

Todas estas magnitudes se encuentran muy ligadas entre sí, y a su vez, con los niveles software, de manera que actuando sobre una de ellas se producen multitud de efectos laterales sobre las demás. En consecuencia, no podemos postular una estrategia ganadora en el diseño de microprocesadores, sino tan sólo aspirar a conocerlas con objeto de explotar todas sus ventajas cuando la tecnología o el software nos indiquen que se dan las circunstancias oportunas para apostar por cada una de ellas.

Múltiplos del SI			Submúltiplos del SI		
Factor	Prefijo	Símbolo	Factor	Prefijo	Símbolo
$10^3$	<b>kilo-</b>	K	$10^{-3}$	<b>mili-</b>	m
$10^6$	<b>mega-</b>	M	$10^{-6}$	<b>micro-</b>	$\mu$
$10^9$	<b>giga-</b>	G	$10^{-9}$	<b>nano-</b>	n
$10^{12}$	<b>tera-</b>	T	$10^{-12}$	<b>pico-</b>	p
$10^{15}$	<b>peta-</b>	P	$10^{-15}$	<b>femto-</b>	f
$10^{18}$	<b>exa-</b>	E	$10^{-18}$	<b>atto-</b>	a
$10^{21}$	<b>zetta-</b>	Z	$10^{-21}$	<b>zepto-</b>	z
$10^{24}$	<b>yotta-</b>	Y	$10^{-24}$	<b>yocto-</b>	y

TABLA 3.18: Principales prefijos para los múltiplos de las unidades del Sistema Internacional (SI).

Nuevos múltiplos binarios del SI					Exceso respecto a $10^x$
Factor	Prefijo	Procedencia	Símbolo	Valor decimal	
$2^3$	<b>kibi-</b>	KIloBInary	Ki	1.024	2.40 %
$2^6$	<b>mebi-</b>	MEgaBInary	Mi	1.048.576	4.85 %
$2^9$	<b>gibi-</b>	GIbiBInary	Gi	1.073.741.824	7.37 %
$2^{12}$	<b>tebi-</b>	TEraBInary	Ti	1.099.511.627.776	9.95 %
$2^{15}$	<b>pebi-</b>	PEtaBInary	Pi	1.125.899.906.842.624	12.59 %
$2^{18}$	<b>exbi-</b>	EXaBInary	Ei	1.152.921.504.606.846.976	15.29 %

TABLA 3.19: Los nuevos prefijos para los múltiplos binarios definidos por la ISO y la IEC



## La anécdota: Magnitudes oficiales y oficiosas



El universo numérico mantiene ciertos desencuentros entre la comunidad científica en general, acostumbrada a la tradicional base decimal del Sistema Internacional (SI), y el gremio de los informáticos, más proclives a utilizar la base binaria originaria del bit.

El azar ha querido que los sucesivos exponentes ternarios de la base decimal ( $10^3 = 1000$ ) que el SI utiliza para articular los prefijos Kilo-, Mega- y Giga-, ostenten valores muy similares a las potencias decimales de la base binaria ( $2^{10} = 1024$ ) (ver [tabla 3.18](#)). Así se ha alentado un tácito redondeo en el que usamos 1 Mbyte para indicar 1.048.576 ( $2^{20}$ ) bytes, cuando en realidad, según el SI, serían sólo 1.000.000 ( $10^6$ ) bytes.

La organización interna de la memoria provoca que cueste lo mismo fabricar 1.048.576 que 1.000.000 de bytes, por lo que a su industria poco le preocupa que en realidad se ofrezca al cliente un 4.85 % de espacio de regalo. Pero en los discos, la cosa cambia: 40 Gbytes son para una marca 40.000.000.000 bytes, ni uno más, ya que comulgar con 1024 en lugar de con 1000 a la altura del Gigabyte supone ya un regalo del 7.37 %, que además, esta vez sí repercute en el coste de fabricación.

Las imprecisiones serán mayores a medida que el tiempo avance y las magnitudes se vayan haciendo más grandes, por lo que de cara a evitarlas en los textos de corte formal y científico, tanto la ISO (International Standardization Organization) como la IEC (International Electrotechnical Commission) definieron entre 1999 y 2000 nuevos prefijos para los múltiplos binarios <sup>5</sup>, con objeto de poder distinguirlos de los decimales (ver [tabla 3.19](#)).

<sup>5</sup> Pueden consultarse sus respectivas publicaciones, "ISO/TC 12, Quantities, units, symbols, conversion factors", y "IEC/TC 25, Quantities, units, and their letter symbols".

En el tiempo que ha transcurrido desde entonces, confesamos no haber visto semejante denominación en ninguna publicación escrita. Nos gustan los formalismos, pero no cuando ya hay consolidado un estándar de facto diferente. Emplear Kibibyte o Gibibit en nuestros libros hubiera supuesto, aparte de arrancar alguna que otra carcajada, navegar contracorriente y dificultar su comprensión al usuario de PC al que también nos debemos. Para nosotros, todos los Megas y Gigas serán implícitamente múltiplos de 1000, excepto en el contexto de las memorias (dinámicas o estáticas, no magnéticas), en que se convertirán en múltiplos de 1024. En este sentido, ya cursamos un aviso de nuestras intenciones en la definición de ancho de banda nada más comenzar el tratamiento de dichos contenidos (ver el pie de página de la [sección 10.4](#)).

← p.18/Vo1.2

## 📖 Cuestionario de evaluación 📖

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

1 🗣️ La magnitud que mide la velocidad de un microprocesador es

- a La frecuencia.
- b La tecnología de integración.
- c El paralelismo a nivel de instrucción.
- d El conjunto de instrucciones.

2 🗣️ La señal de reloj emitida desde el oscilador de cuarzo de un PC describe una onda

- a Senoidal, analógica, periódica y síncrona.
- b Continua, nanocíclica, periódica y asíncrona.
- c Cuadrada, digital, periódica y síncrona.
- d Eléctrica, digital, acíclica y síncrona.

3 🗣️ El ciclo de reloj de un procesador es

- a La magnitud inversa de su frecuencia de reloj.
- b El máximo tiempo de que dispone para realizar una operación atómica, ya sea concurrentemente con otras o no.
- c Inferior a 1 nanosegundo en los microprocesadores actuales.
- d Las tres respuestas anteriores son correctas.

4 🗣️ El período de la señal de reloj de un microprocesador actual delimita

- a El tiempo que tarda en ejecutar completamente una instrucción máquina.

b El tiempo que tarda en ejecutar parcialmente una instrucción máquina.

c El tiempo que tarda en ejecutar parcial o totalmente una instrucción máquina.

d El tiempo que tarda en ejecutar parcial o totalmente varias instrucciones máquina simultáneamente.

5 🗣️ ¿Son verdaderas las siguientes afirmaciones acerca de la frecuencia de un microprocesador? Marque cada una de ellas por separado en caso de que sea necesario y razone sus respuestas.

a “Un computador siempre es más rápido si aumentamos la frecuencia de reloj de su microprocesador y sigue funcionando correctamente”.

b “Un computador cuyo microprocesador funciona a una frecuencia mayor que el de otro computador es siempre más rápido que éste”.

c “La frecuencia en MHz es siempre la magnitud que mejor refleja la presteza con que un microprocesador ejecuta un determinado programa”.

d “La ejecución de un programa sobre un microprocesador de una frecuencia determinada puede bloquearse si ésta última se ralentiza hasta la mitad de su valor original”.

6 🗣️ Ante el diseño de un microprocesador que está condenado a funcionar a baja frecuencia, ¿Cómo suplirías la falta de velocidad?

a Aumentando el número de unidades funcionales e implementando superescalaridad.

b Acelerando la caché interna.

c Incorporando un sistema de refrigeración más eficiente.

**d** Añadiendo más memoria.

**7** ¿Qué magnitud de un microprocesador es mejor cuanto más bajo es su valor?

**a** La frecuencia.

**b** La tecnología de integración.

**c** El paralelismo a nivel de instrucción.

**d** El tamaño de la memoria caché integrada.

**8** Cuando decimos que un microprocesador se fabrica con tecnología de integración de 0.13 micras, queremos reflejar que

**a** Se compone de transistores cuya anchura de puerta es de 0.13 micras.

**b** Utiliza tecnología de semiconductores con distancias de integración para las puertas de sus transistores de 0.13 micras.

**c** 0.13 micras es la mínima resolución de la maquinaria responsable de integrar sus circuitos mediante técnicas de litografía.

**d** Las tres respuestas anteriores son correctas.

**9** Queremos construir un microprocesador que sea a la vez más rápido y más barato. ¿Cuál de las siguientes cuatro bazas utilizarías?

**a** Una ampliación de la caché L2.

**b** Una reducción en la distancia de integración de sus transistores constituyentes.

**c** Un mayor grado de superescalaridad.

**d** Un conjunto de instrucciones más amplio.

**10** ¿Qué parámetro del microprocesador evolucionó con mayor rapidez durante la década de los 90?

**a** La tecnología de integración (micras).

**b** La frecuencia de reloj (MHz).

**c** El tamaño de la caché L1.

**d** El número de transistores.

**11** ¿Cuál de las siguientes implicaciones es cierta con respecto a la integración de un microprocesador?

**a** A menor número de componentes, mayor frecuencia de reloj.

**b** A mayor calidad de integración, mayor voltaje de alimentación.

**c** A mayor número de transistores, mayor funcionalidad en las operaciones atómicas del microprocesador.

**d** Sólo dos de las anteriores son ciertas.

**12** ¿Qué fabricante fue el primero en utilizar la tecnología de integración de 0.18 micras en un microprocesador comercial?

**a** Digital, con el Alpha 21264 a 600 MHz.

**b** IBM, con el Power PC 750 a 300 MHz.

**c** Silicon Graphics, con el R10000.

**d** Intel, con el Merced ó Itanium.

**13** ¿Qué beneficios reporta la integración de un mismo modelo de microprocesador con transistores de 0.13 micras en lugar de hacerlo con otros de 0.18 micras?

**a** Se calentará menos.

**b** Podremos bajarle el voltaje.

**c** Tendrá un área de integración inferior.

**d** Muchos, y los tres anteriores son buenos ejemplos.

**14** ¿Qué beneficios presenta la reducción de la distancia de integración en los transistores de un chip desde las 0.18 hasta las 0.13 micras?

**a** Menor coste por transistor.

**b** Cachés internas más rápidas.

**c** Voltajes más reducidos, y por tanto, temperaturas más bajas y frecuencias más altas.

**d** Todas las anteriores.

**15** Una reducción en la distancia de integración de los transistores constituyentes de un microprocesador, ¿Cuál de las siguientes cuatro variables puede aprovecharla en mayor medida?

**a** El voltaje.

**b** La frecuencia.

**c** La temperatura.

**d** El número de transistores.

- 16** ¿Qué razón puede justificar la asiduidad con la que nuevas arquitecturas irrumpen en el mercado al final del período de vigencia de una determinada tecnología de integración? (por ejemplo, el Pentium vió la luz al final de la integración a 0.8 micras, el K7 al final de las 0.25 micras y el Pentium 4 al final de las 0.18 micras?)
- a** Para las nuevas arquitecturas, tiene la ventaja de que enseguida llega un nuevo proceso de fabricación que permite dedicar un mayor número de transistores a mejorar las principales carencias mostradas durante el proceso inicial de rodaje.
  - b** Porque así sale más barato en sus primeros meses de existencia.
  - c** Para garantizar que se fabrican pocas unidades defectuosas.
  - d** Porque nueva arquitectura y nueva integración representan demasiadas innovaciones para ser acometidas conjuntamente.
- 17** ¿Cómo compensarías la ausencia de una frecuencia de reloj elevada en el diseño de un microprocesador?
- a** Aumentando el paralelismo a nivel de instrucción.
  - b** Incrementando el tamaño de la memoria caché integrada.
  - c** Con un conjunto de instrucciones más complejo.
  - d** Con un conjunto de instrucciones más completo.
- 18** ¿Son la frecuencia y el paralelismo a nivel de instrucción dos mejoras de un procesador incompatibles entre sí?
- a** No, porque el incremento de la frecuencia es una estrategia muy popular dentro de las muchas variantes que incluye el paralelismo a nivel de instrucción.
  - b** Sí. Un incremento en la frecuencia siempre es llevado a cabo a costa de sacrificar paralelismo a nivel de instrucción y viceversa.
  - c** No, pues reduciendo la distancia de integración de los transistores, éstos serán más rápidos y estarán disponibles en mayor número, con lo que podremos aplicar ambas mejoras.
  - d** Sí, y el Pentium 4 es un claro ejemplo comercial: Dispone de una frecuencia elevada pero a costa de reducir su paralelismo a nivel de instrucción.
- 19** ¿Qué ventaja principal tiene apostar por la segmentación frente a la superescalaridad como forma de explotar el paralelismo a nivel de instrucción?
- a** La segmentación apenas requiere la incorporación de circuitería adicional, mientras que la superescalaridad requiere la replicación de toda la unidad funcional correspondiente.
  - b** Le afectan menos las dependencias de datos.
  - c** Le afectan menos las dependencias de control.
  - d** Le afectan menos las dependencias estructurales.
- 20** ¿Qué mecanismos incorpora un microprocesador para defenderse del negativo impacto que tienen las dependencias de un programa en su rendimiento?
- a** BRC (Búfer de Reordenación Circular) para las dependencias de datos.
  - b** BTB (Branch Target Buffer) para las dependencias de control.
  - c** Las dos respuestas anteriores son correctas.
  - d** No, es justo lo contrario: BRC para las dependencias de control y BTB para las dependencias de datos.
- 21** ¿Qué elementos habilita un microprocesador para defenderse de la negativa influencia de las dependencias?
- a** La BTB para las dependencias de datos y el BRC para las dependencias de control.
  - b** La BTB para las dependencias de control y el BRC para las dependencias de datos.
  - c** La BTB para las dependencias de control, el BRC para las dependencias de datos y la separación de la caché de primer nivel en dos (datos e instrucciones) para las dependencias estructurales.
  - d** Las instrucciones de carga retrasada para cualquier tipo de dependencias.
- 22** ¿Tiene algo que ver la ejecución fuera de orden y el BRC (búfer de reordenación circular) de un microprocesador?
- a** Sí, el BRC es parte de la infraestructura que un procesador necesita para implementar la ejecución fuera de orden.
  - b** Sí, se encuentran relacionadas a través de la TLB del procesador.
  - c** Sí, pero sólo en procesadores superescalares.
  - d** No.
- 23** ¿Dónde se implementa el soporte para multiprocesador en un PC?
- a** En el microprocesador, si éste presenta carácter servidor.

- b** En la placa base, que ocasionalmente lo incorpora como valor añadido.
- c** Las dos anteriores son correctas.
- d** Ninguna de las anteriores es válida.
- 24** ¿Qué técnica para aumentar el rendimiento de un microprocesador se incluye dentro del paralelismo a nivel de instrucción?
- a** La segmentación.
- b** La superescalaridad.
- c** La supersegmentación.
- d** Las tres anteriores.
- 25** ¿Cuál de los siguientes silogismos es correcto en relación al diseño de un microprocesador?
- a** Si es superescalar, tiene ejecución fuera de orden.
- b** Si tiene ejecución fuera de orden, es superescalar.
- c** Si es supersegmentado, es segmentado y superescalar.
- d** Sólo dos de las anteriores son ciertas.
- 26** Un programa compuesto por la secuencia de instrucciones A, B, C, se ejecuta fuera de orden en un procesador. Puede ocurrir que
- a** C comience su fase de decodificación antes que B.
- b** C concluya su fase de ejecución antes que B.
- c** Las dos anteriores.
- d** Ninguna de las anteriores.
- 27** ¿Cuál de las siguientes implicaciones es cierta con respecto al paralelismo a nivel de instrucción de un procesador?
- a** Cuantas más etapas de segmentación, mayor impacto de las dependencias en el rendimiento del código.
- b** Cuantas más etapas de segmentación, mayor grado de superescalaridad.
- c** A mayor grado de superescalaridad, más etapas de segmentación.
- d** Dos de las tres anteriores son ciertas.
- 28** ¿Qué relación liga al número de etapas de segmentación (NES) con el factor de superescalaridad (FS) de un microprocesador?
- a**  $NES < FS$ .
- b**  $NES = FS$ .
- c**  $NES > FS$ .
- d** Ninguna.
- 29** ¿Cuál de las siguientes mejoras en el paralelismo a nivel de instrucción conduce a un aumento en la frecuencia de reloj del procesador?
- a** Un mayor número de etapas de segmentación.
- b** Un mayor grado de superescalaridad.
- c** La supersegmentación, puesto que lleva consigo la subdivisión del período de reloj.
- d** Todas las anteriores.
- 30** ¿Qué procesador de entre los siguientes es más eficiente?
- a** Un segmentado en 5 etapas.
- b** Un superescalar de factor 3.
- c** Uno que reúna las dos características anteriores.
- d** Depende de las dependencias de datos que tenga el programa con el que medimos su rendimiento.
- 31** La función de una BTB dentro de un microprocesador es...
- a** Eliminar los conflictos por dependencias de control.
- b** Aumentar la probabilidad de acertar en la predicción de salto.
- c** Reducir la penalización por fallo en una predicción de salto.
- d** Todas las anteriores.
- 32** ¿Tienen algo que ver la ejecución fuera de orden y la TLB de un microprocesador?
- a** Son la misma cosa, puesto que TLB son las iniciales de "Ejecución Fuera de Orden" en inglés.
- b** Ambas se encuentran relacionadas a través del factor de superescalaridad del procesador: Sólo si es superescalar puede adicionalmente disponer de ejecución fuera de orden, y para ello, necesita de la TLB para equilibrar su diseño y alcanzar una eficiencia razonable.
- c** No. La ejecución fuera de orden es cosa del *back-end* o núcleo interno de ejecución, mientras que la TLB se sitúa en el *front-end*, pues se encuentra ligada a la fase de búsqueda de instrucciones y datos.

**d** La respuesta anterior puede completarse apostillando que en el caso de disponerse de caché de traza, como en el Pentium 4, la TLB se dispone más adelante en el cauce segmentado del procesador, interactuando ya de forma directa con las unidades de ejecución (ALU, MMX, FPU, ...).

**33** ¿Qué razones previenen a la segmentación de un microprocesador de utilizar un elevado número de etapas?

**a** El diseño de su Unidad de Control desbordaría por su excesiva complejidad.

**b** Las dependencias de datos y control de los programas que no hubiesen podido ser resueltas por el compilador o el hardware tendrían un enorme impacto negativo sobre su rendimiento.

**c** El Banco de Registros alcanzaría un tamaño de varios Kilobytes.

**d** Las tres respuestas anteriores son correctas.

**34** ¿Qué razones previenen al diseño de un microprocesador de utilizar un factor de superescalaridad muy elevado?

**a** La creciente necesidad de memoria principal que demanda un programa actual.

**b** El excesivo coste de su implementación, al tener que replicar circuitería por cada nuevo desdoble de dicho factor.

**c** La dificultad para encontrar un número grande de flujos de ejecución independientes en un programa que aprovechen los recursos hardware disponibles.

**d** La conjunción de las dos opciones anteriores, que hace que presente una relación rendimiento/coste desfavorable.

**35** En un procesador de 5 GHz de frecuencia, ¿Es posible ejecutar un programa compuesto por 5000 instrucciones en menos de un microsegundo?

**a** Sí, empleando un gran cauce de segmentación y suponiendo que el programa apenas presenta dependencias.

**b** Sí, empleando un gran factor de superescalaridad en todas sus etapas y suponiendo que el programa apenas presenta dependencias.

**c** Sí, pero es necesario aplicar las dos estrategias anteriores.

**d** No, el techo de las optimizaciones para el paralelismo a nivel de instrucción consiste en ejecutar una instrucción por ciclo de reloj, y eso nos deja el tiempo mínimo precisamente en un microsegundo.

**36** ¿Es posible construir un procesador en el que el factor de superescalaridad supere al número de etapas de segmentación?

**a** Sí, aunque no ha ocurrido en ninguno de los modelos comerciales de quinta, sexta y séptima generación.

**b** Sí, aunque lo normal es que sea al contrario porque la superescalaridad conlleva un ingente gasto en hardware, mientras que la segmentación sale casi gratis.

**c** Las dos respuestas anteriores son correctas.

**d** No.

**37** Hemos construido un procesador con 50 etapas de segmentación, lo que nos conduce de forma casi inevitable a

**a** Una elevada frecuencia.

**b** Un gran número de transistores.

**c** Un enorme factor de superescalaridad.

**d** Un conjunto de instrucciones de tipo RISC.

**38** Hemos construido un procesador con un factor de superescalaridad de diez, lo que nos lleva de forma casi irreversible a

**a** Una elevada frecuencia.

**b** Un gran número de transistores.

**c** Un cauce de segmentación muy profundo.

**d** Un conjunto de instrucciones de tipo RISC.

**39** Tomando como referencia cualquier período evolutivo superior a los diez años en el contexto de los microprocesadores para PC, el factor de superescalaridad ha venido creciendo de forma más pausada que el número de etapas de segmentación. Esto se explica porque











**a** La superescalaridad depende de una gran segmentación para progresar adecuadamente.

**b** Aplicar superescalaridad resulta más caro que aplicar segmentación de forma extensiva.

**c** Las dependencias de datos y control afectan mucho más al rendimiento de un procesador superescalar que al de un segmentado.

**d** La incorporación de cachés integradas favorece mucho más a la idea de la segmentación que a la de la superescalaridad.




- 40  En un PC con caché L1 integrada, caché L2 interna y caché L3 externa, los buses para acceder a cada caché reciben, respectivamente, los siguientes nombres:
- a Ninguno, trasero (backside) y frontal (frontside).
  - b Local, frontal (frontside) y trasero (backside).
  - c Local, trasero (backside) y de memoria.
  - d De caché, local y de memoria.
- 41  ¿Cómo podemos hacer una memoria caché más rápida desde el punto de vista del microprocesador?
- a Integrandolo el controlador de caché dentro del chip procesador.
  - b Acelerando la conexión por el bus.
  - c Cambiando la caché de interna a integrada.
  - d Cualquiera de las anteriores puede ser una alternativa válida.
- 42  ¿Qué devuelve una caché L2 como salida frente a una petición de acceso?
- a Una palabra de memoria.
  - b Una línea de caché.
  - c Una palabra de caché.
  - d Una palabra del procesador.
- 43  ¿Puede un microprocesador acceder a la caché L2 sin pasar por la L1?
- a Sí, a través del banco de registros.
  - b Sí, utilizando para ello su bus trasero (backside).
  - c Sí, pero únicamente si la caché L2 es integrada.
  - d Sólo deshabilitando la caché L1 internamente en caso de que el microprocesador así lo permita (por ejemplo, a través de una opción de la BIOS del sistema).
- 44  ¿Por qué cuando la caché L2 de un microprocesador evoluciona de interna a integrada suele ver disminuido su tamaño en Kbytes?
- a Porque al sumar su área de integración a la del chip CPU resultan unas dimensiones conjuntas que plantean problemas de temperatura, retardos, testeo del conjunto, etc.
- b Porque al ser más rápida, ya no hace falta que sea tan grande.
- c Porque se fabrica con transistores más pequeños.
- d Las tres respuestas anteriores son correctas.
- 45  ¿Qué transición representa una mayor ganancia en velocidad para una caché?
- a El paso de externa a interna.
  - b El paso de integrada a interna.
  - c El paso de interna a integrada.
  - d El paso de externa a integrada.
- 46  ¿Qué transición requiere un mayor aumento en el número de transistores necesarios para integrar el área de datos de una memoria caché?
- a El paso de externa a interna.
  - b El paso de interna a integrada.
  - c El paso de externa a integrada.
  - d Ninguna. En todos los casos se requiere el mismo número de transistores.
- 47  ¿Qué bus se esquivo en la transición de un tipo de caché a otro?
- a El bus local en la transición de caché externa a interna.
  - b El bus trasero en la transición de caché interna a integrada.
  - c Las dos respuestas anteriores son correctas.
  - d Depende del nivel de caché a que se haga referencia en las respuestas anteriores.
- 48  ¿Qué configuración de entre las siguientes produce un mayor rendimiento?
- a CPU a 100 MHz, bus local a 100 MHz, caché L2 interna de 512Kb.
  - b CPU a 200 MHz, bus local a 100 MHz, caché L2 interna de 256 Kb.
  - c CPU a 133 MHz, bus local a 133 MHz, caché L2 interna de 256 Kb.
  - d CPU a 300 MHz, bus local a 100 MHz, caché L2 externa de 256 Kb.
- 49  ¿Dónde se encuentran implementados el controlador y directorio caché de una memoria caché interna al procesador?

- a** En un chip aparte.
- b** Junto al chip que contiene el área de datos.
- c** Dentro del procesador.
- d** Cualquiera de las tres opciones anteriores es posible.
- 50** ¿Qué parte de una caché experimenta una mayor metamorfosis cuando ésta cambia de interna a integrada?
- a** El área de datos.
- b** El modo de direccionamiento.
- c** El controlador de caché.
- d** El bus trasero que la conecta al procesador.
- 51** ¿Qué falla en la siguiente proposición? "Si el Pentium con 3 millones de transistores dispone de dos cachés L1 de 8 Kbytes, el Pentium 4 con 42 millones (14 veces más) debiera tener dos cachés L1 de 112 Kbytes"
- a** Que esos 112 Kbytes en realidad están repartidos entre la L1 y la L2 en el Pentium 4.
- b** Que el Pentium 4 dispone de caché de traza. Si no fuera por eso, la opción **a** sería correcta.
- c** Estamos asumiendo que el procesador crece proporcionalmente en todas sus unidades funcionales (incluidas las cachés), cuando en realidad se trata de arquitecturas diferentes.
- d** Los transistores del Pentium 4 son mucho más pequeños, por lo que su capacidad de almacenamiento en Kbytes es menor a igual número de transistores
- 52** ¿Qué nombre recibe la estrategia para reducir la complejidad inherente al repertorio de instrucciones de un microprocesador?
- a** CISC.
- b** MISC.
- c** RISC.
- d** VLIW.
- 53** ¿Qué es VLIW?
- a** "Very Long Instruction Word", una filosofía de diseño de microprocesadores.
- b** "Very Last Input Well", una nueva especificación de bus.
- c** "Visit Longer, Intruder Welcome", una política de gestión de peticiones en el juego de chips de la placa base.
- d** Lo contrario de WILV.
- 54** ¿Cuál de los siguientes aspectos influye en el rendimiento de un microprocesador RISC?
- a** Los criterios de selección de su conjunto de instrucciones.
- b** El diseño de sus compiladores.
- c** El diseño de sus sistemas operativos.
- d** Los tres anteriores son importantes.
- 55** ¿Por qué la caché L2 no está integrada dentro del microprocesador en muchos de los procesadores RISC comerciales?
- a** Sí que lo está.
- b** Porque después de meter la TLB, la BTB, el enorme banco de registros y los búferes de prebúsqueda, no queda espacio libre.
- c** Porque los primeros niveles de la jerarquía de memoria son más eficientes en un RISC y no se hace tan necesario.
- d** Porque el bus local es más rápido en todos estos micros y no hay tanta diferencia entre una L2 interna y una externa.
- 56** Nos piden mejorar un RISC de 300 MHz, superescalar 4, bus local a 200 MHz y caché L2 interna de 512 Kb. ¿Qué característica mejorarías atendiendo al mejor ratio rendimiento/coste?
- a** La frecuencia a 500 MHz (25.000 ptas).
- b** El factor superescalar de 20 (50.000 ptas).
- c** El bus local a 300 MHz (25.000 ptas).
- d** La caché L2 de 1 Mbyte (35.000 ptas).
- 57** ¿Qué familia de procesadores presenta en general un carácter superescalar más acusado?
- a** El Alpha de Digital.
- b** El Power PC de Motorola.
- c** El UltraSparc de Sun.
- d** El R10000 de Silicon Graphics.
- 58** ¿Cuál fue la saga de procesadores que menos evolucionó en la segunda mitad de la década de los 90 respecto a sus competidores?

- a El Alpha de Digital.
- b El Pentium de Intel.
- c El PA8000 de Hewlett-Packard.
- d El R10000 de Silicon Graphics.

59  ¿Qué conjunto de instrucciones es más completo?

- a MMX2.
- b MMX3.
- c SSE2.
- d Enhanced 3DNow!.

60  ¿Qué agente permite que el conjunto de instrucciones multimedia de un microprocesador pueda ser verdaderamente aprovechado durante la ejecución de los programas?


- a El Sistema Operativo, a través de un API (Application Program Interface) como el Direct3D de Windows.
- b El compilador, soportando las nuevas instrucciones en su fase de generación de código (archivo ejecutable).
- c Un ensamblador que reconozca sus códigos de instrucción.
- d Cualquiera de las tres vías anteriores puede habilitar el aprovechamiento de las instrucciones multimedia.

61  ¿Cuál es el rasgo más distintivo de una instrucción multimedia?


- a Tiene un único código de operación.
- b Dispone de múltiples operandos (normalmente, más de dos).
- c Lleva a cabo operaciones de corte aritmético.
- d Tiene un código de operación del tipo VLIW (Very Long Instruction Word).

62  El conjunto de instrucciones MMX se fundamenta en


- a El paralelismo que exhiben los algoritmos multimedia.
- b Las operaciones repetitivas de elevado coste computacional.
- c Los tipos de datos de tamaño reducido.
- d Todas las respuestas anteriores son correctas.

63  ¿Cuál de las siguientes series de palabras describe mejor un acercamiento puntual del mercado de microprocesadores hacia filosofías de tipo CISC?


- a Klamath, Deschutes, Xeon.
- b ISA, PCI, AGP.
- c Silicio, cobre, aluminio.
- d MMX, KNI, SIMD.

64  ¿Cuál de las siguientes series de palabras describe mejor cronológicamente la tendencia actual del conjunto de instrucciones del procesador hacia implementaciones SIMD?

- a MMX, 3DNow!, KNI.
- b Instrucciones en punto flotante, MMX, MMX2.
- c Multimedia, punto flotante, resto de instrucciones.
- d Multimedia, punto flotante, enteras.

65  Ha llegado el anhelado momento en que el bus de direcciones del procesador pasa de tener 32 líneas a disponer de 64. Todas las firmas comerciales deciden producir modelos iguales a los que ya poseen, pero introduciendo, únicamente, esta innovación. ¿Qué unidades funcionales deberemos sustituir?

- a Los controladores o directorios de todas las memorias caché (aunque no sus áreas de datos).
- b El controlador de bus local.
- c Las TLB.
- d Las dos respuestas anteriores son correctas.

66  Un PC tiene una memoria caché L2 de 256 Kbytes, otro PC dispone de dos módulos de memoria principal de 128 Mbytes (total: 256 Mbytes), y un tercer PC cuenta con un disco duro de distintas particiones en el que el espacio total es de 256 Gbytes. Determinar cuál de ellos tiene el procesador con el bus de datos más grande.

- a El primero.
- b El segundo.
- c El tercero.
- d Ninguno de los valores indicados permite calcular un tamaño concreto para el bus de datos del procesador.

En base a un viejo procesador PK1 de 250 nm. (0.25 micras), 2 GHz (gigahercios) y 50 Mt. (millones de transistores), se deciden desarrollar tres nuevas versiones con las siguientes características:

- PK2: 180 nm, 2.6 GHz y 50 Mt.
- PK3: 130 nm, 2 GHz y 200 Mt.
- PK4: 130 nm, 4 GHz y 50 Mt.


Asumiendo la visión simplista del sistema que nos proporciona el solo conocimiento de esos parámetros, se pide elegir el modelo que mejor cumple cada una de las siguientes premisas:

**67**  Mayor potencia disipada.

- a PK2.
- b PK3.
- c PK4.
- d Los tres están muy igualados.

**68**  Menor área de integración.

- a PK2.
- b PK3.
- c PK4.
- d Los tres están muy igualados.

**69**  Mayor rendimiento bruto (toda la frecuencia y todos los transistores revierten sobre él de forma similar).

- a PK2.
- b PK3.
- c PK4.
- d Los tres están muy igualados.

**70**  Menor coste de producción suponiendo amortizada la planta de fabricación de chips.


- a PK2.
- b PK3.
- c PK4.
- d Los tres están muy igualados.

**71**  Mejor predisposición para encarar un diseño de marcado carácter superescalar.

- a PK2.
- b PK3.

c PK4.

d Los tres están muy igualados.

**72**  Mejor predisposición para abordar un diseño supersegmentado suponiendo que PK1 ya era superescalar.

a PK2.

b PK3.

c PK4.

d Los tres están muy igualados.

**73**  Posibilidad de integrarle una caché L3 de 2 Mbytes.

a PK2.

b PK3.

c PK4.

d Los tres están muy igualados.

**74**  Posibilidad de ampliarle el conjunto de instrucciones y/o incorporarle juegos de instrucciones multimedia.

a PK2.

b PK3.

c PK4.

d Los tres están muy igualados.

**75**  Mejor predisposición para encarar un diseño de tipo RISC.

a PK2.

b PK3.

c PK4.

d Los tres están muy igualados.

**76**  Mayor vulnerabilidad a las dependencias de datos y control.

a PK2.

b PK3.

c PK4.

d Tendríamos que conocer en qué medida los transistores se emplean en dotar de generosas BRC y BTB al sistema.

Sean los cuatro procesadores que se presentan en la siguiente tabla, todos ellos con idéntico conjunto de instrucciones.

	Frec. de reloj	Etapas segmentación	Supersegmentación	Factor superescalar
Altair	2 GHz	9	No	3
Polar	2 GHz	6	6 x2	2
Sirio	4 GHz	23	No	No
Vega	500 MHz	No segm.	No	5

Y dos programas compilados para ellos con las siguientes propiedades:

- **Osa Menor.** Compuesto de tan sólo 10 instrucciones, todas ellas miméticas en duración e independientes (esto es, no hay dependencias ni saltos).
- **Osa Mayor.** Compuesto de un millón de instrucciones, igualmente con idéntica duración y en el que las dependencias entre ellas inciden también de forma despreciable.

Bajo estas premisas, responder a las siguientes cuestiones:

77 ¿Cuánto tarda en ejecutarse Osa Menor en Altair?

- a 4 ns.
- b 6 ns.
- c 8 ns.
- d 10 ns.

78 ¿Cuánto tarda en ejecutarse Osa Menor en Polar?

- a 4 ns.
- b 6 ns.
- c 8 ns.
- d 10 ns.

79 ¿Cuánto tarda en ejecutarse Osa Menor en Sirio?

- a 4 ns.
- b 6 ns.
- c 8 ns.
- d 10 ns.

80 ¿Cuánto tarda en ejecutarse Osa Menor en Vega?

- a 4 ns.
- b 6 ns.

- c 8 ns.
- d 10 ns.

81 ¿Qué procesador presenta un mayor potencial respecto al número máximo de instrucciones que puede ejecutar de forma simultánea (mayor grado de paralelismo inherente)?

- a Altair.
- b Polar.
- c Sirio.
- d Vega.

82 ¿Qué procesador presenta un menor potencial respecto al número máximo de instrucciones que puede ejecutar de forma simultánea (menor grado de paralelismo inherente)?

- a Altair.
- b Polar.
- c Sirio.
- d Vega.

83 ¿Qué procesador tarda menos en ejecutar Osa Mayor?


- a Altair.
- b Polar.
- c Sirio.
- d Vega.

84 ¿Qué procesador tarda más en ejecutar Osa Mayor?

- a Altair.
- b Polar.
- c Sirio.
- d Vega.

85 ¿Qué procesador presenta un mayor número medio de ciclos por instrucción ejecutada (parámetro CPI) sobre Osa Mayor?

- a Altair.
- b Polar.
- c Sirio.
- d Vega.

**86**  ¿Qué procesador presenta un menor número medio de ciclos por instrucción ejecutada (parámetro CPI) sobre Osa Mayor?

- a Altair.
- b Polar.
- c Sirio.
- d Vega.

# La quinta generación

## Sumario

<b>4.1. Intel</b> . . . . .	<b>134</b>
4.1.1. Pentium . . . . .	134
4.1.1.1. Frecuencia . . . . .	135
4.1.1.2. Tecnología de integración . . . . .	136
4.1.1.3. Paralelismo a nivel de instrucción . . . . .	137
4.1.1.4. Resolución de dependencias . . . . .	138
4.1.1.5. La caché de primer nivel . . . . .	138
4.1.1.6. El conjunto de instrucciones . . . . .	138
4.1.1.7. Unidades Funcionales . . . . .	139
4.1.2. Pentium MMX . . . . .	140
4.1.2.1. Frecuencia . . . . .	141
4.1.2.2. Tecnología de integración . . . . .	141
4.1.2.3. Paralelismo a nivel de instrucción . . . . .	142
4.1.2.4. La caché de primer nivel . . . . .	143
4.1.2.5. El conjunto de instrucciones . . . . .	143
<b>4.2. AMD</b> . . . . .	<b>143</b>
4.2.1. K5 . . . . .	144
<b>4.3. Cyrix</b> . . . . .	<b>145</b>
<b>Resumen</b> . . . . .	<b>146</b>
<b>Cuestionario de evaluación</b> . . . . .	<b>146</b>

Los microprocesadores de quinta generación se caracterizan fundamentalmente por un énfasis en las familias de microprocesadores y su frecuencia de reloj, donde el máximo exponente es el Pentium de Intel.

Si en la cuarta generación, placa base y microprocesador eran un mismo ente y funcionaban a una misma frecuencia, la quinta generación supone un claro distanciamiento entre ambos, obligando a la habilitación de un espacio exclusivo destinado a albergar el microprocesador dentro de la placa base, que se conoce como zócalo del microprocesador. La diferencia en velocidad entre uno y otro también provoca la introducción del multiplicador de frecuencia (ver [sección 17.2](#)) en la placa base como elemento conversor de una a otra frecuencia.

La quinta generación vive su auge en la primera mitad de los años 90, y con ella la informática personal vive un nuevo boom, no sólo por el salto en velocidad experimentado por los equipos, sino por tres hechos de índole social ciertamente relevantes:

zócalo  
multiplicador  
p. 8/Vol.3

nuevo boom

- amigable interfaz ❶ La facilidad de manejo de los equipos, conseguida gracias a la amigable interfaz que proporciona la capa software de las aplicaciones (ventanas e iconos) y a la introducción de periféricos de manejo simple e intuitivo por parte de la capa hardware (sobre todo, el ratón). Aunque este tipo de facilidades ya se vió en los años 80 en otros países, en España es Microsoft quien lo institucionaliza con su Sistema Operativo Windows.
- bajo coste ❷ El bajo coste de los equipos, provocado por una estandarización en el computador tipo que adquiere un usuario doméstico.
- funcionalidad ❸ La funcionalidad ampliación de la gama de funciones que puede llevar a cabo un PC.

El PC se consagra así como un electrodoméstico más del hogar, convirtiéndose en un elemento imprescindible de nuestras vidas, y buena parte de esta popularidad se la debe al enorme éxito en ventas cosechado por el Pentium.

Para nosotros, el Pentium supone el punto de partida en nuestra singladura por los modelos comerciales de microprocesadores para PC por dos razones básicas:

- ruptura
  - Por un lado, es una arquitectura que rompe con respecto a los modelos anteriores, proporcionándonos la oportunidad de relegarlos a un segundo plano. Si está pensando en renovar su equipo ahora, difícilmente habrá aguantado tanto tiempo con un equipo anterior al Pentium.
- perspectiva
  - Por otro lado, nos proporciona el grado de perspectiva indispensable para entender el estado de la computación actual. Pasar directamente al Pentium 4 y tratar de ponerse al día en veinte páginas saltándose toda la evolución que vamos a desglosar hasta llegar allí es como arruinar la lectura de una buena novela comenzándola por el final. Déjese llevar por la inercia que le va a proporcionar nuestro paso por la quinta y sexta generación, y ya verá cómo cuando llegue al presente, habrá adquirido el fondo suficiente para entender mucho mejor las cosas que allí se explican.

compañías Elegido nuestro punto de partida, había que seleccionar la gama de productos de una serie de compañías, y en este sentido, elegimos a las dos compañías que mayor volumen de ventas atesoran en el segmento de microprocesadores para PC, y dentro de sus diferentes modelos, dedicamos más tiempo a conocer los que mejor se venden, citando sólo de manera testimonial el segmento de los servidores. Ha sido así de sencillo. Antaño analizábamos lo que más nos gustaba, pero durante todo este tiempo hemos percibido que lo que la gente realmente quiere conocer no son los microprocesadores que utiliza la NASA, sino las intimidades del computador que convive con él en su hogar, y adoptando la regla anterior, nuestra sola intención es maximizar la probabilidad de que usted pueda encontrar el suyo en este libro.

pág. 135 ➔ La [tabla 4.1](#) resume los principales modelos de quinta generación por parte de Intel y AMD que nos van a acompañar a lo largo de este capítulo.

## SECCIÓN 4.1

### Intel

#### 1.1 ▶ Pentium

Pentium fue el nombre dado por Intel al microprocesador 80586, cansado del impedimento legal de poder registrar como marca propia un número y de que toda una legión de compañías aprovecharan este hecho para clonar sus diseños precedentes (primero fue NEC con el V20/V30



Firma	Nombre comercial	Fecha lanz.	Tecn. Int. (micras)	Transist. (mill.)	Frecuencias (MHz)	Frec. bus (MHz)	Código de referencia
Intel	Pentium	03/93	0.80	3.1	60/66	60/66	80501
	Pentium	10/94	0.60	3.2	75/120	50/66	80502 ó P54C
	Pentium	03/95	0.35	3.3	133/200	66	P54CS
	MMX	01/97	0.35	4.5	166/233	66	P55C
AMD	K5	06/95	0.60	4.3	75/100	50/66	Am5x86
	K5	03/96	0.35	4.3	100/166	66	Am5x86

TABLA 4.1: Nombre comercial y código de referencia para los procesadores de quinta generación.

- versus 8086/8088, luego AMD y Cyrix con sus modelos del 80286, 80386 y 80486, y finalmente, NexGen con su Nx586 - versión preliminar del propio Pentium, lanzada justo antes de que AMD comprara la compañía).

Pero ésa es sólo una forma de ver las cosas. Una lectura diferente es la siguiente: el Pentium es un diseño tan sumamente enrevesado que deja muy pocos caminos abiertos para mejoras adicionales. Tanto es así que incluso incita a pensar que fuese una jugada premeditada de la compañía para blindarse frente a la posibilidad de que la competencia pudiera derivar versiones optimizadas del mismo.

Desde el punto de vista arquitectural, el Pentium es un diseño terminal. Nadie, ni siquiera la propia Intel, se ha atrevido a reaprovechar su esqueleto para refinamientos posteriores. El resto de esta sección tratará de dar las razones para ello, mientras que el [capítulo 5](#) nos enseñará los caminos que se escudriñaron para continuar evolucionando.

Lo que está claro es que la llegada del Pentium supone la abolición de una socorrida estrategia: Esa de analizar el buque insignia de Intel, descubrir sus debilidades, y tratar de lanzarse al acoso y derribo con secuelas que subsanen estas carencias en mayor o menor medida. La competencia entiende que ahora debe emprender su propia singladura, y esto va a ser bueno para el mercado en su conjunto: Para los fabricantes, porque va a discriminar plagiadores frente a innovadores, poniendo a cada cual en su sitio. Para los consumidores, porque van a disponer de una extensa gama de arquitecturas sobre las que poder elegir. Y para los que nos gusta este mundillo, porque vamos a poder disfrutar observando y juzgando las diferentes alternativas de diseño por las que opta cada fabricante. Ahora bien, si hasta la quinta generación todo está reglado por los pasos de Intel, la frontera con la sexta es más bien difusa, y ya en la séptima cada fabricante va a entender las cosas a su manera.

### 1.1.1 Frecuencia

El Pentium fue un microprocesador fabricado en un amplio rango de frecuencias. Sobre placas base a 60 MHz, se montan las versiones de 60, 90, 120 y 150 MHz; sobre placas de 50 MHz, una versión a 75 MHz; y finalmente, sobre placas a 66 MHz, las posibilidades son 66, 100, 133, 166 y 200 MHz.

Al mantenerse inalterable la frecuencia de la placa base, lo que va creciendo conforme aumenta la del microprocesador es el multiplicador responsable de cubrir este desfase, y que sucesivamente evoluciona sobre las cinco últimas frecuencias mostradas a pasos discretos: 1, 1.5, 2, 2.5 y 3. Esta discretización a intervalos de un factor de 0.5 ha sido posteriormente respetada por Intel para las placas base a 100 MHz del Pentium II y para las placas base a 133 MHz del Pentium III, de tal forma que basándose en estos valores discretos, uno puede normalmente conocer qué frecuencia de bus tolera un procesador atendiendo a la que el fabricante publicita para su microprocesador.

diseño terminal

→ [pág. 149](#)

secuelas

valores

multiplicador

Características	Modelo de procesador Pentium		
	P5 (1993)	P54C (1994)	P54CS, P55 (1995, 1996)
Variables eléctricas:			
Frecuencia de reloj (MHz)	60, 66, 75	100, 120	133, 150, 166, 200
Frecuencia bus local (MHz)	60, 66, 50	66, 60	66, 60, 66, 66
Máx potencia disipada (W)	16	5	11, 22, 29, 35
Voltaje alimentación (voltios)	5/3.3	3.3	3.3, 3.3, 3.3, 3.45
Zócalo al que se acopla	Socket 4/5	Socket 7	Socket 7
Distancia integración (micras)	0.8	0.6	0.35
Tecnología de integración	BiCMOS (Patillaje:Bipolar; Núcleo:CMOS)		
Número de transistores	3.1 mill.	3.2 mill.	3.3 mill.
Formato y dimensiones	Cuadrado de 296 mm <sup>2</sup> de área		
Encapsulado	Ceramic Pin Grid Array 296 pines		
Segmentación:			
Cauce aritmética entera	5 etapas		
Cauce de punto flotante	8 etapas		
Superescalaridad:			
Unidades de ejecución	3 (2 enteras y 1 en punto flotante)		
Factor de superescalaridad	2, con fuertes restricciones		
Memoria caché:			
L1 Datos (integrada)	8 Kbytes, 128 ctos, 2 lns/cto, 32 bytes/lín		
L1 instrucciones (integrada)	8 Kbytes, 128 ctos, 2 lns/cto, 32 bytes/lín		
L2 unificada	Externa		
Conjunto de instrucciones:	Compatible 80x86		

TABLA 4.2: Resumen de las principales magnitudes del microprocesador Pentium a lo largo de las cuatro versiones de su fabricación.

### 1.1.2 Tecnología de integración

El gran recorrido que muestra la variable frecuencia delata también sucesivas mejoras en la distancia de integración. Los primeros diseños ven la luz a 0.8 micras, haciendo un alto en el camino en las 0.5 micras, y concluyendo su viaje en las 0.35 micras. La [tabla 4.2](#) resume todas estas características, donde podremos apreciar el código de referencia que delimita cada distancia y su implicación sobre el resto de variables eléctricas, como la frecuencia, voltaje, potencia disipada, dimensiones y número de transistores.

Pero lo más novedoso de Intel aquí es, por una vez, la tecnología de integración. En efecto, si en la [sección 3.2](#) que dedicamos a esta variable dijimos que asumiríamos CMOS de forma fija y que para darle recorrido la acompañaríamos de la distancia de integración, el Pentium es, junto al Pentium Pro, la excepción que confirma la regla: Estos dos microprocesadores usan BiCMOS, tecnología que trata de beneficiarse de lo mejor de ambas: La rapidez de conmutación de la bipolar y el bajo consumo de la CMOS.

Normalmente, cuando se procede así, es la parte más externa relacionada con el patillaje la que se integra con tecnología bipolar, pues es allí donde se demanda mayor intensidad de corriente. El núcleo interno suele requerir una tensión inferior, y por ello queda predispuesto para ser integrado en CMOS. En el caso del Pentium, el voltaje de alimentación es de 3.3 voltios tanto para la parte interna como para la externa.

← pág. 136

pág. 50 →

BiCMOS

### 1.1.3 Paralelismo a nivel de instrucción

El microprocesador Pentium dispone de un esquema de segmentación que hereda del 80486, y que se escinde en cinco etapas: Búsqueda de instrucción, decodificación, decodificación-2, ejecución de la operación y escritura del resultado.

segmentación

Todas las etapas se encuentran claramente delimitadas excepto la tercera, que es un tanto heterogénea y donde pueden tener lugar (a) la compleción de la decodificación para pasar a señales eléctricas la palabra de control leída en la ROM de proyección y/o (b) los cálculos necesarios para la generación de dirección de los operandos fuente. Tanto (a) como (b) representan parte del precio a pagar por mantener compatibilidad con un conjunto de instrucciones como el del 80x86, pues son dos de sus peores cualidades las que nos están perjudicando aquí:

- ❶ Su carácter CISC, de donde emerge su unidad microprogramada y la lentitud asociada a su decodificación.
- ❷ Sus enrevesados modos de direccionamiento, que obligan a realizar operaciones internamente antes de disponer de los valores de los operandos con los que trabajar en la fase de ejecución del cauce segmentado del microprocesador.

CISC

direccionamiento

Con respecto a la superescalaridad, es el primer diseño de Intel para PC que contempla esta posibilidad. Estamos ante un superescalar de factor 2, aunque con fuertes limitaciones que enseña descubriremos.

superescalar

La [figura 4.2](#) muestra el diagrama de bloques de este procesador. Horizontalmente, pueden establecerse líneas divisorias imaginarias entre los cinco bloques funcionales que, de arriba a abajo, son responsables de la ejecución de las cinco etapas segmentadas mencionadas.

pág. 140

Por otro lado, la superescalaridad se encuentra asociada con la existencia de múltiples cauces de ejecución dispuestos en el sentido vertical: Los dos cauces de ejecución para datos enteros se denominan cauce-u y cauce-v, y en ellos puede apreciarse una asimetría que funcionalmente se traduce en que el cauce-v sólo pueda ejecutar las instrucciones más sencillas, precisamente aquellas que se implementan de forma completamente cableada y que se ejecutan en un sólo ciclo de reloj.

cauce-u  
cauce-v

El cauce de ejecución para datos de punto flotante aparece más a la derecha en la [figura 4.2](#), y dispone de una segmentación más complicada, consecuencia del número de ciclos que se tardan en ejecutar sus operaciones. En total, son ocho las etapas: Búsqueda de instrucción, decodificación, generación de dirección, búsqueda de operandos, ejecución 1, ejecución 2, escritura, e informe de errores. Estas ocho etapas se controlan desde la unidad en punto flotante, aunque las cinco primeras coinciden con las de los cauces enteros y comparten su misma circuitería, con la salvedad de que las instrucciones enteras escriben el resultado cuando las flotantes lo ejecutan. Al compartir ese hardware de las etapas con los cauces enteros y requerir la unidad en punto flotante operandos de 64 bits, las instrucciones en punto flotante no pueden ejecutarse simultáneamente con las de tipo entero, a no ser que las primeras sean extremadamente simples.

pág. 140

8 etapas FPU

Las dos limitaciones de los párrafos anteriores se traducen en una considerable merma de la facultad superescalar del Pentium en su fase de ejecución: Sólo van a poder ejecutarse dos instrucciones a la vez cuando las dos sean enteras y una de ellas se ejecute en un único ciclo, o cuando una entera se acompañe de otra de punto flotante de extremada sencillez.

merma

Si a esto le sumamos que el complejo sistema de decodificación y direccionamiento consume dos ciclos, y que la ejecución de instrucciones complejas del conjunto 80x86 puede consumir hasta 4 ciclos, no es de extrañar que la cadencia neta con que van finalizando las instrucciones enteras del programa quede ya en torno a una por ciclo de reloj. Y aún debemos cuantificar la incidencia negativa que tienen las dependencias de datos y control presentes en el programa fuente.

complejo

### 1.1.4 Resolución de dependencias

Las dependencias de datos sólo penalizan (1 ciclo) cuando una carga de datos en registro preceda al uso del valor de dicho registro, pues es la única ocasión en la que los caminos de anticipación del Pentium no suministran el valor a tiempo.

**BTB** Las dependencias de control penalizan entre 3 y 4 ciclos sólo cuando la BTB falla en la predicción. La probabilidad de ocurrencia para cada uno de estos eventos es variable dependiendo de las características del programa fuente, aunque como valor ilustrativo podemos establecer una horquilla entre el 5 y el 15 % del total de dependencias de cada tipo.

**algoritmo BTB** La BTB del Pentium procede de la siguiente manera: Cuando una instrucción de salto condicional se ejecuta por primera vez, el procesador la guarda en una de las 256 entradas de su BTB, que se encuentran organizadas en 64 conjuntos de cuatro entradas cada uno en clara semejanza con la forma de organizar las líneas de la memoria caché. Posteriormente, cada vez que se decodifica una instrucción de salto condicional, el procesador la busca en la BTB, y si la encuentra, se dispone de dos bits donde se encuentra anotado lo que hizo la instrucción las dos últimas veces que se ejecutó, y que sirven para predecir si se salta (valores 11 ó 10) o si no (valores 01 ó 00). Si la instrucción no se encuentra en la BTB, el Pentium conmuta a predicción estática fija de no saltar, lo que evita la espera hasta que el cálculo de la dirección efectiva se haya realizado.

**búfers** Emparentados con las decisiones de la BTB, el Pentium utiliza dos búfers de prebúsqueda de instrucciones. El primer búfer está asociado a la ejecución secuencial, y precarga en cada momento las instrucciones del programa que siguen secuencialmente a la instrucción en curso, hasta que se llega a una de salto. En ese momento, si la BTB indica que el salto no se va a realizar, la prebúsqueda continúa como normalmente. Por el contrario, si la predicción de la BTB es de saltar, el segundo búfer comienza a precargar instrucciones a partir de la dirección de salto.

**predicción** El resultado de la predicción no se conoce hasta el comienzo de la etapa de escritura de la instrucción de salto. Si se descubre una predicción errónea, los cauces de instrucción se vacían y el subsistema de prebúsqueda comienza a traer instrucciones procedentes de la nueva secuencia. Una mala predicción penaliza la ejecución entre tres y cuatro ciclos, dependiendo del tipo de la instrucción de salto.

### 1.1.5 La caché de primer nivel

El Pentium dispone de dos cachés internas que son gemelas y se utilizan para almacenar datos e instrucciones por separado. Estas cachés tienen un tamaño de 8 Kilobytes, una organización asociativa de 128 conjuntos, con 2 líneas por conjunto y 32 bytes por línea. El algoritmo de reemplazo es LRU y la política de actualización es directa (write-through). Además, cada caché dispone de su propia TLB (Translation Look-Aside Buffer), una especie de caché para acelerar la traducción de direcciones virtuales a físicas, que dispone de 64 entradas en el caso de los datos y 32 en el de las instrucciones, ambas organizadas de forma totalmente asociativa.

**TLB**

### 1.1.6 El conjunto de instrucciones

**80x86** El Pentium es un procesador CISC en su mayoría, legado de la familia de los 80x86 con cuyo conjunto de instrucciones se decide mantener compatibilidad. Este hecho es quizá el que provoca una influencia más perjudicial en su diseño, pero paradójicamente, también es el principal responsable de su arrollador éxito comercial: Con él, Intel emprende una meteórica evolución que la encumbra como una de las diez compañías que más dinero facturan a nivel mundial (en fechas recientes ha escalado hasta la quinta posición).

**éxito comercial**

**rasgos CISC** Sus rasgos CISC más marcados son:

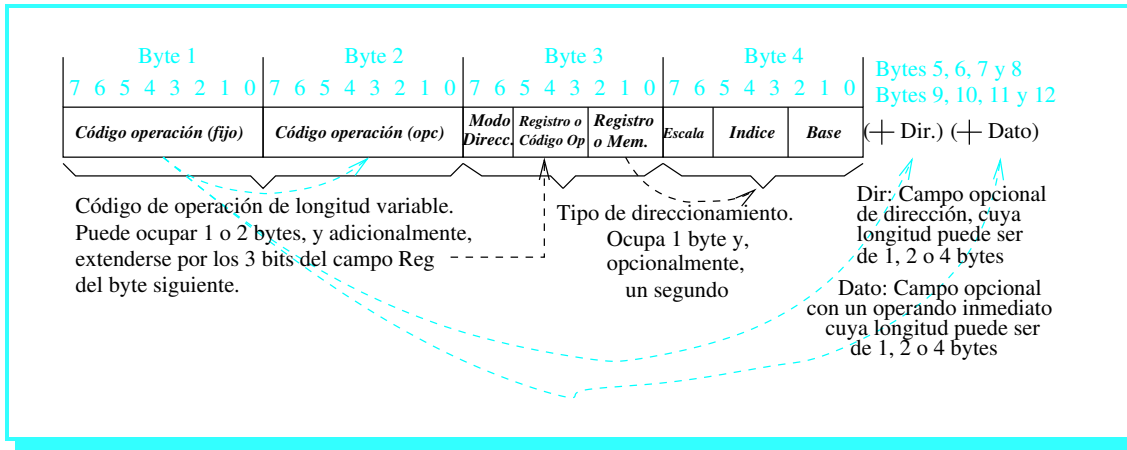


FIGURA 4.1: El formato de instrucción de la familia de microprocesadores x86 de Intel, un magnífico ejemplo de cómo NO debe diseñarse un formato de instrucción.

- ❶ La existencia de una memoria de microprograma en su Unidad de Control (ROM que describe la señales eléctricas que gobiernan el funcionamiento de la circuitería por cada uno de los ciclos de ejecución de cada instrucción).
- ❷ Sus múltiples y complejos modos de direccionamiento.
- ❸ Su escaso banco de registros.
- ❹ Su formato de instrucción de longitud variable de entre 1 y 12 bytes como consecuencia no sólo de un diseño enrevesado, sino de aceptar operandos inmediatos de longitud muy dispar: enteros de 1, 2 y 4 bytes y reales de 4, 9 y 10 bytes).

Este formato de instrucción aparece en la figura 4.1, y resulta clave para entender la complejidad en la decodificación de instrucciones de los microprocesadores para PC, puesto que toda la familia de Intel hasta el Pentium 4 y de AMD hasta el K7 son compatibles con él y por lo tanto lo aceptan como código de entrada interno para posteriormente redefinirlo a microinstrucciones de su arquitectura nativa.

formato de instrucción compatibles

La longitud media de las instrucciones de un programa Pentium se cifra en 3.5 bytes, con una horquilla de un mínimo de 3.2 y un máximo de 3.7 en función del tipo de programa.

Sin embargo, también tiene algunas pinceladas RISC, principalmente derivadas de su carácter segmentado y superescalar, y de la vertiente cableada para el control de sus instrucciones más sencillas.

RISC

### 1.1.7 Unidades Funcionales

El bloque funcional dedicado al procesamiento de enteros se completa con 8 registros de propósito general de 32 bits y dos ALUs de 32 bits.

8 registros

El bloque funcional de punto flotante también cuenta con 8 registros, pero de 80 bits, con objeto de albergar datos de precisión simple, doble o extendida, y que se corresponden con longitudes de 32, 64 y 80 bits de acuerdo a los respectivos formatos del estándar IEEE-754 ya establecidos en el 80486. Dispone de un sumador, un multiplicador y un divisor, que operan todos ellos en los tres formatos mencionados. En el peor de los casos (división sobre formato extendido), el procesador consume 63 ciclos de ejecución, lo que evidentemente tiene un mayúsculo impacto en el rendimiento de este cauce segmentado.

punto flotante

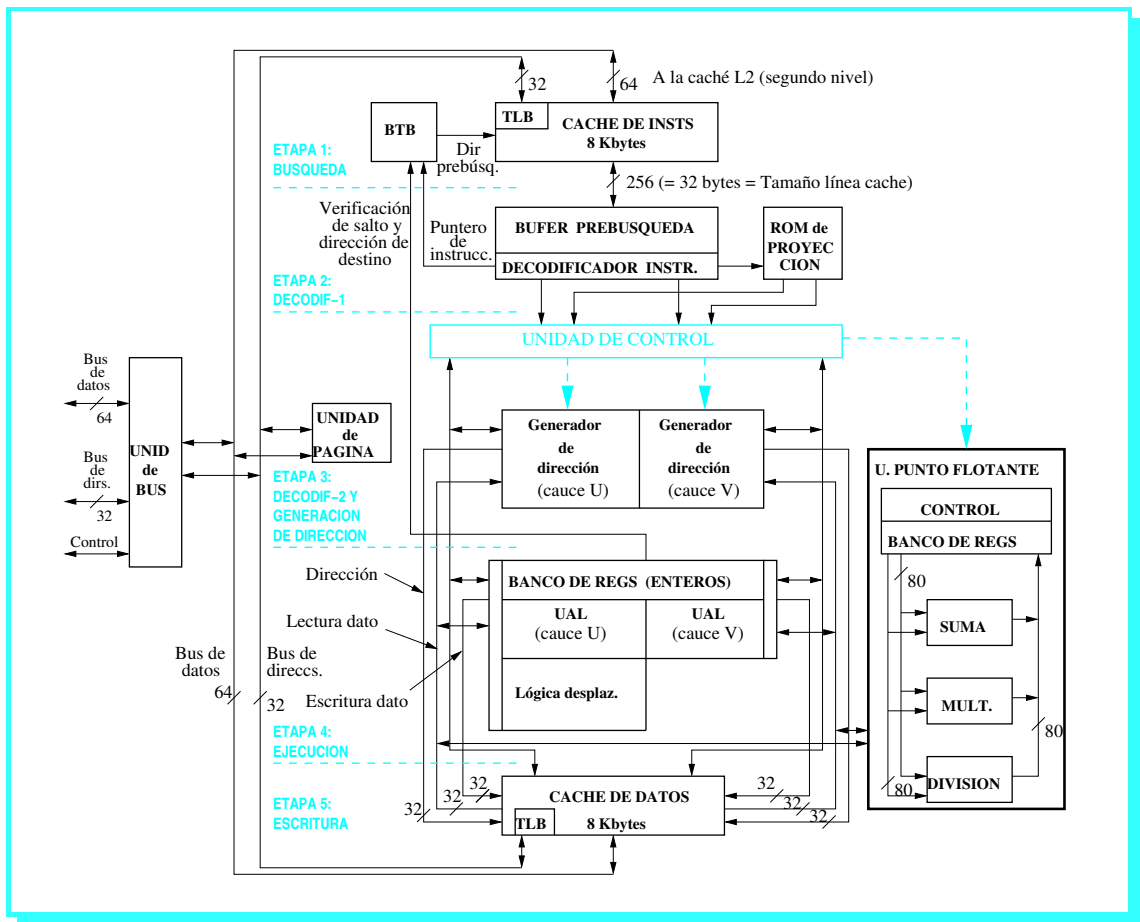


FIGURA 4.2: Diagrama de bloques del microprocesador Pentium de Intel.

## 1.2 ▶ Pentium MMX

solapado

El Pentium MMX entra en escena en 1997, cuando la versión base del procesador estaba ya dando sus últimos coletazos. Resulta curioso observar cómo Intel solapa en el tiempo diseños de distintas generaciones, pues este microprocesador es contemporáneo al Pentium Pro, que en 1995 abre la sexta generación; y algo similar ocurrió al inaugurar el Pentium la quinta generación (1993), pues el 80486 DX4 también vio la luz con posterioridad (1994).

Socket 7

En realidad, el Pentium MMX resistió incluso los embites del Pentium II durante todo 1997, puesto que a Intel le interesa mantener en el mercado una alternativa barata para las arquitecturas de gama más baja. Fue a la tercera, con la salida del Celeron, la secuela de bajo coste del Pentium II, cuando el Pentium MMX perdió su sitio en el mercado y dejó de fabricarse. Aunque el Celeron nunca dió a la compañía los días de gloria proporcionados por el MMX, tampoco era ésa su finalidad; se trataba más bien dar un golpe de efecto a los productos basados en zócalos Socket 7, principalmente placas base y microprocesadores, en los que la competencia, sobre todo VIA y AMD, basaba su línea de producción.

Características	Microprocesador Pentium MMX P55C (1997)
<b>Variables eléctricas:</b>	
Frecuencia de reloj (MHz)	166, 200, 233, 266
Frecuencia bus local (MHz)	66
Máx potencia disipada (W)	14, 15.7, 17, 20
Voltaje alimentación (voltios)	2.8
Zócalo al que se acopla	Socket 7
Tecnología de integración	CMOS a 0.35 micras
Número transistores (mill.)	4.5
Formato y dimensiones	Cuadrado de 141 mm <sup>2</sup> de área
Encapsulado	Ceramic Pin Grid Array 296 pines
<b>Segmentación:</b>	
Cauce aritmética entera	6 etapas
Cauce punto flotante y MMX	8 etapas
<b>Superescalaridad:</b>	
Unidades de ejecución	5 (2 enteras, 2 MMX y una punto flotante)
Factor de superescalaridad	2, con algunas restricciones
<b>Memoria caché:</b>	
L1 Datos (integrada)	16 Kbytes, 128 ctos, 4 líns/cto, 32 bytes/lín
L1 instrucciones (integrada)	16 Kbytes, 128 ctos, 4 líns/cto, 32 bytes/lín
L2 unificada	Externa
Conjunto de instrucciones:	80x86 + 57 instrucciones multimedia

TABLA 4.3: El Pentium MMX en números.

### 1.2.1 Frecuencia

El Pentium MMX comienza a fabricarse a 133 MHz, y avanza sucesivamente por los 166 y 200 MHz para adelantar al Pentium y situarse en los 233 MHz. En Estados Unidos llegó incluso a fabricarse “in extremis” una versión a 266 MHz. Todos estos modelos compartían el controlador de bus del Pentium clásico, donde la frecuencia de trabajo del bus local se situaba en los 66 MHz.

266 MHz

### 1.2.2 Tecnología de integración

El Pentium MMX se fabrica ya exclusivamente sobre CMOS. Si el Pentium distingue un núcleo interno integrado en CMOS y reserva la tecnología bipolar para la zona circundante donde se sitúa el patillaje, el MMX unificará estas dos regiones con CMOS, pero las diferenciará colocando a una tensión inferior de 2.8 voltios la primera de ellas. La razón para reducir el voltaje en la zona interna no es otra que la de esquivar problemas de sobrecalentamiento, lo que posibilitó la posterior consecución de las versiones a 233 y 266 MHz. Por el contrario, el mantenimiento de los 3.3 voltios en la zona externa se explica porque en esa parte del chip se encuentra la memoria caché de primer nivel, y la tecnología de fabricación más usual para ella por aquel entonces descansaba sobre este voltaje. De hecho, no es casualidad que durante la segunda mitad de los 80 predominaran los microprocesadores a 5 voltios, durante la primera mitad de los 90 a 3.3 voltios, y durante la segunda mitad de los 90 a 2.5 voltios, precisamente los tres voltajes elegidos para las tecnologías de caché más extendidas durante esos períodos temporales.

CMOS

voltaje

La dualidad de voltaje del Pentium MMX planteó el problema de que muchas placas carecían de un segundo convertidor de voltaje, y por tanto, aunque el microprocesador podía insertarse en el zócalo de la placa, no funcionaba e incluso corría el riesgo de ser dañado internamente. Para

dualidad de voltaje

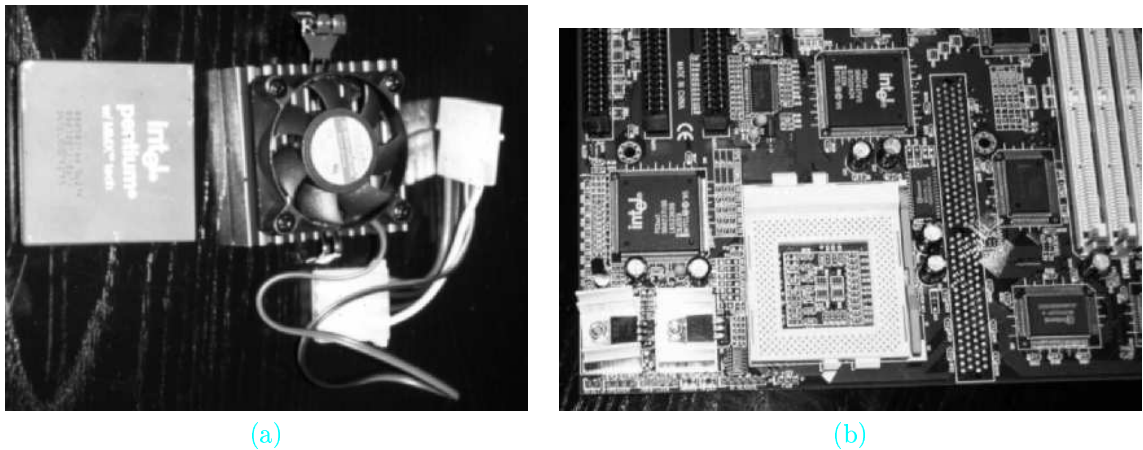


FOTO 4.1: (a) El Pentium MMX y su disipador de calor. (b) El zócalo que la placa base habilita para pincharlo y los dos reguladores de voltaje que requiere (esquina inferior izquierda).

subsanar este problema, los fabricantes de placas suministraron a partir de entonces un segundo convertor a 2.7 voltios. La [foto 4.1](#) muestra en (a) el procesador y la unidad de disipación de calor responsable de su refrigeración, y en (b) el zócalo del procesador con sus dos aparatos convertidores de voltaje a su izquierda.

← pág. 142

### 1.2.3 Paralelismo a nivel de instrucción

Tomando como referente el esqueleto del Pentium, la versión MMX introduce una serie de mejoras para la ejecución simultánea de instrucciones que se resumen en las dos premisas que sintetizan su capacidad segmentada y superescalar.

una etapas más

- ① **Segmentación.** Los dos cauces U y V para el procesamiento de instrucciones enteras se han incrementado en una etapa, completando un total de seis y mejorando su capacidad de procesamiento simultáneo, lo que en la práctica se traduce en una superescalaridad de factor dos más realista.

factor dos

- ② **Superescalaridad.** Se mantiene en este factor dos, a pesar de que ahora se dispone de cinco unidades funcionales para la ejecución de instrucciones, consecuencia de agregar las dos nuevas unidades MMX a las tres que ya se tenían (las dos de tipo entero y la de punto flotante). También se han duplicado (de dos a cuatro) el número de búfers de escritura en las unidades de carga/almacenamiento encargadas de realizar operaciones con memoria.

Con toda esta artillería, queda al alcance del diseño una superescalaridad de factor 3 (la tercera instrucción que acompañaría a las dos provenientes del Pentium podría ser una de corte MMX). Sin embargo, esto no es posible por dos razones. Primero, porque la capacidad para la búsqueda y decodificación de instrucciones no experimenta mejora alguna; y segundo, porque las unidades MMX trabajan exclusivamente con ocho registros de 64 bits (MM0 al MM7) que no constituyen un banco independiente, sino que se mapean sobre la parte menos significativa de los 8 registros de 80 bits de que dispone la unidad de punto flotante para sus cálculos (ver parte derecha de la [figura 6.6](#)).

pág. 217 →

amortizar HW

El planteamiento de Intel es acertado desde el punto de vista de que las aplicaciones multimedia que emplean las operaciones MMX rara vez hacen uso de la aritmética de punto flotante, y viceversa, con lo que en realidad se procura amortizar al máximo la presencia de esta circuitería dándole un doble uso.



### 1.2.4 La caché de primer nivel

El Pentium MMX duplica la capacidad de las cachés de primer nivel, siendo éstas ahora de 16 Kbytes cada una. Los parámetros de esta caché son todos iguales a los del Pentium, con excepción del número de líneas por conjunto, que es ahora el doble. Precisamente es ése el parámetro que más interesa aumentar en una caché, por lo que podemos decir que el aumento del rendimiento es apreciable: no sólo se gana cuantitativa sino también cualitativamente, y en cachés pequeñas como éstas, tan importante es una mejora como la otra, suponiendo cada una una reducción del índice de fallos cercana a la mitad.

duplica la capacidad

La motivación para introducir esta mejora se encuentra en el hecho de que las aplicaciones multimedia manejan normalmente un volumen de datos elevado, siendo las transferencias con memoria una de las cosas que más ralentizan al procesador. El índice de fallos en caché es ahora casi cuatro veces inferior al de un Pentium clásico, y semejante bocado no va a ser paladeado exclusivamente por las aplicaciones MMX, sino por cualquier aplicación de propósito general que utilice la memoria con cierta frecuencia.

### 1.2.5 El conjunto de instrucciones

Si por algo se conoce a este procesador es por el subconjunto de 57 nuevas instrucciones que incorporan 24 nuevas operaciones orientadas exclusivamente a la ejecución eficiente de aplicaciones multimedia (ver [sección 3.5.4](#)). Muchas de estas instrucciones corresponden a una misma operación, ya que el modelo SIMD (Simple Instrucción Múltiple Dato) por el que se rige su diseño contempla hasta seis variantes distintas definidas sobre un mayor o menor volumen de datos dependiendo de su longitud hasta completar la anchura total de 64 bits de almacenamiento de que se dispone en los registros internos. A su vez, las 24 operaciones pueden clasificarse en siete grandes categorías atendiendo a su funcionalidad, tal y como reflejamos en la [tabla 3.14](#).

57 nuevas

☛ pág. 106

☛ pág. 109

## SECCIÓN 4.2

# AMD

En los años 80, veíamos a Intel moverse con extremada agilidad. Sacaba buenos productos al mercado, competía enconadamente con Motorola por el liderazgo de los microprocesadores, y aunque su ensamblador siempre fué mejorable, nos acostumbró a buenas arquitecturas con potentes ALUs. Aun con un comienzo poco acertado arquitecturalmente como el del 8086 (pero brillante desde una perspectiva comercial), fueron sus hermanos mayores los que quitaron a Motorola las ganas de continuar con la rivalidad en el mercado de los PC. Con la llegada del Pentium, Motorola se desmarcó diseñando el Power PC para otro segmento del mercado, dejando momentáneamente sólo a Intel en la fabricación de microprocesadores para las arquitecturas más domésticas.

Pero pronto surgieron nuevos contendientes en la lucha por este apetecible mercado. La década de los 90 nos trajo a AMD (*Advanced Micro Devices* - Micro Dispositivos Avanzados), otra compañía norteamericana afincada en California que decide aprovechar la coyuntura anterior para concentrarse en el segmento del PC, un mercado al que nunca prestó demasiada atención debido a la política de diversificación llevada a cabo por la compañía durante los años 80.

Frente a otros potenciales competidores de Intel como Cyrix o Nexgen, AMD presentaba la ventaja de disponer de su propia planta de fabricación de chips, lo que concedía cierta autonomía a sus diseños. De hecho, en el pasado fue en ocasiones subcontratada por Intel para fabricarle algunos de sus diseños. AMD aprovechó aquella vieja amistad para obtener una licencia de explotación para el microcódigo de los procesadores de Intel, y de esta manera, lanzó al mercado

autonomía

sus propias versiones del 80386 y 80486 de Intel, tratando de ganar cuota de mercado a base de unos precios más atractivos.

**precios** A partir de entonces los productos de AMD van ganando en calidad, ofreciendo en ocasiones prestaciones no incluidas en los modelos de Intel, y a unos precios siempre inferiores a los de su competidor.

Por ejemplo, ya en el tramo final de vida del 80486 vimos versiones de AMD como el DX4 a 120 y 133 MHz, cuando Intel abandonó el diseño en los 100 MHz (aunque con 16 Kbytes de caché L1 frente a los 8 Kbytes de AMD).

La llegada del Pentium de Intel pone al resto de fabricantes en una tesitura: Hay que mantener la compatibilidad de zócalo y patillaje de los modelos de Intel para aprovecharse de sus placas base y demás logística hardware, al tiempo que se debe respetar su mismo conjunto de instrucciones para conseguir ejecutar las aplicaciones desarrolladas para él.

Sin embargo, para ganar credibilidad y solera en el hermético mundo del PC como compañía, hay que desarrollar nuevas arquitecturas. ¿Acaso es esto posible? Desde luego, y la pista clave nos la dió Intel con su sucesor el Pentium Pro, que mantiene hasta los decodificadores de instrucción y redefine la forma de ejecutar estas instrucciones sobre la circuitería. De esta manera, las etapas de búsqueda y decodificación del cauce segmentado del procesador, que son las únicas ligadas al conjunto de instrucciones nativo, quedarán fuertemente condicionadas en todos los procesadores hasta que se rompe la baraja con el Itanium, ya con un nuevo repertorio de instrucciones.

**compatible** AMD acuña esta conducta como propia, y tanto el K5, como el K6, K6-II, K6-III ó K7 aceptarán código nativo del Pentium, aunque siempre tratando de mejorar la forma en que éste se ejecuta. Este cóctel de compatibilidad, buenas prestaciones y bajos precios provocan que la popularidad de AMD aumente a pasos agigantados. A partir de ahí, la compañía muestra también buenas maneras en su vertiente empresarial:

**alianzas** ❶ Desarrolla muy buenas alianzas con fabricantes de mucha solera, como Digital (cuando aún no había sido adquirida por Compaq) o VIA, uno de los fabricantes líderes de juegos de chips para placas base, lo que supone un buen aval para la viabilidad de la empresa en el corto y medio plazo y un claro puntal de futuro.

**beneficios** ❷ Comienza a obtener grandes beneficios derivados de esta actividad. La cuenta de resultados de la compañía estuvo en números rojos hasta 1998, contrastando con la de Intel, que aumenta vertiginosamente sus beneficios y ve como sus acciones mantienen una meteórica evolución durante toda la década de los noventa. A partir de ahí, AMD espabila y su creciente cuota de mercado empieza a darle alegrías que compiten sin rubor con la trayectoria seguida por Intel.

Pero dejemos la contabilidad empresarial atrás y regresemos al mundo de la computación para centrarnos de forma decidida en la aportación de AMD dentro de la quinta generación de microprocesadores.

## 2.1 ▶ K5

**1996** El microprocesador K5 fué el elegido por AMD para competir con el Pentium original. Corría el año 1996, y este microprocesador se ubicó entre medias del Pentium y el Pentium Pro en cuanto a prestaciones, pero su coste era muy inferior.

**Nx586** El diseño del K5 no está inspirado en el Pentium, sino en su precursor, el Nx586 de NexGen, compañía que AMD adquirió para tomar como punto de partida una arquitectura propia. Sin embargo, el principal argumento para sus ventas siguió siendo la compatibilidad con los modelos de Intel, y ésta pasó por mantener su misma infraestructura externa, tanto su conexión a la placas base a través del zócalo Socket 7 como su controlador de bus local.

**Socket 7**

Internamente, el K5 disponía de 4.3 millones de transistores, una cantidad muy similar a la del Pentium MMX, con los que se implementó una arquitectura con cuatro rasgos sobresalientes:

cuatro rasgos

- ❶ Un novedoso sistema de predicción de saltos.
- ❷ Ejecución fuera de orden que no se penalizaba al ejecutar aplicaciones de 16 bits, algo que al Pentium Pro le costó que el mercado le diera la espalda.
- ❸ Factor de superescalaridad de cuatro (frente a dos con fuertes restricciones en el Pentium).
- ❹ Una caché L1 de 24 Kbytes, dividida en 16 Kbytes para instrucciones y 8 Kbytes para datos.

Llamamos la atención sobre las dos últimas características, pues es aquí donde comienza a forjarse lo que va a ser la personalidad propia de AMD en sus arquitecturas venideras:

- ❶ Diseños fuertemente superescalares donde esta faceta se prima frente a la segmentación como forma de explotar el paralelismo a nivel de instrucción. Se persigue así completar una arquitectura interna a salvo de carencias, trasladando los posibles cuellos de botella a la parte más externa del procesador, donde su tratamiento y solución puede encararse mediante actuaciones que no atañen a la constitución del procesador.
- ❷ Las actuaciones anteriores consisten fundamentalmente en la aceleración del bus local y en la incorporación de memorias cachés de mayor tamaño, velocidad y proximidad al procesador. La decisión de primar en el K5 el tamaño de la caché de instrucciones L1I frente a la de datos L1D va precisamente en esta línea, y a partir de aquí siempre veremos cachés de primer nivel más grandes en los modelos de AMD que en sus homólogos coetáneos de Intel.

superescalar

caché

La segmentación del K5 es bastante sencilla: 5 etapas, asemejándose bastante a la descomposición en etapas realizada en la arquitectura Pentium, sólo que en lugar de dedicar dos etapas a la fase de decodificación, utiliza sólo una para ese propósito y dedica la siguiente a labores relacionadas con la ejecución fuera de orden.

segmentación

Del K5 se comercializaron dos versiones: Una a 0.6 micras que comenzó en 75 MHz sobre un bus de 50 MHz con multiplicador de 1.5 y prosiguió hasta los 90 y 100 MHz como techo manteniendo el multiplicador y subiendo el bus a 60 y 66 MHz, y otra a 0.35 micras con ligeras mejoras arquitecturales que partiendo de estos 100 MHz alcanzó sucesivamente especificaciones de PR120, PR133 y PR166 MHz. Estos valores PR simbolizaban Pentium Rated, es decir, configuraciones equivalentes a un Pentium de la frecuencia indicada. Cyrix también utilizaría esta denominación, prolongándola incluso hasta pasada la sexta generación.

versiones

Desgraciadamente para AMD, el K5 llegó al mercado cuando ya Intel lo había sembrado todo de Pentiums. Ese error lo pagó la compañía con una cuota de mercado y un volumen de ventas irrisorio para la calidad que acreditaba aquel diseño. Pero al menos supo aprender del error, y para cuando Intel lanzó el Pentium II, AMD ya tenía preparada su contraoferta: El K6.

error

SECCIÓN 4.3

## Cyrix

Durante la tercera y cuarta generación, Cyrix comercializa algunos diseños interesantes, como el 486DRx2, que puede reemplazar al 80386 de Intel pero poniendo en práctica muchas de las optimizaciones del 80486. O el 486DX2, que alcanza un techo de frecuencia de 80 MHz y ya completa los 8 Kbytes de caché L1 típicos de una arquitectura del tipo 80486.

El valerse de IBM para integrar sus diseños le sirve a Cyrix para obtener cierta independencia de Intel y ponerse a salvo de los problemas de plagio con que ésta acusó a compañías como AMD durante esta época. IBM manufacturó los procesadores de Cyrix hasta la versión inicial del 6x86, dentro ya de la sexta generación.

IBM

En esta quinta generación, Cyrix comercializó con un éxito relativo una saga de microprocesadores de rendimiento similar al del Pentium de Intel. Se trataba del 5x86, un chip compatible en zócalo y patillaje con el 80486 de Intel, pero que puso a disposición del usuario buena parte de las prestaciones utilizadas por su hermano mayor el 6x86 en lo referente al paralelismo a nivel de instrucción. Aunque no dispuso de ejecución superescalar, sí que fue capaz de predecir los saltos para resolver las dependencias de control y anticipar valores para resolver las dependencias de datos.

La compatibilidad con el 80486 le obligó además a hacer uso de un bus de datos e interfaz de 32 bits, pero internamente sus caminos de datos eran de 64 bits. En cuanto a sus frecuencias de trabajo, alcanzaron los 100, 120 y 133 MHz.



## Resumen



Desde el punto de vista eléctrico, la quinta generación supone un énfasis en la frecuencia del procesador. Arquitecturalmente, supone la llegada de los multiplicadores de reloj y el cuello de botella al bus local, acuciado por la presencia de una caché externa con la que debe interactuar.

Comercialmente, esta generación es la última en la que todos los productos de cada firma reciben una única denominación: Pentium en Intel y K5 en AMD. Tan rotundo fue el éxito en ventas del primero como el fracaso del segundo, pero ninguno de ellos nos merece el aprobado, y poco o nada se aprovecharía de ellos en generaciones venideras.

Por salvar algo, nos quedamos con la aportación al conjunto de instrucciones de las extensiones multimedia MMX, aunque no con los mecanismos que se articularon para sacarles partido, que también suspenden nuestro examen. La vitola MMX reluce gracias a una latente mejora en las cachés de primer nivel, y el mercado tomará buena cuenta de ello en la sexta generación.

## Cuestionario de evaluación

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

1 ¿Qué arquitectura está considerada como la más importante de toda la quinta generación de microprocesadores? (toma el volumen de ventas y la repercusión sobre modelos futuros como referencia)

- a El K5 de AMD.
- b El Pentium de Intel.
- c El Pentium MMX de Intel.
- d El K6 de AMD.

2 El Pentium de Intel

a Integra sus cachés L1 de datos e instrucciones en el propio chip del procesador.

b Dispone de una ejecución superescalar y segmentada.

c Integra varias ALUs y una FPU internamente.

d Todas las respuestas anteriores son correctas.

3 ¿Cuáles son los aspectos más negativos de la arquitectura de un procesador Pentium?

- a** Las intrincadas restricciones en su factor de superescalaridad.
- b** La ausencia de cobertura para caché de segundo nivel (L2).
- c** Las dos anteriores.
- d** La frecuencia de reloj.
- 4** ¿Qué aspecto del microprocesador Pentium se encuentra fuertemente acotado en su rendimiento por la caprichosa combinación de complejidad y tipo (entero o de punto flotante) que puedan presentar instrucciones adyacentes del programa fuente?
- a** La memoria caché.
- b** La superescalaridad.
- c** La segmentación.
- d** Las ALUs.
- 5** El K5 fue, comparado con el Pentium, un procesador
- a** Posterior pero más sofisticado.
- b** Anterior y más sofisticado.
- c** Posterior pero menos sofisticado.
- d** Anterior y menos sofisticado.
- 6** ¿Por qué tuvo el K5 una aceptación tan baja en el mercado?
- a** Era una arquitectura demasiado rudimentaria.
- b** Tenía un precio excesivo.
- c** Llegó demasiado tarde.
- d** No era compatible con los programas x86 existentes.

- 7** ¿Por qué razón muchos de los zócalos para Pentium no son válidos para la versión MMX del procesador?
- a** El MMX tiene más patillas.
- b** El MMX funciona a diferente voltaje.
- c** El MMX no es simétrico.
- d** El MMX incluye una caché interna mayor.
- 8** ¿Qué material utilizan los Pentium para conectar los transistores dentro del procesador?
- a** Aluminio (Al).
- b** Cobre (Cu).
- c** Silicio (Si).
- d** Arseniuro de Galio (GaAs).
- 9** La Unidad de Punto Flotante (FPU) del Pentium MMX se encontraba mejorada respecto a la del Pentium clásico.
- a** Falso, era la misma en ambos casos.
- b** Verdadero, la FPU del MMX era más avanzada puesto que también ejecutaba las instrucciones MMX.
- c** Falso, puesto que las operaciones MMX se realizaban en una Unidad Funcional dedicada. Era el banco de registros para operandos en punto flotante lo que se encontraba reutilizado por los operandos MMX.
- d** El matiz señalado en la opción c es correcto, pero eso no contradice la afirmación del enunciado de la cuestión, puesto que la FPU del microprocesador MMX permitía además operar con números reales representados en formato IEEE-1394.



# La sexta generación

## Sumario

<b>5.1. Intel</b> . . . . .	<b>151</b>
5.1.1. Pentium Pro: El esqueleto de la sexta generación . . . . .	151
5.1.1.1. Tecnología de integración . . . . .	152
5.1.1.2. Frecuencia de reloj . . . . .	152
5.1.1.3. Paralelismo a nivel de instrucción . . . . .	152
5.1.2. Predicción de salto . . . . .	158
5.1.2.1. Memoria caché . . . . .	159
5.1.2.2. El conjunto de instrucciones . . . . .	159
5.1.3. Pentium II: Un nuevo formato . . . . .	161
5.1.3.1. Frecuencia . . . . .	161
5.1.3.2. Integración y zócalo . . . . .	161
5.1.3.3. Paralelismo a nivel de instrucción . . . . .	162
5.1.3.4. Memoria caché . . . . .	162
5.1.3.5. Hardware subyacente . . . . .	164
5.1.4. Deschutes: Pentium II Turbo . . . . .	164
5.1.5. Pentium II Xeon: Para servidores . . . . .	165
5.1.6. Celeron: Un paso en falso . . . . .	166
5.1.7. CeleronA: La rectificación . . . . .	167
5.1.8. Celeron Coppermine: Reconversión a Pentium III . . . . .	167
5.1.9. Pentium III: Más instrucciones multimedia . . . . .	168
5.1.9.1. SSE: 70 instrucciones multimedia . . . . .	170
5.1.9.2. Número de identificación para una CPU . . . . .	172
5.1.9.3. Zócalo e integración . . . . .	173
5.1.10. Coppermine: Novedades en la integración . . . . .	173
5.1.10.1. Frecuencia de reloj . . . . .	173
5.1.10.2. Tecnología de integración y zócalo . . . . .	174
5.1.10.3. Memoria interna . . . . .	174
5.1.11. Pentium III Xeon: Más para servidores . . . . .	175
<b>5.2. AMD</b> . . . . .	<b>176</b>
5.2.1. K6: El primer problema para Intel . . . . .	176
5.2.1.1. Frecuencia . . . . .	176
5.2.1.2. Tecnología de integración . . . . .	176
5.2.1.3. Paralelismo a nivel de instrucción . . . . .	176
5.2.1.4. Memoria caché . . . . .	177
5.2.1.5. Conjunto de instrucciones . . . . .	177

5.2.2. K6-2: Comienza el baile multimedia . . . . .	177
5.2.3. K6-III: El ejemplo de configuración equilibrada . . . . .	179
<b>5.3. Comparativa: Intel versus AMD . . . . .</b>	<b>180</b>
5.3.1. Síntesis de la trayectoria de Intel . . . . .	181
5.3.2. Síntesis de la trayectoria de AMD . . . . .	181
5.3.3. Conclusiones . . . . .	182
<b>5.4. Cyrix . . . . .</b>	<b>183</b>
<b>Resumen . . . . .</b>	<b>184</b>
<b>Cuestionario de evaluación . . . . .</b>	<b>185</b>

**SMP** La sexta generación entra en escena con el auge de los multiprocesadores, irrumpiendo a mediados de los años 90. Los sistemas de esta generación se caracterizan en un primer estadio por su vertiente SMP (Symmetric MultiProcessing) o la facultad del microprocesador para entenderse con sus homólogos conectados a una misma placa base.

**servidor** El primer microprocesador de sexta generación para PC es el Pentium Pro, pionero en permitir la conexión de hasta cuatro microprocesadores en la misma placa base. Cuando surgió esta idea, muchos vaticinaron que sería toda una revolución. La perspectiva histórica con que ahora miramos aquello nos dice que no fue así. Hoy en día sigue resultando difícil encontrar este tipo de configuración, salvo que nos metamos en un sistema tipo servidor, esto es, un computador central que proporciona una serie de servicios (disco, impresora, etc) a una multitud de clientes que lo solicitan a través de una red de interconexión, y cuyo precio sólo puede amortizarse en entornos con multitud de equipos que se salen de los ambientes meramente domésticos.

Es por ello que de tener que caracterizar a la sexta generación en un plano más modesto, elegiríamos como rasgos distintivos los dos siguientes:

**PNI** ❶ El uso extensivo del paralelismo a nivel de instrucción. A mediados de los 90, la industria del PC tiene ya a tiro la frontera del Gigahecio tras una increíble progresión de la frecuencia de trabajo del microprocesador a lo largo de la quinta generación. Pero existen otros métodos para aumentar el rendimiento del equipo que aún no han sido suficientemente explotados, como la superescalaridad, la segmentación y la supersegmentación (ver [sección 3.3](#)). Estas técnicas suponen un coste adicional en lo que se refiere al número de transistores que integran el microprocesador, pero el camino hacia su extensiva utilización se encuentra allanado por dos motivos: (1) la simplificación del conjunto de instrucciones que introducen los diseños basados en filosofía RISC (ver [sección 3.5.2](#)), y (2) las progresivas mejoras de la tecnología de integración, que durante la quinta generación han evolucionado desde las 0.8 micras hasta las 0.35.

[pág. 58](#) ➔

[pág. 97](#) ➔

**grandes cachés** ❷ La inclusión de grandes cachés en el interior del chip microprocesador. El rendimiento de una arquitectura PC tipo Pentium se encuentra seriamente limitado por la lentitud del bus que transporta los datos entre el procesador y el resto de elementos del computador. La frecuencia del bus (y por ende, la de la placa base) no resulta fácil de acelerar, ya que a él están conectados todo tipo de dispositivos heterogéneos con unos parámetros internos muy dispares. En consecuencia, la clave para esquivar este cuello de botella es utilizarlo lo menos posible, y puesto que su uso depende más de los accesos a memoria principal que de ninguna otra cosa, implementar grandes cachés dentro del procesador es la mejor solución.



	Nombre comercial	Fecha lanz.	T.l. (micras)	Transist. (mill.)	Frecuencias (MHz)	F. bus (MHz)	Caché L2   L3		Segmento	Código de referencia
I N T E L	Pentium Pro	11/95	0.6/0.35	5.5	150/200	66	✓	✗	SE	P6
	Pentium II	05/97	0.35	7.5	233/300	66	✓	✗	PC	Klamath
	Pentium II	01/98	0.25	7.5	333/450	100	✓	✗	PC	Deschutes
	P. II Xeon	01/98	0.25	7.5	400/450	100	✓	✗	SE	Xeon
	Celeron	06/98	0.25	7.5	266/300	66	✗	✗	PC	Covington
	CeleronA (L2)	01/00	0.25	19	300/533	66	✓	✗	PC	Mendocino
	Celeron	11/00	0.18	19	800/1100	100	★	✗	PC	Coppermine
	CeleronA	05/02	0.13	19	1200/1400	100	★	✗	PC	Coppermine
	Pentium III	08/99	0.25	9.5	450/600	100	✓	✗	PC	Katmai
	Pentium III	10/99	0.18	28.1	500/1000	133	★	✗	PC	Coppermine
	Pentium III	07/01	0.13	N/D	1.13/1.3	133	★	✗	PC	Tualatin
	P. III Xeon	03/99	0.25	9.5	500/550	100	✓	✗	SE	Tanner
	P. III Xeon	04/00	0.18	28.1	600/933	133	★	✓	SE	Cascades
A M D	K6	04/97	.35/.25	8.8	166/333	66	✗	✗	PC	K6
	K6-2/K6-3D	05/98	0.25	9.3	266/500	66/100	✗	✗	PC	Chomper
	K6-2+	04/00	0.18	N/D	450/500	100	★	✗	PO	"
	K6-III	02/99	0.25	21.3	400/600	100	★	✗	PC	Sharptooth
	K6-III+	04/00	0.18	N/D	450/500	100	★	✗	PO	"

TABLA 5.1: Nombre comercial y código de referencia para los procesadores de sexta generación. En la columna dedicada a las cachés, hemos simbolizado con ✗ la ausencia de caché, con ✓ su presencia interna en un chip separado, y con ★ su integración dentro del propio chip de la CPU, en cuyo caso contabilizamos sus transistores en el montante total. En la columna dedicada al segmento, PC indica PC de sobremesa, PO indica PC portátil, y SE indica plataforma de tipo servidor.

## Intel

### Pentium Pro: El esqueleto de la sexta generación

#### ◀ 1.1

El Pentium Pro adopta el código de referencia P6 en sus albores, identificativo inequívoco de que se trata del diseño inaugural de la sexta generación para Intel. El éxito comercial de este procesador fue muy inferior al cosechado por el Pentium, y sin embargo, se trata de un diseño lleno de posibilidades, y que va a permitir a Intel prolongar la agonía de los microprocesadores basados en el conjunto de instrucciones 80x86 a lo largo de toda la sexta generación, con secuelas tan aclamadas popularmente como el Pentium II, el Celeron y el Pentium III. El influjo del diseño del Pentium Pro alcanza incluso al Pentium 4, una nueva arquitectura, pero inspirada en muchos de los conceptos que explicaremos en esta sección.

P6

secuelas

Pentium 4

En consecuencia, si estamos interesados en alguno de estos procesadores, la clave para entenderlos se encuentra aquí más que en ningún otro sitio, y por eso hemos querido ser particularmente generosos en su descripción.

la clave

El Pentium Pro no va a ser sólo un producto doméstico, ya que también constituye una clara referencia en el campo de los servidores hasta la llegada del Pentium II Xeon en la segunda mitad de 1998, producto que coge el relevo en Intel para este segmento del mercado, y que provoca que

servidores

el Pentium Pro con 256 Kbytes de caché L2 deje de fabricarse a finales del 98, y las versiones de 512 Kbytes y 1 Mbyte sigan similares derroteros pocos meses más tarde.

### 1.1.1 Tecnología de integración

**transistores**  
**micras**

Físicamente, el Pentium Pro presenta un aspecto algo más grande que el Pentium, con 5.5 millones de transistores frente a los 3.1 millones de éste último. Este mayor número de transistores se consigue en parte gracias a una rápida mejora en las técnicas de paralelismo a nivel de instrucción. Si el Pentium comienza a fabricarse a 0.8 micras, el Pentium Pro se estrena ya a 0.6, y muy pronto ambos alcanzarán las 0.35 micras. Eso sí, tras esta meteórica carrera, esta variable se va a conceder un período de hibernación que durará tres años, hasta que la necesidad del bus a 100 MHz le saca de su letargo, evolucionando entonces hasta las 0.25 micras del Pentium II Deschutes.

**chip L2**  
**pág. 153**

La mayor disponibilidad de transistores va a ser explotada en el diseño del procesador para desarrollar una agresiva apuesta por el paralelismo a nivel de instrucción, capacidad que se complementa estratégicamente con la presencia de un segundo nivel de caché en un chip aparte (aunque interno al patillaje del propio procesador). Este segundo chip albergará muchos más transistores que el procesador en sí según se refleja en la [tabla 5.2](#), y será responsable en gran medida del elevado coste del conjunto.

### 1.1.2 Frecuencia de reloj

**pág. 47**  
**150-200 MHz**

Si retomamos lo dicho en el [capítulo 3](#) con respecto al negativo impacto que el paralelismo a nivel de instrucción tiene sobre la frecuencia de reloj, la apuesta por el primero debe reflejarse como contrapartida en la segunda, pudiéndose esperar para este microprocesador una evolución casi plana con respecto a sus predecesores. Y así es como puede explicarse que aún siendo bastante más potente que el Pentium, lo viésemos a las mismas frecuencias que éste en el tramo final de su existencia: 150, 166, 180 y 200 MHz, siendo incluso superado por el Pentium MMX de 233 MHz y 266 MHz, que aunque es de una generación anterior, tuvo mayor longevidad.

De las cuatro frecuencias mostradas, 150 y 180 MHz se corresponden con las versiones que adoptan un bus local de 60 MHz, mientras que las otras dos lo tienen a 66 MHz, con el consiguiente aumento del rendimiento. No obstante, este parámetro es menos crítico aquí que en el Pentium debido a la presencia de la caché interna de segundo nivel, que reduce el uso del bus y por tanto la dependencia del exterior.

### 1.1.3 Paralelismo a nivel de instrucción

**supersegm.**  
**11 etapas**  
**pág. 154**

El Pentium Pro es un diseño supersegmentado con ejecución fuera de orden, lo que le lleva a escindir en **once** etapas las cinco originales del cauce segmentado para datos enteros del Pentium. La correlación entre uno y otro ha sido ilustrada en la [figura 5.1](#), donde puede apreciarse que las principales modificaciones se encuentran asociadas a la fase de búsqueda de instrucción, que ahora se prolonga a lo largo de tres etapas, y a la reordenación de instrucciones, que consume otras tres. Esto nos da claros indicios de dónde se complica este diseño.

**14 etapas FPU**

El cauce para la ejecución segmentada de operaciones de punto flotante sufre una transformación muy similar, de tal forma que si el Pentium disponía de ocho etapas, ahora serán **catorce** las que conformen este cauce para el Pentium Pro, esto es, siempre tres etapas más que su cauce entero como consecuencia de la complejidad adicional que sufre la etapa de ejecución en sí.

**superescalar**  
**pág. 61**

Con respecto a la superescalaridad, presenta un factor de **tres**, pudiendo ejecutarse hasta dos instrucciones sencillas y una compleja de forma simultánea, tal y como ya ilustramos en la [figura 3.6](#), aunque teniendo en cuenta ciertas normas y restricciones que ahora comentaremos.

Características	Microprocesador Pentium Pro			
	L2 256 Kb (1995)	L2 512Kb (1996)	L2 1Mb (1997)	
<b>Variables eléctricas:</b>				
Frecuencia de reloj (MHz)	150	180, 200	166, 200	200
Frecuencia bus local (MHz)	60	60, 66	66	66
Máx potencia disipada (W)	29	35	30	30
Número transistores (mill.)	5.5+15.5	5.5+15.5	5.5+31	5.5+62
Formato y dimensiones	306 mm <sup>2</sup>	Rectangular de 195 mm <sup>2</sup>		
Voltaje alimentación (voltios)	3.1	3.3		
Zócalo al que se acopla	Socket 8	Socket 9		
Distancia de integración	0.6	0.35		
Tecnología de integración	BiCMOS			
Empaquetado y envoltura externa	Dual Die Cavity de 387 pines			
<b>Segmentación:</b>				
Cauce aritmética entera	11 etapas			
Cauce punto flotante	14 etapas			
<b>Superescalaridad:</b>				
Unidades de ejecución	4 (2 enteras y 2 de punto flotante)			
Factor de superescalaridad	3			
<b>Memoria caché:</b>				
L1 Datos (integrada)	8 Kb, 128 ctos, 2 líns/cto, 32 bytes/lín			
L1 instrucciones (integrada)	8 Kb, 64 ctos, 4 líns/cto, 32 bytes/lín			
L2 unificada (interna)	De 4 líneas/conjunto y 32 bytes/línea			
Conjunto de instrucciones:	80x86 + 4 multiprocesador			

TABLA 5.2: El Pentium Pro en números. Desglose para las tres configuraciones de memoria caché L2 interna que se comercializaron.

El problema para entender todos los mecanismos del Pentium Pro radica en su dualidad CISC/RISC. Recordemos que el procesador acepta el código ejecutable del Pentium, y que también es compatible con la familia 80x86, lo cual quiere decir que su etapa de búsqueda actúa sobre instrucciones CISC, con un formato de instrucción y una duración en su ejecución completamente variables, según indicamos a nuestro paso por el Pentium. Pero a diferencia de éste, la fase de decodificación se encargará de dar un espectacular lavado de cara a este vetusto código ejecutable para transformarlo en otro más moderno con instrucciones de longitud fija de 118 bits y ejecución uniforme, que pueden ser consideradas como instrucciones RISC a todos los efectos.

decodificación

Conceptualmente, estas instrucciones RISC son las microinstrucciones que salen de la memoria de microprograma al final de la etapa de decodificación en una máquina CISC, sólo que aquí esto se va a producir únicamente cuando la instrucción original sea tan compleja como para requerir su conversión a más de cuatro microinstrucciones. Para cuatro o menos, existe un decodificador complejo que deriva el número de microinstrucciones consecutivas de forma directa a partir del código de operación, y para las instrucciones más sencillas que pueden transformarse en una única microinstrucción, se dispone de otros dos decodificadores más simples, tal y como queda reflejado en la [figura 5.2](#).

microprograma

☛ pág. 155

De las tres unidades de decodificación anteriores, las dos últimas operarán simultáneamente y de forma excluyente con la primera (la memoria de microprograma). Y precisamente para enganchar todo este engranaje con la fase de búsqueda necesitamos descomponer ésta en tres etapas: La primera (etiquetada como BUS1 en la [figura 5.1](#)) leerá un bloque o línea de 32 bytes de la caché de instrucciones, la segunda (BUS2) lo analizará para derivar los límites de separación de

BUS1

☛ pág. 154

BUS2

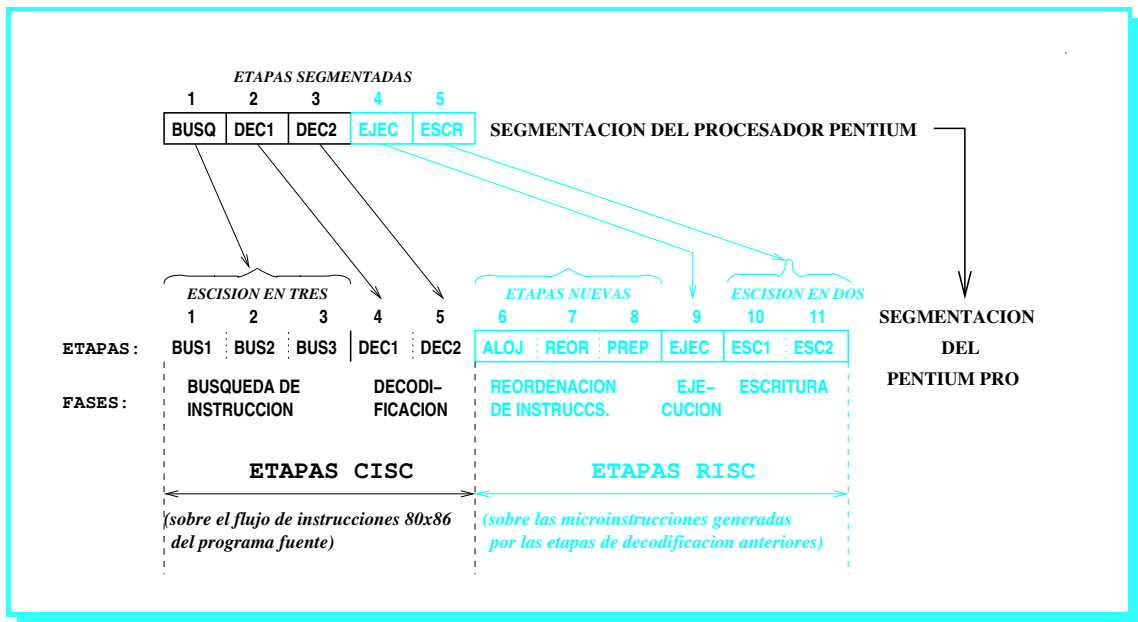


FIGURA 5.1: Derivación de las etapas de segmentación del microprocesador Pentium en las de su sucesor, el Pentium Pro. Este esquema segmentado será posteriormente reutilizado en los hermanos mayores de la familia (el Pentium II, el Celeron y el Pentium III).

BUS3 las instrucciones que lo componen, y la tercera (BUS3) tratará de alinear éstas para enviarlas a la decodificación agrupadas de las siguientes formas posibles:

- ❶ Dos simples y una compleja de las que se escinden en hasta cuatro microinstrucciones, y que derivan seis microinstrucciones que proseguirán su ejecución de forma agrupada.
- ❷ Tres simples (no se ha podido encontrar una compleja para que las acompañe).
- ❸ Una compleja o una compleja y una simple (no se han podido encontrar las simples que completen la terna).
- ❹ Una compleja de las que se transforman en más de cuatro microinstrucciones, pudiendo llegar hasta varios cientos de ellas, y que se fragmentará sucesivamente en grupos de seis microinstrucciones agrupadas.

Por lo tanto, hasta un máximo de **seis** microinstrucciones pueden ser generadas ahora en la fase DEC1 de forma simultánea desde la memoria de microprograma (casos 1 y 4), aunque en realidad éstas proceden de la decodificación de un máximo de tres instrucciones máquina (casos 1 y 2), que es lo que realmente establece el factor de superescalaridad del microprocesador. En general, cada instrucción ocupa por término medio 3.5 bytes, y genera a su vez 1.5 microinstrucciones, lo que representa 177 bits. Esto está bastante lejos de cualquier RISC del mercado en el que el formato de instrucción no excede de 32 bits, y es sin duda una de las grandes rémoras del diseño de la saga P6.

pág. 155

empaquetado

Nótese en la [figura 5.2](#) que la etapa DEC1 acepta un máximo de 16 bytes, por lo que en ese tamaño deberán caber las tres instrucciones 80x86 que se pretenden decodificar de forma simultánea. Como la fase BUS3 no tiene la facultad de reordenar las instrucciones para maximizar el número de casos en que esto ocurra, simplemente sacrificará instrucciones en el alineamiento hasta cumplir con el tamaño establecido. Por ello, de cara a maximizar el rendimiento del procesador, una receta casera sería tratar de escribir el código objeto por ternas de instrucciones que generen 4, 1

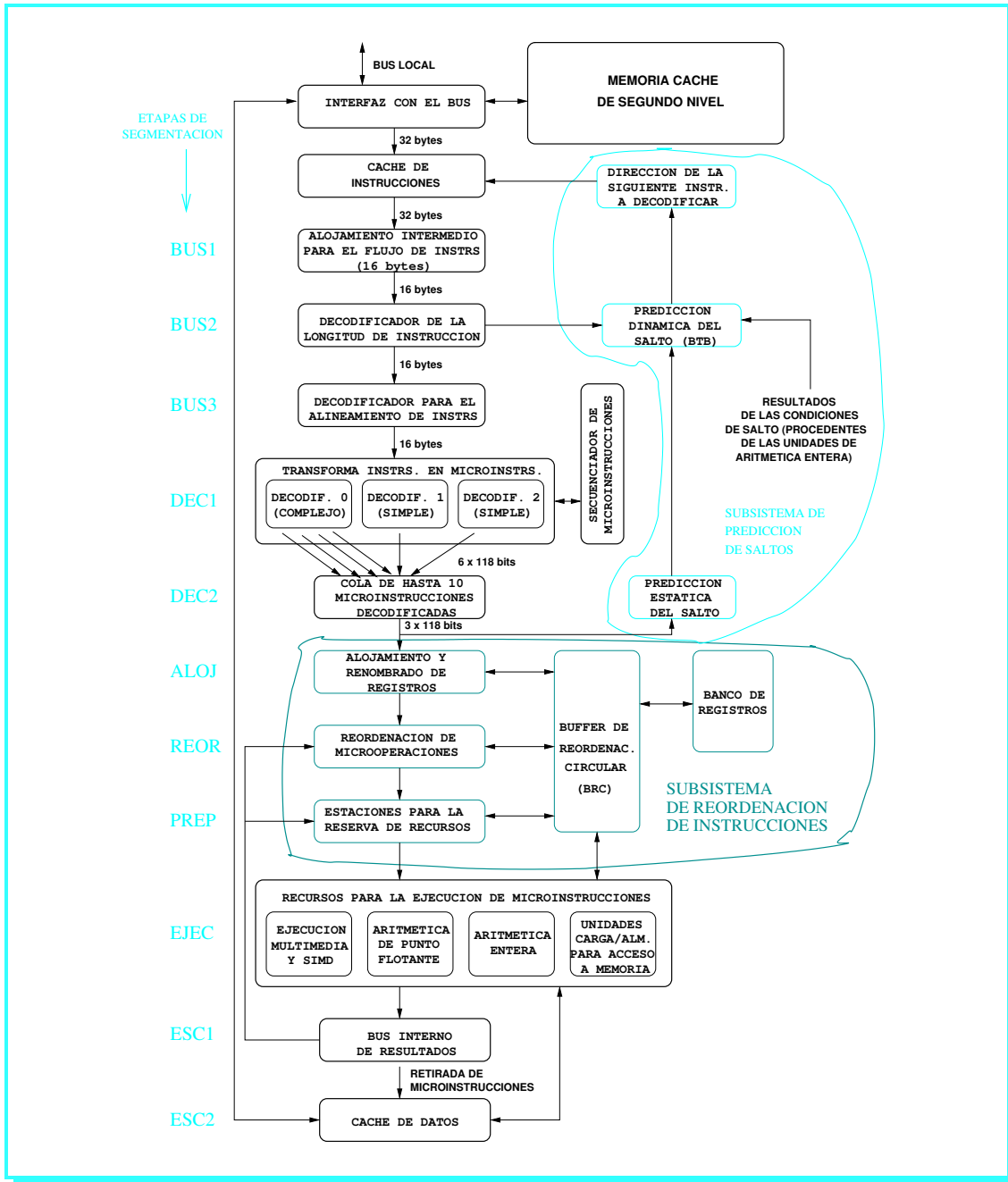


FIGURA 5.2: Diagrama general de funcionamiento del microprocesador Pentium Pro. La única diferencia con respecto al Celeron A y a los Pentium II y III se encuentra en la etapa de ejecución, que aunque siempre conserva el factor superescalar de tres, incorpora sucesivas mejoras que mostraremos separadamente para cada microprocesador. El búfer de reordenación y el resto de la circuitería para implementar la ejecución fuera de orden es idéntica en todos los casos, al igual que la correspondiente a las etapas de búsqueda y decodificación que transforman el código 80x86 en otro RISC de longitud uniforme que es el que terminará ejecutándose en el núcleo más interno del microprocesador.

y 1 microinstrucciones, respectivamente. Algunos compiladores sofisticados son también capaces de realizar estas optimizaciones por sí solos.

**DEC2** La segunda etapa de decodificación (DEC2) concatena los grupos de microinstrucciones obtenidos anteriormente componiendo una secuencia de longitud uniforme de 118 bits, bastante más larga que los 32 bits que suele tener un buen RISC, pero es el tributo a pagar por tomar como partida un código con la complejidad de operandos y los modos de direccionamiento inherentes al Pentium y a las secuelas del código 80x86 en general; la duración, de seis etapas más hasta su finalización, completa un cauce segmentado bastante profundo, lo que pone al diseño en una comprometida situación frente al impacto que sobre él tendrán las dependencias de datos y control que no puedan resolverse eficientemente.

Las buenas noticias son que a partir de este momento el código a ejecutar corresponde a un RISC puro, que progresivamente se irá enviando a una cola con capacidad para albergar hasta 10 microinstrucciones y que constituye el punto de partida para una ejecución más optimizada.

Continuando con la evolución de las instrucciones por el cauce segmentado, las tres etapas siguientes nos descubren las estrategias utilizadas para resolver las dependencias de datos, cimentadas sobre los pilares del renombrado de registros y la ejecución fuera de orden. Respecto a las dependencias de control, serán abordadas por separado una vez terminemos la descripción del cauce segmentado. La primera de estas tres etapas es **AL0J**, que acepta hasta un máximo de tres microinstrucciones por ciclo de reloj para realizar sobre ellas el renombrado de registros, técnica por la que se disponen varios bancos de registros virtuales que se terminarán alojando sobre el banco de registros físico, pero cuya separación provisional permitirá eliminar las dependencias de datos en que incurre el código objeto por la utilización de un número limitado de registros para el almacenamiento de resultados parciales.

Concluida esta operación, las microinstrucciones se envían a un búfer de reordenación circular (BRC) de 40 entradas. Cada entrada albergará una microinstrucción, manteniendo para ella los siguientes cuatro campos:

- ❶ **Dirección de memoria.** La dirección de la instrucción Pentium de la que procede esta microinstrucción.
- ❷ **Micro-op.** Descripción de la operación a realizar en la fase de ejecución (etapa EJEC).
- ❸ **Registro virtual.** Si la operación a realizar actúa sobre registros del banco, se incluyen aquí los registros virtuales sobre los que se han renombrado en la fase AL0J.
- ❹ **Estado.** Indica en qué etapa de la ejecución se encuentra la microinstrucción. Las posibilidades, en base a las etapas que restan en el cauce, son tres: PREP, EJEC y ESC.

El campo Estado es necesario porque aunque cada microinstrucción evolucione en el cauce desde el BRC a la unidad funcional en que es finalmente ejecutada, continuará registrada en el BRC hasta su conclusión, puesto que cada etapa restante usará y/o actualizará esta información para desarrollar su labor.

**RE0R** La etapa RE0R es la responsable de decidir, en base a la información del BRC, el mejor orden de ejecución de las microinstrucciones. Decidido esto, cada microinstrucción se encamina a la estación de reserva de la unidad funcional que requiere para su ejecución, el puerto o registro temporal en el que la microinstrucción permanecerá hasta que se den las condiciones necesarias para su ejecución. Nótese que las microinstrucciones alcanzan el BRC según el orden de prelación establecido en el programa fuente, y son las estaciones de reserva las que permiten el adelantamiento de unas microinstrucciones a otras en el cauce segmentado y la ejecución fuera de orden.

**PREP** La etapa PREP es la que se encarga de comprobar las condiciones para la ejecución de las instrucciones ubicadas en las estaciones de reserva. Estas condiciones son básicamente dos: La unidad funcional está libre (se han resuelto las dependencias estructurales que se producen por disponer de recursos hardware limitados), y los registros con los que quiere operar contienen los resultados esperados (se han resuelto las dependencias de datos presentes en el código objeto).

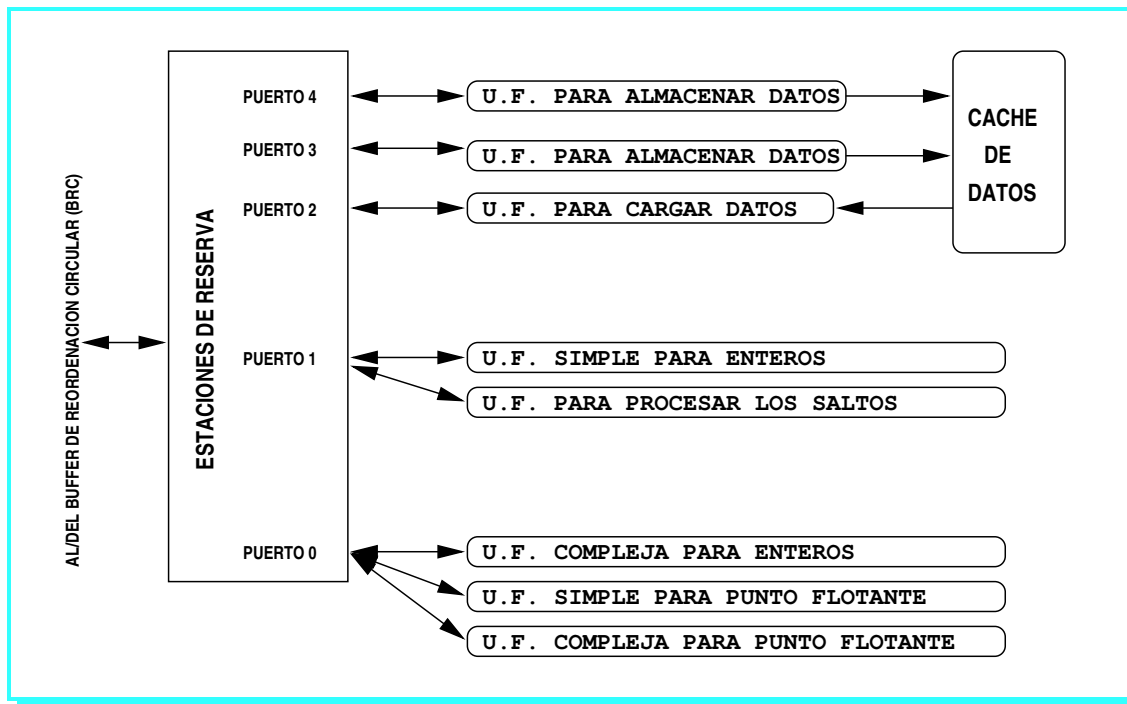


FIGURA 5.3: Detalle de las unidades funcionales de que dispone el Pentium Pro para la ejecución de instrucciones. Las de carga y almacenamiento de datos contienen principalmente la aritmética necesaria para la generación de la dirección absoluta en direccionamientos complejos.

La etapa EJEC recibe a todas las instrucciones que satisfacen las condiciones anteriores, y que en el mejor de los casos serán una por cada unidad funcional de que dispone la circuitería del microprocesador. En el caso del Pentium Pro, existen siete unidades funcionales, pero sólo cinco estaciones de reserva, ya que como se aprecia en la figura 5.3, existen tres unidades funcionales (FPU simple, FPU compleja y ALU compleja) que comparten una misma estación.

EJEC

pág. 157

Cabe preguntarse por qué estas tres unidades no se encuentran desacopladas con objeto de permitir un funcionamiento independiente y explotar un mayor grado de paralelismo a nivel de instrucción. La respuesta la tenemos en la etapa de decodificación, de donde arrastramos ciertas restricciones que debemos considerar: Si allí se decodificaban seis microinstrucciones por ciclo de reloj en el mejor de los casos, y posteriormente se renombraban los registros de un máximo de tres por ciclo de reloj (aunque algunas puedan no referenciar a registro alguno), no resulta lógico pensar que aquí nos van a llegar siete microinstrucciones. Una media realista es que traspasen la decodificación una media de cinco, y por eso es éste el mejor número a considerar para la gestión de un control simultáneo no exento de complejidad <sup>1</sup>.

restricciones

Completada ya la ejecución de una microinstrucción, la unidad funcional queda libre para ejecutar otra microinstrucción, y la etapa EJEC concluye actualizando la entrada del BRC, donde quedarán reflejados los resultados obtenidos, pendientes de ser trasladados al banco de registros.

Esto es precisamente lo que se hace en las dos últimas etapas, ESC1 y ESC2 . La segunda de ellas vuelca la información al banco de registros y borra la entrada ocupada por la microinstrucción en el BRC, mientras que la primera comprueba la licitez de esta operación. Recordemos que estamos en un procesador con (1) ejecución fuera de orden y (2) predicción de salto, por lo que antes de escribir el resultado de una operación debemos asegurarnos de que su microinstrucción no pertenece a una instrucción que (1) escriba sus resultados antes que cualquier microinstruc-

ESC1  
ESC2

<sup>1</sup>Suele ocurrir en no pocas ocasiones - y ésta es una de las más claras - que el problema no es ejecutar una decena de cosas a la vez, sino controlar todas ellas de forma simultánea.

Valor actual del contador	Predicción a efectuar	Nuevo valor del contador si el salto:	
		(a) Se realiza	(b) No se realiza
00	No saltar	01	00
01	No saltar	10	00
10	Saltar	11	01
11	Saltar	11	10

TABLA 5.3: Valor que pueden tomar cada uno de los 16 contadores de dos bits que se habilitan por cada entrada de la BTB y su relación con la predicción de salto. Nótese que los contadores se actualizan sumando o restando una unidad en función de que el salto termine o no efectuándose, realizándose estas operaciones con saturación, esto es,  $3+1=3$  y  $0-1=0$ .

ción perteneciente a una instrucción que le antecede en el código objeto, y (2) corresponda a una secuencia de instrucciones introducida en el cauce gracias a una predicción de salto errónea. Sólo así garantizaremos que (1) las dependencias de datos y (2) las dependencias de control han tenido un tratamiento correcto desde el punto de vista de la semántica del código objeto que aceptó como entrada el microprocesador.

Concluida la descripción de la segmentación y superescalaridad del Pentium Pro, abordaremos ahora los mecanismos de que dispone el procesador para la predicción de salto, y que de nuevo son un fiel calco de lo que se incorporó con posterioridad a diseños más contemporáneos como los Pentium II y III.

## 1.2 ► Predicción de salto

BTB Al igual que el Pentium, la versión Pro utiliza un esquema de predicción dinámica de la condición de salto basado en una BTB, que en este caso es de 512 entradas organizadas en 128 conjuntos de 4 entradas cada uno. El concepto de conjunto nace ligado a la organización de la memoria caché, para posteriormente trasladarse a unidades funcionales de más reciente incorporación (la BTB es una de ellas; la TLB, otra). En todos los casos, condiciona la forma de buscar la información en ella tomando siempre como entrada una dirección de memoria que, en el caso que nos ocupa, será la de la instrucción de salto a localizar en la BTB. Los bits 2 al 8 de esta dirección seleccionan el conjunto en el que cotejar una por una sus 4 entradas para ver si los bits restantes de la dirección que almacenan coinciden con la dirección buscada:

- En caso afirmativo, los bits que almacenan el comportamiento histórico determinarán la predicción a realizar. Frente al Pentium, se sustituye el algoritmo de un único nivel con dos bits de predicción por otro más complejo de dos niveles: el primero, de cuatro bits de predicción, que almacena el comportamiento de cada instrucción de salto las últimas cuatro veces que se ejecutó, y el segundo, de 16 contadores de dos bits (un contador por cada combinación de los cuatro bits anteriores), que se incrementan cuando el salto tiene lugar y se decrecientan cuando no, y cuyo valor determina la predicción a realizar según se observa en la [tabla 5.3](#).

36 bits Todos los detalles sobre este algoritmo se encuentran en una publicación de sus autores, Yeh y Patt, que referenciamos en la bibliografía correspondiente al presente capítulo. En total, se utilizan 36 bits por cada entrada de la BTB ( $4 + 2 \cdot 16$ ), pero mientras mejore notablemente el acierto de las predicciones, su coste estará justificado. Piénsese que estamos en un diseño con cauces de segmentación mucho más profundos que los del Pentium normal, por lo que la penalización en caso de fallo es bastante mayor: Si en el Pentium es de tres ciclos en el mejor de los casos, aquí va a ser de un mínimo de seis.

- En caso negativo, la instrucción de salto se incorporará a la BTB sustituyendo a la que se encuentre más en desuso dentro del conjunto, y el microprocesador conmutará a predicción



estática que en lugar de ser fija como en el Pentium es algo más inteligente: Se predicen como saltos realizados todos los incondicionales y aquellos condicionales en los que la dirección de destino sea inferior a la de la instrucción de salto, tratando con ello de captar el movimiento de los numerosos bucles o lazos de que consta un programa. El hecho de que la predicción estática actúe tras la dinámica le hacen partir de un retardo de seis ciclos, pero si a partir de aquí la predicción estática falla, la penalización total puede incluso sobrepasar los doce ciclos.

La predicción dinámica se lleva a cabo en la etapa de segmentación BUS2, mientras que la estática lo hace en la etapa DEC2, siendo esta ubicación más avanzada en el cauce la que explica su mayor penalidad en caso de fallo. Pero no queda más remedio que colocar la predicción estática más adelante, ya que según acabamos de ver, la dirección de salto es tenida en cuenta para establecer la predicción, y si esta dirección viene dada en función de un modo de direccionamiento complicado, es necesario esperar unos ciclos hasta que se ha calculado la dirección final.

dinámica  
estática

La [figura 5.2](#) muestra la relación entre la predicción estática y dinámica y el funcionamiento general del Pentium Pro. Si hasta el momento las etapas BUS2 y DEC2 parecían casi nimias, la complejidad del mecanismo de predicción con que se sobrecarga cada una de ellas justifica sobradamente su presencia en el cauce segmentado del microprocesador.

☛ pág. 155

### 1.2.1 Memoria caché

Ya anticipamos que el Pentium Pro incorpora como novedad un segundo nivel de caché (L2). El primero consta de dos cachés de 8 Kbytes, una para datos y otra para instrucciones.

Se sigue manteniendo un tamaño de línea de 32 bytes, pero estas cachés ya no son gemelas: la de datos mantiene el nivel de asociatividad en dos líneas por conjunto, pero la de instrucciones aumenta este nivel hasta cuatro. Esta modificación es una consecuencia lógica del mayor paralelismo a nivel de instrucción del Pentium Pro: Si el procesador tiene capacidad para procesar más instrucciones por ciclo de reloj, debemos mejorar las unidades funcionales responsables de suministrarle las instrucciones para que lo hagan al ritmo que su ejecución exige. Lo que posibilita el aumento del número de líneas por conjunto en una caché es una notable reducción del índice de fallos, siendo esta relación causa-efecto casi lineal cuando los conjuntos son de pocas líneas. Es decir, se produce el mismo efecto que doblando su capacidad.

asociatividad  
caché L1

La caché L2 se integra en un segundo chip dentro del encapsulado conjunto, y se compone de 15.5 millones de transistores para la versión de 256 Kbytes, originalmente a 0.6 micras, y 31 y 62 millones para las de 512 Kbytes y 1 Mbyte, respectivamente, ambas ya a 0.35 micras. Esta caché funciona a la velocidad del microprocesador, con rendimiento y coste elevados. Aún así, la tardanza de Intel en desarrollar el Pentium II Xeon como recambio del Pentium Pro, le obligó a comercializar fugazmente la versión con 1 Mbyte de L2 para los clientes más exigentes.

caché L2

La [figura 5.4](#) muestra la relación entre las cachés del Pentium Pro y las unidades funcionales más relacionadas con su paralelismo a nivel de instrucción.

☛ pág. 160

### 1.2.2 El conjunto de instrucciones

El conjunto de instrucciones máquina aceptado por el Pentium Pro es en esencia el 80x86 del Pentium, al que se le agregan 4 nuevas instrucciones dedicadas a la comunicación interprocesador que debe establecerse en los sistemas multiprocesador orientados normalmente al campo de los servidores.

80x86

Aunque los servidores Pentium multiprocesador son factibles, la funcionalidad mencionada no se encuentra en los microprocesadores, sino que es incorporada por la placa base multizócalo

multizócalo

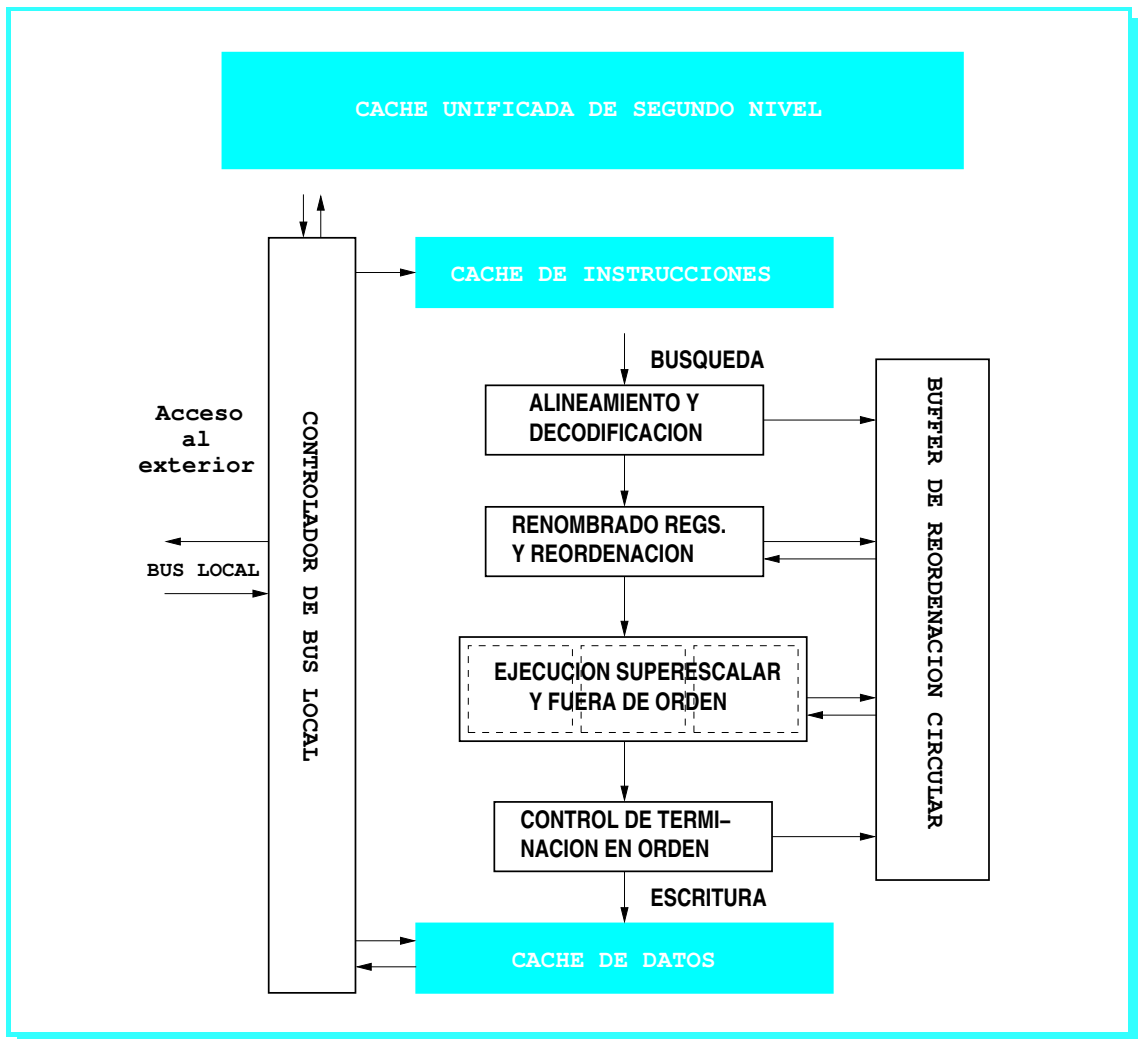


FIGURA 5.4: Relación entre las cachés de primer y segundo nivel del Pentium Pro y las unidades de ejecución y reordenación de instrucciones a través de las etapas que componen su cauce segmentado. El diagrama es válido también para otros procesadores que comparten este mismo esqueleto, como el Pentium II, el Celeron A y el Pentium III.

en la que se montan. El Pentium Pro, en cambio, crea diseños multiprocesador a mayor escala y autosuficientes, ya que soporta multiprocesamiento simétrico independiente de cuatro vías (los Pentium sólo pueden ser montados por parejas) en los propios microprocesadores, con lo que el servidor Pentium Pro se monta simplemente estableciendo el conexionado apropiado para los pines de las distintas CPUs. En la práctica, no obstante, es un proceso mucho más complejo, porque hay que equilibrar correctamente las cargas eléctricas y proporcionar el abundante flujo de aire necesario para mantener frescos a todos los procesadores.

Desde el punto de vista del rendimiento, el enfoque multiprocesador del Pentium Pro dice mucho en su favor. Las cachés individuales del procesador son más eficaces que el diseño de caché externa compartida en la placa que se emplea en las configuraciones dual-Pentium.

Además, gracias al bus orientado a transacciones de las placas base fabricadas por Intel para las plataformas multi-Pentium-Pro, la agrupación de varios servidores se soporta con bastante facilidad, lo que dejó la puerta abierta para la creación de redes de servidores multiprocesador para aplicaciones avanzadas de alto nivel (simplemente utilizando enlaces de datos a alta velocidad

como ATM para interconectarlos).

Estas dos características, unidas al elevado coste del Pentium Pro, llevaron a Intel a comercializar este chip como una alternativa de gama alta, en lugar de como un sucesor del Pentium clásico, que quedó instalado cómodamente al nivel más doméstico (PC).

La lástima es que Intel fallase en su previsión acerca de los derroteros que iba a tomar la capa software. Se apostó desde el principio por las aplicaciones de 32 bits, optimizando mucho más el procesador en el tratamiento de los conflictos en que incurrían los registros extendidos de 32 bits (EAX, EBX, ...) y descuidando un tanto los ocasionados por los registros normales de 16 bits (AX, BX, ...) o incluso los parciales de 8 bits (AH, AL, ...), que por aquella fecha continuaban de plena vigencia. Pero sin un Sistema Operativo de 32 bits que hiciese de puente entre éstas y el hardware, lo que llegaba a los aledaños del Pentium Pro era código de 16 bits, resultando una ejecución más lenta incluso que la del propio Pentium. Intel afirmó entonces que consideró la alteración del diseño del procesador para permitir un mejor tratamiento del código de 16 bits, pero obtuvo como conclusión que realizar los cambios no resultaba práctico o acaso excesivamente caro. Procesadores de empresas competidoras como AMD o Cyrix no parecieron encontrar esas mismas limitaciones, y supieron entenderse mucho mejor con una capa software de 16 bits, conservando un mejor talante de compatibilidad.

En realidad, la jugada de Intel era evolucionar hacia las aplicaciones y los sistemas operativos de 32 bits adelantándose a lo que ellos pensaban que sería el camino natural en la evolución de los PC. De esta forma, esperarían al mercado de PC en el hito de los 32 bits, y cuando su capa software avanzara hasta allí, ellos ya tendrían un microprocesador consagrado que ofrecer y con el que batir en toda regla a la competencia. Pero Windows 95 nunca fue el Sistema Operativo de 32 bits que de él se esperaba, y tampoco Windows 98 satisfizo demasiado. Tan sólo Windows NT o Unix, sabiendo conjugarlos con unos cuantos paquetes de ofimática, ingeniería y gráficos de lo más avanzado del mercado, hicieron las delicias del Pentium Pro.

## Pentium II: Un nuevo formato

### ◀ 1.3

A la hora de hablar del Pentium II cabe distinguir sus dos versiones: La inaugural, cuyo código de referencia es Klamath, fabricada a 0.35 micras en 1997, y la terminal, con código de referencia Deschutes, y fabricada un año más tarde a 0.25 micras. Continuando con nuestro barrido cronológico, trataremos el Klamath ahora, y sobre él, cubriremos el Deschutes ([sección 5.1.4](#)).

El Klamath irrumpe en el mercado en Junio de 1997, e introduce pocos cambios sobre lo que es el Pentium Pro en sí. En realidad, Intel ya había decidido tiempo atrás que la evolución de sus modelos se desarrollaría sobre la base de la arquitectura P6, aparcando la P5 del Pentium original.

### 1.3.1 Frecuencia

El primer Pentium II trabaja a 233 MHz, tomando el relevo de los Pentium y Pentium Pro de 200 MHz. La frecuencia progresa a una cadencia de 33 MHz, signo inequívoco de que el procesador continúa instalado en un bus local de 66 MHz. Las tres configuraciones del Klamath son 66 x 3.5 (233MHz), 66 x 4 (266 MHz) y 66 x 4.5 (300MHz).

### 1.3.2 Integración y zócalo

La tecnología de integración del Pentium II Klamath fue CMOS de 0.35 micras, aboliendo definitivamente BiCMOS, y su voltaje de alimentación, 2.8 voltios. A estas alturas, las 0.35 micras están ya en su período final, y el hecho de que sólo se dispusiese de tres frecuencias de trabajo así lo corrobora.

gama alta

32 bits

penalidad

2 versiones

Klamath

Deschutes

← pág. 164

P6, no P5

233-300 MHz

0.35 micras



FOTO 5.1: El microprocesador Pentium II envuelto en su cartucho cerámico SECC y con su disipador de calor y ventilador acoplado por la parte superior.

El Pentium II rompe estéticamente con todo lo que Intel había hecho hasta el momento. Su forma física es la de un prisma rectangular, el contorno de una placa de circuito impreso donde el chip que alberga al procesador en sí se ubica en su parte central y los chips en los que se implementa la caché L2 interna se colocan en uno de los laterales. El zócalo recibió el nombre de Slot1.

La foto 5.1 muestra el aspecto de este procesador, donde nada de lo descrito puede apreciarse por la presencia del disipador de calor y su ventilador que vienen ya acoplados al conjunto (en la foto en la parte superior), y porque el conjunto del procesador se envuelve en un cartucho cerámico como solución térmica y apantallado frente al ruido eléctrico. Es como si el procesador fuese una caja de cerillas y se conectase a la placa por uno de sus rascadores.

Según Intel, el cartucho cerámico reducía costes y conseguía aumentar las posibilidades de un diseño que ya no permitía progresos. Si bien es cierto que consiguieron hacer un producto más rápido, no lo es menos que tres años más tarde no había ni un solo procesador fabricándose con esta estrategia, a pesar de haber desfilado por ella también AMD con su Slot A para K7. Y es que siempre es más complicado electrónicamente diseñar dispositivos que funcionen a frecuencias elevadas en disposición perpendicular a la de la placa base, ya que sus pistas tienden a comportarse como antenas que irradian y reciben radiaciones. y que son la fuente de multitud de interferencias y ruidos.

### 1.3.3 Paralelismo a nivel de instrucción

El Pentium II tiene esencialmente la misma organización segmentada y superescalar que ya vimos para el Pentium Pro, con la única salvedad de las unidades MMX, que son incorporadas aquí a la fase de ejecución de instrucciones como ilustra la figura 5.5.

pág. 164 ➔

### 1.3.4 Memoria caché

Al igual que el Pentium Pro, el Pentium II también incorpora dos niveles de memoria caché, el primero integrado y separado en una parte para datos y otra para instrucciones, y el segundo

Características	Microprocesador Pentium II		
	Klamath (1997)	Deschutes (1998)	
Variables eléctricas:			
Frecuencia de reloj (MHz)	233, 266, 300	333	350, 400, 450
Frecuencia bus local (MHz)	66	66	100
Número transistores (mill.)	7.5M procesador + 15M caché = 22.5M		
Potencia disipada(W)	34.8, 38.2, 43.0	23.7	21.5, 24.3, 27.1
Voltaje del núcleo (voltios)	2.8	2.0	2.0
Voltaje de la caché interna	3.3		
Voltaje de terminación del bus	1.5		
Zócalo al que se acopla	Slot 1		
Distancia de integración	0.35	0.25	
Tecnología de integración	CMOS		
Encapsulado	SECC	SECC	SECC/SECC2
Segmentación:			
Cauce aritmética entera	11 etapas		
Cauce punto flotante y MMX	14 etapas		
Superescalaridad:			
Unidades de ejecución	7 (2 enteras, 2 punto flot, 3 MMX)		
Factor de superescalaridad	3		
Memoria caché:			
L1 Datos (integrada)	16 Kb, 128 ctos, 4 líns/cto, 32 bytes/lín		
L1 instrucciones (integrada)	16 Kb, 128 ctos, 4 líns/cto, 32 bytes/lín		
L2 unificada (interna)	512 Kb, 4 líneas/conjunto, 32 bytes/línea		
Conjunto de instrucciones:	Compatible 80x86 + 57 MMX		

TABLA 5.4: Resumen de los principales parámetros del Pentium II en sus dos versiones: Klamath, de 0.35 micras, y Deschutes, de 0.25 micras. La diferencia entre el formato SECC y el SECC2 es la cobertura completa o parcial del cartucho cerámico; en el primer caso, el disipador de calor se aplica al conjunto, y en el segundo, de forma directa e individual sobre el chip procesador.

interno y unificado. Las mejoras no vienen aquí por el tipo de caché, sino por sus respectivos tamaños.

La caché de primer nivel (L1) integrada, que en los Pentium y Pentium Pro es de 8 Kbytes para datos y 8 Kbytes para instrucciones, en los Pentium MMX y en todos los modelos de Pentium II dobla su capacidad, esto es, 16 Kbytes para datos y 16 Kbytes para instrucciones (ver [tabla 5.7](#)).

Con respecto a la caché de segundo nivel (L2) interna, en los Pentium II se sitúa en los 512 Kbytes, frente a los 256 Kbytes de la versión base del Pentium Pro (y la carencia de este elemento tanto en el Pentium clásico como en el MMX).

No obstante, el acceso a la caché L2 del Pentium II es algo más lento que el que proporciona el Pentium Pro. Esto es así porque la L2 del Pentium II se sitúa dentro de su cartucho cerámico pero en otro chip más lejano al microprocesador en sí. En realidad, la solución adoptada en la implementación de la caché L2 del Pentium II es un compromiso de rendimiento y coste frente a la L2 del Pentium Pro, sincronizada con el procesador, pero con un coste excesivo (cada 256 Kb. de caché del Pentium Pro encarecía el procesador en unos 600€).

L1

➔ [pág. 170](#)

L2 interna

coste

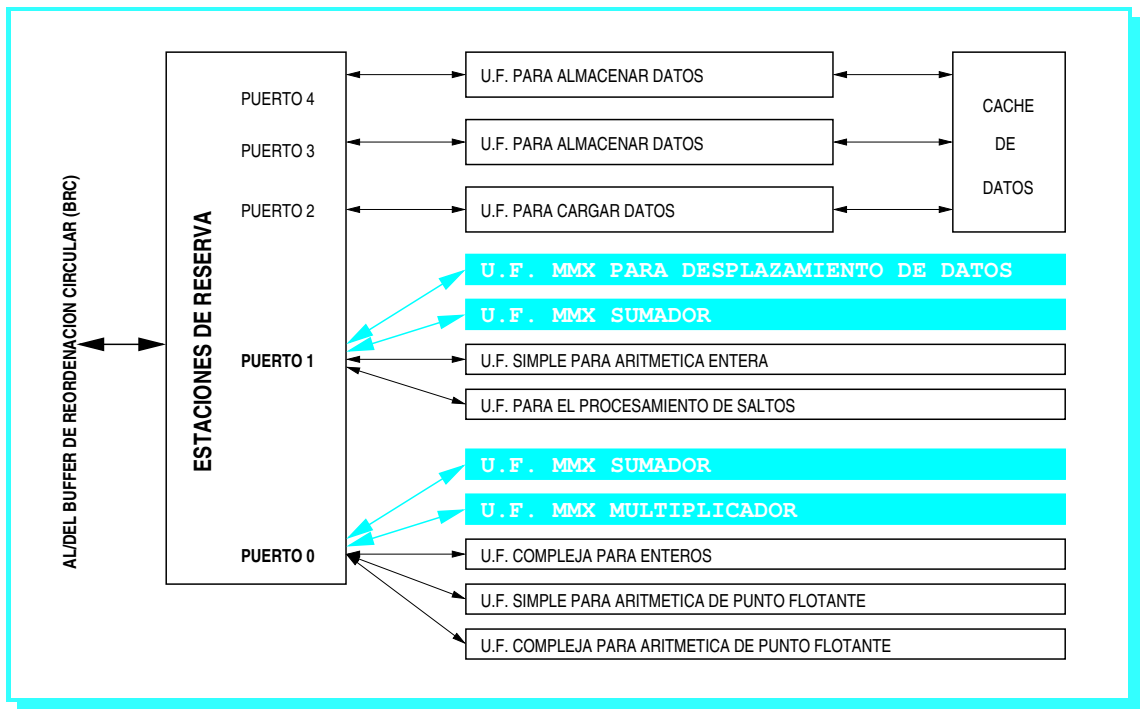


FIGURA 5.5: Detalle de las unidades funcionales de que dispone el Pentium II para la ejecución de instrucciones. Las mostradas en el segundo color constituyen el añadido con respecto a la arquitectura del Pentium Pro.

### 1.3.5 Hardware subyacente

Algunas mejoras introducidas por el Pentium II no se deben al microprocesador en sí, sino a la flota de placas base sobre las que se monta. En concreto, después de un fugaz paso por el juego de chips 440FX de Intel reutilizado del Pentium Pro, se dota al sistema de juego de chips propio: El 440LX. En él cabe destacar la arquitectura Dual Independent Bus (DIB), que desacopla los buses trasero para el tráfico con caché L2 y frontal para los diálogos con memoria y el resto de dispositivos externos. Esto ya lo disfrutó el Pentium Pro. Lo que no disfrutó fue el nuevo bus AGP para el tráfico de datos gráficos, la posibilidad de conectar memoria principal SDRAM y el nuevo formato ATX de la placa, que mejora la disposición física de los componentes haciendo más fácil su montaje y ventilación.

## 1.4 ▶ Deschutes: Pentium II Turbo

0.25 micras  
origen

La llegada del proceso de fabricación de 0.25 micras lleva a Intel a integrar un nuevo modelo de Pentium II, que recibe el código de referencia Deschutes (nombre del río que atraviesa la vega de Santa Clara, el área californiana en la que está emplazada la planta de fabricación de chips más emblemática de Intel).

frecuencia

bus 100 MHz

El paso de 0.35 micras a 0.25 micras es aprovechado por Intel en el Deschutes para mejorar tan sólo un aspecto: Su frecuencia. Salió al mercado en los primeros meses de 1998, tomando el relevo del Klamath y llevando la frecuencia hasta los 333 MHz, aún con el bus local situado en los 66 MHz, para posteriormente alcanzar 350, 400 y 450 MHz, ya con bus a 100 MHz.

zócalo

El zócalo del procesador es el mismo que el del Klamath, el Slot 1, por lo que no hubo problemas para reutilizar las mismas placas base que ya existían para él. Eso sí, esto sólo era aconsejable en el único modelo bajo bus a 66 MHz, ya que en el resto se desaprovechaba el bus de 100 MHz

Período	Pentium		Pentium II		
	MMX	Pro	Original	Xeon	Celeron
Jul-Sep'97	133-200 L2 ext	150-200 256,512Kb	233,266 512Kb		
Oct-Dic'97	233-266 L2 ext	200 1Mb	300 512Kb		
Ene-Jun'98			333,350 512Kb	350,400 512Kb,1Mb	266,300 0Kb
Jul-Dic'98			400,450 512Kb	450 512Kb,1Mb	300-333 128Kb
Ene-Jun'99				450 2Mb	366-466 128Kb
Jul-Dic'99					500-600 128Kb
2000					633-766 128Kb
2001					800 128Kb

**TABLA 5.5:** Evolución temporal de la saga del procesador Pentium y Pentium II con respecto a la frecuencia (MHz: Megahercios) y al tamaño de la caché L2 interna (Kb:Kilobytes ó Mb:Megabytes). El Pentium clásico dejó de fabricarse con la salida del MMX, y éste lo hizo a su vez poco después de la salida del Celeron, siendo las versiones más rápidas las últimas en extinguirse (233-266 MHz). La columna Original representa al Klamath hasta la versión de 300 MHz incluida, y al Deschutes a partir de ahí. La columna Celeron refleja la versión Covington de 0.35 micras hasta los 300 MHz, la versión Mendocino de 0.25 micras hasta los 533 MHz, y la versión Coppermine de 0.18 micras desde ahí en adelante, momento en que pasó a ser una versión menor obtenida a partir de la fabricación del Pentium III en lugar del Pentium II.

y con ello casi todas sus mejoras. En placa base, el primer juego de chips con bus a 100 MHz es el 440BX, posteriormente reutilizado también por los modelos de Pentium III. Más adelante, el juego de chips 440ZX también sería compatible para estos dos modelos.

Externamente, el Klamath y el Deschutes no presentan diferencia alguna, y sólo pueden distinguirse echando un vistazo a su número de serie y decodificándolo de manera oportuna (el manual *Pentium II Processor Specification Update* editado por Intel y accesible a través de su página Web refleja esta información).

El cartucho cerámico se desnudaría un poco más tarde por uno de sus laterales para mejorar la disipación de calor, dando lugar a la envoltura SECC2, pero esto no es un rasgo distintivo del Deschutes, pues hay procesadores Katmai bajo SECC2 y Deschutes bajo SECC (esos casos se corresponden con las versiones de frecuencia más baja de cada modelo).

SECC2

## Pentium II Xeon: Para servidores

## 1.5

Poco después de su salida al mercado, se diseñó una variante del Pentium II de alto rendimiento y coste elevado dirigida al segmento de los servidores. Se denominó comercialmente Pentium II Xeon, y vio la luz en los primeros meses de 1998, esperando para arrancar ya a 0.25 micras.

1998

Este microprocesador resulta difícil de encontrar en las arquitecturas domésticas, dada la escasa presencia de sistemas servidores a este nivel que justifiquen un desembolso de dinero tan generoso. Se trata por tanto de una especie poco frecuente en el hábitat natural en el que se mue-

ve la gente de la calle, como en su día le ocurrió al Pentium Pro y en el futuro le ocurrirá a otros modelos de la gama Xeon.

**frecuencias** El rango de frecuencias en el que se movió este procesador se sitúa en la parte alta de la horquilla asociada a los Pentium II más avanzados, esto es, 350, 400 y 450MHz, habiendo sido éste el primer procesador de Intel en alcanzar cada una de ellas. El bus local pasa a una frecuencia de 100 MHz, con lo que los multiplicadores coinciden con los mismos tres valores del Pentium II Klamath: 3.5, 4 y 4.5.

**fabricación** Intel aprovechó el logro de las 0.25 micras en el Xeon para retomar la fabricación de cachés más grandes y rápidas, algo que ha constituido siempre el sello distintivo de las configuraciones de la compañía en el segmento de servidores.

**PNI** El núcleo de ejecución de este procesador es el mismo que el de la versiones Klamath y Deschutes, con lo que la arquitectura no altera los valores de segmentación y superescalaridad ya conocidos, por lo que seguimos amortizando el haber desarrollado extensivamente esta arquitectura cuando pasamos por el primer modelo de la saga, el Pentium Pro.

**caché** La primera versión del Xeon sólo disponía de 512 Kbytes de caché L2, pero ésta se encontraba mejor integrada en su chip, pudiéndose emplear una nueva tecnología, CSRAM, más rápida y con menos circuitos que la BSRAM utilizada en el Pentium II convencional. El voltaje fue de 2.5 voltios.

**coste** La rapidez de la CSRAM permitió a la caché sincronizarse con el procesador a frecuencias de hasta 450 MHz, mientras que su menor complejidad puso al alcance de la mano las versiones de 1 Mbyte de caché L2 (logro conseguido a finales de 1998) y de 2 Mbytes pocos meses después (ver [tabla 5.5](#)). Eso sí, el precio de lanzamiento de la versión de 1 Mbyte de caché L2 fue de unos 2.000€ euros al cambio, y el de 2 Mbytes, 3.700€.

**Slot 2** La reducción en el voltaje llevó a Intel a cambiar el zócalo de las placas base, que para este procesador se denominó Slot 2. Frente al zócalo SECC Slot 1 que comparten el Pentium II y algunos de sus sucesores (Deschutes y Celeron), el nuevo zócalo tiene más contactos, y éstos se encuentran dispuestos en tres filas en lugar de dos, aunque ocupando un espacio físico similar.

**450NX** Así, mientras que las placas base con zócalo Slot 1 utilizan el chip controlador 440BX, el Xeon Slot 2 se instala sobre la placa 450NX, una adaptación de las placas base que Intel fabricó para los Pentium Pro que permiten direccionar hasta 4 Gbytes de memoria principal y montar configuraciones de hasta cuatro CPU. Sin embargo, las placas 450NX no disponen de soporte para AGP (ver [sección 16.1](#)). Esto se debe al diseño antiguo de la placa de la que proceden y a que en el segmento de los servidores la potencia de proceso es mucho más prioritaria que la capacidad gráfica de la máquina.

## 1.6 ► Celeron: Un paso en falso

**Covington** A mediados de 1998 salió al mercado el Celeron, un Pentium II sin memoria caché L2, y con un coste inferior en un 70% a la versión completa del procesador. El primer Celeron, cuyo código de referencia es Covington, es mucho más lento que cualquier Pentium II. A quién se le ocurre prescindir de la caché L2 cuando la arquitectura muestra un claro cuello de botella en el bus local a 66 MHz.

**tres gamas** La estrategia era muy clara: La versión Socket 7 del zócalo de los Pentium tradicionales y MMX sería suplantada por el Celeron, y las versiones Socket 8 y 9 del Pentium Pro, por el Pentium II Xeon. Intel conforma de esta manera tres líneas o gamas de productos: Una de corte doméstico preocupada mayoritariamente por adquirir un computador de bajo coste (o low end), otra de gama media más preocupada por obtener un rendimiento aceptable (o mid end), y una tercera de gama alta orientada a estaciones de trabajo y servidores (o high end).



La arquitectura estrella camina por el segmento intermedio, que es la que presenta un mayor volumen de ventas, y sobre ésta, se sacrifican ciertos elementos para abaratar costes y generar la de gama baja (Celeron), y se incorporan aditivos para producir la de gama alta (Xeon). Así, con una sola arquitectura puede cubrirse todo el espectro de usuarios y los gustos más variopintos. La idea es todo un logro empresarial para maximizar la productividad, y unos años más tarde no sólo sigue de plena vigencia, sino que ha sido apadrinada por otras muchas compañías del sector (en AMD, por ejemplo, su Celeron es el Duron, y su Xeon los nuevos diseños SMP del K7 y el K8).

A pesar de su gran carencia, el Celeron tiene un aspecto estupendo para los usuarios neófitos de informática, que suelen fijarse principalmente en dos variables cuando compran un microprocesador: Frecuencia y coste. Así que se coge un modelo ya consolidado como el Pentium II, se le quita la generosa caché L2 que es lo que más transistores necesita y en la que pocos reparan, y nos queda una configuración de elevada frecuencia y bajo coste.

frecuencia  
y coste

Menos mal que vivimos en la sociedad de la información. Pronto el Celeron quedaría en evidencia al coleccionar pírricas estadísticas sobre todo tipo de programas de evaluación del rendimiento. Y es que, claro está, frente a un Pentium II no tenía nada que hacer, y frente a un MMX, llegó a rozar el ridículo. La excelente relación calidad/precio de los productos que AMD y Cyrix diseñaron mientras tanto hizo el resto, y los grandes beneficiados de todo esto fueron los usuarios que permanecieron fieles al Socket 7 de la competencia (compradores del K6 o K6-2, por poner dos buenos ejemplos).

ridículo

Socket 7

## CeleronA: La rectificación

◀ 1.7

El proceso de fabricación de 0.25 micras acudió rápidamente al rescate de Intel. Apenas seis meses más tarde del estreno del Celeron, en Septiembre de 1998, se puso en marcha una nueva versión, con código de referencia Mendocino, y que nosotros distinguiremos como CeleronA.

Mendocino

La única novedad arquitectural es la inclusión de una caché L2 integrada de 128 Kbytes, que por tanto funciona ya a la misma velocidad del procesador. Externamente, el procesador migra de cartucho tipo SECC a otro denominado SEPP (Single Edge Pin Package) que ya deja al descubierto sus dos laterales. La [sección 34.4](#) describe las propiedades de todos estos cartuchos e incluye fotografías en las que puede apreciarse el aspecto de los mismos.

caché L2

SECC  
SEPP

◀ Vol.5 en Web

En sus versiones finales, el CeleronA se desmarcó del Pentium II no sólo internamente en su área de integración, sino también externamente: En su formato (de placa de circuito impreso a PGA), envoltura (de cerámica a plástica) y zócalo de conexión a placa base (de Slot 1 a Socket 370, el que luego sería reaprovechado por la versión de 0.18 micras y el propio Pentium III).

## Celeron Coppermine: Reconversión a Pentium III

◀ 1.8

Posteriormente, Intel reforzó la línea de productos de su gama baja con el estreno en Enero de 1999 de una nueva versión del procesador para el zócalo Socket 370 que ya se vió en el tramo final del Mendocino.

Socket 370

La [foto 5.2](#) contrasta las diferencias entre el aspecto de este modelo más reciente de Celeron y su versión inaugural. Este nuevo Celeron es idéntico externamente al Pentium III de 0.18 micras (formato FC-PGA), y de hecho deriva de su mismo proceso de fabricación (por estas fechas, Intel ha dejado ya de fabricar el Pentium II).

◀ pág. 168

FC-PGA

El nuevo Celeron desarrolla una frecuencia muy superior a sus predecesores, llegando hasta los 800 MHz y afianzándose como la alternativa de bajo coste al Pentium III. Las diferencias con respecto a éste se localizan en dos puntos clave:

800 MHz

- 1 El bus local, en lo referente a su velocidad (MHz), que es de 66 MHz en todos los Celeron A, mientras que en los Pentium II de 0.25 micras es de 100 MHz y en los Pentium III de 0.18

bus local

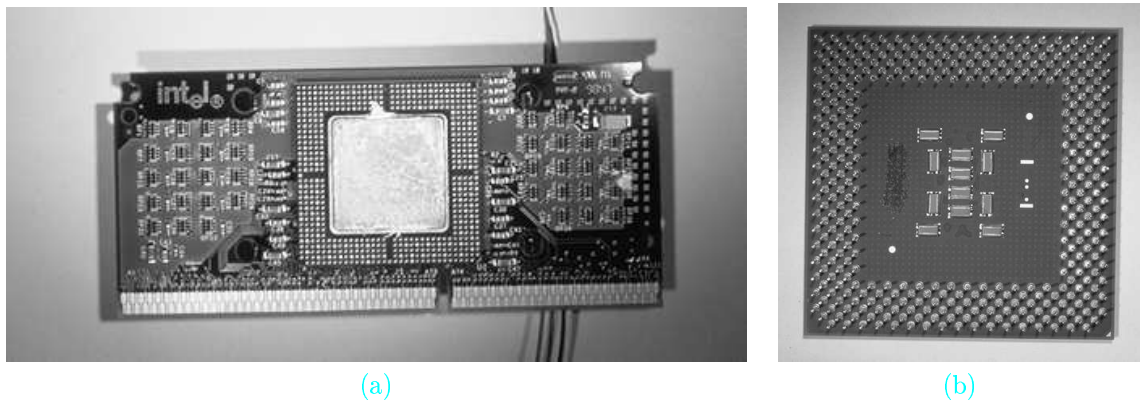


FOTO 5.2: (a) El primer procesador Celeron que salió al mercado, en formato Slot 1 pero carente de caché L2 interna. (b) La última versión, de 0.18 micras, que es ya un Pentium III reducido bajo su mismo formato Socket 370 y que contiene 128 Kbytes de caché integrada a la misma velocidad del procesador.

micras de 133 MHz. Puesto que la frecuencia del procesador crece y crece, estamos obligados a utilizar multiplicadores de reloj muy altos, delatando a una arquitectura excesivamente taponada en el bus local. Sólo en el último modelo, el de 800 MHz que data ya del año 2001, el bus se coloca en 100 MHz. Hasta él, el modelo de 766 MHz, por ejemplo, tiene un multiplicador de 11.5, algo inverosímil.

caché L2

- ② La memoria caché L2, en lo que respecta al tamaño y su frecuencia de funcionamiento. En este Celeron, sigue siendo de 128 Kbytes a la misma velocidad del procesador, pero ya integrada dentro del chip procesador, lo que en la práctica le hace ser bastante más rápida como ya demostramos en la [sección 3.4](#). En los Pentium II de 0.25 micras, esta caché es interna de 512 Kbytes a la mitad de velocidad, y en el Pentium III de 0.18 micras es de 256 Kbytes integrada a la velocidad del procesador, es decir, todo igual pero doblando el tamaño.

pág. 69 ➔

rendimiento

En cuanto una aplicación se deslice por encima de los 128 Kbytes de caché de este Celeron, los accesos al bus van a ser numerosos para traer los datos de memoria, y el rendimiento de nuestra aplicación software va a verse notablemente degradado. Además, este bus no sólo era lento, sino que además impedía sacar partido de los diseños de memoria principal más rápidos del mercado, que trabajaban ya sobre buses de 100 MHz o superiores.

Socket 370

El nuevo zócalo Socket 370 que Intel ha desarrollado para las nuevas versiones del procesador Celeron tiene el mismo aspecto cuadrado y plano que el Socket 7 ya conocido de los Pentium y MMX. A pesar de ello, es incompatible con su predecesor Socket 7, debido a que tiene una fila más de patillas que completan el total de 370 pines. Aunque no se pudo adaptar el Celeron a este viejo zócalo, sí que se facilitó una conversión al zócalo Slot 1 del Pentium II (el modelo MS-6905 de MSI, por ejemplo), pues aunque estructuralmente son muy diferentes, se ha dado la circunstancia de que ambos han confluído en otros dos factores: El voltaje de alimentación y el protocolo de bus, GTL+, heredado del Pentium Pro.

adaptador

## 1.9 ▶ Pentium III: Más instrucciones multimedia

El 26 de Febrero de 1999, Intel lanzó el Pentium III, abanderado por la firma como el procesador ideal para el apetitoso campo multimedia y de Internet. El código de referencia utilizado para este modelo inaugural es Katmai. Se integra a 0.25 micras, retoma la frecuencia donde la había dejado el Pentium II (450 MHz), y utiliza su misma jerarquía de caché L1 integrada de 32 Kbytes

Katmai  
frecuencia  
jerarquía de  
caché

Características	Microprocesador Celeron		
	Covington (1998)	Mendocino (1998)	Coppermine (1999)
Variables eléctricas:			
Frecuencia de reloj (MHz)	266, 300	300-533(*)	533-800(*)
Frecuencia bus local (MHz)	66 MHz (**)		
Potencia disipada (vatios)	16.6, 18.4	18.4-28.3	17.1-20
Voltaje del núcleo (voltios)	2.0	2.0	1.5
Zócalo al que se acopla	Slot1	Slot1/Socket370	Socket370
Distancia de integración	0.35	0.25	0.18
Tecnología de integración	CMOS		
Empaquetado y envoltura externa	SEPP	SEPP/PPGA	PPGA/FC-PGA
Segmentación:			
Cauce aritmética entera	11 etapas		
Cauce punto flotante y MMX	14 etapas		
Superescalaridad:			
Unidades de ejecución	7 (2 enteras, 2 punto flot, 3 MMX)		
Factor de superescalaridad	3		
Memoria caché:			
L1 Datos (integrada)	16 Kb, 128 ctos, 4 líneas/cto, 32 bytes/lín		
L1 instrucciones (integrada)	16 Kb, 128 ctos, 4 líneas/cto, 32 bytes/lín		
L2 unificada (integrada)	No lleva	128 Kbytes, 4 líneas/cto, 32 bytes/lín	
Conjunto de instrucciones:	Compatible 80x86 + 57 MMX		

TABLA 5.6: Resumen de los principales parámetros del Celeron en sus tres versiones: Covington, de 0.35 micras, Mendocino, de 0.25 micras y Coppermine, de 0.18 micras. El formato PPGA se incorpora a partir de los modelos de 300 MHz, y el formato FC-PGA entra a partir de los modelos de 533 MHz, donde el procesador deriva ya del Pentium III. (\*) Se cubre todo el rango de frecuencias mostrado a incrementos de 33 MHz. (\*\*) A excepción del modelo de 800 MHz, el único que dispone de un bus a 100 MHz.

escindida en dos y L2 interna de 512 Kbytes a la mitad de velocidad del procesador.

Arquitecturalmente, el Katmai es una secuela más del Pentium Pro, siendo las únicas novedades respecto al Pentium II Deschutes la inclusión de un código de identificación para la CPU y la extensión del conjunto de instrucciones con 70 nuevas instrucciones del tipo SIMD que utilizan la misma idea subyacente que dió lugar a la extensión MMX.

Esto de los nombres resulta la mar de curioso. Si tomamos la secuencia de modelos Pentium II Klamath, Pentium II Deschutes, Pentium III Katmai y Pentium III Coppermine y los vamos comparando cada uno con su anterior, los dos Pentium II son bastante diferentes y los dos Pentium III también, y sin embargo, en ambos casos se usa la misma denominación comercial. El mayor parentesco se produce entre el Pentium II Deschutes y el Pentium III Katmai, precisamente la pareja en la que se produce el cambio de nombre.

El Pentium III era el último paso de sexta generación de Intel, y como tal, estaba destinado a entretener al mercado hasta que la compañía finalizase el diseño del Pentium 4 o se decidiese a desenmarañar el culebrón del Itanium.

El mercado no necesita un nuevo microprocesador cada seis meses, pero las compañías saben que tienen que dar una imagen de constante innovación. Así que se cambia el nombre con la excusa de un cambio menor como el de las instrucciones multimedia, se completa la jugada con una campaña de publicidad de coste cercano a los 300 millones de dólares, y el que quiera, ya

CPU\_ID

paradoja en los bautismos

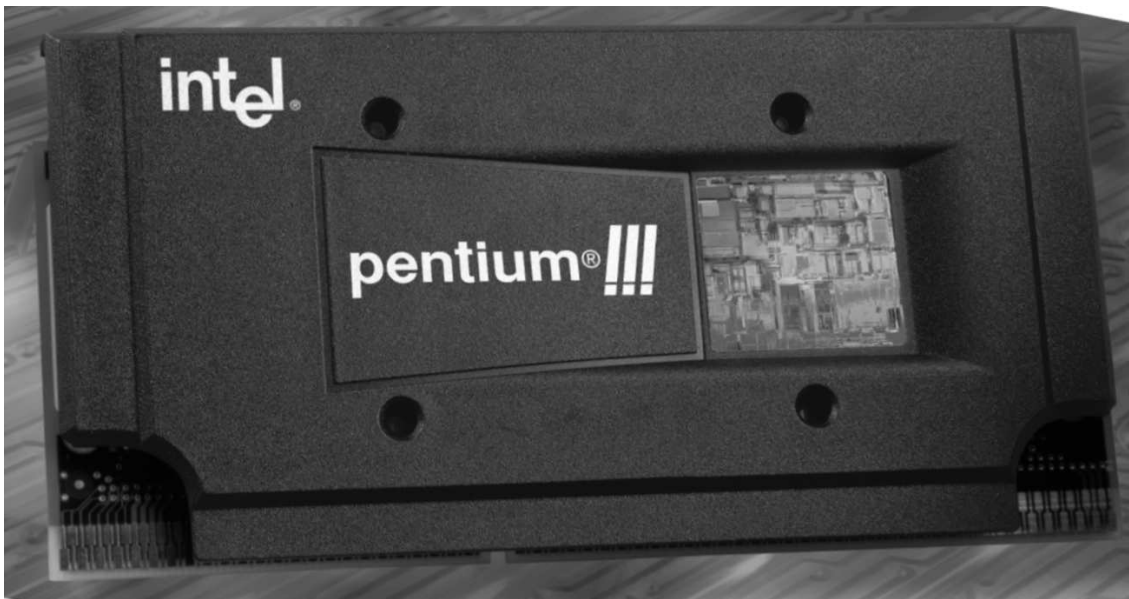
	Código de referencia	Año	Frec. (MHz)	F.bus (MHz)	Cache (Kb)		Micras	Trans. (Mill.)	Zócalo
					L1	L2			
P E N T I U M	Original ó P5	1993	60-90, 100,120, 133,150, 166,200	50, 60, 66	16	(X)	0.8 0.6 0.5, 0.35	3.3	Soc.5, Soc.7
	Pro ó P6	1995	150,166, 180,200	60, 66	16	(✓)256, 512, 1024	0.6 0.5 0.35	5.5	Soc.8, Soc.9
	MMX	1996	133,150, ...,266	66	32	(X)	0.35	3.3	Soc.7
P E N T I U M II	Klamath	1997	233,266 300	66	32	(✓)512	0.35	7.5	Slot1
	Deschutes	1998	333,350, 400,450	66, 100	32	(✓)512	0.25	7.5	Slot1
	Xeon	1998	400,450	100	32	(✓)512, 1M,2M	0.25	7.5	Slot2
	Celeron ó Covington	1998	266, 300	66	32	(✓)	0.25	7.5	Slot1
	Celeron A ó Mendocino	1998	300,333, ...,533	66	32	(★)128	0.25	22	Soc370
P E N T I U M III	Celeron	1999	533,... ...,800	66	32	(★)128	0.18	22	Soc370
	Katmai	1999	450,500, 550,600	100,	32	(✓)512	0.25	9.5	Slot1
	Tanner (Kat.Xeon)	1999	500 600	100	32	(✓)512, 1M,2M	0.25	9.5	Slot2
	Coppermine	1999	600,666, ...,1GHz	133	32	(★)256	0.18	28.1	Soc370
	Cascades (Cop.Xeon)	2000	600 1GHz	133	32	(✓)512, 1M,2M	0.18	28.1	Slot2
	Tualatin	2001	1GHz	133	32	(★)512	0.13	?	Soc370

TABLA 5.7: La familia de microprocesadores Pentium de quinta y sexta generación resumida en sus principales parámetros. En la columna correspondiente a la caché L2 hemos utilizado nuestra particular simbología para denotar con X la caché externa, con ✓ la caché interna y con ★ la caché integrada. Todas las cachés integradas funcionan a la velocidad del procesador, y de las internas, también lo hacen las del Pentium Pro y el Celeron-A. En el modelo Xeon derivado del Coppermine, la caché mostrada en la columna L2 es realmente la L3 interna puesto que ya dispone de la misma L2 integrada que el Coppermine. La contabilización del número de transistores se refiere únicamente al chip de la CPU, esto es, no refleja los transistores de la caché interna que van en un chip aparte y son aproximadamente unos 15 millones por cada 512 KBytes.

puede estrenar procesador nuevo. Así se batían récords de ventas todos los años y se obliga a los usuarios a renovar los equipos cada 2-3 años.

### 1.9.1 SSE: 70 instrucciones multimedia

Este conjunto de instrucciones se denominó en primera instancia MMX2, posteriormente KNI (Katmai New Instructions), y finalmente, SSE (Streaming SIMD Extensions), nombre que parece



(a)



(b)



(c)

Foto del Xeon. Cortesía de Intel

FOTO 5.3: Las tres versiones de Pentium III que conviven en el mercado. (a) La versión inaugural, el Katmai, de 0.25 micras con caché interna de 512 Kbytes. (b) La versión Xeon para plataformas de tipo servidor, alias Tanner, donde la caché interna dispone de mayor tamaño y velocidad. (c) El Coppermine de 0.18 micras, ya en formato Socket y con caché L2 integrada de 256 Kbytes. Puesto que las tres fotos tienen una escala diferente, podemos utilizar los logotipos Pentium III para comparar las fotos (a) y (b), ya que en ambas representa la misma distancia real.

haberse consolidado ya.

El objetivo de Intel con SSE es superar las prestaciones conseguidas por AMD con su 3DNow!, la versión ampliada al MMX de la competencia. Para ello, las nuevas instrucciones, 70 en total, no sólo tratan de acelerar las funciones más usadas en la gestión de gráficos 3D y vídeo, sino que también incluyen soporte para las aplicaciones de reconocimiento de voz, la computación de

nuevas  
funciones

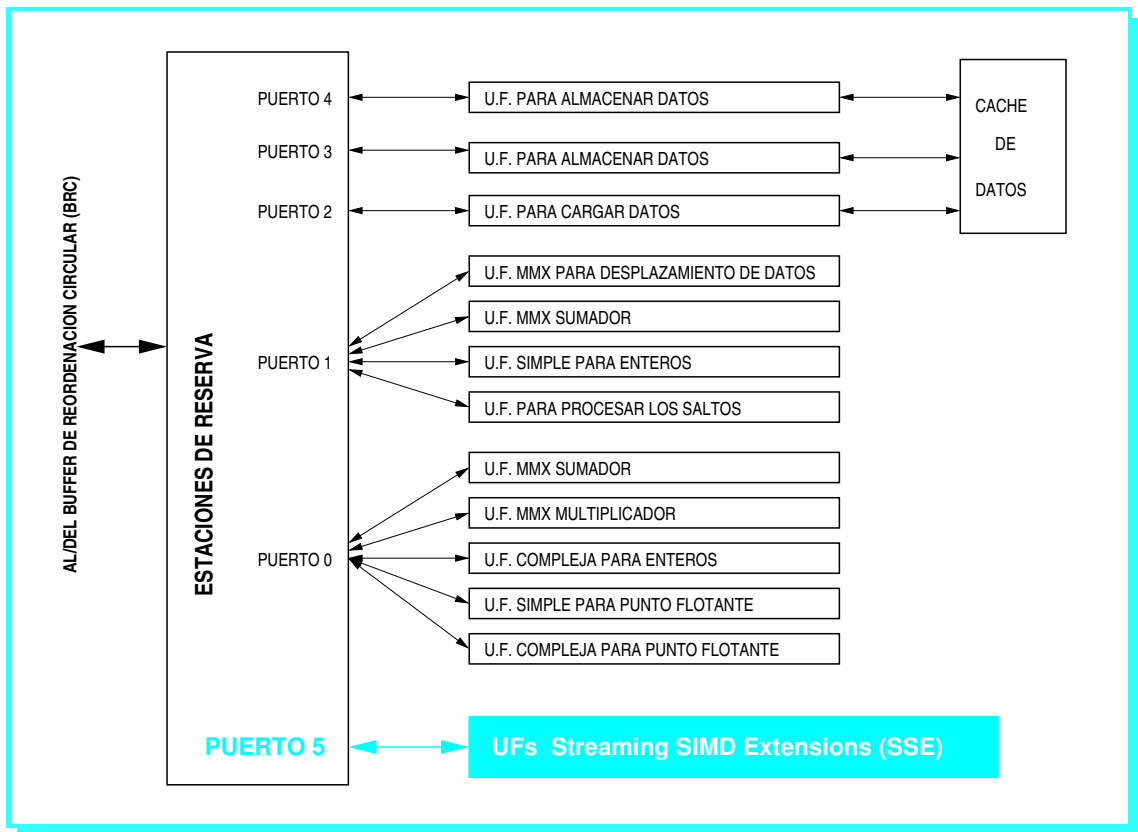


FIGURA 5.6: Detalle de las unidades funcionales de que dispone el Pentium III para la ejecución de instrucciones. La que se encuentra ubicada en el puerto 5 (resaltada) es la única novedad respecto al Pentium II.

☛ pág. 106

punto flotante y la gestión de memoria caché. La [sección 3.5.4](#) trata en mayor profundidad todo lo referente a las instrucciones multimedia en general y al conjunto SSE en particular.

Como ya ocurriese con la extensión MMX, Intel incorpora a partir de aquí el conjunto SSE a todos sus nuevos procesadores, como el Pentium 4 o el Itanium.

### 1.9.2 Número de identificación para una CPU

hackers

Con el Pentium III, Intel también anunció la implementación de un número de identificación que llevan grabado internamente todos los microprocesadores fabricados a partir del Katmai. Con ello se pretendía identificar de forma unívoca cada máquina del universo informático con objeto de que las transferencias realizadas por Internet puedan llevarse a cabo de una forma segura sin que otro usuario o *hacker* pueda suplantar a uno mismo.

Para impedir que alguien pueda descubrir nuestro número de identificación, éste se codificará utilizando otros números completamente impredecibles que se conseguirán a partir de un generador de números aleatorios interno a la CPU, implementado a partir del valor del ruido térmico procedente del semiconductor.

seguridad

intimidad

Así, este número trata de poner a salvo el agujero de seguridad actualmente existente en las transacciones comerciales a través de Internet. El aspecto negativo de este número viene por el hecho de que el usuario pierde toda su intimidad, y cualquiera puede saber lo que uno está haciendo en la red. Por ello, Intel permite deshabilitar este número a voluntad del usuario, operación fácil

de llevar a cabo si ha sido contemplada en los menús de la BIOS de su sistema.

### 1.9.3 Zócalo e integración

El zócalo de conexión a la placa base para el Katmai sigue siendo el Slot 1, lo que permite usar para los primeros modelos Pentium III las placas base ya existentes para Pentium II.

zócalo

Lo que cambió pronto fue el cartucho cerámico del procesador. A mediados de 1999 comenzó a montarse el Katmai bajo cartucho SECC2, que frente al tradicional SECC introduce como novedad que sólo una de las caras del circuito impreso va realmente recubierta, dejando los chips expuestos al exterior por la otra cara para facilitar la ventilación. El núcleo del procesador carece de cualquier tipo de cobertura metálica, lo que le hace entrar en contacto directo con el disipador de calor, algo más grande que el del Pentium II. La placa de aluminio del disipador se engancha al procesador a través de 8 agujeros presentes en la placa de circuito impreso, donde también se colocan dos chips de caché L2 de Samsung a 4 ns. (250 MHz, la mitad de frecuencia del procesador, para el caso de una frecuencia de 500 MHz).

SECC2

La [foto 5.3.a](#) muestra el aspecto de este procesador comparándola con su versión Xeon, el modelo orientado a servidores con caché L2 más grande y más rápida, y su versión Coppermine, que trataremos seguidamente.

pág. 171

El voltaje es de 3.3 voltios en la parte externa dedicada a la caché L2 y de 2 voltios para el núcleo interno, en clara semejanza con el Deschutes y el Celeron que comparten el zócalo Slot 1.

voltaje

Según fuentes de la propia compañía, el aumento con respecto a la misma configuración bajo Pentium II es de un 42 %. Fuentes menos interesadas tasan la mejora en un 5 %. Visto lo visto, esta última cifra parece bastante más creíble.

## Coppermine: Novedades en la integración

◀ 1.10

La llegada del proceso de fabricación de 0.18 micras provoca ciertas metamorfosis en el Pentium III. Se trata de sacar partido de las dos principales vertientes de mejora que permite: Frecuencia más elevada y mayor número de transistores. En frecuencia, flanquearemos el arraigado listón del Gigahercio. En número de transistores, sobrepasaremos los 25 millones en un solo chip.

frecuencia y transistores

La versión del Pentium III en la que se repercuten todas estas mejoras recibe el código de referencia Coppermine.

Coppermine

### 1.10.1 Frecuencia de reloj

El rango de frecuencias del Coppermine comienza en los 533 MHz, evolucionando a pasos de 66 MHz: 600, 667, 733, 800, 866, 933, 1000, 1066 y 1133 MHz (1.13 GHz). Algo más tarde vendría la versión de 0.13 micras, con código de referencia Tualatin, para superar estos valores e integrar una caché L2 de 512 Kbytes.

Tualatin

Los saltos de 66 MHz delatan multiplicadores de reloj de 0.5 y frecuencia para el bus local de 133 MHz. Esta última supone una notable mejoría con respecto al bus local de 100 MHz del Katmai: La memoria principal respalda este incremento en frecuencia con su implementación PC133, y puesto que el bus local sigue siendo el cuello de botella de la arquitectura, una mejora de 33 % en este elemento revierte en un porcentaje similar sobre el conjunto.

cuello de botella

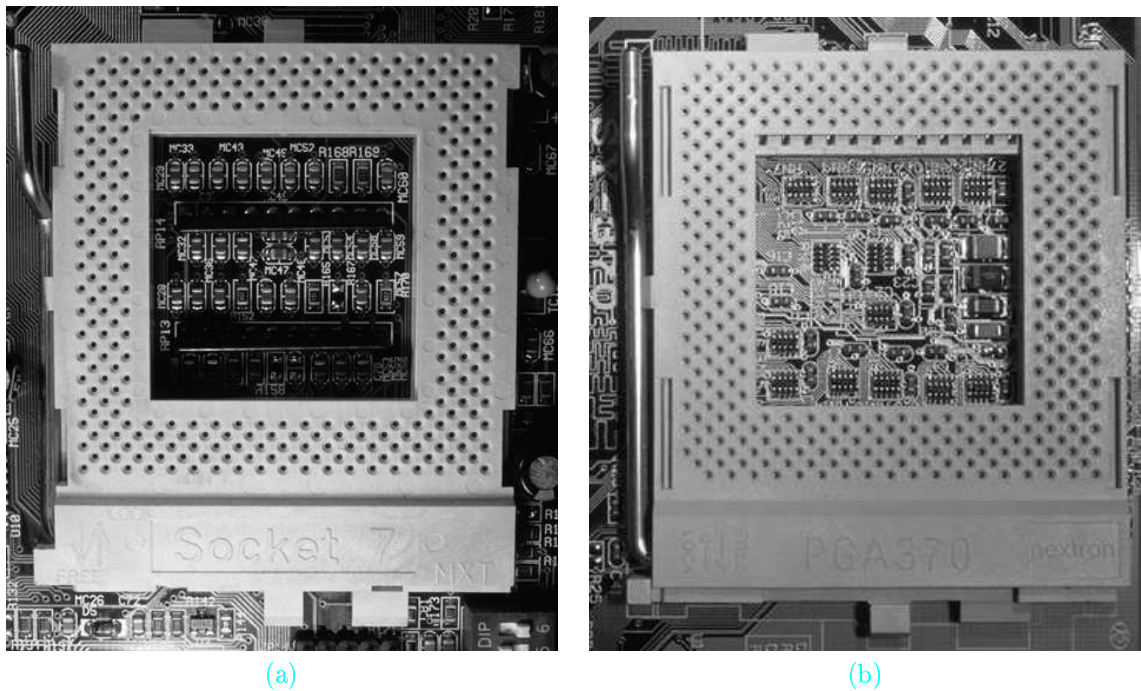


FOTO 5.4: La vuelta al zócalo tipo Socket por parte de Intel con la llegada de su Pentium III Coppermine supone retomar una solución similar al Socket7 del Pentium MMX (a) pero de mayor patillaje: El Socket 370 (b). Puede apreciarse que el perímetro del Socket 370 dispone de seis filas de pines (370 en total), mientras que el del Socket7 sólo tiene cinco filas y 296 pines. Los dos zócalos albergan como inquilino a un procesador fabricado en formato PGA.

### 1.10.2 Tecnología de integración y zócalo

Las 0.18 micras son el elemento clave del Coppermine. Aprovechando la migración al nuevo proceso de fabricación, Intel retorna al formato PGA (Pin Grid Array) que dejó atrás con el Pentium MMX. El zócalo vuelve a ser de tipo Socket, recibiendo esta vez la denominación Socket 370 en honor al número de pines que a él se conectan. El zócalo es compatible con la versión Celeron de 0.18 micras ya vista. La [foto 5.4](#) contrasta estos dos zócalos.

Se completa de esta manera un camino de ida y vuelta entre los zócalos ZIF, que montaron las versiones PGA del procesador bajo Socket 5, 7, 8, 9 y 370, y los zócalos Slot 1 y 2 sobre los que se acoplaron los cartuchos cerámicos de tipo SECC, SECC2 y SEPP.

El diseño tipo SECC pasa así a la historia como una solución provisional, válida sólo para la fase transitoria en la que la tecnología de integración no permite integrar el núcleo de un microprocesador y su caché L2 en la misma pastilla de silicio, pero la arquitectura necesita tener a la L2 lo más cerca posible. Lograda la integración conjunta con las 0.18 micras, dado que las soluciones SECC incrementan el coste de fabricación del procesador, éstas quedan marginadas al uso por parte de las versiones Xeon del procesador, donde el enorme tamaño de la caché L2 aún supone un desafío para la tecnología.

### 1.10.3 Memoria interna

El Coppermine invierte buena parte de su mayor patrimonio en integrar una caché L2 dentro del mismo área de integración del chip procesador. Esta caché tiene una configuración espléndida: 256 Kbytes y 8 líneas por conjunto, manteniendo los 32 bytes por línea ya tradicionales.

PGA  
Socket

← pág. 174

ZIF  
Slot

solución  
transitoria

buena  
configuración



Características	Microprocesador Pentium III	
	Katmai (1999)	Coppermine (1999)
<b>Variables eléctricas:</b>		
Frecuencia de reloj (MHz)	450, 500, 550, 600	533, 600, ..., 1133
Frecuencia bus local (MHz)	100	133
Potencia disipada (vatios)	25.3, 28.0, 30.8, 34.5	17.4, 19.6, ..., 35
Voltaje del núcleo (voltios)	2.0	1.65
Zócalo al que se acopla	Slot1	Slot1/Socket370
Tecnología de integración (micras)	CMOS 0.25	CMOS 0.18
Empaquetado y envoltura externa	SECC	FC-PGA
<b>Segmentación:</b>		
Cauce aritmética entera	11 etapas	
Cauce punto flotante y MMX	14 etapas	
<b>Superescalaridad:</b>		
Unidades de ejecución	8 (2 enteras, 2 punto flot, 3 MMX, 1 SSE)	
Factor de superescalaridad	3	
<b>Memoria caché:</b>		
L1 Datos (integrada)	16 Kb, 128 ctos, 4 líns/cto, 32 bytes/lín	
L1 instrucciones (integrada)	16 Kb, 128 ctos, 4 líns/cto, 32 bytes/lín	
Tipo L2	Interna	Integrada
Velocidad L2	La mitad del proc.	La misma del proc.
Tamaño L2	512 Kbytes	256 Kbytes
Organización L2	4 líneas/conjunto	8 líneas/conjunto
Conjunto de instrucciones:	Compatible 80x86 + 57 MMX + 70 SSE	

TABLA 5.8: El Pentium III resumido en sus dos modelos para gama PC: El Katmai de 0.25 micras y el Coppermine de 0.18 micras.

Hasta este momento, los diseños de Intel estaban en su mayoría en las 4 líneas por conjunto, y los de AMD en 2 líneas por conjunto. A partir de aquí, esta variable despierta y comienzan a verse valores de hasta 16. Esto aumenta el índice de aciertos a caché para un mismo tamaño, al crecer la flexibilidad en la ubicación de los datos procedentes de memoria principal.

## Pentium III Xeon: Más para servidores

### ◀ 1.11

El Pentium III Xeon es otra versión de Intel para el segmento de servidores. Se trata básicamente de un Pentium II Xeon con el agregado del Katmai, esto es, el juego de instrucciones SSE y el número de identificación.

SSE  
CPU\_ID

Ocurre sin embargo que la mayoría de los sistemas servidor existentes en la actualidad, como los servidores de ficheros y de impresora, no van a hacer uso de ninguna de estas prestaciones. Sólo a servidores más especializados, como los de video, podrá serle útil esta versión del procesador. Y quizá también a las estaciones de trabajo que lo monten.

El código de referencia para esta CPU es Tanner, y después de él, Intel aumenta su escalabilidad desarrollando la versión de 0.18 micras y bus local a 133 MHz, cuyo código de referencia es Cascades.

Tanner  
Cascades

En estas configuraciones, lo más destacado vuelve a ser el carácter multiprocesador SMP y la presencia de una caché interna, la L2 en el caso del Tanner y ya la L3 en el caso del Cascades, con tamaños que incluso pueden superar el Megabyte, siempre a la misma velocidad del procesador.

SMP  
L2

## SECCIÓN 5.2

## AMD

## 2.1 ► K6: El primer problema para Intel

**madrugador** AMD inauguró la sexta generación en abril de 1997, un mes antes de que Intel lanzara su primer Pentium II. Su arquitectura fue capaz de diseñarse con tanta celeridad porque en realidad era una modificación del procesador 686 de NexGen, la compañía que adquirió antes de la salida del K5, y cuyo 586 también le sirvió para desarrollar aquél.

**novedades** Sobre el núcleo del 686, el K6 incorporó un módulo MMX, aumentó el tamaño de la caché de primer nivel, y potenció algunas características RISC. El objetivo inicial era ofrecer un chip con un rendimiento similar al Pentium II, pero siempre con un coste inferior. No sólo porque el chip en sí era más barato, sino porque se mantenía el zócalo Socket 7 de los Pentium normales, permitiendo pinchar al procesador en las placas base para Pentium, que siempre fueron más baratas que las de Pentium Pro o Pentium II.

**Super 7** Sobre esta plataforma como base, desarrollarían posteriormente la especificación de zócalo Super 7 para diálogos a 100 MHz, y establecerían una acertada estrategia comercial basada en una alianza con VIA para que les desarrollara juegos de chips que aprovecharan estas prestaciones.

**cueillos de botella** Estos chips constituyeron la saga Apollo VP, VP2, VP3, ..., que proporcionaban un rendimiento bastante superior al de los juegos de chips de Intel para su zócalo Socket 7, y ya muy próximo a los que encontramos bajo zócalo Slot 1. Esto reducía el negativo impacto de los cuellos de botella del equipo, por aquel entonces ubicado en el bus local. Esta fue, junto a su coste, una de las principales razones para decantarse por este procesador.

En términos de potencia de proceso bruta, el K6 se encuentra entre medias de un Pentium Pro y un Pentium II. Pasamos a desglosar sus características más distintivas.

## 2.1.1 Frecuencia

El K6 comenzó su andadura a tres frecuencias: 166, 200 y 233 MHz. Posteriormente aprovecharía la utilización de las 0.25 micras para inaugurar otro tramo de 266, 300 y 333 MHz, manteniendo siempre la frecuencia del bus en los 66 MHz.

## 2.1.2 Tecnología de integración

**voltaje** CMOS a 0.35 micras, que apenas un año más tarde migraría hasta las 0.25 micras reduciendo el voltaje hasta los 2.2 voltios y el área de integración hasta los  $68 \text{ mm}^2$ , permitiendo a la empresa iniciar su singladura en el segmento de los computadores portátiles. El diseño consta de 8.8 millones de transistores.

## 2.1.3 Paralelismo a nivel de instrucción

**6 etapas** El K6 dispone de un esquema segmentado en 6 etapas: (1) Búsqueda de instrucción, (2) Conversión de código x86 a código nativo (microinstrucciones) y decodificación, (3) emisión de instrucciones, donde tiene lugar la reordenación y planificación de instrucciones fuera de orden, (4)

Ejecución-1, (5) Ejecución-2, y (6) Escritura del resultado de la operación.

Como se observa, estamos ante una arquitectura con un núcleo RISC dotado de ejecución fuera de orden donde se procesan las instrucciones aceptadas en formato compatible 80x86. Una vez traducido el conjunto de instrucciones compatible 80x86 de entrada, se transforma en microinstrucciones RISC de forma similar a como ya vimos que operaba el Pentium Pro.

conversión

El procesador acepta dos instrucciones x86 a su entrada, decodifica hasta cuatro microinstrucciones, y la ejecución es incluso más potente, pudiendo realizar hasta seis operaciones de forma simultánea gracias a que dispone de ocho unidades funcionales de ejecución: Dos ALU enteras, dos ALU MMX, un multiplicador MMX, carga en memoria, almacenamiento en memoria y FPU. Estas seis operaciones son: Una de carga en memoria y otra de almacenamiento, una entera (6 MMX) sencilla y otra compleja, una de punto flotante y una evaluación de la condición de salto.

unidades  
funcionales

Puesto que para las arquitecturas anteriores hemos medido el factor de superescalaridad en función de las microinstrucciones nativas, en este caso debemos quedarnos con el menor valor de entre la capacidad de decodificar (4) y ejecutar (6), con lo que nos deja un factor de superescalaridad de 4 para este procesador.

superescalar

#### 2.1.4 Memoria caché

El K6 dispone de caché L1 de 64 Kbytes, separada en dos cachés gemelas de 32 Kbytes para datos e instrucciones, respectivamente. El tamaño de la línea de caché es de 32 bytes, y su organización es asociativa con 512 conjuntos de 2 líneas cada uno.

sólo L1

Frente a la caché de primer nivel de los Pentium MMX, II y III, el K6 es cuantitativamente superior pero cualitativamente inferior. Dispone del doble de espacio, pero no lo gestiona de forma tan eficiente como la familia Pentium.

#### 2.1.5 Conjunto de instrucciones

Compatible x86 como viene siendo tradición en AMD, pero esta vez contando ya con la incorporación del conjunto de instrucciones MMX.

80x86+MMX

## K6-2: Comienza el baile multimedia

## ◀ 2.2

El modelo K6-2 (o K6-3D) de AMD es la versión ampliada del procesador K6. Ve la luz en Mayo de 1998, presentando como principales novedades respecto a su hermano menor las tres siguientes:

3 novedades

- ❶ Frecuencia de reloj: Las primeras versiones de mediados de 1998 se lanzaron a 266, 300 y 333 MHz. Pero para finales de ese mismo año ya teníamos en el mercado los modelos a 350 y 400 MHz. Y el procesador llegó a tocar los 500 MHz.
- ❷ Frecuencia de bus: Las versiones de 266, 300 y 333 MHz se montaron sobre un bus local a 66MHz, pero todas las frecuencias posteriores admitieron ya frecuencias de bus a 100 MHz y dispusieron de placas base con estas prestaciones. Un ejemplo es el juego de chips VP/3 de VIA, que proporciona también bus AGP a 66 MHz y un zócalo Socket 7 con la especificación Super 7. Los modelos de 300MHz son los únicos que coexisten en la modalidad de bus a 66MHz y bus a 100MHz. Si disponemos de un K6-2 300MHz y queremos saber a qué versión corresponde, debemos fijarnos en la etiqueta del procesador y buscar los caracteres AFR66. Si los encontramos, nuestro K6-2 sólo funcionará de forma fiable sobre placas con bus a 66 MHz y gran parte de las ventajas esgrimidas se habrán esfumado.

500 MHz

AFR66

Características	La familia del K6 de AMD al completo		
	K6 (1997)	K6-2 (1998)	K6-III (1999)
Variables eléctricas:			
Frecuencia de reloj (MHz)	166-333	266-500	400-600
Frecuencia bus local (MHz)	66	66/100	100
Voltaje del núcleo (voltios)	2.9	2.2	2.2/2.4
Tecnología de integración (micras)	0.35/0.25	0.25	0.25
Millones de transistores	8.8	9.3	21.3
Zócalo al que se acopla	Socket7 bajo especificación propia Super7		
Empaquetado y envoltura externa	CPGA (Ceramic Pin Grid Array)		
Superescalaridad:			
Unidades de ejecución	8. 2 enteras, 3 MMX, 1 FPU, (2 3DNow!)		
Factor de superescalaridad	4		
Memoria caché:			
L1 Datos (integrada)	32 Kb, 512 ctos, 2 líneas/cto, 32 bytes/línea		
L1 instrucciones (integrada)	32 Kb, 512 ctos, 2 líneas/cto, 32 bytes/línea		
Tipo L2	Externa	Externa	Integrada
Organización L2	256 Kb, 4 líneas/cto		
Conjunto de instrucciones:	80x86 + 57 MMX (+ 24 3DNow!)		

TABLA 5.9: Resumen de los principales parámetros de la saga arquitectural K6 de AMD. Los dos rasgos de 3DNow! colocados entre paréntesis simbolizan la presencia únicamente en los modelos K6-2 y K6-III.

conjunto de instrucciones

- ③ Notables mejoras en el conjunto de instrucciones. En un primer nivel de actuación se mejoraron las instrucciones en punto flotante y MMX, dotándolas de ejecución superescalar, reduciendo las distancias que le separaban del Pentium II en lo que a la potencia de cálculo se refiere.

Posteriormente, se amplió el conjunto de instrucciones MMX con 24 nuevas instrucciones multimedia, que dotaron al procesador de nuevas implementaciones gráficas y sonoras. Nos estamos refiriendo al conjunto de instrucciones 3DNow! (ver [sección 3.5.4](#)).

3DNow!  
pág. 106

rendimiento

El rendimiento de un procesador AMD K6-2 utilizando aplicaciones ofimáticas es un poco más bajo que el de un Pentium II funcionando a su misma velocidad, ya que la caché L2 de éste último desequilibra la balanza a su favor.

Sin embargo, cuando se utilizan aplicaciones multimedia optimizadas para la tecnología 3DNow!, o cuando se emplean programas que llaman a Direct3D o a la versión DirectX se aprecian velocidades de hasta 70 imágenes por segundo en la generación de escenas tridimensionales, lo que supera a un Pentium II de sus mismas características.

éxito comercial

El éxito comercial del K6-2 fue más que notable. De hecho, corresponde a este modelo el honor de ser el primero de la década de los 90 que realmente hizo mella en la cuota de mercado de Intel para microprocesadores de corte más doméstico. Sus sucesores bajo Socket 7, que describiremos seguidamente, continuaron esta tendencia, y fue así como los fabricantes de placas base y juegos de chip comenzaron también a preocuparse por arropar convenientemente a los productos de AMD en lugar de fijarse de forma exclusiva en Intel. Con el tiempo, las buenas prestaciones a nivel hardware se complementaron con la compatibilidad a nivel software. Todo este caldo de cultivo hizo subir las ventas de AMD, y a su vez, a favorecer la bajada de precios de sus productos para plantar cara al gigante Intel: La batalla por el mercado de microprocesadores para PC comenzaba así su fase más cruenta.

competencia

## K6-III: El ejemplo de configuración equilibrada

## ◀ 2.3

En el mismo horizonte temporal en que entró en escena el Pentium III, pero una vez más un mes antes que Intel (Febrero de 1999), AMD lanzó al mercado el procesador K6-III, cuyo código de referencia es Sharptooth (del inglés, *diente afilado*).

Sharptooth

Este procesador presentó como única novedad frente al K6 la presencia en el interior del procesador de una caché L2 integrada de 256 Kbytes que opera a la misma frecuencia del microprocesador. Se incurrió para ello en un gasto considerable de transistores, sumando 21.3 millones en el total del diseño, un 45% más que el K6-2. La nueva tecnología de integración a 0.25 micras permitió situar el voltaje en los 2.4 voltios, e incluso incorporar internamente el controlador para la caché externa de tercer nivel (L3), a la misma velocidad del procesador, y preparado para manejar cachés externas L3 de entre 512 Kbytes y 2 Mbytes (aunque en la práctica éstas apenas sí se vieron).

caché L2

El área de integración creció hasta los 135 milímetros cuadrados, y sus frecuencias de trabajo iniciales, hasta los 400 y 450 MHz, ofreciendo ambas un rendimiento superior al de un Pentium II a su misma frecuencia. Posteriormente, saldrían versiones a 500 y 600 MHz que ruborizaron al propio Pentium III.

frecuencias

Y es que, si nos ceñimos a Intel y AMD, no ha habido en toda la década de los noventa decisión arquitectural más acertada que integrar una caché L2 al K6-2. La comparativa de sexta generación que realizamos en la [sección 5.3](#) justificará esta afirmación.

acierto  
histórico  
→ [pág. 180](#)

El éxito arquitectural del K6-III instauró en el mercado una tendencia a integrar la L2 dentro del procesador. El fue el primero que lo hizo a 0.25 micras, pero cuando las 0.18 micras concedieron una nueva oportunidad, los demás fabricantes acudieron prestos a imitarle sin vacilación.

copia

El K6-III trajo además consigo un pequeño misterio. Los diseñadores del procesador cambiaron la organización de la caché con respecto a la de primer nivel, y volvieron a hacerlo para el controlador de la caché L3 que también venía integrado en el procesador. La caché L1 tenía una asociatividad de dos líneas por conjunto, mientras que la L2 pasaba a tener cuatro y la L3 sólo una (es lo que se conoce como *caché de organización directa* - para más detalles sobre la asociatividad de una caché, consultar la [sección 11.2](#)). El objetivo que está detrás de todos estos cambios no es otro que tratar de barrer un área disjunta de memoria desde cada caché, proporcionando un efecto global de sumar sus tamaños para aumentar su efectividad.

asociatividad

→ [p.123/Vol.2](#)

El K6-III fue el último diseño de AMD bajo formato Socket 7 antes de migrar a zócalo propio con el Slot A del K7 (ver [sección 6.1](#)). Eso sí, con posterioridad al lanzamiento del K7, AMD repercutiría el proceso de fabricación de 0.18 micras de su Thunderbird sobre sus diseños K6-2 y K6-III, dando lugar a los K6-2+ y K6-III+, cuya principal novedad fue un voltaje más reducido, con la intención de destinar ambos productos al segmento de los portátiles. En el caso del K6-2, además, se aprovechó para integrar una caché L2 de 128 Kbytes, la mitad de la disponible en el K6-III, pero que, visto lo visto, era suficiente para rectificar entonando aquello de “más vale tarde que nunca”.

→ [pág. 192](#)

K6-2+  
K6-III+

Todos los modelos que comparten la arquitectura K6 pueden montarse sobre una misma placa base bajo Socket 7 sin más que realizar una simple actualización de la Flash-BIOS que adecúe voltaje de entrada, frecuencia de bus local, y multiplicador de reloj para el procesador, en línea con el esquema de configuración que más adelante se repetiría para el Athlon XP.

Procesadores	Número máximo de instrucciones procesadas en fase de:			
	Búsqueda	Decodificación	Reordenación	Escritura
Pentium Pro, Pentium II, Celeron y Pentium III	Tres del tipo x86	Seis nativas	Tres nativas	Tres nativas
K6, K6-2 y K6-III	Dos del tipo x86	Cuatro nativas	Cuatro nativas	Seis nativas

TABLA 5.10: El esqueleto de las dos principales familias de microprocesadores de sexta generación frente a frente en su dimensión más arquitectural. Hemos sombreado en cada caso la fase en la que se sitúan sus principales carencias, esto es, la más proclive a convertirse en el cuello de botella de la arquitectura.

### SECCIÓN 5.3

## Comparativa: Intel versus AMD

A la hora de establecer una comparativa entre las arquitecturas de sexta generación de Intel y AMD, nos encontramos con que hemos desmenuzado la del Pentium Pro, el esqueleto utilizado por Intel, pero no la del K6, que es lo que le infunde el carácter a los diseños de AMD.

Creemos que el diseño del K6 se explica mejor contrastándolo con el del Pentium Pro en lugar de abordarlo de manera independiente. Si hemos comprendido cómo trabaja el Pentium Pro, la explicación sobre la forma de proceder del K6 puede simplificarse de manera extraordinaria.

El K6 es un diseño segmentado y superescalar con ejecución fuera de orden, esto es, dispone los mismos ingredientes que el Pentium Pro en su arquitectura, sólo que ponderándolos de manera diferente:

- boca

❶ Búsqueda y decodificación: A la hora de aceptar instrucciones x86, el Pentium Pro toma hasta tres, que reconvertía a un máximo de seis microinstrucciones nativas, mientras que el K6 toma un máximo de dos, que traduce a cuatro microinstrucciones nativas a lo sumo.
- estómago

❷ Fase de reordenación y ejecución: El BRC (búfer de reordenación de instrucciones del Pentium Pro) sirve tres instrucciones por cada ciclo de reloj y ejecuta otras tres, lo que lo otorga ese grado de superescalaridad. El BRC del K6 puede servir cuatro instrucciones por cada ciclo, y ejecutar hasta un máximo de seis.

La [tabla 5.10](#) resume todos estos números. Recopilando, podemos observar que el K6 es más torpón en la búsqueda y decodificación, pero más ágil en la reordenación y ejecución de instrucciones. El Pentium Pro, en cambio, es un diseño más compensado, y a simple vista, parece más conveniente.

Sin embargo, conviene no extraer conclusiones precipitadas. Hemos de adentrarnos un poco más en el estudio de cada modelo para conocer sus ventajas frente al rival. Acudiremos a una socorrida analogía para abreviar nuestro análisis técnico y así poder ver de una manera inmediata aquello que nos interesa mostrar.



### Analogía 5.1: K6: EL MODELO DE BOCA ESTRECHA Y ESTÓMAGO GLOTÓN

Asociemos el esqueleto de un procesador con la anatomía humana. La búsqueda de instrucciones va a ser la operación de llevarse los alimentos a la boca, la decodificación es su masticación, la reordenación es el paso por el esófago, y la ejecución es la digestión en el estómago (vamos a ahorrarnos la escritura, intrascendente en ambos casos).

De esta manera, el Pentium Pro tiene una boca generosa y mastica los alimentos con gran celeridad, pero luego se atraganta un poco en el conducto del esófago. Como resultado de ello, suele tener la boca llena de comida, y su velocidad real viene condicionada por la capacidad digestiva de su estómago.

Por el contrario, el K6 tiene una boca más estrecha, y mastica también a un ritmo más cándido, pero todo lo que mastica puede pasarlo de forma inmediata al esófago, y desde ahí, el estómago que espera más abajo es aún más voraz, con ocho unidades funcionales de ejecución cooperando en la ejecución de hasta seis operaciones de forma simultánea. En definitiva, se comporta como un polluelo hambriento: Siempre tiene la boca abierta esperando comida, ya sea de papá banco de registros o mamá memoria caché.

Ante este panorama, la evolución marcada por cada compañía arroja un balance diferente que sintetizamos indicando las dos pinceladas que nos han parecido más sobresalientes en cada una de ellas.

## Síntesis de la trayectoria de Intel

◀ 3.1

Intel muestra una inclinación por incorporar nuevas unidades funcionales a cada nuevo modelo gestado: Partió con cuatro en el Pentium Pro, incorporó tres más en el Pentium II (las MMX), y finalmente una más en el Pentium III (la SSE), habilitando también nuevos puertos de conexión para ellas. En definitiva, terminó prácticamente doblando la capacidad de su estómago en relación a como empezó, una evolución ciertamente plausible.

potencia el estómago

El otro rasgo que hemos observado en Intel es su denodado interés por mejorar el subsistema de predicción de salto, algo igualmente loable. La arquitectura P6 prima la segmentación por encima de la superescalaridad, y tomados estos derroteros, la sombra que proyectan las dependencias de control es muy alargada. Como el déficit de la arquitectura está en el estómago, no pasa nada si se entretiene un poco más a la boca masticando las cosas que más adelante pueden atragantar. Esta tendencia a aumentar la segmentación y perfeccionar al unísono la predicción de salto es algo que luego proseguirá incluso de forma un tanto ya exagerada en la séptima generación por parte de Intel.

predicción de salto

## Síntesis de la trayectoria de AMD

◀ 3.2

A lo largo de la sexta generación, hemos visto a AMD ofrecer una de cal y otra de arena:

- La de arena. El punto débil de la arquitectura K6 era la entrada al procesador, y el K6-2 debió dar cuenta del mismo. En su lugar, se optó por potenciar el conjunto de instrucciones 3DNow! e incorporar sus respectivas unidades funcionales, algo que ni era tan urgente, ni en el mejor de los escenarios posibles hubiera igualado a la productividad de una caché

error inicial

Posición	Operaciones normales	Operaciones multimedia
1	Pentium III	K6-III
2	K6-III	Pentium III
3	Pentium II	K6-2
4	Celeron A	Pentium II
5	K6-2	Celeron A
6	K6	K6
7	MMX y Celeron	MMX y Celeron
8	K5	Pentium y K5
9	Pentium	

TABLA 5.11: Ránking orientativo de los microprocesadores de Intel y AMD de quinta y sexta generación destinados al PC doméstico.

L2 integrada. El escenario que dibujó el paso del tiempo fue bastante más desolador, ya que 3DNow! apenas sí fue utilizado en la práctica. AMD llegó a dotar al procesador de una amplia caché externa, e incluso le subió la frecuencia del bus hasta los 100 MHz, pero llevando al procesador hasta los 500 MHz, ésa era mucha velocidad para luego tener que llevarle la comida desde tan lejos como refleja su multiplicador de reloj de cinco. Como resultado, el polluelo siguió pasando hambre.

- La de cal. Con el K6-III, se asestó un golpe certero: Esa caché L2 integrada de 256 Kbytes es una bomba que impulsa la comida a la boca del procesador prácticamente sin concederle tregua, y como resultado, el diseño alcanza un perfecto equilibrio: Sus fauces tragan sin parar lo que el estómago después digiere con singular presteza, y así, un K6-III de 450 MHz trabaja más del doble de rápido que un K6-2 de 300 MHz.

acierto  
posterior

### 3.3 ► Conclusiones

- 1 AMD se lo había puesto muy fácil a sí mismo. Construyó un modelo en el que el tapón estaba a la entrada de la arquitectura, con la ventaja de que sus mejoras más significativas podían sucederse trabajando desde el exterior, sin necesidad de alterar un ápice el diseño del procesador. Era la forma natural de luchar frente a un gigante como Intel, pues sus cambios no requerían apenas recursos humanos. Sin embargo, parece que esto lo entendieron un poco tarde.
- 2 La caché es menos importante en el modelo arquitectural de Intel que en el de AMD. Un Celeron-A de 450 MHz (caché L2 integrada de 128 Kbytes) se comporta en torno a un 50-60% mejor que un Celeron de 300 MHz (sin caché), mientras que en un K6-III acabamos de ver que esta mejora se traduce en incrementos de rendimiento superiores al 100%. La diferencia lo dice todo.
- 3 El punto de partida del Pentium Pro y el K6 está muy parejo. En la primera avanzadilla, Intel supo dotar de cambios más necesarios a su arquitectura, y el Pentium II para nosotros queda por encima del K6-2. En la siguiente confrontación, Intel vuelve a acertar, pero AMD se recupera a lo grande, y siempre que se comparen modelos de la misma frecuencia, el K6-III tiene una ligera ventaja frente al Pentium III Katmai. Al Coppermine lo vemos ya como un competidor del K7, así que a estos dos los compararemos al finalizar la cobertura del K7 (ver [tabla 6.3](#)).

la entrada

caché

comparativa

pág. 201

A modo de resumen comparativo entre Intel y AMD, hemos incluido en la [tabla 5.11](#) un ránking final orientativo en el que hemos posicionado todos los modelos visitados en la quinta y sexta



generación atendiendo exclusivamente a los criterios de rendimiento ya comentados, y distinguiendo las aplicaciones de usuario convencional y usuario multimedia.

## SECCIÓN 5.4

## Cyrix

El primer microprocesador que Cyrix comercializó en el marco de la sexta generación fue el 6x86 o M1. Este sería el último modelo manufacturado por IBM, ya que en los primeros meses de 1996, Cyrix materializó su fusión con National Semiconductor, y el resto de modelos de este diseño y los sucesivos fueron ya integrados por esta compañía.

6x86  
M1

El M1 era un supuesto competidor del Pentium Pro. Disponía de un doble canal superescalar y ejecución fuera de orden, soporte multiprocesador, y caché L1 de 16 Kbytes unificada para datos e instrucciones.

L1

Los rangos de frecuencias de reloj del M1 fueron un poco extraños. Coexistieron en el mercado versiones de 100, 110, 120, 133 y 150 MHz, siempre con multiplicadores de 2 y algunos con multiplicadores de 3, pero ninguno con multiplicadores 1.5 ó 2.5.

frecuencias

En el debe de este diseño anotamos su excesivo calentamiento, pero a pesar de ello, el rendimiento del M1 era en torno al 25 % más rápido que el Pentium clásico, en parte debido a la mayor frecuencia de su bus local (hasta 75 MHz).

calentamiento

La compañía se fijaba tanto en Intel que sus productos solía etiquetarlos con un código PR-XXX que especificaba: *Pentium Rated XXX MHz*, es decir, procesador equivalente a un Pentium de Intel de XXX MHz, algo que continuó utilizando hasta después de la extinción del Pentium.

PR-XXX

El problema de la temperatura se solucionó con la versión 6x86L (L de *low-voltage, bajo voltaje*), que baja hasta los 2.8 voltios la alimentación del zócalo de forma similar a como lo hizo la versión MMX de Intel. El coste del 6x86L era bastante inferior al de éste último, pero también lo era su rendimiento, sobre todo en lo concerniente a las prestaciones de su unidad de punto flotante o FPU: A frecuencias de 166 MHz quedaba a la altura de un Pentium a 90 MHz, problema éste que no ha podido ser solventado hasta la llegada del MXi.

6x86L

Posteriormente, tenemos el 6x86MX o M2, procesador que ya incorpora la tecnología MMX y puede trabajar a frecuencias de 166 MHz, 200 MHz y 233 MHz. Con posterioridad, el fabricante ampliaría la gama con modelos a frecuencias incluso superiores a los 300 MHz.

6x86MX  
M2

El M2 es, al igual que el K6, compatible con zócalos y placas base para Pentium clásico, e incorpora caché L1 de 64 Kbytes y L2 de 512 Kbytes.

Los microprocesadores de Cyrix siempre han sido tildados de tener una FPU bastante floja, lo que repercute negativamente en aplicaciones como las de diseño asistido por computador (CAD), donde el porcentaje de cálculos en punto flotante es elevado. La compañía quiso lavar su imagen con el lanzamiento del Cayenne o MXi durante el tercer trimestre de 1998. Se trata de una versión más avanzada del procesador M2 con las siguientes prestaciones:

MXi

- ❶ Tecnología de integración de 0.25 micras para 1998, reduciendo el tamaño del chip hasta los 88 mm<sup>2</sup>, y posteriormente migrando a las 0.18 micras en Abril de 1999.
- ❷ Frecuencia de reloj entre 300 MHz y 400 MHz para la versión a 0.25 micras, y superior a los 500 MHz para las versiones a 0.18 micras.
- ❸ 64 Kbytes de caché L1.
- ❹ Bus local a 100 MHz.

nuevas  
operaciones

- ⑤ Mejora al conjunto MMX con 15 nuevas instrucciones orientadas a aritmética en punto flotante (MMXFP) y reproducción de discos DVD (MPEG2).
- ⑥ FPU avanzada capaz de doblar e incluso triplicar las prestaciones actuales.
- ⑦ Prestaciones 3D similares a las de muchas tarjetas 3D actuales, soportando filtros bilineal y trilineal, operaciones de MIP-mapping y alpha-blending, búfer Z para el procesamiento de video, y una velocidad de proceso de 2 millones de triángulos por segundo. Además, es compatible con los lenguajes de diseño gráfico más populares, como Direct3D y OpenGL.

Slot

A partir de aquí, el camino que Cyrix tenía trazado pasaba por utilizar también zócalos de tipo Slot. Para ello, pretendió aprovecharse de un intercambio de patentes firmado con Intel en 1998 que le permitía clonar el bus local del Pentium II y desarrollar así procesadores equivalentes a más bajo coste. Estos modelos podrían además aprovechar las innovaciones y los productos que Intel sacara al mercado como fabricante de placas base, algo que terminaría sucediendo a más largo plazo.

quiebra

A corto plazo, lo que ocurrió fue que las ventas de Cyrix fueron descendiendo desde que se alejara de la órbita de IBM y materializara su fusión con National Semiconductor, durando esta alianza tan sólo tres años. En Mayo de 1999, Cyrix cerraba sus puertas. Unos meses más tarde, la compañía sería adquirida por VIA, retomando ésta la línea de sus diseños actuales (ver [sección 6.7.1](#)).

pág. 245



## Resumen



fin del  
monopolio

Si la quinta generación supuso el boom comercial del PC, la sexta nos regala el final del monopolio Intel. Esto va a traer aires renovados al mercado, y la riqueza formativa de ver en acción nuevas arquitecturas.

pág. 7

caché

En el resumen del [capítulo 1](#) indicamos que las generaciones pares habían focalizado sus optimizaciones sobre el sistema de memoria del PC, y la sexta generación es el mejor exponente de ello: La batalla decisiva entre Intel y AMD se libra en torno a la caché.

modelos

La serie de Intel se compone de su embrión, Pentium Pro, y tres sucesores, Pentium II, Celeron y Pentium III. La serie de AMD queda conformada por los K6, K6-2 y K6-III. En ambos casos, el aspecto más dinamizador es la caché L2, que encontramos en sus tres formas posibles: Externa (Celeron, K6 y K6-2), interna (Pentium Pro, II y los primeros Pentium III) e integrada (en el modelo terminal de cada saga - Pentium III y K6-III).

comparativa

En general, la saga de Intel posee un mejor *front-end* o boca (las fases iniciales de búsqueda y decodificación de instrucciones), mientras que AMD gana la partida en el *back-end* o estómago (las etapas terminales de ejecución y escritura) y necesita más de la caché para encontrar un equilibrio.

tres gamas

Pero es Intel quien descubre aquí un filón que le reporta pingües beneficios: Cómo servirse de una única arquitectura matriz para producir procesadores dirigidos a tres gamas hasta entonces disjuntas: Los servidores, donde emerge el modelo Xeon dotado de una L2 más grande y rápida, la gama media, donde se sitúa el procesador convencional, y la gama baja, en la que la versión Celeron sacrifica caché como solución de bajo coste.

Cuando AMD se da cuenta de ello, la batalla se libraré en estos tres frentes. Pero esa historia pertenece ya a la séptima generación.

## Cuestionario de evaluación

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

1 🗨️ La sexta generación de microprocesadores adopta esta denominación porque

- a Su bus local es de 66 MHz.
- b Sus arquitecturas son todas segmentadas en 6 etapas.
- c Incorpora una unidad funcional más sobre las 5 de que disponen los diseños de quinta generación.
- d Ninguna de las anteriores.

2 🗨️ Tenemos ante nosotros un procesador de quinta generación y otro de sexta generación. A la hora de distinguir cada uno, sabemos que

- a Si vemos uno de ellos en formato cartucho cerámico y que se monta sobre zócalo Slot, ése es el de sexta porque no puede ser de quinta.
- b Si vemos uno de ellos en formato PGA y que se monta sobre zócalo Socket, ése es el de quinta porque no puede ser de sexta.
- c Si vemos uno de ellos con ventilador y disipador de calor acoplados, ése es el de sexta porque no puede ser de quinta.
- d Nunca va a ser posible distinguirlos por su aspecto externo.

3 🗨️ ¿Qué microprocesador de sexta generación fue el primero en incluir instrucciones multimedia?

- a El Pentium Pro.
- b El Pentium MMX.
- c El K6.
- d El K6-2.

4 🗨️ ¿Qué rasgo de un microprocesador para PC se ve por primera vez en el contexto de su sexta generación?

- a Un cauce segmentado en 20 etapas.
- b Una caché integrada L2.
- c Un multiplicador interno para la frecuencia de reloj.

d La extensión multimedia para el conjunto de instrucciones.

5 🗨️ ¿Por qué razón abandonó Intel el zócalo Socket 7 con la llegada de la sexta generación de microprocesadores?

- a Porque es un diseño rudimentario y obsoleto.
- b Porque quería fabricar placas incompatibles con las de sus rivales.
- c Porque introducía interferencias al comportarse sus pistas como antenas.
- d Dos de las respuestas anteriores son ciertas.

6 🗨️ ¿Qué relación guardan los términos SECC y Slot?

- a Son sinónimos.
- b Son antónimos.
- c Un microprocesador que se presenta en formato SECC puede acoplarse sobre un zócalo de tipo Slot.
- d Un microprocesador que se presenta en formato Slot puede acoplarse sobre un zócalo de tipo SECC.

7 🗨️ ¿Cuál es el punto fuerte de los microprocesadores de Intel de sexta generación con respecto a sus competidores AMD y Cyrix?

- a La potencia bruta de cálculo.
- b La mayor caché interna de segundo nivel.
- c La mayor disponibilidad de placas base sobre las que montarlos.
- d En las tres variables anteriores Intel llevaba ventaja.

8 🗨️ ¿Cómo resumirías una arquitectura AMD de sexta generación?

- a Mayor capacidad para la decodificación que para la ejecución de instrucciones.

**b** Mayor capacidad para la ejecución que para la decodificación de instrucciones.

**c** Empleo muy acusado de la segmentación.

**d** Empleo de cachés integradas de segundo nivel (L2).

**9** ¿Qué fabricante de microprocesadores para PC consiguió en la sexta generación una arquitectura en la que el rendimiento del equipo mostraba poca dependencia respecto a la frecuencia del bus local?

**a** Intel.

**b** AMD.

**c** Cyrix.

**d** IDT.

**10** ¿Qué compañía cruzó por primera vez la barrera de los 500 MHz en un microprocesador?

**a** Intel.

**b** AMD.

**c** Digital.

**d** Motorola.

**11** “El Pentium Pro es más rápido que el Pentium clásico”.

**a** No, depende de la frecuencia de reloj de cada procesador.

**b** Sí, siempre que se utilice un código de 32 bits.

**c** Sí, siempre que el código no tenga dependencias de datos.

**d** Todas las respuestas anteriores son correctas.

**12** ¿Por qué le llevó tanto tiempo a Intel desarrollar el Pentium Pro y tan poco tiempo a AMD desarrollar el K6?

**a** El Pentium Pro era una arquitectura nueva que partió prácticamente de cero, mientras que el K6 se basó en otra ya existente.

**b** El Pentium Pro contiene muchos más transistores.

**c** El Pentium Pro es una arquitectura mucho más completa.

**d** El K6 no es realmente un microprocesador, sino un coprocesador.

**13** ¿Cuál fue la razón de ser del primer microprocesador Celeron de Intel?

**a** Ofrecer mayor rendimiento a menor coste.

**b** Precios más bajos para extinguir los zócalos Socket 7.

**c** Eliminar la caché interna para reducir el tamaño del microprocesador.

**d** Las tres anteriores.

**14** ¿Qué distingue a un microprocesador Celeron de un Celeron A?

**a** La frecuencia de reloj.

**b** El tamaño de caché interna.

**c** La frecuencia del bus local.

**d** El zócalo al que se acopla.

**15** Disponemos de un modelo de procesador Celeron. Sabemos que es...

**a** Un Pentium II sin caché.

**b** Un Pentium III con menos caché.

**c** Un Pentium III con el bus local más lento.

**d** Un procesador derivado de otro de gama más alta al que se recorta el tamaño de caché y adicionalmente la velocidad del bus local.

**16** ¿Por qué el Celeron es el procesador de Intel que más tiempo lleva conviviendo con nosotros (desde comienzos de 1998 hasta finales de 2001 van casi cuatro años)?

**a** Porque a Intel le ha interesado desde entonces mantener siempre un modelo de bajo coste para responder al perfil de usuario menos exigente.

**b** Porque es el producto que permite dar salida a las muestras de otros chips que presentaron algún error en su integración.

**c** En realidad no es un único procesador, sino versiones reducidas de la arquitectura estrella de Intel en cada período temporal, y por lo tanto, se encuentra sujeto a los cambios producidos en éstas.

**d** Todas las respuestas anteriores son correctas.

**17** ¿Qué alternativa abarata más el precio de un Pentium II 300MHz?

**a** Subir su tecnología de integración de 0.25 a 0.35 micras.

**b** Prescindir de su caché interna de tercer nivel (L3).

**c** Bajar su frecuencia a 233MHz.

**d** Esperar seis meses para comprarlo.

**18** ¿Qué distingue internamente al Pentium II 333 MHz del Pentium II 350 MHz?

- a** El controlador del bus local.
- b** El controlador de la memoria caché interna.
- c** El diseño del cauce segmentado.
- d** El factor de superescalaridad.

**19** ¿Qué rasgo diferencia al Pentium II Xeon de un Pentium II Deschutes?

- a** El bus local a 100 MHz.
- b** El zócalo Slot 2.
- c** La caché L2 interna de mayor tamaño y sincronizada a la frecuencia del micro.
- d** La tecnología de integración a 0.25 micras.

**20** Resume la configuración de un Pentium III 500 MHz (buses de datos y dirs, caché L1, caché L2, frec. bus, micras).

- a** 64 y 32, 16Kb y 16Kb, 512Kb, 66MHz, 0.35.
- b** 32 y 64, 32Kb y 32Kb, 512Kb, 100MHz, 0.25.
- c** 64 y 32, 16Kb y 16Kb, 512Kb, 100MHz, 0.25.
- d** 32 y 64, 32Kb y 32Kb, 512Kb, 133MHz, 0.18.

**21** ¿En qué evolucionaron los Pentium III respecto a los Pentium II?

- a** En que pueden ejecutar hasta tres instrucciones por ciclo de reloj, frente a dos en el Pentium II.
- b** En que disponen de tres unidades MMX en lugar de dos.
- c** En que tienen tres cachés internas en lugar de dos.
- d** En ninguna de las tres características anteriores.

**22** ¿Qué diferencia hay entre un Pentium III 450 MHz y un Pentium III 500 MHz?

- a** El número de transistores que contiene.
- b** Lo bien que se han integrado sus transistores en la planta de fabricación.
- c** La distancia de integración.
- d** El voltaje nominal.

**23** ¿Dónde se encuentra la caché L2 del microprocesador Pentium III?

- a** Dentro del chip en el que se integra su núcleo de ejecución, como la L1.
- b** Fuera de ese chip, pero dentro de la placa de circuito impreso recubierta a modo de cartucho cerámico que constituye el producto comercial.
- c** En la placa base.
- d** c seguro que es falsa. Elegiría a o b dependiendo del formato de presentación del microprocesador (FC-PGA o SECC2, respectivamente).

**24** Si Intel fabricase su viejo Pentium III de 500 MHz con el nuevo proceso de integración de 0.18 micras, ¿Cómo le resultaría comparativamente atendiendo a la variable coste por unidad fabricada?

- a** Más caro, puesto que cada transistor a 0.18 micras es de una calidad superior.
- b** Igual, pues ambos tienen el mismo número de transistores.
- c** Igual, pues ambos tienen la misma frecuencia.
- d** Más barato, pues el espacio ocupado es inferior y, por tanto, el gasto de materia prima también lo será.

**25** Matiza la frase siguiente: El microprocesador K6-2 dispone de caché propia.

- a** Si se está refiriendo a la caché L1, es cierta.
- b** Si se está refiriendo a la caché L2, es falsa.
- c** Las dos respuestas anteriores son correctas.
- d** Todo depende de si la caché interna de un procesador se considera caché propia o no.

**26** Resume la configuración de un K6-III 400 MHz (zócalo, voltaje, caché L2, frec. bus, micras).

- a** Socket 7, 2.7v, 512Kb, 100MHz, 0.35.
- b** Slot A, 2v, 512Kb, 100MHz, 0.25.
- c** Slot A, 1.8v, 512Kb, 133MHz, 0.25.
- d** Super 7, 2.4v, 256Kb, 100MHz, 0.25.

**27** ¿Qué microprocesador mantiene el mismo núcleo arquitectural respecto a su predecesor?

- a** El Pentium del K5.

**b** El K6-2 del K6.

**c** El Itanium del Pentium III.

**d** El K7 del Pentium III.

**28** 🗨️ Si comparamos un Pentium III frente a un K6-III y un Pentium II frente a un K6-2, ¿Qué arquitecturas resultan más recomendables?

**a** Los Pentium, porque resultan más fiables.

**b** Los Pentium, porque ofrecen una compatibilidad superior con los modelos futuros.

**c** Los K6, porque tienen una frecuencia superior.

**d** Los Pentium para operaciones normales y los K6 para operaciones multimedia.

**29** 🗨️ ¿Son parecidas las innovaciones introducidas por el Pentium II y el K6-2 respecto a sus predecesores y las del Pentium III y el K6-III respecto a éstos?

**a** Depende de qué modelos de Pentium II y III estemos hablando. Si se trata del Deschutes y el Coppermine, sí.

**b** Sí. Los primeros mejoran el conjunto multimedia, y los segundos, la caché L2, que pasa a ser integrada de 256 Kbytes.

**c** La respuesta anterior es correcta, aunque conviene matizar que en el caso del K6 se pasa directamente de caché externa a integrada.

**d** No.

**30** 🗨️ ¿Cuál ha sido la progresión de la saga de procesadores Pentium al respecto de las instrucciones multimedia (tómese como referencia el primer modelo aparecido de cada versión)?

**a** El Pentium II tiene 57 más que el Pentium, y el Pentium III, 70 más que el Pentium II.

**b** Todos aceptan el mismo conjunto de instrucciones.

**c** El Pentium II tiene el mismo conjunto de instrucciones que el Pentium, y el Pentium III dispone de 70 más que éstos.

**d** El Pentium II dispone de 57 instrucciones más que el Pentium, y el Pentium III incorpora 13 más sobre éstas hasta completar las 70.

**31** 🗨️ ¿Qué pareja de microprocesadores presenta mayores similitudes atendiendo a su núcleo interno (decodificación y ejecución de instrucciones)?

**a** El K5 y el K6.

**b** El K6 y el K6-2.

**c** El K6-2 y el K6-III.

**d** El K6-III y el K7.

**32** 🗨️ ¿Qué pareja de microprocesadores presenta mayores divergencias atendiendo a su arquitectura interna?

**a** El Pentium y el Pentium MMX.

**b** El Pentium Pro y el Pentium II.

**c** El Pentium II y el Pentium III.

**d** El Pentium III y el Itanium.

**33** 🗨️ ¿Qué unidad funcional requiere un mayor número de transistores a la hora de su integración en silicio?

**a** La unidad de ejecución 3DNow! del K6 original.

**b** La caché L2 integrada del K6-2.

**c** El controlador de bus local de hasta 200 MHz en el K6-III.

**d** Ninguno de los tres procesadores anteriores incluye las unidades de ejecución mencionadas en sus respectivas respuestas.

**34** 🗨️ ¿Qué unidad funcional requiere un mayor número de transistores a la hora de su integración en un chip?

**a** La unidad de ejecución 3DNow! del K6-2.

**b** La caché L2 integrada del K6-III.

**c** El controlador de bus local de 200 MHz del K7.

**d** Ninguno de los tres procesadores anteriores contiene las unidades funcionales que se le presuponen.

**35** 🗨️ Uno de los escasos ejemplos que conocemos de paso directo de caché L2 externa a integrada en dos modelos consecutivos de una misma familia de procesadores se produce en la transición

**a** Del K5 al primer K6.

**b** Del primer Pentium II al último Pentium III.

**c** Del K6-2 al K6-III.

**d** Del K7 Athlon al K7 Thunderbird.

# La séptima generación

## Sumario

<b>6.1. K7 de AMD</b>	<b>192</b>
6.1.1. Frecuencia de reloj	192
6.1.2. El bus local	192
6.1.3. Tecnología de integración	193
6.1.4. Zócalo de la placa base	193
6.1.5. Memoria principal	194
6.1.6. Memoria caché	194
6.1.7. El entendimiento con el sistema de memoria	194
6.1.8. El conjunto de instrucciones	195
6.1.9. Segmentación	197
6.1.10. Superescalaridad	197
6.1.11. Resolución de dependencias	197
6.1.11.1. Dependencias de datos	197
6.1.11.2. Dependencias de control	198
6.1.12. Instrucciones multimedia	198
6.1.13. Unidades de ejecución	198
6.1.14. Valoración final	198
6.1.14.1. Lo mejor	199
6.1.14.2. Lo peor	199
<b>6.2. Thunderbird y Duron: Las 0.18 micras en AMD</b>	<b>199</b>
6.2.1. Thunderbird	199
6.2.2. Duron	200
6.2.3. Duron con aditivos	201
<b>6.3. Athlon XP</b>	<b>202</b>
6.3.1. Frecuencia de reloj	202
6.3.2. El bus local y sus diálogos con memoria principal	204
6.3.3. Tecnología de integración	204
6.3.4. El jeroglífico de la frecuencia	206
6.3.5. Paralelismo a nivel de instrucción	206
6.3.6. Memoria caché	206
6.3.7. Conjunto de instrucciones	207
6.3.8. Valoración final	207
6.3.8.1. Lo mejor	207

6.3.8.2. Lo peor . . . . .	208
<b>6.4. Pentium 4 de Intel . . . . .</b>	<b>208</b>
6.4.1. Nueva arquitectura . . . . .	209
6.4.2. Frecuencia de reloj . . . . .	210
6.4.3. Tecnología de integración . . . . .	210
6.4.4. El bus local . . . . .	211
6.4.5. Diálogos con memoria principal . . . . .	212
6.4.5.1. Sobre SDRAM y DDRAM . . . . .	213
6.4.5.2. Sobre RDRAM . . . . .	215
6.4.6. El conjunto de instrucciones y su decodificación . . . . .	215
6.4.7. Memoria interna . . . . .	216
6.4.7.1. Bancos de registros . . . . .	216
6.4.7.2. Caché de traza, o L1I . . . . .	216
6.4.7.3. Caché de datos (L1D) . . . . .	219
6.4.7.4. Caché de segundo nivel (L2) . . . . .	219
6.4.8. Segmentación . . . . .	219
6.4.9. Superescalaridad . . . . .	221
6.4.10. Resolución de dependencias . . . . .	221
6.4.11. Instrucciones multimedia . . . . .	222
6.4.12. Unidades de ejecución . . . . .	224
6.4.13. Valoración final . . . . .	224
6.4.13.1. Lo mejor . . . . .	224
6.4.13.2. Lo peor . . . . .	225
<b>6.5. Northwood: Las 0.13 micras en Intel . . . . .</b>	<b>226</b>
6.5.1. Mejoras introducidas . . . . .	226
6.5.2. Los diálogos con memoria principal . . . . .	227
6.5.2.1. El interfaz con DDRAM . . . . .	228
6.5.2.2. El interfaz con RDRAM . . . . .	228
6.5.3. HyperThreading . . . . .	229
<b>6.6. Comparativa: K7 versus Pentium 4 . . . . .</b>	<b>231</b>
6.6.1. Ejecución de una instrucción . . . . .	232
6.6.1.1. Fase de búsqueda de instrucción . . . . .	232
6.6.1.2. Fase de decodificación de instrucción . . . . .	232
6.6.1.3. Fase de planificación y reordenación de instrucciones . . . . .	234
6.6.1.4. Fase de búsqueda de operandos . . . . .	235
6.6.1.5. Fase de ejecución . . . . .	235
6.6.1.6. Fase de escritura . . . . .	236
6.6.2. Frecuencia . . . . .	237
6.6.3. Paralelismo a nivel de instrucción . . . . .	237
6.6.4. Incidencia de las dependencias . . . . .	239
6.6.4.1. Dependencias de datos . . . . .	239
6.6.4.2. Dependencias de control . . . . .	241
6.6.5. Acceso a memoria . . . . .	241
6.6.6. Valoración final . . . . .	242
6.6.7. Comparativa a 0.13 micras . . . . .	244
<b>6.7. Otros fabricantes . . . . .</b>	<b>245</b>
6.7.1. Cyrix . . . . .	245
<b>Resumen . . . . .</b>	<b>245</b>
<b>La anécdota: El curso cíclico de la historia . . . . .</b>	<b>246</b>
<b>Cuestionario de evaluación . . . . .</b>	<b>247</b>



	Nombre comercial	Fecha lanz.	T. l.	M. T.	N. M.	Frec. (GHz)	F. bus (MHz)	Caché		Segmento	Código de referencia
								L2	L3		
AMD	K7/Athlon	08/99	250	22	6	0.5/1	200	Sí	No	Media	Argon
	K7/Athlon	06/00	180	37	6	0.65/1.4	200/266	256	No	Media	Thunderbird
	Duron	06/00	180	25	6	0.6/1	200	64	No	Baja	Spitfire
	Duron+DDR	08/01	180	25.2	6	.95/1.3	200	64	No	Baja	Morgan
	Athlon XP	10/01	180	37.5	7	1.33/1.7	266	256	No	Media	Palomino
	Athlon XP	06/02	130	37.2	8	1.33/1.8	266	256	No	Media	Thoroughbred A
	Athlon XP	08/02	130	37.6	9	2.0/2.4	266/333	256	No	Media	Thoroughbred B
	Athlon XP	03/03	130	53.9	9	2.2/2.5	333	512	No	Media	Barton
	Athlon MP	04/02	180	37.5	7	1.33/1.8	266	256	No	Alta	Palomino
Intel	Pentium 4	11/00	180	42	7	1.3/2	400	256	No	Media	Willamette
	Pentium 4	08/01	130	55	7	2/3.2	400/533	512	No	Media	Northwood
	Pentium 5 (*)	09/03	90	100	7	4	667/800	1024	No	Media	Prescott
	Pentium 5 (*)	03/04	90	100	7	5.6	1066	1024	No	Media	Tejas
	P. 4 Celeron	05/02	180	N/D	7	1.7/1.8	400	128	No	Baja	Willamette
	P. 4 Celeron	09/02	130	N/D	7	2/2.2	400	256	No	Baja	Northwood
	P. 4 Xeon	05/01	180	42	7	1.4/2	400	256	Sí	Alta	Willamette
	P. 4 Xeon	01/02	130	55	7	1.8/2.8	400	512	Sí	Alta	Prestonia
	P. 4 Xeon MP	03/02	180	108	7	1.4/2	400	256	1024	Alta	Foster
	P. 4 Xeon MP	11/02	130	108	7	2/2.8	533	512	1024	Alta	Gallatin
	P. 4 Xeon MP	2005	90	N/D	7	>3	666/800	1024	N/D	Alta	Nocona

**TABLA 6.1:** Nombre comercial y código de referencia para los procesadores de séptima generación. La casilla T.I. indica el proceso de fabricación en nanómetros para los transistores, cuyo número en millones se adjunta en la casilla contigua, M.T.; la casilla siguiente, N.M., muestra los niveles de metalización empleados para sus interconexiones. Para la caché, damos el tamaño sólo cuando es integrada (el “Sí” significa “Interna”), en cuyo caso contabilizamos sus transistores en el montante total del chip. La gama baja corresponde a los PC de precio en torno a los 500-600 €, mientras que la gama alta comienza en los 3000 €. (\*) El cambio de numeración no está confirmado aún.

Hasta el final del siglo pasado, las distintas generaciones de procesadores estuvieron sincronizadas entre sus fabricantes, con una vigencia media de entre tres y cuatro años. La fecha de inicio venía marcada por Intel, que fue siempre el primero en golpear durante los cinco asaltos que duró su combate con Motorola (consultar fechas en la [tabla 2.2](#)).

El cambio de púgil provocó cierta convulsión sobre este método *de facto* al que la industria se había acomodado. Nada más tomar el relevo, AMD ya avisó de sus intenciones anticipándose al Pentium II en el que fue el primer producto de ambas compañías para el PC doméstico de sexta generación.

En la séptima generación, el adelanto de AMD es ya de 18 meses, lo que supone casi la mitad del período de vigencia de esta generación. Esto rompe la sincronización temporal y dificulta nuestra cobertura de las diferentes arquitecturas, puesto que ya no es posible establecer directrices generales comunes a cada generación: El K7 se va a pasar la mitad de su vida luchando frente al Pentium III, y la otra mitad frente al Pentium 4. ¿Cuál de ellos es realmente su rival?

Tan injusto para Intel resulta confrontar su Pentium III con el K7, como para AMD rivalizar con el Pentium 4. Y tan difícil le será a éste encajar los golpes que le asesta el K8 <sup>1</sup> en 2003, como a éste someterse al que luego acuda en su defensa. Estos episodios recuerdan a una pelea barriobajera

<sup>1</sup>El código de referencia del K8, *Sledgehammer*, cuyo significado literal es *maza rompedorras*, trata de poner en aviso a Intel acerca de lo que supuestamente se le viene encima.

conductas previas

☛ pág. 38

ese alumno aventajado

asincronismo temporal

enfrentando a AMD e Intel

que empieza entre chavales, prosigue con primos ya curtidos, y acaba enzarzando a los patriarcas de cada familia. Al final, pierden los dos, y ganamos todos los demás, esto es, el mercado, que se beneficia de la feroz competencia adquiriendo potentes PC a precio de saldo.

Lo que no ha cambiado hasta ahora es que cada firma presenta una sola arquitectura en cada generación para vestir sus PC, y que todos sus modelos se derivan de ella, con retoques más o menos livianos. Cierto es que en esta generación Intel dispone de una segunda arquitectura, el Itanium, un procesador VLIW (*Very Long Instruction Word* - ver [sección 3.5.3](#)) con conjunto de instrucciones propio, el IA-64. Pero es una creación pensada para coexistir con el Pentium 4 en otro nicho de mercado diferente: El mundo de las estaciones de trabajo y los servidores, razón por la cual no tiene cabida en un libro como éste.

un estandarte  
por bando.  
Itanium  
pág. 104

pág. 191

visión de  
conjunto

Volviendo ya al universo de nuestros PC, las dos arquitecturas de esta generación son el K7 por parte de AMD y el Pentium 4 por parte de Intel. La [tabla 6.1](#) nos presenta un resumen preliminar de todos los modelos que cada firma ha derivado de ellos, como siempre aderezados con divertidos códigos de referencia o *codenames* (por cierto, que Palomino y Thoroughbred son variedades equinas en el contexto de los pura sangre). Establecida esta visión de conjunto, procederemos analizando primero la arquitectura base de cada firma, que actuará como eje principal para luego desmenuzar todas las variantes que se han sucedido.

## SECCIÓN 6.1

### K7 de AMD

Argon Durante el verano de 1999, AMD lanzó al mercado el que se ha convertido en el emblema de la compañía: Su K7 o Athlon, con código de referencia Argon.

cuota de mercado AMD planeaba obtener con este procesador una cuota de mercado en torno al 30 % para el año 2001, y el objetivo fue conseguido en bastantes países, consolidando a la firma a escala mundial y abriéndole un buen número de mercados.

como opción de compra Ya apuntamos en nuestra primera edición de *Arquitectura del PC* (1999), que este procesador sería durante algún tiempo la mejor opción de compra atendiendo al ratio rendimiento/coste. Lo que no podíamos imaginar era que semejante venia del mercado duraría hasta 2002.

#### 1.1 ▶ Frecuencia de reloj

dos relevos El K7 fue lanzado al mercado a una frecuencia inicial de 500 MHz, ampliándose sucesivamente de 50 en 50 MHz hasta alcanzar el gigahercio. En ese hito toma el relevo la versión de 0.18 micras, Thunderbird, que acelera a pasos de 66 MHz hasta alcanzar los 1.4 GHz, y posteriormente, el Athlon XP de 0.13 micras para llevarlo en volandas hasta los 3 GHz.

multiplicadores El problema que plantea toda esta evolución son los multiplicadores tan elevados que necesita para cubrir el desfase con el bus local de su placa base. Cuando surgió el procesador, el controlador de su bus local permitía trabajar a 200 MHz, a razón de una doble respuesta en flanco de subida y bajada sobre un reloj de 100 MHz. En la práctica, sin embargo, pocas placas base incorporaron semejante prestación, pues se encontraban varadas en los 133 MHz del Pentium III. Con Intel y la memoria SDRAM ceñidas a esa frecuencia, pocos fabricantes de juegos de chips iban a pasarse a los 200 MHz dando de lado a la memoria y al 80 % de la cuota de mercado de Intel.

2x100 MHz

#### 1.2 ▶ El bus local

mejoras en el bus En Febrero de 2001, la compañía anunció mejoras en la frecuencia del bus, a razón de 2x133 MHz. A éste sí se le pudo sacar todo su jugo, pero porque Intel había lanzado ya su Pentium 4

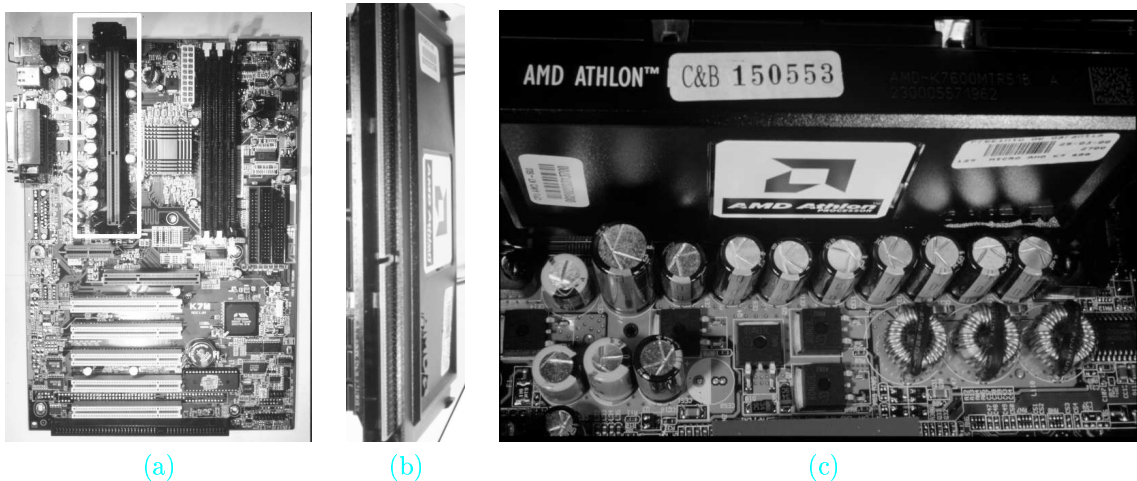


FOTO 6.1: El procesador Athlon de 0.25 micras de AMD. (a) Una placa base, donde hemos recuadrado su zócalo Slot A. (b) Vista del procesador desde un plano inferior, donde se aprecian los contactos con el zócalo. (c) El procesador acoplado a su zócalo y rodeado de su sempiterna flota de condensadores.

con el bus local a 4x100 MHz y los fabricantes de memoria principal (con DDRAM - ver [sección 10.13.5](#)) y tarjetas gráficas (con AGP 2.0 - ver [sección 16.1](#)) habían emprendido la estampida detrás de él.

☛ p. 68/Vol. 2

☛ p. 220/Vol. 2

Con esta ayuda, y el inestimable apoyo de una hermosa caché L2 integrada de 256 Kbytes, se pudo alimentar al K7 hasta con 72 instrucciones del tipo 80x86 de forma simultánea, en lo que es su techo interno por lo que respecta al paralelismo a nivel de instrucción, la cualidad más infrutilizada del procesador hasta ese momento.

conurrencia  
máxima

## Tecnología de integración

◀ 1.3

Las primeras versiones del K7 se fabricaron a 0.25 micras con interconexiones de aluminio, para posteriormente pasar a 0.18 micras, y casi de forma inmediata, también al cobre, trayendo consigo la migración del formato Slot A al formato Socket A en el subsiguiente Thunderbird. El posterior Athlon XP es también muy parecido en aspecto a este último (ver [foto 6.2](#)).

☛ pág. 202

El área de integración fué bastante reducida, tan sólo 184 mm<sup>2</sup> para la versión de 0.25 micras, con un total de 22 millones de transistores. En estos números no contabilizamos la caché L2, que al ser interna va en un chip aparte.

área de  
integración

## Zócalo de la placa base

◀ 1.4

El aspecto externo del primer K7 fue muy parecido al de un Pentium II, con un zócalo mecánicamente idéntico al Slot 1 denominado **Slot A** (ver [foto 6.1](#)), pero que integraría la especificación de bus que montaban los microprocesadores Alpha de Digital fabricados bajo tecnología EV6. De hecho, al frente del diseño de este procesador, estuvo el propio Dirk Meyer, ingeniero jefe del mítico procesador Alpha, fichado de Compaq (la empresa propietaria de Digital).

Slot A

Respecto al protocolo de bus, es muy diferente al AGTL+ de Intel, pues habilita modos de transferencias en ráfaga de 64 bytes, código de corrección de errores ECC de 8 bits para datos e instrucciones por bloques de transferencia de 64 bits, señales de bajo voltaje, y la habilidad para direccionar más de 8 Terabytes de memoria física habilitando las líneas de paridad para direccionamiento.

protocolo  
de bus

**incompatibilidad** Este diferente protocolo de bus hizo que ningún microprocesador bajo Slot 1 de la familia Pentium funcionase sobre Slot A y viceversa. La jugada, pensada para obligar a los usuarios que quisieran aprovechar esta característica a comprar el microprocesador también a AMD, salió justo al revés de como se pronosticó, pues en la práctica resultó que el usuario seleccionaba el Athlon antes que la placa base, pero luego se encontraba que era complicado dar con una que sacara partido a sus prestaciones.

## 1.5 ▶ Memoria principal

**p. 73/Vo1.2**  
**ancho de banda** En el rendimiento de los buses de la arquitectura del K7 tenía un peso enorme la relación que el procesador tuviese con memoria principal. Con un bus de 200 MHz, se pudo haber sacado mucho partido a memorias como la RDRAM (ver [sección 10.13.6](#)), puesto que su primera versión estaba ya en el mercado a 800 MHz y 16 bits de anchura, dando el mismo ancho de banda que admitía el K7: 1.6 Gbytes/sg. (el bus del procesador es cuatro veces más lento, pero también cuatro veces más ancho).

Comparativamente, el bus de 100 MHz de los últimos Pentium II y primeros Pentium III de Intel sólo recogía la mitad de este aluvión de datos (800 MBytes/sg.), y el bus local a 133 MHz del Pentium III Coppermine se quedaba en los 1066 Mbytes/segundo.

**intereses corporativos** Pero si AMD renunció a ese incremento del rendimiento fue porque Intel estaba ya posicionada junto a Rambus, la firma que diseña la memoria RDRAM. Y también porque no recibió el apoyo que se esperaba por parte de los fabricantes de placas base y juegos de chips: La gran mayoría de configuraciones de K7 que datan de 1999 y 2000 están montadas sobre juegos de chips que funcionan a 133 MHz sincronizados con una memoria principal SDRAM también de 133 MHz, y así se desmerecen las tremendas cualidades que la arquitectura atesora para las comunicaciones externas.

## 1.6 ▶ Memoria caché

**L1**  
**pág. 196** El K7 dispone inicialmente de 128 Kbytes de caché L1, divididos en 64 Kbytes para datos y 64 Kbytes para instrucciones. Ambas son gemelas, con una organización asociativa de 1024 conjuntos, 2 líneas por conjunto y 32 bytes por línea. En este aspecto, el procesador también es un calco de lo que implementaba el procesador Alpha de Digital. La [figura 6.2](#) muestra el diagrama de bloques de la arquitectura.

**tamaño** El gran tamaño de la caché L1 es uno de los principales requerimientos para conseguir una configuración compensada a elevadas frecuencias de reloj. Ninguno de los diseños de Intel, Pentium 4 incluido, ha conseguido igualar este tamaño, alcanzando como mucho la cuarta parte de esta cantidad.

**L2**  
**pág. 199**  
**variantes** La caché L2 se ubica de forma interna (tan sólo el controlador de caché L2 es integrado), aunque todo esto cambiará con la llegada de las 0.18 micras (ver [sección 6.2.1](#)). Se contempla un enorme abanico de posibilidades para la L2 que combinan distintas velocidades y tamaños: Comenzando con un tamaño de 512 Kbytes y sincronización a 1/3 de la velocidad del procesador, posteriormente se desarrollaron cachés más veloces que apenas tuvieron calado en el mercado.

## 1.7 ▶ El entendimiento con el sistema de memoria

**pág. 195** La interrelación entre el procesador y los distintos niveles de la jerarquía de memoria es la siguiente (ver [figura 6.1](#)):

- CPU**
- ① El procesador tiene una anchura de 128 bits tanto en el bus interno que lo conecta con la caché L1 de instrucciones (el código de operación es de esta longitud), como en el que lo conecta con la caché L1 de datos (se habilita un doble puerto en su conexión con las unidades

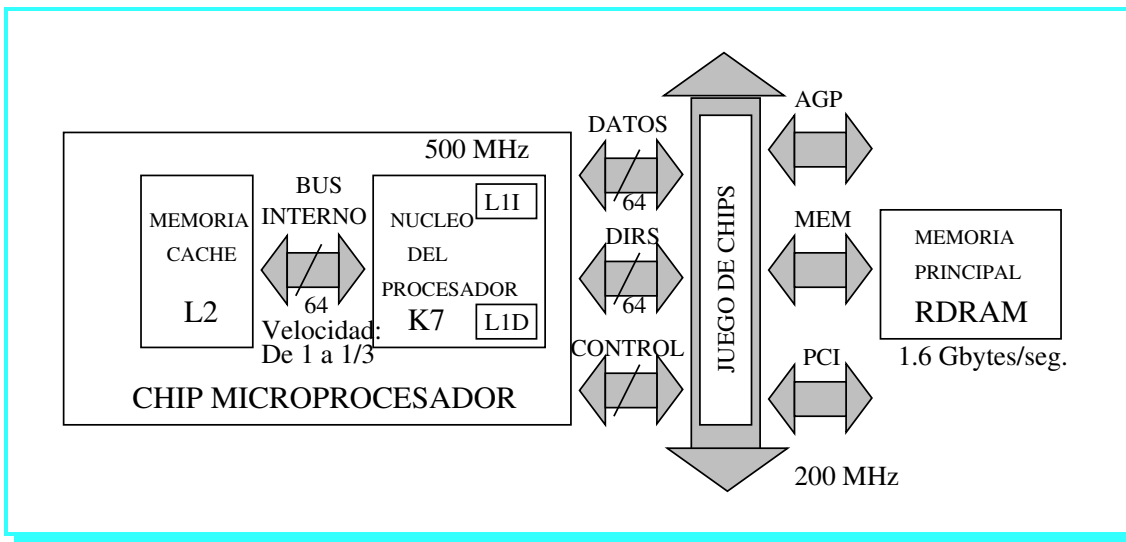


FIGURA 6.1: Relación entre el procesador K7 y los diferentes niveles que componen su jerarquía de memoria.

funcionales que permite la lectura/escritura simultánea de dos datos de 64 bits).

- ② Las dos cachés L1 tienen un tamaño de línea de 32 bytes (256 bits), al igual que la caché L2, por lo que la correspondencia entre ambas es inmediata. L1
- ③ La caché L2 se conecta a memoria principal por el bus de datos del sistema, que tiene una anchura de 64 bits. Entre caché y memoria principal, la unidad lógica de transmisión de datos es siempre el tamaño de línea (256 bits), mientras que la línea física de transmisión sobre la que se implementa es de 64 bits (anchura del bus). En estas condiciones, la mejor estrategia que puede implementar el controlador de memoria presente en la placa base es un de factor cuatro para los módulos de memoria principal. La técnica del entrelazado se describe con detalle en la [sección 10.11](#). L2

entrelazado  
 p. 48/Vol. 2

### El conjunto de instrucciones

### 1.8

El K7 acepta código máquina compatible x86, con un código de instrucción de 128 bits (16 bytes) en el que pueden estar presentes hasta tres instrucciones del formato de instrucción variable que vimos en el Pentium. Esta instrucción múltiple recibe el nombre de macroinstrucción.

código de instrucción

Se establece así una taxonomía en la que cada macroinstrucción contiene hasta tres instrucciones y éstas a su vez hasta tres microinstrucciones, resultando un total de nueve por cada ciclo de reloj en este último nivel. Cada una de las macroinstrucciones está compuesta por una instrucción aritmética seguida de otra de carga de datos procedentes de memoria, y eventualmente, una tercera encargada del almacenamiento en memoria.

macroinstr.  
 microinstr.

Todo esto aboga por un sistema de decodificación de instrucciones ciertamente complejo, que funciona de la siguiente manera:

decodificación

- ① **Direccionamiento del código.** El contador de programa referencia a una posición de memoria virtual que se traduce a dirección física ayudándose de una TLB (Translation Look-Aside Buffer - caché conteniendo los pares de traducciones de este tipo más recientemente utilizados por el procesador) que presenta una implementación multinivel de 24 y 32 entradas en el primer nivel para instrucciones y datos, y de 256 en el segundo nivel para cada caso. TLB

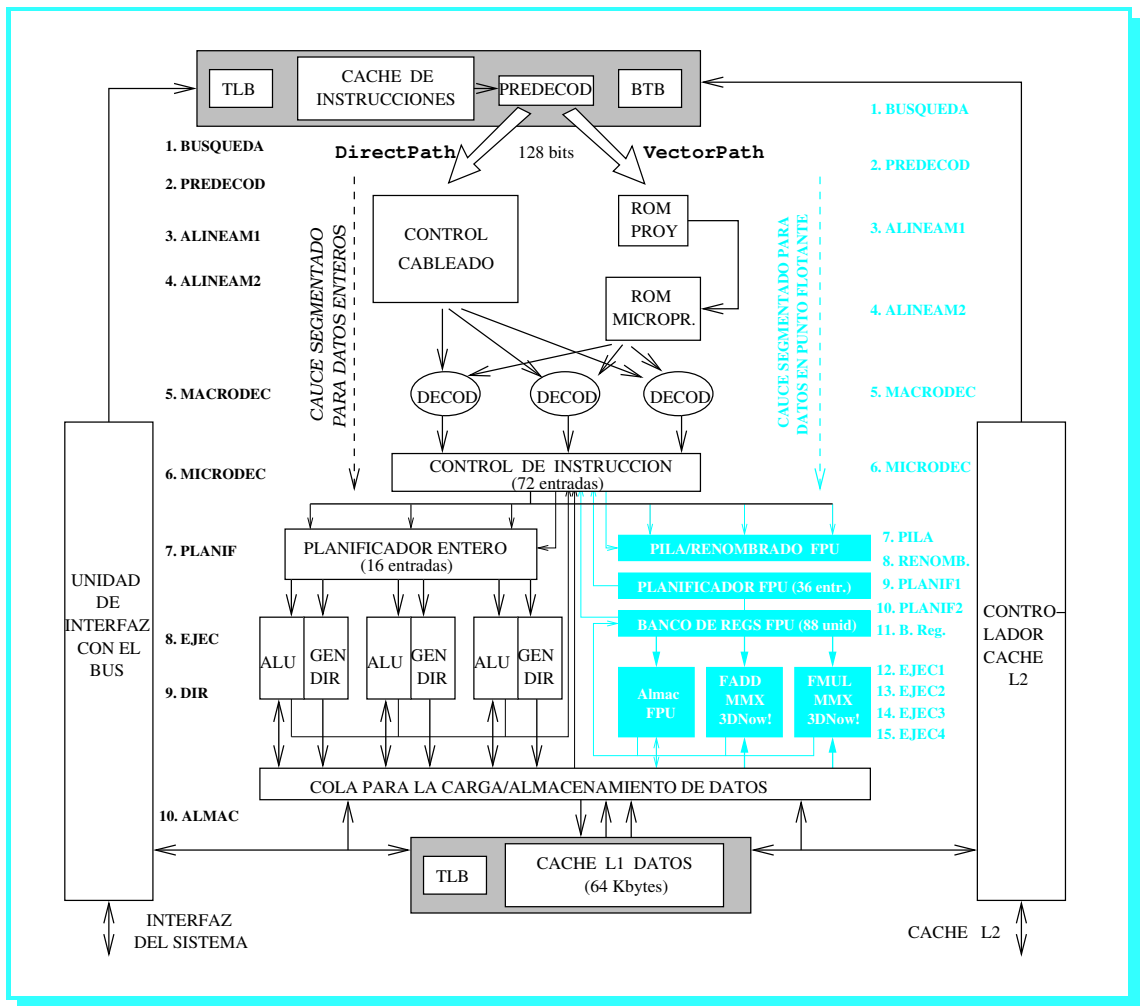


FIGURA 6.2: Diagrama de bloques del microprocesador K7 de AMD, con su complejo sistema de decodificación en paralelo y sus dos cauces de ejecución - uno para enteros y otro para números de punto flotante.

- ❷ **Obtención del código de instrucción.** Con esta dirección física se accede a la caché de instrucciones, de donde se obtiene el código de instrucción de 128 bits.
- ❸ **Predecodificación.** Se pasa el código a un cauce segmentado denominado DirectPath si éste contiene instrucciones sencillas, o a otro denominado VectorPath si contiene instrucciones más complejas.
- ❹ **Macrodecodificación.** En el primer caso, la decodificación y el control se realizan de forma rápida y cableada como en un procesador RISC; en el segundo, la realiza una ROM de proyección que direcciona a su vez a una ROM de microprograma (ver parte superior de la figura 6.2).
- ❺ **Microdecodificación.** A partir de aquí, tres decodificadores de instrucción funcionan en paralelo para traducir individualmente y de forma simultánea los códigos de control de cada una de las tres instrucciones que componen el código de 128 bits. A la conclusión de esta etapa, se han obtenido las microinstrucciones que se encargan de controlar la ejecución de la instrucción en lo sucesivo. Estas microinstrucciones llegan a la unidad de control de instrucción, cuya misión es situar cada instrucción en el cauce segmentado correspondiente según su tipo: Entero o multimedia/punto-flotante.

L1I

DirectPath  
VectorPath

ROM de  
microprograma

microinstr.

Unidad Control

El proceso de decodificación es ciertamente costoso en términos de las unidades involucradas y el tiempo de proceso que lleva, pero va a tener su recompensa más adelante como enseguida veremos.

## Segmentación

◀ 1.9

Uno de los aspectos más novedosos de la arquitectura K7 es el tratamiento que hace de los cauces de segmentación. En todos los procesadores de sexta y séptima generación, las instrucciones enteras y de punto flotante comparten todas las etapas hasta las proximidades de la fase de ejecución, en las que la escisión es inevitable por la complejidad de los cálculos de punto flotante.

novedad

En el K7, en cambio, las instrucciones de cada tipo se separan nada más finalizar la etapa de decodificación, y eso posibilita un tratamiento más personalizado ya en la fase de reordenación. La principal mejora que esto posibilita es la conclusión del cauce entero al cabo de tan sólo 10 etapas, la longitud más corta de toda la sexta y séptima generación.

10 etapas

En la [figura 6.2](#) hemos numerado las etapas de segmentación en la parte izquierda junto a la unidad física donde tiene lugar su ejecución. El cauce que transcurre por la derecha es el de punto flotante, donde se contempla un mínimo de 15 etapas.

◀ [pág. 196](#)

15 etapas

## Superescalaridad

◀ 1.10

El K7 nos pone en un aprieto a la hora de dictaminar su factor de superescalaridad. Puesto que existen tres unidades funcionales de ejecución en cada cauce, el procesador presenta un factor de superescalaridad de **tres** tanto para instrucciones enteras como de punto flotante.

factor

Ahora bien, dijimos que el procesador puede generar hasta nueve microinstrucciones por ciclo de reloj, y nuestra vara de medir la superescalaridad se ha fijado siempre en las microinstrucciones, principalmente porque son éstas las que realmente se encaminan a la arquitectura interna del procesador una vez realizada la conversión a su representación nativa.

Dado que interiormente el posterior procesamiento se reduce a un máximo de seis microinstrucciones (tres por el lado entero y tres por el otro), una media algo más realista la estableceríamos en cinco, en vista de que las dependencias en la capa software impiden utilizar los recursos a plena capacidad.

cinco

## Resolución de dependencias

◀ 1.11

Terminada la decodificación, el procesador se dispone a resolver las dependencias de datos y control del código con anterioridad a su ejecución.

### 1.11.1 Dependencias de datos

El procesador habilita para ello un planificador de instrucción para las secciones entera y de punto flotante. En él mantiene una serie de instrucciones (16 en el caso entero y 36 en el otro) para reorganizarlas y distribuir las de la manera más independiente posible entre las tres unidades funcionales desacopladas que aparecen en cada cauce, siguiendo para ello las normas propias de la ejecución fuera de orden.

planificador de instrucción

Esta unidad cumple la misma función del BRC (Buffer de Reordenación Circular) utilizado por Intel en sus diseños de sexta generación. Sin embargo, el procesador apenas consume medio ciclo de reloj en esta tarea, pasando por ella de forma un tanto simplificada. Apenas faltan cuatro etapas para concluir la ejecución, y el daño que las dependencias pueden hacer sobre el código es limitado, así que se opta por no perder tiempo y asumir las penalizaciones derivadas de los

BRC

riesgos

conflictos que queden sin resolver.

### 1.11.2 Dependencias de control

BTB versus Intel	Para el tratamiento de las dependencias de control se dispone de una BTB ( <i>Branch Target Buffer - unidad de predicción de salto</i> ) con capacidad para retener la dirección de salto más probable de las 2048 últimas instrucciones no secuenciales que se ejecutaron en el código. Esto supone una cantidad cuatro veces superior a la del Pentium III, pero sólo el 50 % si tomamos como referencia el Pentium 4.
predicción de salto	El algoritmo de predicción dinámica de salto tiene reminiscencias del utilizado en el procesador Alpha, acertando en el destino del 90 % de las instrucciones de salto que contiene el programa. Existen en el mercado algoritmos de predicción mejores que éste, pero también es cierto que los han montado procesadores más vulnerables a una predicción errónea de lo que lo es el K7.

## 1.12 ▶ Instrucciones multimedia

Enhanced 3DNow!	El conjunto de instrucciones extiende el repertorio multimedia incorporando a las instrucciones MMX y 3DNow! 24 instrucciones multimedia más, denominadas genéricamente Enhanced 3DNow!. Este conjunto es cuantitativamente menos numeroso y cualitativamente menos potente que el SSE incluido en el Pentium III, aunque sumando el 3DNow! que no lleva éste último, podemos conceder una equivalencia entre ambos conjuntos.
FPU	Al igual que con las extensiones anteriores, el banco de registros, la unidades de ejecución y los cauces de segmentación para las instrucciones multimedia comparten sus recursos con las instrucciones de punto flotante.

## 1.13 ▶ Unidades de ejecución

enteras	Por la parte de aritmética entera, el K7 dispone de tres ALU para cálculos aritméticos y tres unidades de generación de la dirección virtual de acceso a memoria utilizadas en las instrucciones de carga y almacenamiento. Estas unidades consumen un ciclo en el caso entero y dos en el caso MMX y 3DNow!.
de punto flotante	Respecto a la unidad de punto flotante, el K7 ofrece tres cauces de ejecución segmentados independientes. Uno de ellos es exclusivo para operaciones de aritmética de punto flotante, y los otros dos comparten sumas y productos de este tipo con los cálculos multimedia. La etapa de ejecución de números reales se lleva un mínimo de 4 ciclos de reloj.

## 1.14 ▶ Valoración final

Nos encontramos frente a un diseño de ciertos contrastes, donde remarcaríamos los tres siguientes:

- ❶ **Frecuencia.** Se busca más el factor cualitativo que el cuantitativo, esto es, el aprovechamiento de cada ciclo de reloj frente a la rápida sucesión de éstos. Ahí queda ese sorprendente séptimo ciclo en el que se realiza la planificación de instrucciones y la búsqueda de operandos.
- ❷ **Paralelismo a nivel de instrucción.** Se premia la superescalaridad (factor 5 tomando las microinstrucciones como referencia) frente a la segmentación (sólo 10 etapas enteras).
- ❸ **Acceso a memoria.** Se apuesta por el primer nivel de memoria caché en detrimento del segundo nivel y la conexión con memoria principal.



### 1.14.1 Lo mejor

- La separación de los cauces de ejecución en las fases más tempranas del procesamiento de instrucciones, lo que posibilita un tratamiento más personalizado de éstas.
 separación de cauces
- El equilibrio que se consigue entre frecuencia y paralelismo a nivel de instrucción. Estamos ante un diseño CISC sin medias tintas, en el que no se busca un ciclo de reloj rápido, sino su óptimo aprovechamiento. La propia AMD sabe que aquí ha conseguido algo sobresaliente, porque más adelante, cuando toque diseñar el K8, le trasplantará toda esta parte del procesador sin apenas retoques (ver [sección 7.1.4](#)).
 equilibrio
- El primer nivel de memoria caché, el más grande de los modelos para PC que conocemos hasta la fecha.
 caché L1
- El atractivo precio del producto. Aún en su fase preliminar de comercialización, que es cuando la relación rendimiento/coste se encuentra más en contra del usuario, representaba una mejor elección frente al Pentium III. La versión de 0.18 micras que posteriormente visitaremos, el Thunderbird, conservó esta ventaja frente al Pentium 4.
 precio

➔ [pág. 258](#)

### 1.14.2 Lo peor

- La complejidad del sistema de decodificación. Cinco ciclos son muchos para la que es una de las tareas más sencillas en un procesador.
 cuesta decodificar
- Una cierta despreocupación por el tratamiento de las dependencias, que se refleja en el poco tiempo dedicado a la resolución de los riesgos de datos (confiando en exceso en la corta longitud de los cauces segmentados) y en la ocultación de la latencia inherente a las operaciones de memoria (confiando quizá en demasía en la eficiencia de la gran caché de primer nivel).
 dependencias obviadas
- La velocidad de la caché L2 interna, que en la mayoría de modelos es de tan sólo un tercio de la frecuencia del procesador. Como ya hemos advertido, esto se repararía a partir del Thunderbird.
 velocidad L2

SECCIÓN 6.2

## Thunderbird y Duron: Las 0.18 micras en AMD

### Thunderbird

◀ 2.1

Algo muy parecido a lo que ocurre con el Pentium III Coppermine de Intel acontece en el K7 unos meses más tarde cuando AMD instaura su proceso de fabricación a 0.18 micras: Se gesta un procesador en el que las mejoras en la tecnología de integración se aprovechan en sus mismas tres vertientes:

- ❶ **Frecuencia.** Se aprovechan las mejoras en el proceso de fabricación para subir la frecuencia al procesador, alcanzándose ya los 1.4 GHz, y también el bus local, cuyo techo es ahora 2x133 MHz (frecuencia de 133 MHz y multiplicador interno de 2x).
- ❷ **Cambio de formato.** Se vuelve a migrar al zócalo Socket, en este caso el Socket A propio de AMD, bajo un empaquetado CPGA (Ceramic Pin Grid Array) muy similar al FC-PGA usado por Intel. El patillaje aumenta hasta los 462 pines, buscando una mejor distribución de la alimentación para minimizar el impacto sobre la variable térmica.

Características	El microprocesador K7 de AMD	
	Athlon (1999)	Thunderbird (2000)
Variables eléctricas:		
Frecuencia de reloj (MHz)	500 - 1000	700 - 1400
Frecuencia bus local (MHz)	2x100	2x133
Potencia disipada a 1 GHz (vatios)	65	54
Voltaje del núcleo (voltios)	1.6-1.8	1.7
Zócalo al que se acopla	Slot-A	Socket-A
Tecnología de integración (micras)	CMOS 0.25/0.18	CMOS 0.18
Transistores (millones)	22	37
Interconexiones metálicas	Aluminio	Cobre
Área de integración (mm <sup>2</sup> )	184 (0.25), 102 (0.18)	117
Empaquetado y envoltura externa	SECC	CPGA
Segmentación:		
Cauce aritmética entera	10 etapas	
Cauce punto flotante y MMX	15 etapas	
Superescalaridad:		
Unidades de ejecución	9 (3 ent., 2 MMX, 2 3DNow!, 1 Enh, 1 FPU)	
Factor de superescalaridad	Entre 3 y 6	
Memoria caché:		
L1 Datos (integrada)	64 Kb, 1024 ctos, 2 líns/cto, 32 bytes/lín	
L1 instrucciones (integrada)	64 Kb, 1024 ctos, 2 líns/cto, 32 bytes/lín	
Tipo L2	Interna	Integrada
Tamaño L2 (Kbytes)	512 Kbytes	256 Kbytes
Velocidad L2 (divisor resp a proc.)	1/3, 1/2, 1	1
Organización L2	2 líneas/cto	16 líneas/cto
Conjunto de instrucciones:	80x86 + 57 MMX + 24 3DNow! + 21 Enhanced	

TABLA 6.2: El microprocesador K7 de AMD resumido en sus dos modelos para gama PC: El Athlon de 0.25 micras y el Thunderbird de 0.18 micras.

- ③ **Caché L2 integrada.** La caché interna de 512 Kbytes del formato Slot del procesador se reconvierte a una caché integrada de 256 Kbytes. Esta caché se lleva un patrimonio de 15 millones de transistores en un área de integración de 15 mm<sup>2</sup>. En todos los aspectos es mimética a la del Coppermine de Intel excepto en uno: Su organización es más asociativa: 16 líneas por conjunto.

pág. 201

La [tabla 6.2](#) resume sus prestaciones contrastándolas con el Athlon, mientras que la [tabla 6.3](#) establece una comparativa con respecto a su homónimo de la competencia, el Pentium III de Intel.

## 2.2 ► Duron

Una vez más, y no sabemos cuántas van ya, AMD adopta para sí la misma estrategia que presencia en Intel.

Ya vimos que del proceso de producción del Pentium III se deriva un modelo Celeron de gama más baja con la mitad de caché L2 y la mitad de velocidad en el controlador de bus local. Pues aquí tenemos la versión AMD de esos hechos: El Duron, al que se le extirpan doce millones de transistores respecto al Thunderbird.

No obstante, AMD pondera de forma diferente los sacrificios que han de hacerse a la arquitec-

Características	Intel versus AMD en inauguración 0.18 micras	
	Pentium III Coppermine (1999)	K7 Thunderbird (2000)
Variables eléctricas:		
Frecuencia de reloj (MHz)	533 - 1133	700 - 1400
Frecuencia bus local (MHz)	133	2x133
Potencia disipada a 1 GHz (vatios)	33	54
Voltaje del núcleo (voltios)	1.65	1.7
Zócalo al que se acopla	Socket 370 pines	Socket-A 462 pines
Interconexiones metálicas	Aluminio	Cobre
Transistores (millones)	28	37
Area de integración (mm <sup>2</sup> )	186	117
Empaquetado y envoltura externa	FC-PGA	CPGA
Segmentación:		
Cauce aritmética entera	11 etapas	10 etapas
Cauce punto flotante y MMX	14 etapas	15 etapas
Superescalaridad:		
Unidades de ejecución	8	9
Factor de superescalaridad	3	5
Memoria caché:		
L1 Datos (integrada)	16 Kbytes	64 Kbytes
L1 instrucciones (integrada)	16 Kbytes	64 Kbytes
Organización L1	4 lins/cto	2 lins/cto
Tamaño L2 integrada (Kbytes)	256 Kbytes	256 Kbytes
Organización L2	8 líneas/cto	16 líneas/cto
Extensiones multimedia a MMX:	70 SSE	21 3DNow! + 24 Enhanced

TABLA 6.3: Comparativa entre las prestaciones del primer producto de 0.18 micras de nuestras dos principales compañías: Por Intel, el Pentium III Coppermine, y por AMD, el K7 Thunderbird.

tura para gestar este procesador de gama baja: El bus local no sufre alteración en el Duron, pero en cambio, la caché L2 integrada se reduce hasta la cuarta parte: Sólo 64 Kbytes, la mitad que el primer nivel de caché del procesador, que suma un total de 128 Kbytes.

mismo bus,  
menos caché

Arquitecturalmente, resulta de difícil entendimiento que el primer nivel en la jerarquía de cachés sea más grande que el segundo. Pero para resolver este aparente contrasentido no hay que mirar a la cantidad de caché, sino a su calidad, esto es, a su organización interna. La L1 de la arquitectura K7 tiene sólo dos líneas por conjunto, mientras que su L2 dispone de dieciséis. Cuando esto ocurre, los dos niveles de caché se comportan como si hubiera un único nivel que fuese suma de ambos, puesto que el espacio de direcciones que barre cada uno es casi disjunto. Al final, lo que ocurre a efectos prácticos del rendimiento es que la L2 promociona un nivel, complementando a los niveles de caché L1.

L2 más efectiva

El código de referencia de este procesador es Spitfire, y su diferencia con el K7 Thunderbird es apenas perceptible externamente, tal y como podemos apreciar en la [foto 6.2](#).

← pág. 202

## Duron con aditivos

## 2.3

Tras el Duron, los planes de AMD eran sacar otro procesador similar con código de referencia Morgan, integrando en el mismo chip procesador el controlador para las tarjetas de gráficos y sonido tratando de conseguir un ahorro adicional en el coste conjunto del equipo.

Morgan

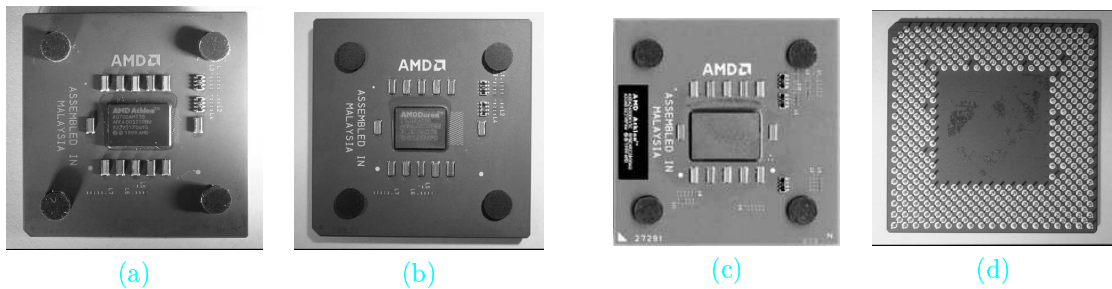


FOTO 6.2: Las 0.18 micras en AMD: (a) Thunderbird. (b) Duron. (c) Athlon XP. (d) El patillaje de 462 pines (dorso de cualquiera de las fotos anteriores) es idéntico en las tres versiones del procesador. Tan sólo podemos apreciar una leve diferencia del Duron con los otros dos en las dimensiones del área de integración, debido a que cuenta con 25 millones de transistores (frente a 37 millones en los otros dos modelos), al disponer de tan sólo la cuarta parte de caché L2 (64 Kbytes).

Timna

Esta jugada estuvo también copiada de un proyecto de Intel, el procesador Timna, que estaba anunciado para Noviembre de 2000 integrando esos mismos elementos en el Celeron, pero que en última instancia Intel decidió cancelar argumentando que la demanda en el mercado para este tipo de productos era más bien escasa.

controlador  
DDR integrado

Unos meses más tarde, AMD tomó una decisión similar y suprimió el proyecto Morgan. Sin embargo, conservó el nombre para un proyecto posterior de cierto parecido con la idea original: Integrar dentro del procesador el controlador para memoria principal DDRAM, algo que más tarde hemos visto repetir nada menos que en el K8 (ver [sección 7.1.7](#)), aunque con una perspectiva mucho más ambiciosa.

pág. 266

el más barato

Apenas doscientos mil transistores más para crear un producto que normalmente se combina con un juego de chips en placa base que lleva todo integrado: Tarjeta gráfica, tarjeta de sonido, y algún que otro periférico más. Como resultado, un completo PC que puede adquirirse por unos 500€ prescindiendo de su extraordinaria modularidad.

### SECCIÓN 6.3

## Athlon XP

Coincidiendo con la salida al mercado del sistema operativo Windows XP a finales de 2001, AMD publicitó el lanzamiento de un, aparentemente, nuevo procesador: El Athlon XP (las iniciales XP provienen de *eXtreme Performance*).

un falso  
señuelo

Si el nombre ya prometía cosas que no eran, sus especificaciones van a ser incluso perversas. Para los que tratamos de formalizar con un mínimo de rigor las características técnicas de los procesadores, campañas de marketing como ésta nos obligan a duplicar esfuerzos: Primero, hemos que extirpar todos los bulos que han extendido, y sólo cuando hemos logrado regresar al punto cero, podemos contar lo que hay de realidad en el nuevo modelo.

### 3.1 ► Frecuencia de reloj

confusión

El Athlon XP es toda una sutileza a la hora de expresar la frecuencia del procesador. Su valor real se encuentra eclipsado por otro, el que se publicita, que es confuso como pocos. En un Athlon XP 1800+, la frecuencia no es 1800 MHz, sino 1533 MHz. Para otros valores, consultar la frecuencia real en la [tabla 6.4](#).

pág. 203

Nombre comercial	Frecuencia real	Frecuencia del bus	Nombre comercial	Frecuencia real	Frecuencia del bus
XP 1500+	1.333 MHz	2x133 MHz	XP 2400+	2.000 MHz	2x133 MHz
XP 1600+	1.400 MHz	2x133 MHz	XP 2600+	2.133 MHz	2x133 MHz
XP 1700+	1.466 MHz	2x133 MHz	XP 2700+	2.166 MHz	2x166 MHz
XP 1800+	1.533 MHz	2x133 MHz	XP 2800+	2.250 MHz	2x166 MHz
XP 1900+	1.600 MHz	2x133 MHz	XP 3000+	2.416 MHz	2x166 MHz
XP 2000+	1.666 MHz	2x133 MHz	XP 3200+	2.500 MHz	2x166 MHz
XP 2100+	1.733 MHz	2x133 MHz	XP 3400+	2.666 MHz	2x166 MHz
XP 2200+	1.800 MHz	2x133 MHz	XP 3600+	2.833 MHz	2x166 MHz

TABLA 6.4: Frecuencia del Athlon XP equivalente a sus especificaciones comerciales.

Con el valor 1800 se pretende dar una medida de rendimiento para una serie de aplicaciones software. Si todas las firmas usaran esta misma medida, la idea no sería mala; pero a cada cual le interesa utilizar unos programas diferentes, pues se ha revelado como un potente mecanismo para adular el rendimiento frente a la competencia. El valor 1800+ está sesgado en favor de AMD, y dicho esto, que ellos resalten que esa medida supone hasta un 20 % de mejora sobre un Pentium 4 de 2 GHz carece de significación.

descrédito

El sufijo + de 1800+ es otra fuente de confusión. El plus puede interpretarse como que la frecuencia efectiva del procesador es superior a 1800 MHz, o como que sobrepasa en rendimiento a un Pentium 4 de esa misma frecuencia. La realidad es que evidencia el complejo de AMD por no poder flanquear los 2 GHz al tiempo que su rival, e irradia la sensación de que la sombra de Intel continúa siendo muy alargada para ellos. En el contexto de la quinta generación se vivió una historia similar, en la que bajo las letras PR (Pentium Rated), la legión de competidores de Intel trataba de indicar una frecuencia equivalente para sus procesadores respecto al Pentium.

complejo

PR

A estas alturas, todos sabemos ya que la frecuencia es sólo una más de un ramillete de virtudes de un procesador, y si no se apuesta por ella es porque la arquitectura ha priorizado otras. Uno debe ser consecuente con sus diseños, asumir sus carencias para lograr credibilidad, y explotar sus bondades para conseguir adeptos. Venimos de ensalzar de forma superlativa esa clara apuesta del K7 por la filosofía CISC, con la que consigue un perfecto equilibrio entre paralelismo y frecuencia. El Pentium 4 es más RISC y tiene más etapas de segmentación, por lo tanto, su diseño se cimenta sobre una elevada frecuencia (y para comprenderlo no hay más que regresar a la [figura 3.1](#), con esa pirámide con la que empezamos a explicarlo todo), así que tiene todas las cartas para ir por delante en esta faceta.

→ pág. 49

La frecuencia de un procesador está definida para ser empleada como tal, y no para subjetivizarla a conveniencia. Hay mucho personal en las tiendas de informática que es comercial mucho antes que técnico, y para el que un Athlon 1800+ apunta un reloj de 1.8 GHz, y muchos usuarios que compran el modelo con idéntica percepción. Explicarles a ellos que rendimiento y frecuencia son cosas distintas puede llevar un tiempo del que a lo peor no disponen.

el perfil de usuario

Para un perfil algo más exigente como el nuestro, la lección es diferente: Es un error dar un valor de frecuencia equivalente, puesto que el rendimiento del hardware está mediatizado por el software, y a medida que se avanza en el diseño de procesadores, sus arquitecturas cursan apuestas más agresivas en las que el grado de sensibilidad ante las dependencias de datos y control, los accesos a memoria, etc, es casi extremo. Elijiendo oportunamente el programa de evaluación dentro de la séptima generación, los bandazos que puede dar una comparativa sobrepasan el 50 % en favor de una u otra arquitectura, lo que evidencia un potente mecanismo de manipulación.

el perfil técnico

Por eso nuestras frecuencias equivalentes para el K7 y el Pentium 4 que mostramos en la comparativa de la [tabla 6.12](#) fueron obtenidas sobre un modelo analítico que considera un programa insesgado con valores promedio para cada ingrediente a que es sensible el rendimiento de una

modelo analítico  
→ pág. 244

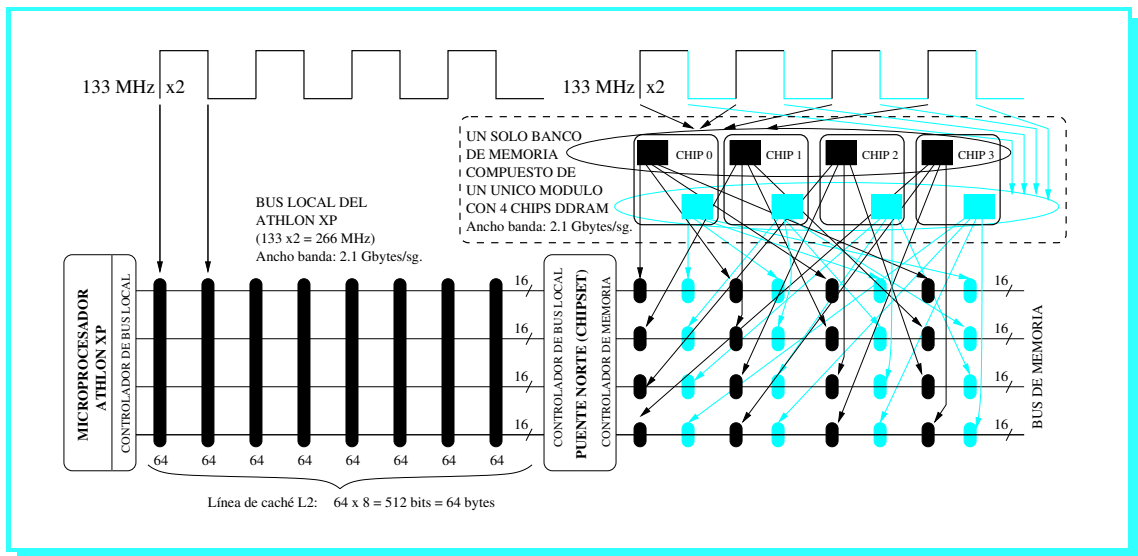


FIGURA 6.3: El perfecto entendimiento entre el Athlon XP y la memoria DDRAM de 2x133 MHz. Se muestra el llenado de una línea de caché L2 desde un módulo de memoria principal a través del controlador de memoria ubicado en el juego de chips de la placa base. El diagrama es perfectamente válido para entender el diálogo entre la nueva versión de bus 2x166 MHz y la memoria DDRAM 2x166 MHz (333), con solo imaginar que la señal de reloj mostrada en ambas partes corresponde a 166 MHz.

arquitectura (y tras 14 páginas de análisis). Dar el valor 1.8+ permite a AMD subir 300 MHz a sus procesadores sin tener que dar más explicaciones. Su camino es más sencillo, pero representa toda una temeridad. Si lo que quieren es generar bulos con los que manipular el mercado a su antojo, ya deberían saber que la era de la información es una saludable democracia que dilapida al dictador.

### 3.2 ► El bus local y sus diálogos con memoria principal

2x133 MHz

Inicialmente, el Athlon XP lleva el mismo bus local de 2x133 MHz de los últimos modelos de Thunderbird, sólo que ahora este ingrediente se complementa con cobertura adicional en el puente norte del juego de chips, donde se ubica un controlador de memoria principal que soporta DDRAM de 266 MHz (esto es, la SDRAM de 133 MHz trabajando con un multiplicador de 2x - ver sección 10.13.5). Esto permite aprovechar todo el ancho de banda del bus local desde la arteria que comunica con memoria principal, tal y como hemos ilustrado en la figura 6.3.

p. 68/Vol. 1.2

2x166 MHz

Tras sucesivas mejoras en el proceso de fabricación que descubriremos enseguida, se consigue una aceleración del bus hasta 2x166 MHz, en buena sincronización con la llegada al mercado de la memoria DDRAM de esa misma frecuencia. La forma en que ambas partes se entienden es exactamente la misma que la ya mostrada en la figura 6.3, con sólo considerar que los relojes mostrados son de 166 MHz en lugar de 133 MHz. Esta prima en velocidad está disponible en las versiones Athlon XP 2700+ o superiores (ver tabla 6.4).

pág. 203

### 3.3 ► Tecnología de integración

La primera versión del Athlon XP dispone de 37.5 millones de transistores sobre un área de integración de 128 mm<sup>2</sup>, lo que supone apenas medio millón de transistores y 11 mm<sup>2</sup> más de silicio que el Thunderbird (ver foto 6.2). Ambos comparten el zócalo Socket A de 462 pines <sup>2</sup>,

pág. 202

<sup>2</sup>Tan sólo es necesaria una actualización de la BIOS de la placa base de los Thunderbird para reutilizarlo.

Característica	Athlon	Athlon XP				Pentium 4 Northwood
	Thunderbird	Palomino	Thoroughbred A	Thoroughbred B	Barton	
Fecha lanzam.	Junio 2000	Octubre 2001	Junio 2002	Agosto 2002	Marzo 2003	Agosto 2001
T.I. (micras)	0.18	0.18	0.13	0.13	0.13	0.13
Area integr.	117 mm <sup>2</sup>	128 mm <sup>2</sup>	84 mm <sup>2</sup>	80 mm <sup>2</sup>	115 mm <sup>2</sup>	146 mm <sup>2</sup>
Niveles metal	6	7	8	9	9	7
Encapsulado	Cerámico	Orgánico	Orgánico	Orgánico	Orgánico	Orgánico
Obleas (diám.)	20 cm.	30 cm.	30 cm.	30 cm.	30 cm.	30 cm.
Frec. (GHz)	0.7 - 1.4	1.33 - 1.7	1.33 - 1.8	2 - 2.4	2.2 - 2.5	2 - 3.2
F. bus (MHz)	2x133	2x133	2x133	2x133/166	2x166	4x133
Zócalo	Socket A	Socket A	Socket A	Socket A	Socket A	mPGA478B
Patillaje	462 pines	462 pines	462 pines	462 pines	462 pines	478 pines
Mill. trans.	37	37.5	37.2	37.6	53.9	55
Voltaje (volt.)	1.7	1.75	1.50-1.65	1.5	1.65	1.5
Caché L1	64 Kb.	64 Kb.	64 Kb.	64 Kb.	64 Kb.	8Kb+Traza
Caché L2	256 Kb.	256 Kb.	256 Kb.	256 Kb.	512 Kb.	512 Kb.
Cto. instrs.	Enh. 3DNow!	3DNow! Professional				SSE2

TABLA 6.5: Evolución tecnológica del Athlon XP a lo largo de las cuatro versiones que se suceden en sus dos años de vigencia. La familia se contrasta con su ancestro más inmediato, el Thunderbird, y su rival más claro, el Pentium 4 Northwood.

voltaje de 1.75 voltios, distancia de integración de 0.18 micras e interconexiones de cobre.

integración

Las principales diferencias nos esperan en el encapsulado y en la fabricación. El encapsulado sigue siendo PGA (Pin Grid Array), pero utiliza un material orgánico en sustitución de la cobertura cerámica del Thunderbird, empleando además la fibra de vidrio para dar consistencia al conjunto, material tradicionalmente utilizado en la placa base para este mismo menester.

encapsulado

La fabricación utiliza obleas de 30 cm. de diámetro siguiendo el camino adoptado por Intel en sus Pentium 4 en la primera mitad del año 2001. Frente a las tradicionales obleas de 20 cm., esto supone un abaratamiento del producto, puesto que el coste es fijo por oblea producida, pero en una de 30 cm. caben aproximadamente el doble de procesadores (706 mm<sup>2</sup> frente a 314 mm<sup>2</sup>, despreciando los contornos).

fabricación

Otro aspecto novedoso en la fabricación del Athlon XP es su metalización, esto es, la etapa de interconexión de sus transistores. La figura 3.4 nos enseñó la presencia de capas de metal apiladas en estratos sobre los transistores, responsables de agruparlos en puertas lógicas, y éstas a su vez en unidades funcionales. El número de capas es una variable que evoluciona de forma cadenciosa. Por ejemplo, en 1993, el proceso de 0.5 micras de Intel contaba con tres niveles, y diez años más tarde, el de 0.13 micras dispone de siete, y seguirá ahí un par de años en las 0.09 micras según ha anunciado la firma. Sin embargo, AMD va a coger cierta carrerilla con su Athlon XP, y en apenas dos años pasará de los seis niveles con que se hereda el Thunderbird hasta alcanzar los nueve del último XP, con código de referencia Thoroughbred B.

metalización

← pág. 51

La tabla 6.5 muestra la inquieta evolución tecnológica del Athlon XP en sus cuatro versiones hasta la fecha, las tres últimas ya sobre 0.13 micras, comparadas con su ancestro, el K7 Thunderbird, y su rival, el Pentium 4 Northwood.

Las mejoras del chip al nivel tecnológico se completan con una disipación de potencia en torno a un 20% inferior a la del Thunderbird, y la incorporación de un diodo térmico interno que monitoriza la temperatura de chip para informar de posibles alarmas por uno de sus pines, algo que ya vimos en Intel desde el Pentium III Coppermine con su línea THERMTRIP.

mejoras  
térmicas

THERMTRIP

### 3.4 ▶ El jeroglífico de la frecuencia

Una vez conocemos los detalles de fabricación del XP, llega el momento de deshacer el jeroglífico al que parece querer someternos la [tabla 6.4](#):

☛ pág. 203

dos enigmas

de 8 a 9 capas

de 133  
a 166 MHz

- Si nos fijamos en su contenido, cada salto de 100 en el nombre comercial se va correspondiendo con un incremento de 66 MHz en la frecuencia del chip, hasta llegar a la especificación 2200+ (1800 MHz), donde hay un salto hasta los 2400+, que se replica también con un incremento neto de 200 MHz en la frecuencia (2000 MHz). Esa extraña transición coincide con la incorporación del noveno nivel de metalización en el chip.
- Un poco más adelante, en el paso de 2600+ a 2700+, la frecuencia sólo sube 33 MHz, y a partir de ahí, cada incremento de 100 en el nombre comercial pasa a significar una subida de 83 MHz. Esta transición irregular se corresponde con el momento en el que el bus local se acelera, pasando de 2x133 MHz a 2x166 MHz. Podemos observar así que el incremento por cada valor de 100 en el nombre comercial del chip se corresponde siempre en sus MHz con la mitad de la frecuencia nominal del bus empleado en cada modelo.

### 3.5 ▶ Paralelismo a nivel de instrucción

QuantiSpeed

A pesar de lo que rezaba la publicidad de AMD, el núcleo arquitectural del XP apenas sufre cambios. QuantiSpeed enumera una serie de cualidades a bombo y platillo, pero acabarían antes indicando que son las mismas que en el Athlon original, con tan sólo leves retoques:

- Las cachés extienden la estrategia de prebúsqueda a la L1I y L2 (la L1D ya incluía esta faceta desde el conjunto 3DNow!).
- Las TLB cuentan con tres mejoras: (1) Se dotan de acceso exclusivo, (2) se amplían desde las 32 hasta las 40 entradas en la parte de datos, y (3) se anticipa la traducción en la TLB de instrucciones.

segmentación

En definitiva, QuantiSpeed es mucho ruido, pero pocas nueces, y eso puede percibirse con tan sólo comprobar que el procesador apenas si cuenta con medio millón de transistores más. Nos quedamos, por tanto, con 10 y 15 etapas en los cauces segmentados y un factor de superescalabilidad de 5, los ingredientes de que ya disponíamos en el primer K7.

superescalar

### 3.6 ▶ Memoria caché

tamaño y  
organización

Con tan pocos transistores adicionales, la caché integrada no dispone de margen alguno para su mejora, y así, en el XP se calca el tamaño y la organización ya conocidos del Thunderbird: Dos cachés L1 gemelas de 64 Kbytes con 2 líneas por conjunto cada una en las que prima la cantidad, y una caché L2 de 256 Kbytes y 16 líneas por conjunto en la que prima la calidad.

espacio  
solapado

No obstante, en la publicidad de la firma hemos encontrado un tamaño de caché de 384 Kbytes para este procesador. El valor se obtiene mediante la suma de los tres tamaños anteriores, pero esa suma no es lícita. El espacio total existe, pero las posiciones de caché L1 replican un subconjunto de caché L2, de la misma manera que L2 contiene una réplica de parte de la memoria principal.

más claro

más general

Más claro: Un hipotético PC con 256 Mbytes de memoria principal, 1 Mbyte de caché L3, 512 Kbytes de caché L2 y dos cachés L1 gemelas de 256 Kbytes no dispone de 258 Mbytes de espacio total de almacenamiento; mantiene, exclusivamente, 256 Mbytes. Más general: El papel de la caché no es aumentar el espacio de memoria disponible, sino *acelerar el que ya existe*.

En favor de AMD aclararemos que es en el XP donde el área de memoria en el que incide cada nivel de caché es más disjunta, ya que mediante una organización interna muy dispar (2



y 16 líneas por conjunto en L1 y L2) se consigue un mayor desacoplo de las celdas de memoria presentes en cada nivel. Pero este efecto desde el prisma del rendimiento, que ya comentamos en el caso del Duron, no otorga licencia para tergiversar la composición interna del chip.

hay cierto  
desacoplo

Finalmente, el modelo que tiene el honor de cerrar el telón del XP, con código de referencia Barton, trae una gran novedad que recoge los frutos de los logros conseguidos en el proceso de fabricación durante estos dos años de vigencia del procesador: La caché L2 duplica su tamaño hasta los 512 Kbytes, alcanzando el chip los sesenta millones de transistores. Pero una vez más, el camino lo enseña Intel, que dos años atrás movió ficha en esta misma dirección en sus primeros modelos de 0.13 micras, tanto para Pentium III (Tualatin) como para Pentium 4 (Northwood).

Barton dobla  
la caché L2

## Conjunto de instrucciones

### 3.7

El Athlon XP no amplía el conjunto de instrucciones nativo del K7, aunque sí incorpora nuevas instrucciones multimedia.

extensiones  
multimedia

Llegamos así al aspecto que supone la innovación que da nombre al producto, pero decir que el Athlon XP es un procesador optimizado para ejecutar aplicaciones bajo Windows XP es otra verdad a medias, porque las cosas han acontecido justo al revés: Microsoft ha reescrito su API DirectX 8.0, el interfaz para las aplicaciones multimedia de su Windows XP (consultar [sección 3.5.4.4](#), y más concretamente la [tabla 3.15](#)), con objeto de aprovechar las nuevas instrucciones multimedia de que dispone este procesador, y que tratan de aprovechar el grueso de ese medio millón de transistores adicional de que dispone el XP.

☛ pág. 112  
☛ pág. 114

Recordemos que el Athlon original ya disponía de tres extensiones multimedia: MMX, 3DNow! y Enhanced 3DNow! (ver [tabla 3.16](#)), pero cuyo principal enemigo fue precisamente navegar de espaldas a Intel, el camino natural que sigue la industria del software para PC. La nueva capa que se articula ahora, denominada 3DNow! Professional, incorpora 72 instrucciones, pero 52 de ellas son compatibles con las SSE del Pentium III. Es decir, lo que ocurre es que en el diagrama que adjuntamos en la [figura 3.19](#), AMD bascula hacia la derecha buscando el espacio ocupado por Intel. Y como es allí donde los programas están más posicionados, incluidos los Windows de Microsoft, AMD puede esgrimir que articula recursos para darles cobertura. La [figura 3.20](#) sintetiza toda la secuencia de pasos, donde apreciamos en primer lugar el trazado de caminos divergentes por las compañías implicadas, para posteriormente inclinarse por la convergencia en un punto común.

☛ pág. 118  
caminos  
enfrentados,

☛ pág. 115

caminos  
encontrados  
☛ pág. 119

## Valoración final

### 3.8

El Athlon XP representa una paradoja que difícilmente volveremos a ver. Lo que más se publicita de él son precisamente sus debilidades, y lo que más se esconde son sus grandes virtudes.

### 3.8.1 Lo mejor




- ☛ La fabricación del procesador, vanguardista como pocas. Esto revierte en una variable fundamental del chip, que trajo de cabeza a los Thunderbird en su fase terminal de producción: La temperatura. Pasar menos calor significa gozar de mucha mayor fiabilidad en nuestro PC (cuelgues espúreos), al tiempo de se garantiza una mayor longevidad al chip y se abre la puerta a una posible sobreaceleración.
- ☛ La compatibilidad con el zócalo Socket A de modelos anteriores, pudiéndose incluso reutilizar la misma placa base con una simple actualización de su Flash-BIOS.
- ☛ El acceso a memoria, que gana bastante con el incremento de frecuencia del bus a un valor que se entiende perfectamente con los chips de memoria DDRAM más populares del mercado, y obtiene un nuevo plus en velocidad con la llegada del modelo Barton, en el que la caché L2 se amplía a 512 Kbytes.

temperatura

compatibilidad

acceso a  
memoria

### 3.8.2 Lo peor

- frecuencia  Esa forma de especificar la velocidad del chip tan confusa y enigmática.
- QuantiSpeed  La arquitectura QuantiSpeed, donde no hay ingredientes de relevancia.
- 3DNow! Prof.  El conjunto de instrucciones 3DNow! Professional, que aunque tiene de bueno el acercamiento a un estándar que reclama la comunidad, no supone una mejora de rendimiento que justifique las expectativas creadas.

## SECCIÓN 6.4

### Pentium 4 de Intel

pág. 209   
P67

Ya durante el año 2000, Intel comenzó a dar las primeras pistas de que el Itanium no iba a ser el buque insignia en la estrategia comercial de la compañía. En efecto, Andy Grove (ver [foto 6.3.a](#)), no se lo había jugado todo a una carta, y en una sabia política de diversificación de riesgos, puso en marcha otro proyecto en paralelo: El P67. Al mismo tiempo, Grove consagraba su idea ya apuntada con el Itanium: Reducir el ciclo de producción de una nueva arquitectura a la mitad, desde los cuatro años precedentes a sólo dos, intercalando en el tiempo los productos concebidos por cada grupo de trabajo. Esta idea es una de las claves para entender la posición dominante de Intel en el mercado en el presente, pues obliga a la competencia a acelerar un ritmo ya de por sí frenético, y que pocas compañías han demostrado ser capaces de aguantar.

actitud


El comportamiento de Intel es siempre sorprendente: Para algunas cosas, se comporta lentamente como el gran peso pesado del sector que es, pero para otras, tiene una conducta más propia de una compañía emergente que acaba de llegar y quiere abrirse paso en un mercado dominado por otras. Y esta actitud tiene parte de la culpa de que se haya mantenido durante tantos años ahí arriba.

P6

pág. 263 

repliegue  
a 32 bits:  
Willamette

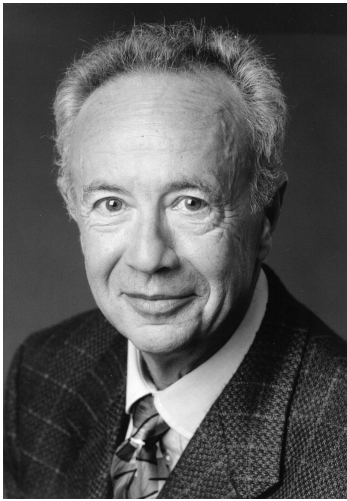
Arquitecturalmente, el P67 es una continuación del P6, el proyecto que dió lugar al Pentium Pro, siendo su objetivo la extensión a 64 bits del conjunto de instrucciones x86. Como se observa, la concepción era parecida a la que años más tarde llevaría a la práctica AMD para desarrollar su K8 a partir del K7 (ver [sección 7.1.6.3](#)), pero quizá también algo prematura. Ya hemos dicho en alguna ocasión aquí que en tecnología resulta tan peligroso ir demasiado adelantado como demasiado rezagado. Quizá por ello Intel redefinió el objetivo del P67 hacia un horizonte algo menos pretencioso: Desarrollar una nueva arquitectura de 32 bits basándose aún en el conjunto de instrucciones del 80x86. El proyecto se hizo finalmente público en Octubre de 1998, recibiendo el código de referencia Willamette.

pág. 7   
Intel menos  
pretencioso

Alguien puede pensar que si las intenciones iniciales del P67 se hubieran preservado, Intel tendría ahora en el mercado un excelente rival para el K8, mientras que la situación actual describe una confrontación desigual. A los que así opinen, me gustaría trasladarles a 1997 con todas las soberanas lecciones históricas que aprendimos con la lectura del [capítulo 1](#). Intel se pasó siete pueblos con el i432 en los años 70, otros siete con el i860 en los años 80, y ya se había merendado tres o cuatro más con el Itanium en los años 90. Probablemente Andy Grove pensó que ya estaba bien por esa década con un fiasco, que ya habían cumplido con el cupo de su media histórica, y que esta vez tocaba por fin aprender la lección, levantar el pie del acelerador, y en todo caso, pecar por defecto.

pág. 209 

Nada más ver la luz el procesador (Navidad de 2000), Paul Otellini (ver [foto 6.3.b](#)), entonces jefe de la división de arquitectura de Intel y ahora presidente de la firma, espetó que a mediados de 2002 el Pentium 4 sería líder de ventas en su segmento y el eje principal de los productos de Intel; el tiempo no ha hecho sino darle la razón.



(a) Andy Grove.



(b) Paul Otellini.



(c) Craig R. Barrett.

Fotos cortesía de Intel

FOTO 6.3: (a) Andy Grove, Presidente y Director Ejecutivo de Intel desde 1979 hasta Mayo de 1998 y mentor del proyecto P67 que dió origen al microprocesador Pentium 4. (b) Paul Otellini, Presidente y Manager General de Intel, director de la división de arquitectura de la compañía durante toda la década de los 90 y principal impulsor del procesador Pentium y sus sucesores. Otellini es en los años 90 el sustituto de alguien insustituible en Intel: Robert Noyce, fallecido en 1990. (c) Craig R. Barrett, el cuarto presidente en la historia de Intel tras Noyce, Moore y Grove, ahora en labores de Director General.

## Nueva arquitectura

### ◀ 4.1

El Pentium 4 es una nueva arquitectura, lo que significa un punto y aparte: Nuevo zócalo y una nueva placa base incompatible con los modelos predecesores de Pentium II, III y Celeron.

Pero por el nombre elegido, Intel parece querer decirnos que estamos ante una secuela arquitectural más del Pentium Pro. Mejor generar ese equívoco que cambiar un nombre que está considerado como el tercero que más personas del planeta admiten conocer, tras Coca-Cola y Mc-Donalds. El gasto en márketing para lograr semejante notoriedad sería muy cuantioso, y continuando la saga la compañía amortiza las campañas de publicidad de los hermanos menores (sólo en el lanzamiento del Pentium III, ya se gastaron más de 300 millones de dólares).

Si Intel no se ha deshecho de la denominación Pentium ahora que tocaba, muy probablemente es porque llegaremos a ver Pentium de dos dígitos en el futuro. Esta predicción la formulamos ya en nuestra edición de 2001, y a pesar de que en dos años no han avanzado los ordinales, seguimos apostando por ella. Lo de usar números latinos en vez de romanos es otra historia, amparada en un márketing un tanto retorcido de diferenciarse del rival hasta en los detalles más nimios. No obstante, será algo que respetaremos dado su carácter de marca registrada.

El Pentium 4 está diseñado por el grupo de trabajo que en 18 meses creó aquel P6 sobre el que se cimentaron los principales productos de la compañía en los últimos cinco años, y está llamado a ser la arquitectura sobre la que recaiga la responsabilidad de mantener a Intel donde está en los cinco siguientes. Liderado por Doug Carmean, el grupo se encuentra ahora definiendo la arquitectura del que será el microprocesador rival del K8 de AMD en la octava generación.

La estructura básica del procesador podemos resumirla en tres premisas básicas:

popularidad

etimología

grupo de trabajo

estructura básica

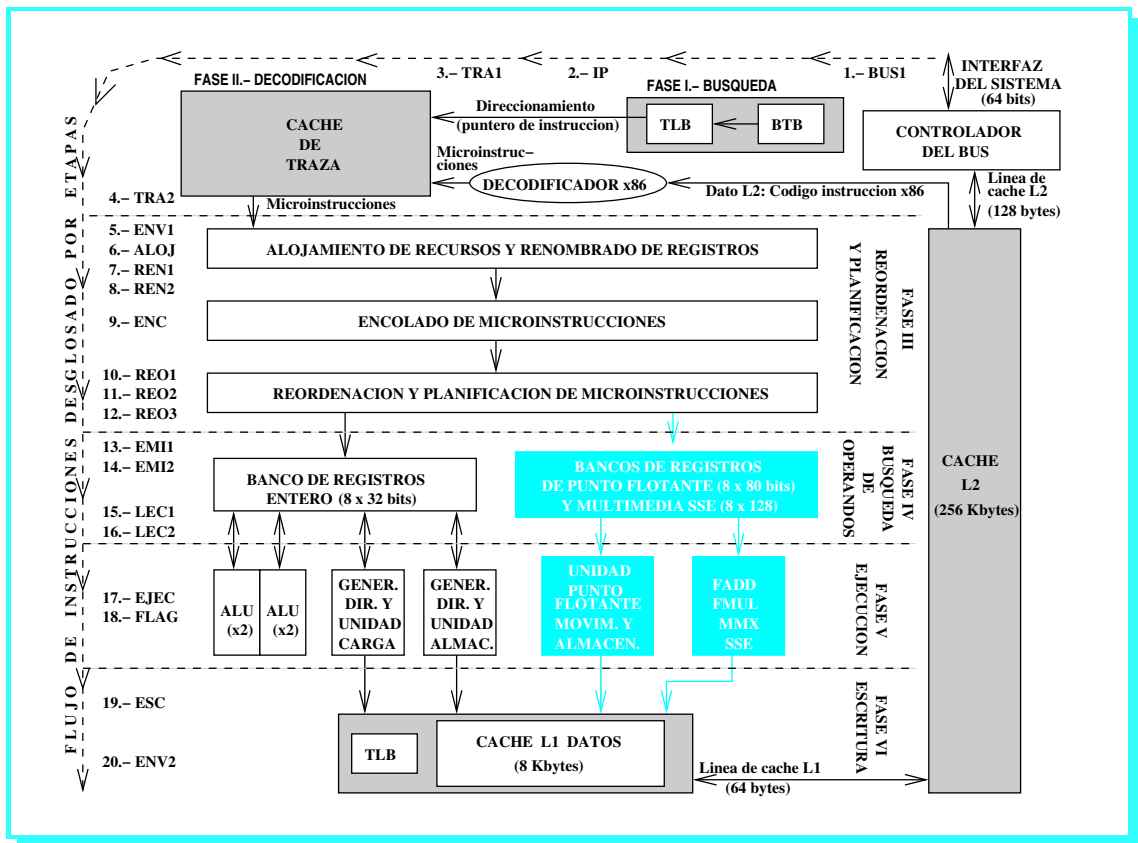


FIGURA 6.4: Diagrama de bloques del microprocesador Pentium 4 de Intel.

- 1 Una boca excepcional: Su bus local.
- 2 Una garganta deficiente: Esa representación y decodificación de instrucciones basada una vez más en el indigesto conjunto de instrucciones compatible con el x86.
- 3 Un estómago arriesgado, con 20 etapas de segmentación.

## 4.2 ▶ Frecuencia de reloj

un ciclo muy corto

La frecuencia del Pentium 4 comienza en 1.3 GHz, avanzando a pasos de 100 MHz hasta superar los 2 GHz, siempre sobre un bus de 4x100 MHz. En 2.27 GHz aparece la versión con el bus de 4x133, que progresa a pasos de 133 MHz hasta superar los 3 GHz.

etapas estériles  
pág. 219

El procesador dispone de un profundo cauce de segmentación, y eso le otorga cierta facilidad para alcanzar frecuencias elevadas. En el lado negativo, un ciclo de reloj tan corto ha obligado a sus diseñadores a introducir algunas etapas un tanto estériles. Cuando analicemos el cauce de segmentación del procesador (sección 6.4.8), justificaremos mejor esta observación.

## 4.3 ▶ Tecnología de integración

fiabilidad

Cada cual explota las virtudes que tiene, e Intel tiene una bien merecida fama a la hora de integrar sus productos en la planta de fabricación. Sus procesadores siempre se han calentado menos que los de la competencia, permitiéndole su venta a frecuencias superiores y un plus de fiabilidad.

La apuesta inicial de la compañía por los 1.5 GHz para un diseño de 42 millones de transistores, que compartió la ya vieja tecnología de 0.18 micras y conexiones de aluminio del Pentium III, habla muy a las claras de la confianza que depositaron en su infraestructura de fabricación.

número de transistores

La [foto 6.4](#) muestra la disposición de sus 42 millones de transistores en un área de integración de 247 mm<sup>2</sup>, donde hemos delimitado la ubicación de sus principales bloques funcionales.

☛ [pág. 212](#)  
área de integración

Algunos analistas estimaron el coste de producción del Pentium 4 entre los 90 € y 100 €, el doble que su predecesor, mientras que el precio inicial de salida (27/11/00) se fijó en 800 € para el modelo de 1.5 GHz y 600 € para el de 1.4 GHz. El elevado margen de beneficio del producto no delataba sino la enorme devaluación que sufriría en sus primeros meses de existencia: Justo dos meses después del lanzamiento, Intel rebajaría los precios en torno a un 25 %, y un año más tarde su valor apenas rebasaba los 200 €. Tenemos una lección que aprender aquí, pues se repite de forma sistemática: Nunca compre un procesador nada más pisar el mercado. En todo caso, aproveche para adquirir a precio de saldo el modelo al que éste reemplaza.

devaluación

Con el Pentium 4, Intel también renueva su apuesta por el formato Socket, donde esta vez el patillaje se compone de 423 pines y el zócalo acoge la denominación obvia: Socket 423. La [foto 6.5](#) muestra el aspecto del empaquetado FC-PGA y el patillaje del procesador, muy parecido al de los Pentium III Coppermine, aunque algo más aparatoso.

zócalo  
☛ [pág. 213](#)

## El bus local

## ◀ 4.4

“El ancho de banda del bus local es el cuello de botella de una arquitectura PC.” Después de pasarnos toda la década de los 90 en clase repitiendo esta cantinela, nos vimos gustosamente obligados a desterrarla tras la llegada de este procesador.

En el Pentium 4, la frecuencia de trabajo del bus local triplica a la de su predecesor, el Pentium III. Como ambos tienen una anchura del bus de datos de 64 bits, los 4x100 MHz del bus local del Pentium 4 se convierten en un ancho de banda de 3.2 Gbytes/sg, frente a los 1.06 Gbytes/sg del Pentium III y el techo de 2.7 Gbytes/sg alcanzado por el más reciente de los Athlon XP (nacido ya en 2003 bajo el código de referencia Barton - ver [tabla 6.5](#)).

4x100 MHz

Esto descongestiona la entrada al procesador, y al mismo tiempo, le permite aprovechar todas las prestaciones de la memoria RDRAM (ver [sección 10.13.6](#)), que se encontraba infrautilizada sobre las arquitecturas anteriores. No obstante, tras la ampliación de caché L2 que tiene lugar con la llegada del Northwood (ver [sección 6.5](#)), el disfrute de esta conexión ya no es el ansiado bocado de años atrás.

☛ [pág. 205](#)

☛ [p. 73/Vol. 2](#)

☛ [pág. 226](#)

Echando un vistazo algo más en profundidad a la especificación del bus uno descubre cosas aburridas entremezcladas con otras más interesantes. La monotonía: Se sigue la especificación AGTL+ para el establecimiento de los niveles de voltaje de las señales y su tolerancia al ruido, esto es, la misma que se diseñó para el bus local a 100 MHz del Pentium II. Lo vanguardista: El bus se bifurca en tres ramales separados:

AGTL+

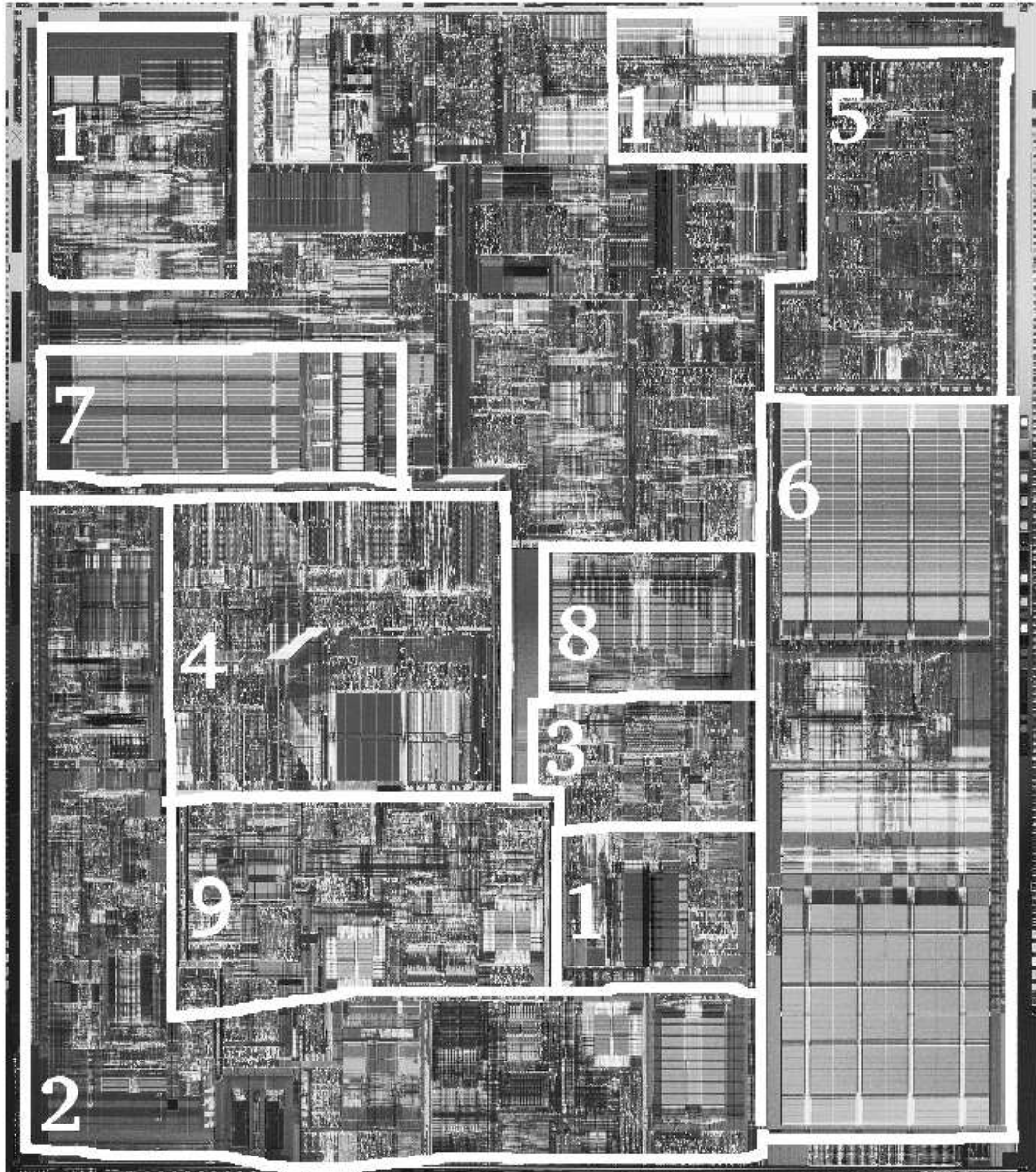
tres caminos

- ❶ Una parte del bus de control se desacopla para dar cobertura y salida a las interrupciones de forma asíncrona. Estas líneas hacen su vida al margen de la señal de reloj del bus.
- ❷ El bus de direcciones (y la parte del bus de control que lo gestiona) funciona sincronizado por un multiplicador de 2x, emitiendo dos direcciones por cada señal de reloj de 100 MHz.
- ❸ El bus de datos (y su parte del bus de control), es el único que trabaja a un régimen de 4x, y sobre él descansa el ancho de banda de 3.2 Gbytes/sg que se atribuye al trasiego de datos.

control

direcciones

datos



Area de integración sin editar: Cortesía de Intel

**FOTO 6.4:** Area de integración del Pentium 4 de 0.18 micras con sus principales bloques funcionales: 1.- Predicción de salto (BTB) y reordenación de instrucciones (BRC). 2.- Control segmentado de 20 etapas. 3.- ALU. 4.- FPU con su banco de registros en la parte inferior derecha. 5.- Controlador de bus local a 400 MHz. 6.- Caché de segundo nivel de 256 Kbytes (L2). 7.- Caché de traza de 12 Kmicroinstrucciones = 180 Kbytes (L1I). 8.- Caché de datos de 8 Kbytes (L1D). 9.- Banco de registros y unidades de ejecución multimedia. Como se aprecia, la BTB se implementa de forma distribuida con objeto de minimizar el retraso en la notificación y confirmación de las predicciones de salto a L1I por un lado (cosa que se realiza en la esquina superior izquierda) y a L2 por otro (parte central inferior).

#### 4.5 ► Diálogos con memoria principal

bus de direcciones

Para entender el funcionamiento del bus de direcciones hay que mirar a lo que hay al otro

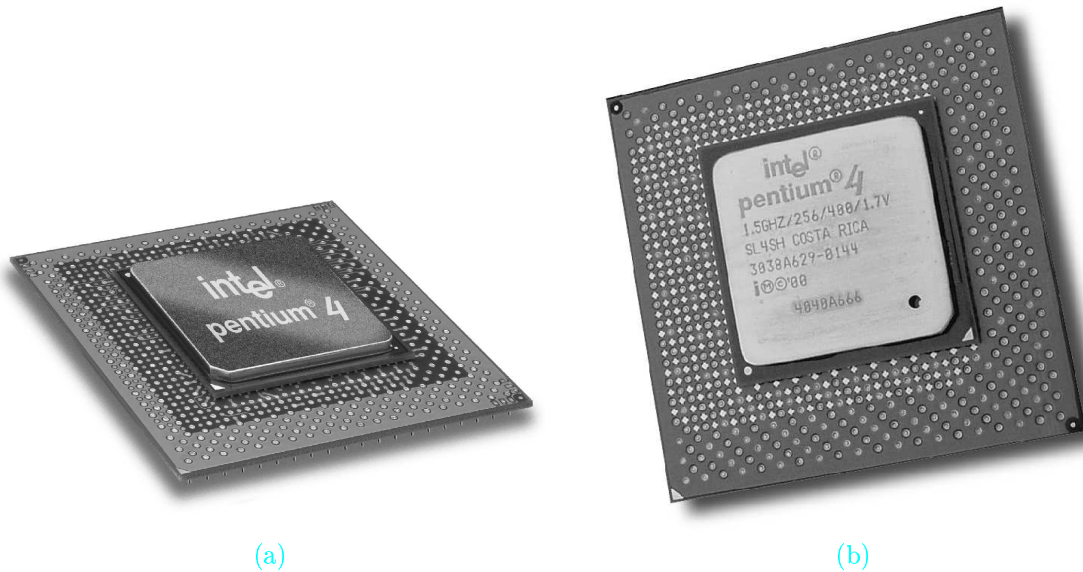


Foto cortesía de Intel

FOTO 6.5: Empaquetado y patillaje del Pentium 4. (a) Parte superior del chip (por la que disipa el calor). (b) Parte inferior (por la que se conecta a la placa base).

lado del procesador: La memoria. Llevamos ya un tiempo en el que la caché ha desaparecido de la placa base, y se encuentra toda dentro del chip del procesador. Esto significa que **siempre** que el procesador sale al bus es porque no ha encontrado dentro el dato que busca, esto es, **se ha producido un fallo en la caché L2**.

Debemos desterrar la vieja idea de que es el procesador quien pide los datos a la memoria principal, porque ya no es así. Quien realmente los pide es el controlador de la caché L2, y ésta no se conforma con una palabra de memoria: Arrastra una línea de caché entera, que en el caso del Pentium 4 es de 128 bytes. Como el bus de datos tiene una anchura de 8 bytes, por cada dirección emitida por el bus de direcciones hay que realizar 16 viajes por el bus de datos. La primera conclusión que sacamos de todo esto es que un bus de direcciones a la mitad de velocidad que su homólogo de datos no sólo no es malo, sino que es bastante más de lo que se necesita.

Para comprender mejor la manera en que trabaja el bus de datos, debemos dejar de mirar a la memoria en abstracto como una suministradora de palabras sueltas, y verla según sus propiedades intrínsecas.

16 viajes

bus de datos

#### 4.5.1 Sobre SDRAM y DDRAM

Ya desde la llegada de los módulos FPM a finales de los años 80 (ver [sección 10.13.1](#)), las memorias responden emitiendo ráfagas de cuatro palabras consecutivas, esquema que se respeta hasta los módulos SDRAM y DDRAM, apoyándose precisamente en el hecho de que al otro lado del bus espera una caché. ¿Cómo se llena entonces una línea de caché? Componiendo las 4 palabras de la ráfaga proveniente de los módulos de memoria con el factor 4x del bus de datos. Así salen los 16 viajes necesarios.

p. 55/Vol. 2

ráfagas

¿Hay que solicitar entonces 4 ráfagas a memoria de cuatro datos cada una? Eso depende de la memoria de que dispongamos:

- 1 Si tenemos una SDRAM de 100 MHz, la respuesta es sí.

SDRAM 100 MHz

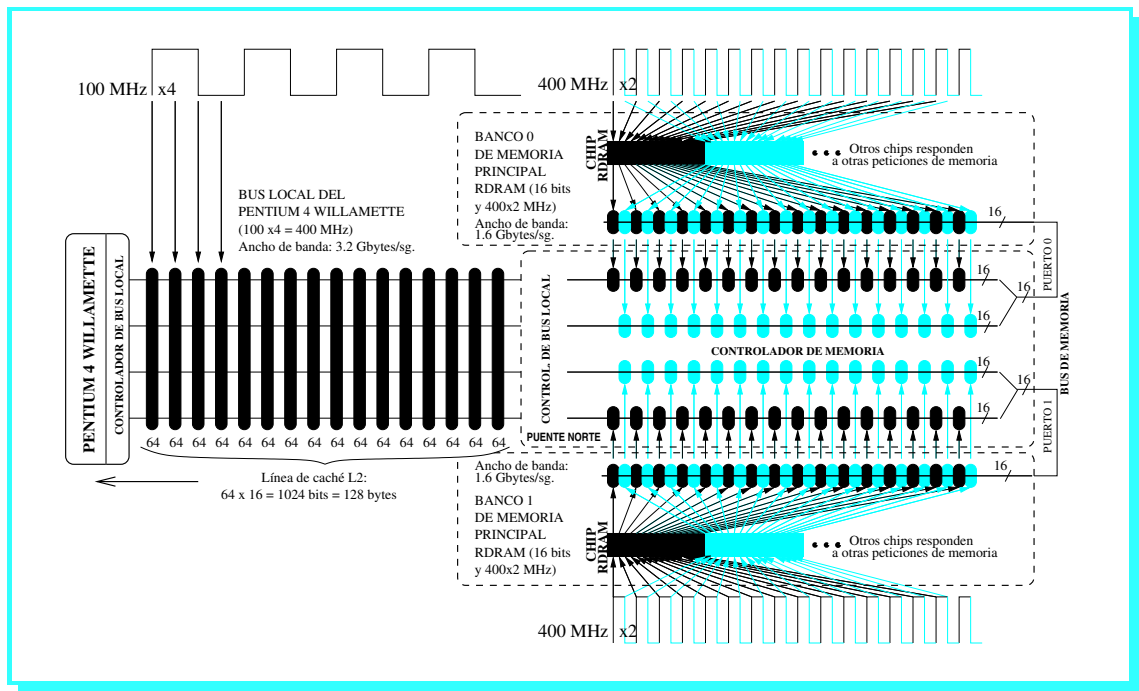


FIGURA 6.5: El perfecto entendimiento del Pentium 4 con la versión base de la memoria RDRAM habilitando un doble puerto que duplica el ancho de banda pero obliga a llenar los zócalos RIMM por pares.

DDRAM 2x100 MHz

- ❷ Si tenemos una DDRAM de 2x100 MHz, la respuesta es no: Sólo hace falta solicitar dos ráfagas, pues las otras dos se intercalan en los flancos de bajada de esa señal de reloj.

DDRAM 2x200 MHz

- ❸ Si disponemos de una DDRAM de 200x2 MHz, sólo haría falta un acceso, y en ese caso el sistema funcionaría **exactamente igual** que en un Pentium II Deschutes o Pentium III Katmai (los dos con bus a 100 MHz): Realizando cuatro viajes por el bus de 400 MHz, ya que la DDRAM de 200x2 funciona como una SDRAM 100 MHz a 4x, y este multiplicador se alinearía con el del bus local para sincronizar las cuatro ráfagas. Fijándose ahora sólo en la primera de ellas y olvidando el 4x, lo que nos queda es una SDRAM a 100 MHz sobre un bus de su misma frecuencia.

analogía con  
bus 100 MHz

Si nos quedamos en este último escenario, la línea de caché L2 del Pentium 4 se llena ahora exactamente en el mismo tiempo que tardaban sus antepasados, e incluso siguiendo un protocolo bastante similar, a diferencia de que en el nuevo procesador el tamaño de línea es cuatro veces superior, que es en lo que se ha empleado la mejora del ancho de banda del bus en esa misma proporción.

localidad  
asimetría

Así se consigue aumentar el índice de aciertos a caché a poco que el algoritmo exhiba una mínima localidad, y de paso, diferenciar la organización de la L1, consiguiendo una provechosa asimetría en los contenidos de cada una de ellas. Esto permite al procesador trabajar de forma más autónoma respecto al exterior, que ha sido tradicionalmente el gran lastre de las arquitecturas PC.

problema  
temporal

p. 72/Vo1.2

todo queda en  
4x66 MHz

El único problema fue que la memoria DDRAM de 2x200 MHz no vió la luz en forma de chips hasta 2002 (ver [tabla 10.13](#)), e integrada en módulos hasta 2003. Hasta entonces, lo que se hizo fue montar DDRAM 133x2 MHz, lo que permite trabajar exactamente de la misma manera que hemos contado, sólo que el bus local debe convertirse en un 4x sobre 66 MHz en lugar de sobre 100 MHz, desvaneciéndose el 33% de su rendimiento. Por eso, nuestro mejor candidato en términos de rendimiento en este estadio de la evolución del Pentium 4 era la RDRAM, a pesar de que no había tenido toda la difusión comercial que su homóloga en los equipos comercializados bajo



Pentium 4. Más adelante, con la llegada del procesador de 0.13 micras, el interfaz con memoria DDRAM mejora ostensiblemente (ver [figura 6.10](#)).

pág. 228

#### 4.5.2 Sobre RDRAM

Intel no sólo se ha fijado en la operativa de la memoria DDRAM para implementar su bus local, sino que ha encontrado un denominador común con el que cubrir también la otra vertiente del mercado: La RDRAM. Cuando lleguemos a la [sección 10.13.6](#) advertiremos que la forma en que la RDRAM se comunica con el bus de memoria guarda una sospechosa relación con el interfaz que se establece en el bus local del Pentium 4 para llenar una línea de caché de 128 bytes.

p. 73/Vol. 2

Por ahora simplificaremos diciendo que los 1.6 Gbytes/sg. de ancho de banda que proporciona la versión base de RDRAM a 400x2 MHz se enganchan al bus local por el puente norte i850 del juego de chips, donde Intel ha habilitado sabiamente un doble puerto que permite llenar el caudal de 3.2 Gbytes/sg. del bus local del Pentium 4. El entendimiento es perfecto, y el rendimiento, óptimo, tal y como hemos ilustrado en la [figura 6.5](#). El único inconveniente es que se exige llenar los zócalos RIMM por pares.

doble puerto

pág. 214

La idea es tan sencilla, y a la vez tan potente, que AMD se fijaría en ella más adelante para dotar de un espectacular ancho de banda a su K8 (ver [sección 7.1.7.4](#)).

pág. 269

### El conjunto de instrucciones y su decodificación

#### 4.6

Tras las buenas noticias, llegan las malas: El Pentium 4 sigue aceptando como entrada código nativo del conjunto de instrucciones x86. Si con el tapón del bus llevábamos diez años, con éste llevamos veinticinco, y en vista de cómo están las cosas y la decepción que nos ha causado el conjunto de instrucciones x86-64 de octava generación (ver [sección 7.1.6.3](#)), no parece que vayamos a perderlo de vista tan fácilmente.

malas noticias:  
x86

pág. 263

El esquema de decodificación de instrucciones es mimético al ya comentado para el Pentium Pro (ver [sección 5.1.1](#)): Se toma una instrucción x86 compleja de 3.5 bytes de longitud media, se descompone en microinstrucciones nativas de la nueva arquitectura (tres de media), y a partir de ahí, se trabaja como si el código de entrada hubiera sido un RISC puro. Esto ocasiona un doble problema:

decodificación

pág. 151

conversión

dos problemas:

- ❶ El código que generan los compiladores y ensambladores es x86, con lo que cualquier optimización software no puede realizarse sobre el código nativo del procesador, sino sobre otro que no aprovecha las cualidades de la arquitectura sobre la que se ejecuta. Es decir, la conversión de código compatible x86 es una barrera que impide a la capa software conectar con la capa hardware y unir sinergias.
- ❷ En el proceso de conversión interno que tiene lugar se pierden varios ciclos por cada instrucción: Las etapas de búsqueda y decodificación del procesador se dilatan a lo largo de 2 y 3 ciclos de reloj, respectivamente, cuando en un código nativo no tardarían más de uno cada una. Se esfuman tres ciclos sobre veinte que consume cada instrucción, lo que supone el 15% del tiempo útil del microprocesador.

efectividad

penalización

Nos parece bastante más dañino el primero de los dos factores apuntados, aunque sólo podría subsanarse sacrificando la compatibilidad con todas las aplicaciones software desarrolladas para los procesadores anteriores, algo que Intel sólo se ha atrevido a intentar una vez en 25 años (con el Itanium, y ni eso, porque a éste le incorporaron finalmente un modo de trabajo compatible x86).

atrevimiento

Intel parece haberse fijado más en el segundo factor. Y para solventarlo, ha ideado un nuevo concepto: La caché de traza, que es objeto de nuestro estudio a continuación.

## 4.7 ► Memoria interna

La caché de traza completa el conjunto de cachés internas al procesador con la caché de datos de primer nivel (L1D) y la caché de datos unificada de segundo nivel (L2), amén de los diferentes bancos de registros. Empezaremos describiendo éstos.

### 4.7.1 Bancos de registros

Junto al registro de estado EFLAGS para operaciones enteras y al registro MXCSR para operaciones de punto flotante, ambos de 32 bits, el Pentium 4 dispone de los siguientes bancos de registros:

EAX, EBX, ...	❶ 8 registros de propósito general para datos enteros, con una anchura de 32 bits cada uno y referenciados por los nombres EAX, EBX, ECX, EDX, EBP, ESI, EDI y ESP.
MM0-MM7	❷ 8 registros MMX para datos de 64 bits y referenciados por los nombres MM0-MM7. Sobre ellos se realizan las 57 operaciones multimedia MMX, y sobre su extensión solapada de 80 bits, las operaciones de punto flotante, conservando el esquema implementado en el procesador MMX original.
XMM0-XMM7	❸ 8 registros XMM para datos de 128 bits y referenciados por los nombres XMM0-XMM7. Sobre ellos se realizan las 70 operaciones del conjunto SSE que introdujo el Pentium III y las 144 nuevas instrucciones del conjunto multimedia SSE2 originales del Pentium 4.

pág. 217

La [figura 6.6](#) muestra todos los bancos de registros del procesador Pentium 4.

### 4.7.2 Caché de traza, o L1I

12 Kinstrs. La caché de traza es la caché de instrucciones de primer nivel del Pentium 4, alias L1I, pero en lugar de tener 16 Kbytes como en la arquitectura P6, cuenta con un espacio para 12 K-microinstrucciones, donde lo que se guardan son las traducciones de las instrucciones x86 del programa en ejecución.

evita redescodificar Así, cuando el procesador se encuentra ejecutando el bucle de un programa y la caché ha retenido todas las instrucciones del mismo, no es necesario volver a realizar su conversión a microinstrucciones, sino que se toman directamente estas últimas, ahorrándose los ciclos de decodificación desde código x86, que dicho sea de paso, forman parte del camino crítico del procesador y conforman ahora su principal cuello de botella.

pág. 218

ubicación

La [figura 6.7](#) muestra la transformación que produce la presencia de la caché de traza en la nueva arquitectura con respecto a los modelos precedentes.

Acabamos de llevarnos por delante el concepto de línea de caché como bloque de transporte con memoria principal. Ahora, estas dos memorias guardan información semánticamente equivalente, pero en distinto formato, y el transporte de una a otra que tiene lugar cuando se produce un fallo a caché comporta realizar la conversión de uno a otro.

propósito específico En definitiva, lo que ha hecho Intel es aprovecharse de que las cachés llevan ya un tiempo dentro del procesador para concebirlas como algo de propósito específico para su arquitectura, en lugar de como algo general que funciona de la misma manera en todos los procesadores y no aprovecha sus rasgos distintivos.

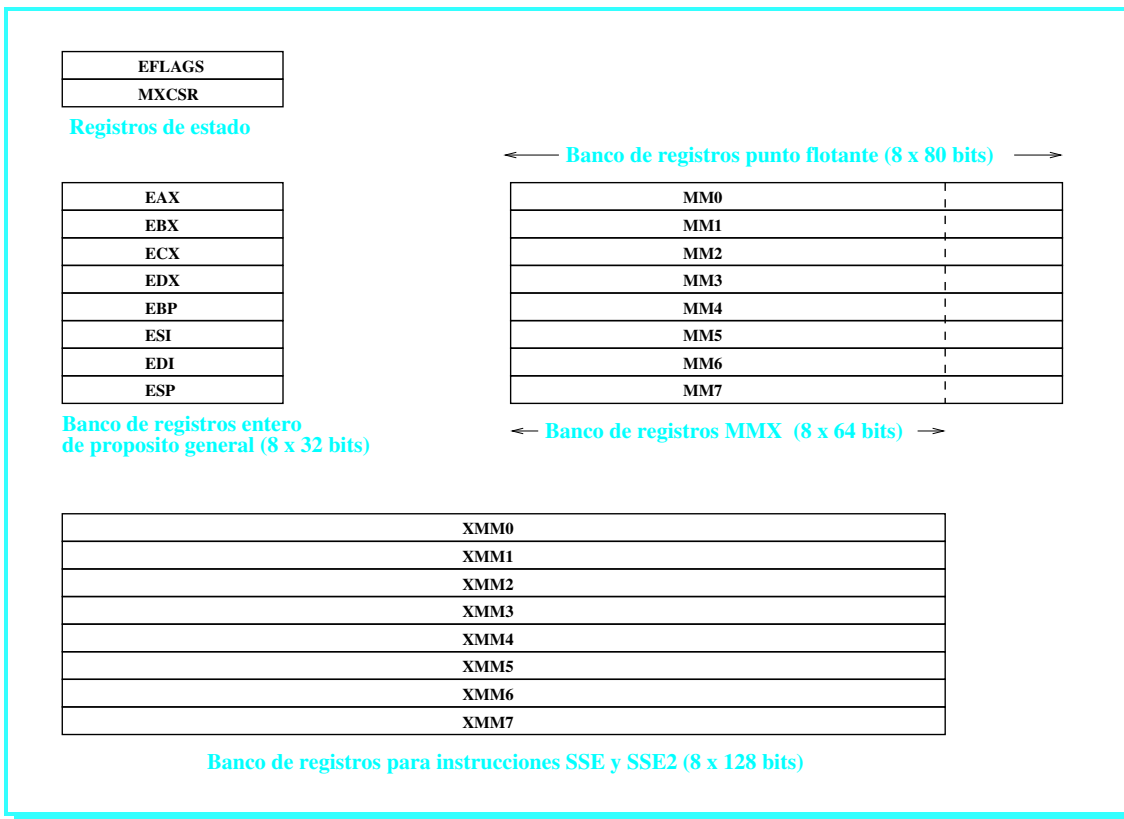


FIGURA 6.6: Los diferentes bancos de registros del Pentium 4.



**Analogía 6.1: IDEAS PRECURSORAS DE LA CACHÉ DE TRAZA**

Encontramos cierto parecido entre la caché de traza y un concepto más antiguo, el de caché direccionada virtualmente, donde la información se direcciona de la forma que mejor relación guarda con el procesador (dirección virtual) en lugar de utilizar la que conecta con el exterior (dirección física). La traducción que aquí se ahorra es la de dirección virtual a física, ya sea directamente o a través de una TLB (Translation Look-aside Buffer - o caché para la traducción de direcciones).

Más recientemente, hemos visto a Transmeta incorporar algo muy parecido a sus microprocesadores Crusoe un año antes de que Intel propugnara su caché de traza como poco menos que una revolución. La decodificación del Crusoe es más lenta aún que la del Pentium 4, y por tanto, este mecanismo tiene allí más sentido si cabe. Y a pesar de ello, Transmeta ha sido muy sigiloso con esta singularidad de sus procesadores.

Sea como fuere, vamos a tratar de buscar una equivalencia de esta caché de traza con una caché de instrucciones convencional, al menos en su tamaño y organización, con objeto de poder tener una impresión más certera acerca de su rendimiento.

equivalencia en tamaño

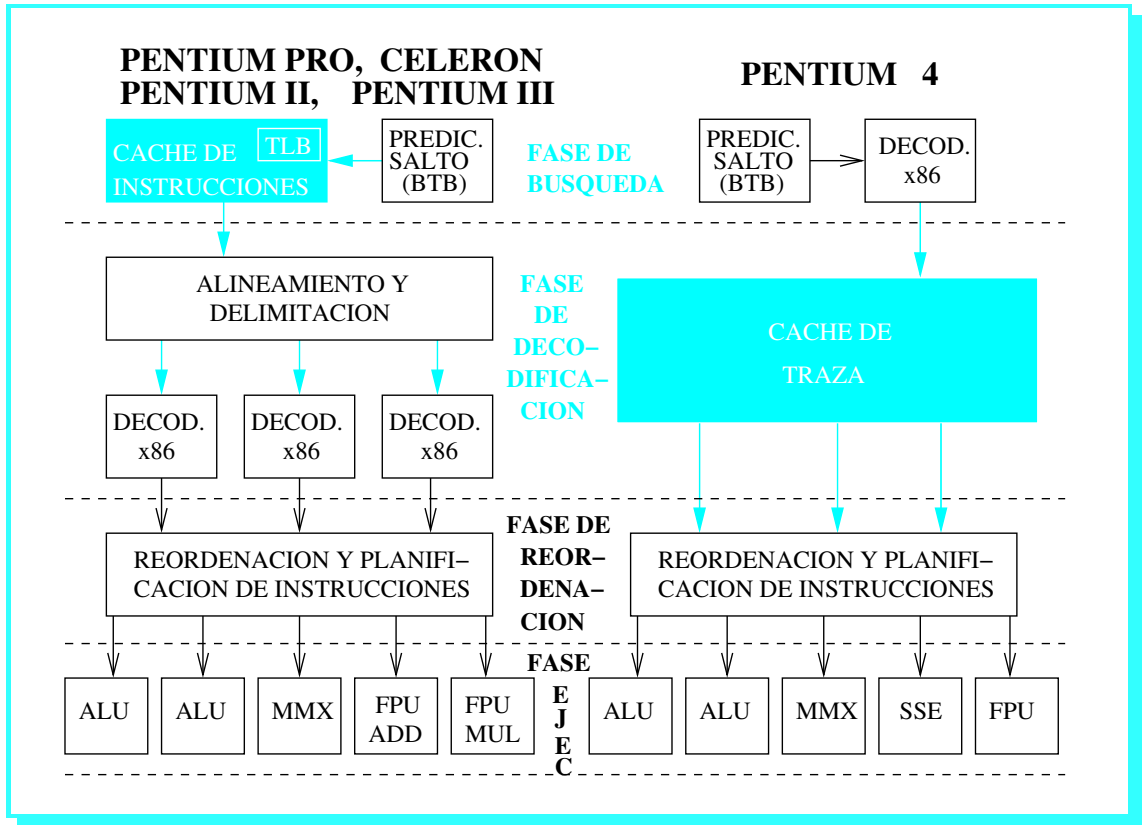


FIGURA 6.7: Una de las principales metamorfosis que produce la nueva arquitectura del Pentium 4 con respecto a la de sus antecesores es la presencia de la caché de traza en lugar de la tradicional caché de instrucciones en el camino de decodificación de instrucciones del procesador.

Cada una de las microinstrucciones del Pentium 4 ocupará en torno a 120 bits (15 bytes). Este número no ha sido proporcionado por Intel; es una estimación nuestra fruto de recordar los 118 bits de anchura de las microinstrucciones de la arquitectura P6 y observar que éstas deben extenderse con la codificación de las nuevas instrucciones multimedia del conjunto extendido SSE2, para las que dos nuevos bits son más que suficientes. Como el tamaño de la caché de traza era de 12 Kmicroinstrucciones, esto nos da un total de 180 Kbytes de caché L1I. Respecto a la organización de esta caché, las líneas coinciden en tamaño con los 15 bytes, y éstas a su vez se agrupan en conjuntos de 4 líneas, es decir, el mismo grado de asociatividad que en el Pentium III y sus antecesores.

180 Kb L1I  
hereda asociatividad

Ahora bien, la capacidad para almacenar información no es tan grande como se desprende de estos números. De cada instrucción del Pentium III que se alojaba en su caché L1I sale una media de 3 microinstrucciones, que serían las que se alojarían en la caché de traza del Pentium 4. El Pentium III tiene un formato de instrucción variable que oscila entre 1 y 12 bytes, siendo su media de una longitud de 3.5 bytes. Esto nos dice que 28 bits alojados en la L1I del Pentium III representan la misma cantidad de información que  $3 \times 120 = 360$  bits en la caché de traza. De aquí resulta un factor de compensación de tamaños de 1:12.85, lo que nos deja una L1I equivalente de 14 Kbytes. Es algo inferior a la del Pentium III, pero mucho peor es que, a la hora de la implementación los que cuentan son los 180 Kbytes. Es decir, integrar esta caché le ha costado a Intel más del doble de lo que le costó a AMD colocar a su K7 su hermosa caché L1I de 64 Kbytes.

equiv: 14 Kb.

La caché de traza alivia mucho el camino crítico del procesador, situado en su vía de decodificación, pero también encarece mucho el coste del diseño. Con la cantidad de transistores que se lleva una caché interna, ya tenemos un presunto culpable para justificar los 42 millones de

coste

transistores del Pentium 4 frente a los 28 del Pentium III. Este sobre coste sólo queda parcialmente mitigado porque la presencia de la caché de traza simplifica mucho el hardware de decodificación de instrucciones tan complejo que arrastramos del P6.

### 4.7.3 Caché de datos (L1D)

Se compone de un total de 8 Kbytes organizados en 32 conjuntos, 4 líneas por conjunto y 64 bytes por línea, con un tiempo de acceso de un solo ciclo del procesador.

8 Kbytes

La organización y la velocidad nos parecen buenas, pero no así el tamaño, que reduce a la mitad el de los Pentium II y III, cuando el Pentium 4 está destinado a manejar un volumen de memoria muy superior. Desde los tiempos del K5 de AMD y otras arquitecturas minoritarias de Cyrix a mediados de los años 90 no veíamos un procesador en entorno PC con un primer nivel de caché en el que las secciones de datos e instrucciones no fueran gemelas <sup>3</sup>.

no gemelas

Sólo vemos dos justificaciones posibles para que esto no resienta de manera importante el rendimiento del procesador:

- ❶ Los numerosos bancos de registros de que se dispone suplen parcialmente estas carencias.
- ❷ El paralelismo a nivel de instrucción del procesador se ha desplazado hacia la segmentación en lugar de sobre la superescalaridad, cuando es esta última la que mayor demanda de datos provoca.

### 4.7.4 Caché de segundo nivel (L2)

Dispone de un total de 256 Kbytes organizados en 256 conjuntos, 8 líneas por conjunto y 128 bytes por línea, capaz de transferir 32 bytes en cada ciclo de reloj, lo que determina un ancho de banda de 44.8 Gbytes/segundo para un procesador de 1.4 GHz. De cara a la comunicación con la caché L1 de datos en la que las líneas son de 64 bytes, las líneas de la L2 se dividen internamente en dos bloques de 64 bytes que tardan dos ciclos en llenarse. Estamos ante una caché muy parecida a la del Pentium III, con la salvedad de que las líneas de caché son aquí cuatro veces más grandes, lo que favorecerá al software secuencial que muestre una fuerte localidad en las referencias a memoria.

256 Kbytes  
organización

## Segmentación

### ◀ 4.8

La segmentación del Pentium alcanza una longitud inesperadamente larga: 20 etapas, casi el doble que su arquitectura predecesora P6. A este diseño, Intel lo bautiza con el nombre de **Hyper-pipelined technology**. Esto representa una apuesta arriesgada: Cuando el procesador encuentre código secuencial escrito sin dependencias, alcanzará una velocidad veinte veces superior, pero cuando le llegue algún conflicto de datos o control, la penalización será mayúscula porque un número creciente de instrucciones se verán afectadas. Esto explica que en el cauce segmentado haya tanta dedicación a resolver las dependencias, tal y como se observa en la [figura 6.8](#). Una descripción más detallada acerca de cada etapa de segmentación se encuentra más adelante, en la [sección 6.6](#).

20 etapas

Hyper-pipelined  
alto riesgo

coberturas

☛ [pág. 220](#)☛ [pág. 231](#)

Observando las etapas de segmentación en sí, vemos que algunas de ellas presentan cierta ingenuidad:

- Las etapas 5 y 20 no hacen otra cosa que transportar un dato, desde la caché de instrucciones

transporte

<sup>3</sup>De hecho, diversos estudios preliminares acerca del procesador erraron en esta variable, al dar por supuesta una configuración de 16K+16K, repitiendo el esquema del P6.

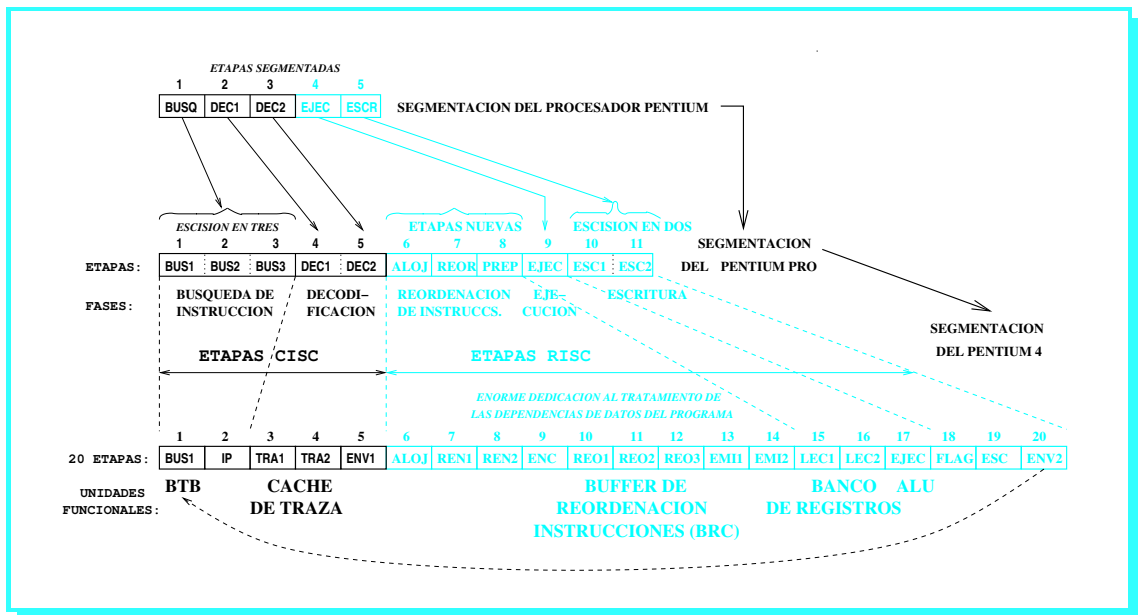


FIGURA 6.8: El cauce de segmentación de la familia P67 (Pentium 4) y su analogía con las familias P6 (Pentium Pro, II y III) y P5 (Pentium). Podemos apreciar la cantidad de tiempo que se dedica a reducir las dependencias de datos del programa.

al búfer de reordenación de instrucciones en el primer caso y desde la ALU a la BTB de la caché de traza en el segundo (ver foto 6.6).

☛ pág. 221

ubicación

demora

- Otras dos etapas (la 2 desde la BTB a la caché de traza y la 14 de emisión a las ALU) consumen el grueso de su tiempo en el transporte y la ubicación de datos.
- Tres unidades funcionales extienden su ejecución a dos o tres ciclos de forma inusual. Nos referimos a las unidades de renombrado de registros virtuales (2 ciclos), planificación de la ejecución (3 ciclos) y lectura del banco de registros (2 ciclos).

Invitamos al lector a que, basándose en las conexiones que deben habilitarse entre las distintas unidades funcionales como consecuencia del secuenciamiento de etapas mostrado en la figura 6.8, ingenie una disposición física para las unidades funcionales del Pentium 4 que de lugar a comunicaciones más cercanas en las que pudiera haberse evitado alguno de los ciclos muertos comentados (sin incurrir en otros nuevos, claro está). La solución aportada por Intel es la que apareció en la foto 6.4, y está llena de ingenio, dando a su vez muchas pistas sobre las premisas básicas que deben respetarse.

disposición física

pág. 212 ☛

El breve análisis realizado para la acumulación de ciclos en las tareas de transporte y procesamiento para este procesador no persiguen otra cosa que la didáctica de una hermosa lección: Para que la frecuencia haya podido saltar de 1 GHz del Pentium III a 1.5 GHz del Pentium 4 ha sido necesario introducir siete ciclos más en el nuevo procesador que no estaban en el cauce segmentado de su predecesor. El paso de 13 a 20 etapas de ejecución para cada instrucción supone ya un retraso superior al 50 %, esto es, más que la ganancia en velocidad conseguida a través del aumento en 500 MHz de la frecuencia de reloj de 1 GHz del Pentium III.

7 ciclos extra

pág. 49 ☛

Ya avisamos en la sección 3.1 que la frecuencia de reloj, aún siendo muy importante para un procesador, no es un fiel reflejo de lo que podemos esperar de él. Situaciones como la descrita la convierten en un parámetro engañoso para los usuarios y traicionero para los diseñadores, y mientras para unos estas situaciones pasarán completamente inadvertidas por mucho que quieran fijarse, para los otros serán perfectamente perceptibles sin necesidad de fijarse porque querrán

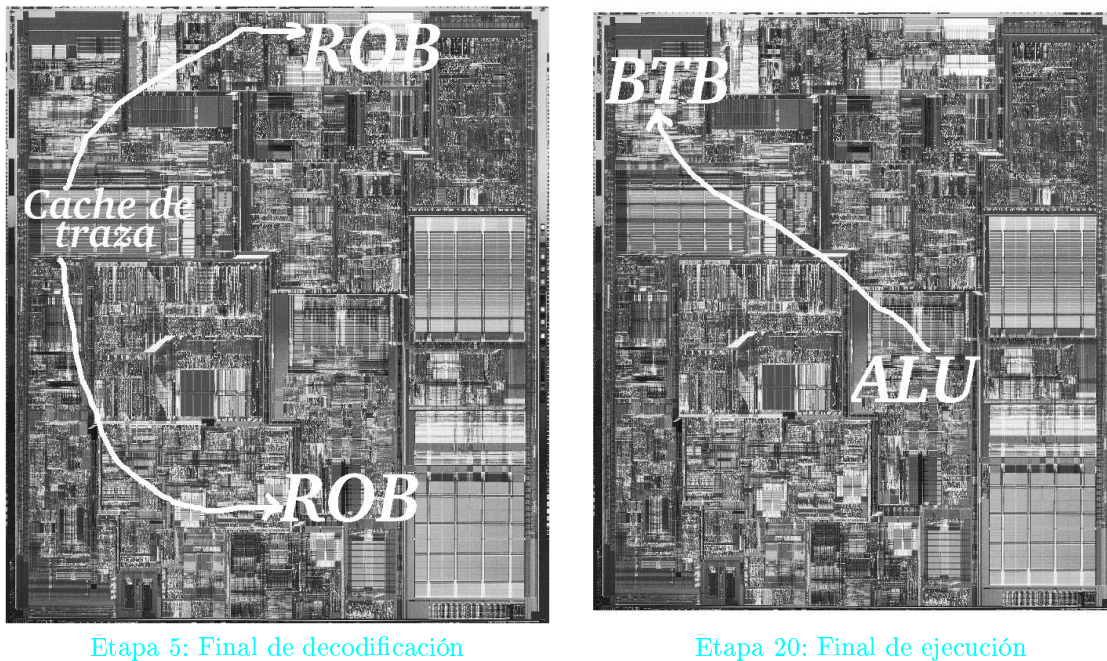


FOTO 6.6: Las dos etapas de transporte que consumen un ciclo de reloj completo en el Pentium 4: La etapa 5, que envía la información de la caché de traza al sistema de planificación y reordenación de instrucciones al final de la fase de decodificación, y la etapa 20 final, donde el resultado de la condición de salto es enviado desde la ALU a la BTB para detectar posibles predicciones de salto erróneas.

mirar para otro lado.

## Superescalaridad

◀ 4.9

El talante superescalar del Pentium 4 es una de las cosas que más desapercibida pasa en el diseño del procesador. Con una elevada segmentación y aceptando instrucciones a una elevada frecuencia, el procesador ya tiene suficiente paralelismo en cada ciclo como para encima desdoblar agresivamente sus cauces de ejecución. Intel opta de esta manera por aceptar una sola instrucción por ciclo de reloj, y habilitar la superescalaridad únicamente cuando ésta se descompone en microinstrucciones. Al igual que en sus hermanos menores con los que comparte el código fuente, esto supone transformar una media de tres microinstrucciones por ciclo de reloj, con lo que éste sería su factor superescalar.

factor tres

Es importante no confundir este factor superescalar de tres con el dado para el K7, pues en este último, el tres iba referido al número de instrucciones fuente, y asumido después un desdoble en microinstrucciones de similares características, el potencial de superescalaridad máximo sube hasta nueve, quedándose la media en torno a cinco.

versus K7

## Resolución de dependencias

◀ 4.10

Una segmentación tan grande hace que el rendimiento del procesador dependa extraordinariamente de la BTB o sistema dinámico de predicción de salto y las unidades de anticipación de datos y planificación de instrucciones, ya que éstas son las responsables de solventar las dependencias de datos y control, respectivamente.

En consecuencia, el rendimiento del Pentium 4 depende más que ningún otro de la bondad

del programa con el que se le ponga a prueba. En este sentido, el procesador tendrá que pasar dos exámenes bien diferentes:

- |             |  |
|-------------|--|
| tipo entero | ❶ Sobre aplicaciones de ofimática y de tipo entero (procesadores de texto, hojas de cálculo, bases de datos, ...), que suelen tener un elevado número de saltos difíciles de predecir y bastante trasiego de datos. Todo esto perjudica más las características del Pentium 4. |
| multimedia  | ❷ Sobre aplicaciones multimedia y de punto flotante (autocad, juegos, cálculos científicos, ...). Suelen contener saltos muy predecibles y poco volumen de datos, con lo que aquí el procesador debe mostrar su mejor cara.  |

Veamos lo que ha dispuesto Intel para solventar el peor de los escenarios posibles:

- |     |   |
|-----|---|
| BTB | ❶ Para los conflictos de control: Una BTB (Branch Target Buffer) de 4096 entradas, ocho veces superior a la de la arquitectura P6 y capaz de almacenar información histórica más detallada acerca de cómo se comportó cada instrucción de salto las últimas veces que se ejecutó. En base a esta información, el Pentium 4 construye un algoritmo más preciso en las predicciones que según sus creadores reduce un 33 % el número de predicciones erróneas. Esto lo expresaremos de otra manera más realista: Si el Pentium acertaba en un 80 % de los saltos (por término medio) y la arquitectura P6 lo hacía en un 90 %, el Pentium 4 acertará un 93 %. Visto así no parece mucho, pero ya adelantamos que cada salto que obligue a vaciar un cauce de 20 etapas hará mucho daño a la ejecución de instrucciones. |
| BRC | ❷ Para los conflictos de datos: Un BRC (Búfer de Reordenación Circular) con una ventana de 126 instrucciones para encontrar la permutación de instrucciones más adecuada (frente a las 44 del P6), y dos unidades para la carga y el almacenamiento retrasados en memoria de 48 y 24 instrucciones, respectivamente (16 y 12 en el P6).   |

pág. 151 ➔

ventana de instrucciones

El cometido de esta unidad BRC también fue explicado ya (ver [sección 5.1.1](#) dedicada a la arquitectura P6): Reordenar las instrucciones para alejar las dependencias de datos con objeto de minimizar esperas y ocultar latencias derivadas de la resolución de conflictos. El tamaño de la ventana de instrucciones hace referencia a la longitud del fragmento de código en el que el procesador puede fijarse para buscar la reordenación más óptima, aumentándose la probabilidad de encontrarla cuanto mayor sea el ámbito de búsqueda.

Respecto a las unidades de carga y almacenamiento retrasados, permiten al procesador intercalar instrucciones útiles en los ciclos muertos que aparecen cuando se está esperando la llegada de operandos de memoria ante un fallo en caché interna.

- |              |  |
|--------------|--|
| complementos | Estos mecanismos complementan su eficacia con las cachés internas, responsables de minimizar esta espera asociada a las operaciones de acceso a memoria. El conjunto es un sistema bastante complejo pero al mismo tiempo eficaz frente al que es el gran enemigo de un diseño como el del Pentium 4: La resolución de dependencias. |
|--------------|--|

## 4.11 ▶ Instrucciones multimedia

- |            |   |
|------------|---|
| 144 nuevas | El Pentium 4 extiende el conjunto de instrucciones multimedia del Pentium III con 144 nuevas instrucciones de tipo SIMD denominadas SSE2 ( <i>Streaming SIMD Extensions 2</i> ). Estas nuevas instrucciones pueden clasificarse en tres grandes grupos: |
|------------|---|

- |                |   |
|----------------|---|
| enteras        | ❶ Operaciones de aritmética enteras definidas sobre múltiples operandos enteros alojados en porciones de 32 bits o inferiores de cada nuevo registro XMM de 128 bits. |
| punto flotante | ❷ Operaciones de punto flotante sobre datos de doble precisión alojados en porciones de 32 bits o superiores de cada nuevo registro XMM de 128 bits.                  |
| memoria        | ❸ Operaciones específicas de gestión de memoria y caché.  |



Característica	Pentium III 1GHz	Pentium 4 1.4 GHz	Mejora	
<b>Frecuencia</b>				
Interna (núcleo)	0.45-1 GHz	1.3-1.5 GHz	+40% ▲	👍 👍 👍
Externa (bus local)	133 MHz	400 MHz	+200% ▲	
Multiplicador	8	3.5	-56% ▲	
<b>Fabricación</b>				
Tecnología de integración	Transistores: CMOS 0.18 micras		0% =	
	Interconexiones: Aluminio		0% =	
Area de integración	106 mm <sup>2</sup>	247 mm <sup>2</sup>	+134% ▼	
	28.1 Mtransist.	42 Mtransist.	+50% ▼	
Zócalo	Socket370	Socket423	0% =	
<b>Variables eléctricas</b>				
Potencia disipada	33 vatios	68 vatios	+57% ▼	👎 👎
Máxima temperatura	70 °C	70 °C	0% =	
Voltaje	1.70-1.76 voltios	1.56-1.70 voltios	-6% ▲	
<b>Paralelismo a nivel de instrucción</b>				
Longitud del cauce	11 etapas	20 etapas	81% ◆	👎
Reordenación instrs.	40 instrs.	126 instrs.	+215% ▲	👍
Carga retrasada	16 instrs.	48 instrs.	+200% ▲	
Almacen. retrasado	12 instrs.	24 instrs.	+100% ▲	
Tamaño BTB	512 saltos	4096 saltos	+700% ▲	👍
<b>Búsqueda de microinstrucciones (<math>\mu</math>i.)</b>				
Velocidad máxima	6 $\mu$ i/ciclo (1 dec)	12 $\mu$ i/ciclo (2 dec)	+100% ▲	
Ancho de banda	3000 M $\mu$ i/sg.	4200 M $\mu$ i/sg.	+40% ▲	
<b>Ejecución de sumas enteras</b>				
Tiempo ejec. ALU	1 nseg.	<0.36 nseg.	> +177% ▲	
Ancho de banda	2000 Mill./seg.	5600 Mill./seg.	> +177% ▲	
<b>Memoria caché integrada (todas a la velocidad del procesador)</b>				
Tamaño total L1D	16 Kbytes	8 Kbytes	-50% ▼	👎 👎
Tamaño total L1I	16 Kbytes	aprox. 14 Kbytes	-12.5% ▼	
Asociatividad L1D y L1I	4 líneas/cto	4 líneas/cto	0% =	
Tamaño línea L1D	32 bytes	64 bytes	+100% ◆	
Ancho de banda L1D	16 Gbytes/sg.	44.8 Gbytes/sg.	+180% ▲	
Tamaño total L2	256 Kbytes	256 Kbytes	0% =	
Asociatividad L2	8 líneas/cto	8 líneas/cto	0% =	
Tamaño línea L2	32 bytes	128 bytes	+300% ◆	
Ancho de banda L2	16 Gbytes/sg.	44.8 Gbytes/sg.	+180% ▲	
<b>Conjunto de instrucciones</b>				
Base	x86	x86	0% =	👎 👎
Extendido	MMX + 70 SSE	PIII + 144 SSE2	▲	👍
<b>Infraestructura externa</b>				
Memoria principal	SDRAM, DDRAM	RDRAM, DDRAM	▲	👍
Placa base (chipset)	440BX, i820	i850, Brookdale	▲	

**TABLA 6.6:** Resumen de las prestaciones del Pentium 4 de 1.4 GHz y comparativa frente al Pentium III de 1 GHz. Todos los datos y porcentajes toman estas frecuencias como referencia. El tamaño de la caché L1I es el equivalente que calculamos para la caché de traza. Hemos señalado con ▲ aquellos aspectos que suponen una mejora en el Pentium 4, con ▼ los que indican un receso (ambos cuantificados porcentualmente), y con ◆ los que resultan más difícil de valorar. En el margen derecho, destacamos lo más positivo (👍) y negativo (👎).

Al ser los registros XMM el doble de anchos que los MMX, ciertas operaciones repetitivas podrán quedar definidas en el Pentium 4 con tan sólo la mitad de instrucciones utilizando el

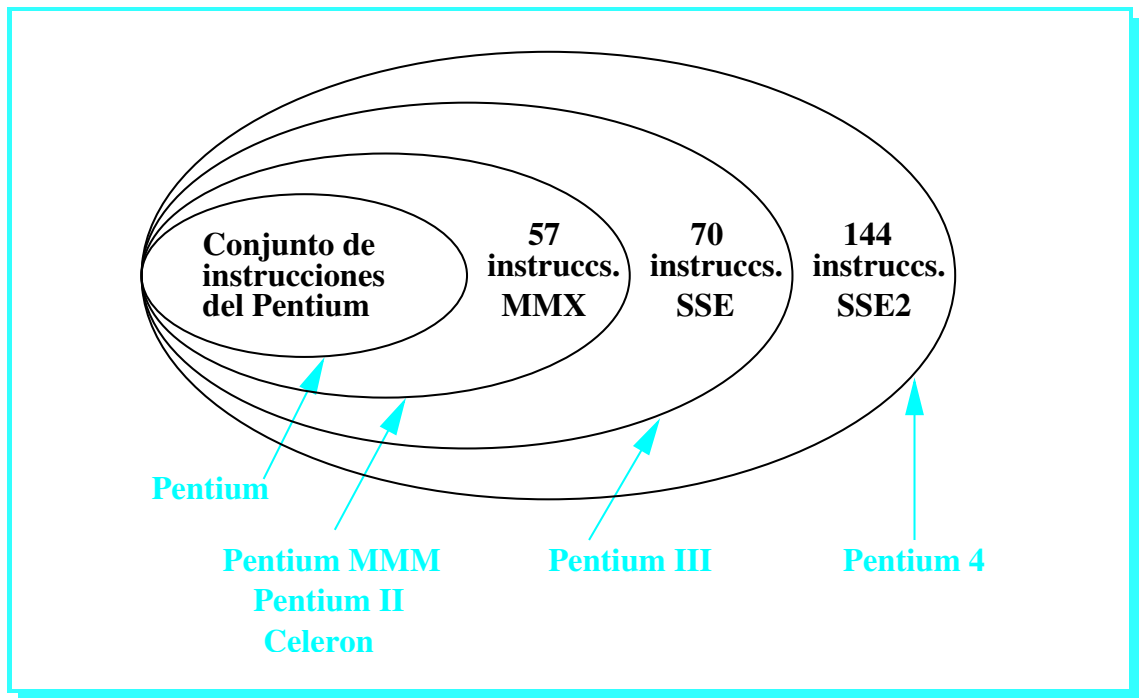


FIGURA 6.9: Los sucesivas extensiones multimedia realizadas por Intel sobre sus diferentes procesadores hasta llegar al Pentium 4.

nuevo conjunto XMM.

La figura 6.9 muestra las sucesivas extensiones multimedia desde una perspectiva evolutiva hasta llegar al Pentium 4.

## 4.12 ► Unidades de ejecución

Se dispone de sólo dos ALU enteras, pero funcionando al doble de velocidad del procesador. Funcionalmente, esto es como disponer de cuatro ALU, y mucho mejor que habilitar sólo dos y segmentarlas, pues el impacto negativo de las dependencias de datos en este caso sería muy superior. También se dispone de una unidad de generación de dirección para operaciones de carga en memoria y otra similar para operaciones de almacenamiento, cada una con su unidad de carga y almacenamiento para albergar las operaciones de carga/almac. retrasado. En total, seis unidades en el cauce entero, cada una con puerto de acceso propio.

Por el lado multimedia se dispone de una unidad MMX y otra SSE para las nuevas instrucciones. Y por el lado de punto flotante, contamos con un sumador, un multiplicador, una unidad de carga/almacenamiento y otra para el movimiento de datos.

## 4.13 ► Valoración final

### 4.13.1 Lo mejor

Estamos ante un diseño que trae novedades a todos los niveles:

**frecuencia** 📈 A más bajo nivel, un notable incremento de la frecuencia. Hemos basado nuestro análisis en 1.5 GHz, pero el procesador supera los 3 GHz en versiones más recientes.

- 👉 A nivel arquitectural, una clara apuesta por la segmentación frente a la superescalaridad, donde lo más acertado es ese énfasis en la eficiente predicción de salto, tanto en el algoritmo de resolución en sí como en la propia BTB y esa amplia ventana de reordenación de instrucciones. Se ha asumido un riesgo excesivo, pero al menos se ha hecho examen de conciencia, y asumidos los puntos débiles de la nueva arquitectura, se ha pasado a la acción tratando de amortiguar su negativo impacto. En nuestra sección comparativa frente al K7, veremos que el procesador sale bastante bien parado de ésta después de haberse adentrado en terreno espinoso.

segmentación
- 👉 Al nivel de unidades funcionales, dos ALU extremadamente rápidas y una caché de traza que trata de ocultar la rémora que supone la aceptación de un conjunto de instrucciones vetusto y caduco.

ALU
- 👉 Al nivel de la interacción externa, un majestuoso bus local de 400 MHz tan anhelado como necesario que permite por fin sacar cumplido partido del ancho de banda de las nuevas tecnologías de memoria principal.

bus local

#### 4.13.2 Lo peor

- 👉 El conjunto de instrucciones. Ya que se diseñaba una nueva arquitectura, se debió abolir todo vicio histórico del pasado con certera cirugía: Se ha extirpado la obstrucción por el bus local, pero se ha enquistado el conjunto de instrucciones x86, y no sabemos qué es peor. Parece mentira que se siga confiando la computación actual a esa horrenda traducción hardware de un repertorio de instrucciones que data de los años 70 y cuyo diseño es tan malo que lo último que merece es pasar a la historia. Ya sabemos que la compatibilidad es una cualidad clave para muchos usuarios, pero también nos dolió al principio desprendernos de nuestros viejos discos de vinilo, y ahora estamos todos encantados disfrutando de la calidad musical del CD. La baraja hay que romperla, y asumido esto, cuanto más se tarde en dar el paso, más traumático terminará resultando, y más conscientes seremos del grave error histórico cometido.

x86
- 👉 En el [capítulo 3](#) debió quedar claro que un fabricante puede optar por primar la frecuencia frente al número de etapas o viceversa. El tirón en la frecuencia del Pentium 4 es mayúsculo, pero se encuentra camuflado con una excesiva escisión en las etapas que toda instrucción debe atravesar hasta concluir su ejecución. Al margen de que esta decisión haya podido ser tomada para vestir al procesador con los hábitos que más gustan al cliente, nuestra principal crítica no viene por este lado, sino porque somos amantes de los diseños equilibrados, y el del Pentium 4 nos parece descompensado en este sentido.

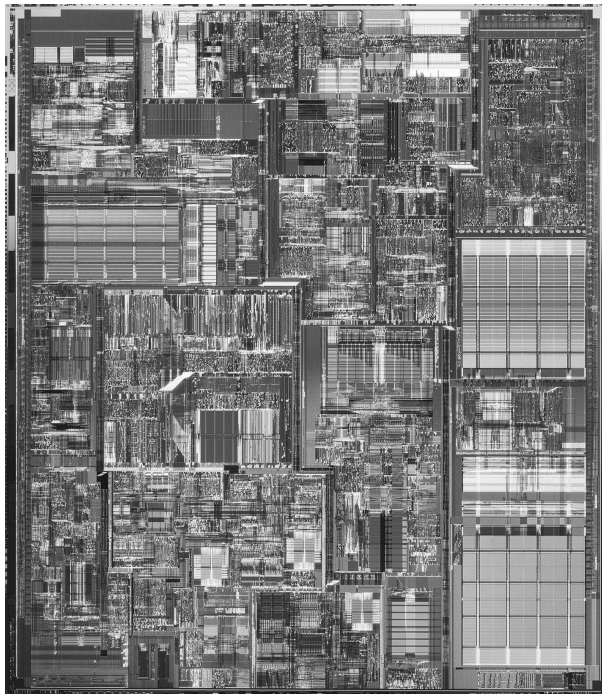
→ [pág. 47](#)

descompensado
- 👉 La memoria caché no se encuentra a la altura de la frecuencia de reloj del procesador. Bien es cierto que la dependencia del exterior es mucho menor ahora, pero creemos que eso no justifica recortar el tamaño del primer nivel de caché en sus dos vertientes (datos e instrucciones). También echamos de menos una caché L2 de 512 Kbytes.

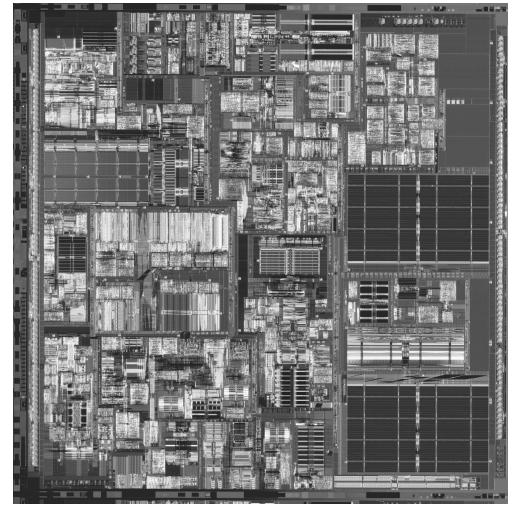
caché

Remarcaremos para concluir que tanto el conjunto de instrucciones como la segmentación son pilares de una arquitectura, y nos gusten o no, van a quedar ahí unos cuantos años. En cambio, la caché admite profundas metamorfosis de forma individual, y por eso sospechábamos que la cicatería que Intel mostraba al respecto de su tamaño en este estadio de desarrollo de la arquitectura, venía condicionada por el uso de las 0.18 micras y el aluminio a tan elevadas frecuencias.

El tiempo no tardó en darnos la razón. A la primera oportunidad que la tecnología le concedió, Intel gastó buena parte del patrimonio adicional disponible en fortalecer los niveles de caché integrada de la arquitectura. El resultado: El Pentium 4 Northwood.



(a) Pentium 4 Willamette



(b) Pentium 4 Northwood

Foto cortesía de Intel

FOTO 6.7: Áreas de integración a escala del Pentium 4 Willamette (247 mm<sup>2</sup> para 42 millones de transistores a 0.18 micras) y Northwood (146 mm<sup>2</sup> para 55 millones de transistores a 0.13 micras). Puede apreciarse el mayor espacio de silicio ocupado por la caché L2 del Northwood respecto al total de su área de integración (cuadrante inferior derecho en ambas imágenes), al contar éste con 512 Kbytes frente a los 256 Kbytes del Willamette.

## SECCIÓN 6.5

## Northwood: Las 0.13 micras en Intel

Tualatin

El primer microprocesador que Intel fabricó utilizando su tecnología de 0.13 micras e interconexiones de cobre no fue el Pentium 4, sino el Pentium III, con código de referencia Tualatin. Aconteció en Junio de 2001, y allí ya se avisó de cuáles eran las intenciones en un futuro inmediato: Incluir en el chip una caché L2 de 512 Kbytes. Dos meses más tarde entró en escena el

Northwood

Northwood, cumpliendo con las expectativas para alcanzar los 55 millones de transistores.

## 5.1 ► Mejoras introducidas

El Northwood incorpora a la arquitectura Pentium 4 las siguientes novedades:

bus 4x133 MHz

- 1 Un bus de 4x133 MHz, números que como enseguida veremos están muy ligados a la memoria principal DDRAM.

L2 512 Kbytes

- 2 La ampliación de caché L2 integrada desde los 256 Kbytes hasta los 512.

zócalo mPGA478

- 3 Fabricación exclusivamente en formato PGA de 478 pines, en un cambio de zócalo que aumenta el patillaje y reduce el espacio físico ocupado por el chip (ver foto 6.8). Con las nuevas mejoras, el chip disipa casi la mitad de potencia, lo que le permite apuntar hacia frecuencias más elevadas, flanqueando ya la barrera de los 3 GHz.

pág. 227

frecuencia

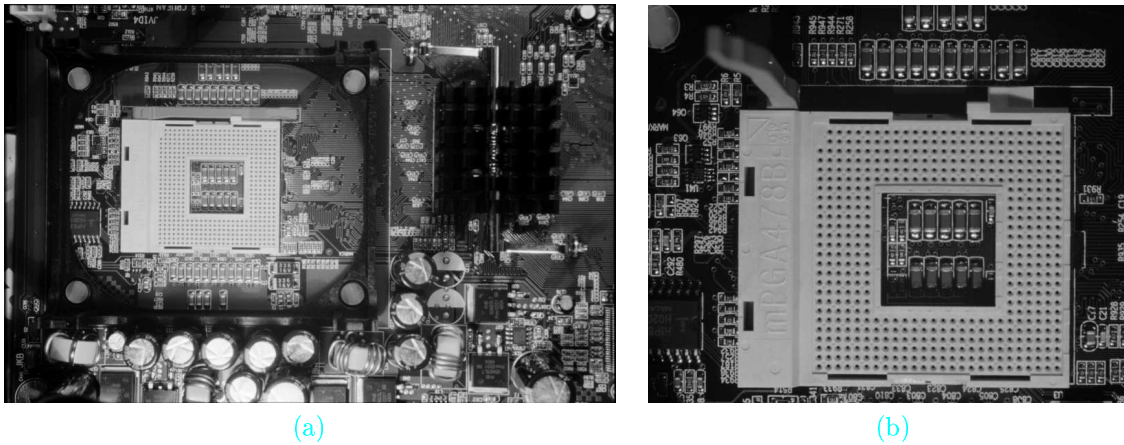


FOTO 6.8: El zócalo mPGA478 del nuevo Pentium 4 Northwood. (a) Aspecto general dentro de la placa base. (b) Visión más cercana, con la varilla del zócalo abierta para insertar el procesador.

Característica	Willamette-423	Willamette-478	Northwood-478
Frecuencia	1.3-2 GHz	1.5-2 GHz	1.6-3.2 GHz
Frecuencia de bus	4x100 MHz	4x100 MHz	4x100/133 MHz
Tecnología de integración	0.18 micras	0.18 micras	0.13 micras
Área de integración	247 mm <sup>2</sup>	247 mm <sup>2</sup>	146 mm <sup>2</sup>
Millones de transistores	42	42	55
Potencia disipada a 2 GHz	75 W	75 W	44.3 W
Zócalo y patillaje	FC-PGA/423 pines	FC-PGA y mPGA	mPGA/478 pines
Tamaño caché L2	256 Kbytes	256 Kbytes	512 Kbytes

TABLA 6.7: Diferencias entre las distintas versiones de Pentium 4.

La foto 6.7 contrasta el área de integración del Northwood, de 146 mm<sup>2</sup>, con la del Willamette, de 247 mm<sup>2</sup>, donde puede apreciarse en el cuadrante inferior derecho el espacio ocupado por la caché L2 en ambos casos.

Conviene aclarar que de las tres novedades expuestas, tan sólo la caché L2 caracteriza al Northwood, ya que hay procesadores Willamette fabricados bajo zócalo mPGA478, y Northwood fabricados con bus de 100 MHz. La tabla 6.7 muestra los rasgos de las tres versiones de Pentium 4 aparecidas hasta la fecha.

área de integración

versiones

## Los diálogos con memoria principal

## 5.2

Coincidiendo con la irrupción de este procesador en el mercado, aparece también un nuevo juego de chips: el i845, con código de referencia Brookdale, que proporciona cobertura para combinar el procesador con la memoria DDRAM, reservada al Pentium III hasta ese momento.

Brookdale

A partir de este momento, hay que decantarse por una de las dos variantes, para acompañar en cada caso al procesador con la memoria y la placa base correspondientes:

- 1 Memoria DDRAM (ver sección 10.13.5) con juego de chips i845 (Brookdale).
- 2 Memoria RDRAM (ver sección 10.13.6) con juego de chips i850 (Tehama).

• p. 68/Vol. 2

• p. 73/Vol. 2

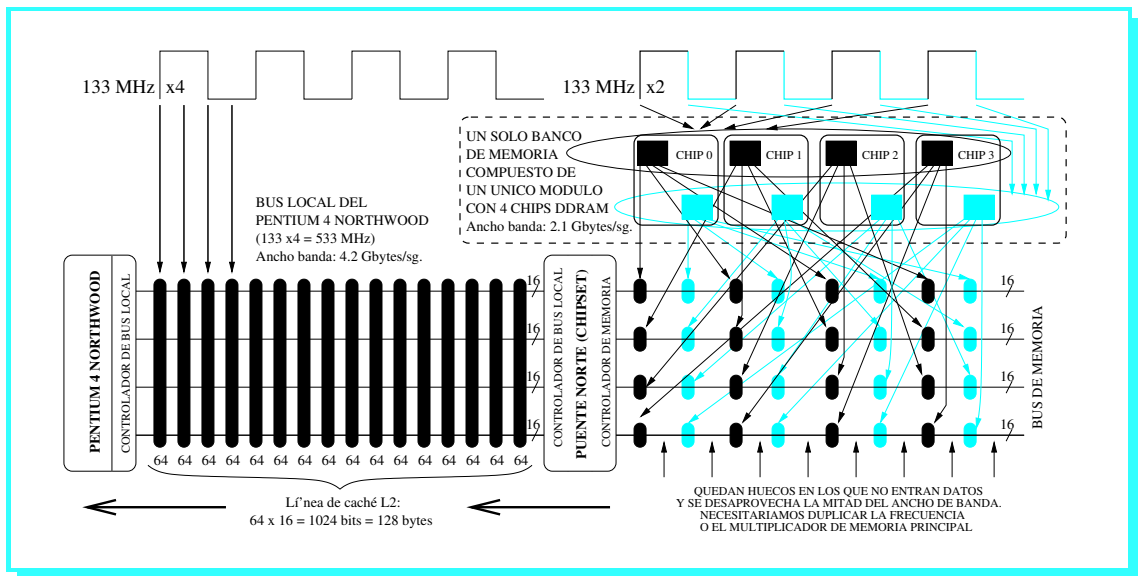


FIGURA 6.10: Forma en que el Pentium 4 dialoga con la memoria DDRAM de 2x133 MHz. Se desaprovecha la mitad del ancho de banda, con lo que el procesador dispone de potencial para absorber todo el caudal de datos hasta la versión 2x266 MHz de esta tecnología de memoria.

Vamos a analizar las dos variantes para sopesar la mejor elección.

### 5.2.1 El interfaz con DDRAM

El Pentium 4 Willamette se entendía perfectamente con una memoria DDRAM 2x133 MHz, aunque con el problema de que su bus local debía funcionar a 4x66 MHz, desperdiciando el ancho de banda hasta 4x100 MHz (ver [sección 6.4.5.1](#)). Al Northwood le va a ocurrir otro tanto similar, sólo que ahora se desaprovecha justamente la mitad de su ancho de banda, tal y como indicamos en la [figura 6.10](#).

pág. 213

Para subsanar esta pérdida de rendimiento existen dos caminos:

- ① Actuar en anchura, desdoblado la entrada procedente del bus de memoria en un doble puerto, siguiendo las directrices que ya narramos para el Willamette al respecto de la RDRAM (ver [sección 6.4.5.2](#)), y que completaremos enseguida. Esta es la variante que sale más barata.
- ② Actuar en velocidad, porque la memoria DDRAM no se queda en su versión base, sino que lleva un tiempo evolucionando hacia frecuencias más elevadas. En la [tabla 10.13](#) se lista la gama de productos del fabricante Micron para 2002, donde se pone de manifiesto que sus miras están en los 600 MHz (esto es, 2x300 MHz). Si nos conformamos con un poco menos, 2x266 MHz, tendremos justo la memoria DDRAM que se sincroniza de forma perfecta con nuestro procesador, alcanzándose un ancho de banda óptimo para el bus local del Northwood: 4.2 Gbytes/sg.

pág. 215

p. 72/Vol. 2

### 5.2.2 El interfaz con RDRAM

A la hora de optimizar el ancho de banda en la memoria RDRAM sobre el procesador Northwood, también existen dos caminos principales, pero a diferencia de la DDRAM, los dos suponen

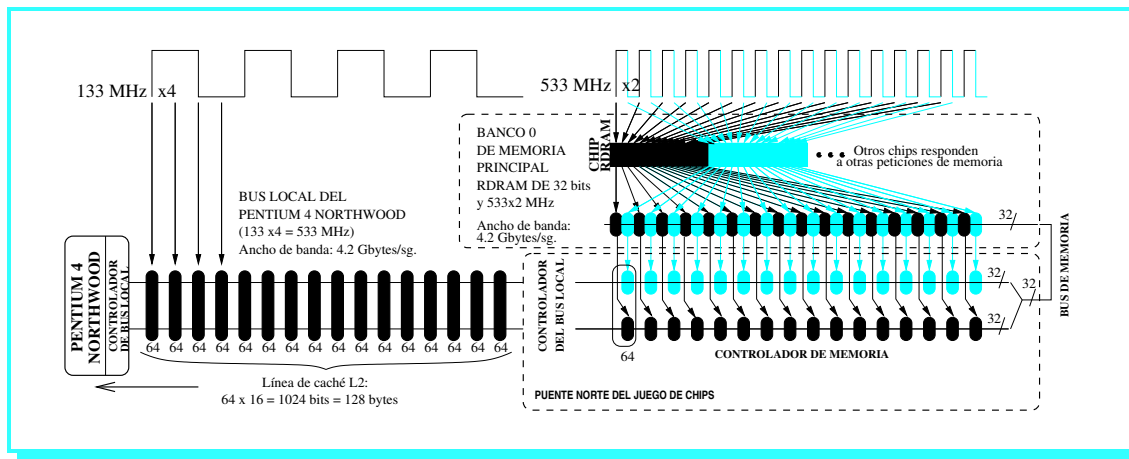


FIGURA 6.11: El perfecto entendimiento del Pentium 4 con la memoria RDRAM de 2x533 MHz y 32 bits de anchura.

actuar en anchura, dado que en velocidad RDRAM lleva bastante delantera al Pentium 4 (en la [tabla 10.15](#) pueden consultarse sus versiones, que alcanzan 2x667 MHz para anchuras de 16, 32 y 64 bits):

p. 83/Vol. 1.2

- ❶ La idea del doble puerto en memoria RDRAM puesta en marcha para dialogar con el Willamette sigue siendo muy válida, porque además representa una vez más la opción más barata. En ese caso, el diálogo tiene lugar exactamente como se muestra en la [figura 6.5](#), con la salvedad de que el bus que allí se dibuja tendría una frecuencia de 4x133 MHz (en lugar de 4x100), y que la memoria RDRAM sería la de 2x533 MHz (y no la 2x400 MHz allí mostrada).
- ❷ Entra en escena la posibilidad de montar módulos RDRAM de 32 bits, esto es, la ampliación de esta memoria desde los 16 hasta los 32 bits de anchura, versión que ya ha visto la luz. La forma de establecer estos diálogos elimina la necesidad del doble puerto, tal y como indicamos en la [figura 6.11](#).

pág. 214

## HyperThreading

5.3

Coincidiendo con el lanzamiento del Pentium 4 Northwood de 3.06 GHz, Intel incorporó en el chip un nuevo rasgo arquitectural: HyperThreading. El concepto proviene del grupo de trabajo del procesador Alpha 21464 de Compaq, que tras la fusión con Hewlett-Packard redefinió su estrategia vendiendo el grupo de arquitecturas avanzadas a Intel. Un par de años más tarde, comienzan a trascender sus logros en los diseños comerciales de la saga Pentium, pues la del Alpha quedó desmantelada después de aquellos acontecimientos.

Alpha 21464

La denominación comercial nos parece desacertada, dada su semejanza con otro bautismo coetáneo, el HyperTransport del K8 (ver [sección 7.1.7.2](#)), concepto con el que no tiene nada que ver. Por ello, para evitar confusiones, utilizaremos el nombre conceptual original de la idea: **SMT** (*Simultaneous MultiThreading* - los *thread* o *hebras* de un proceso aparecen como consecuencia de la escisión de un programa en varios flujos de instrucción concurrentes).

pág. 267  
SMT

SMT consiste en trasladar desde la capa hardware a la capa software del PC la ilusión de que se dispone de dos procesadores. Es decir, que tiene lugar en la capa lógica del PC algo que físicamente sólo ocurre en los sistemas SMP *Symmetric Multi-Processing* (ver [sección 2.2.3](#)), o de multiprocesamiento simétrico dotados de varios zócalos, esto es, las plataformas de tipo servidor con procesadores Xeon o K8.

el concepto  
SMP  
pág. 42

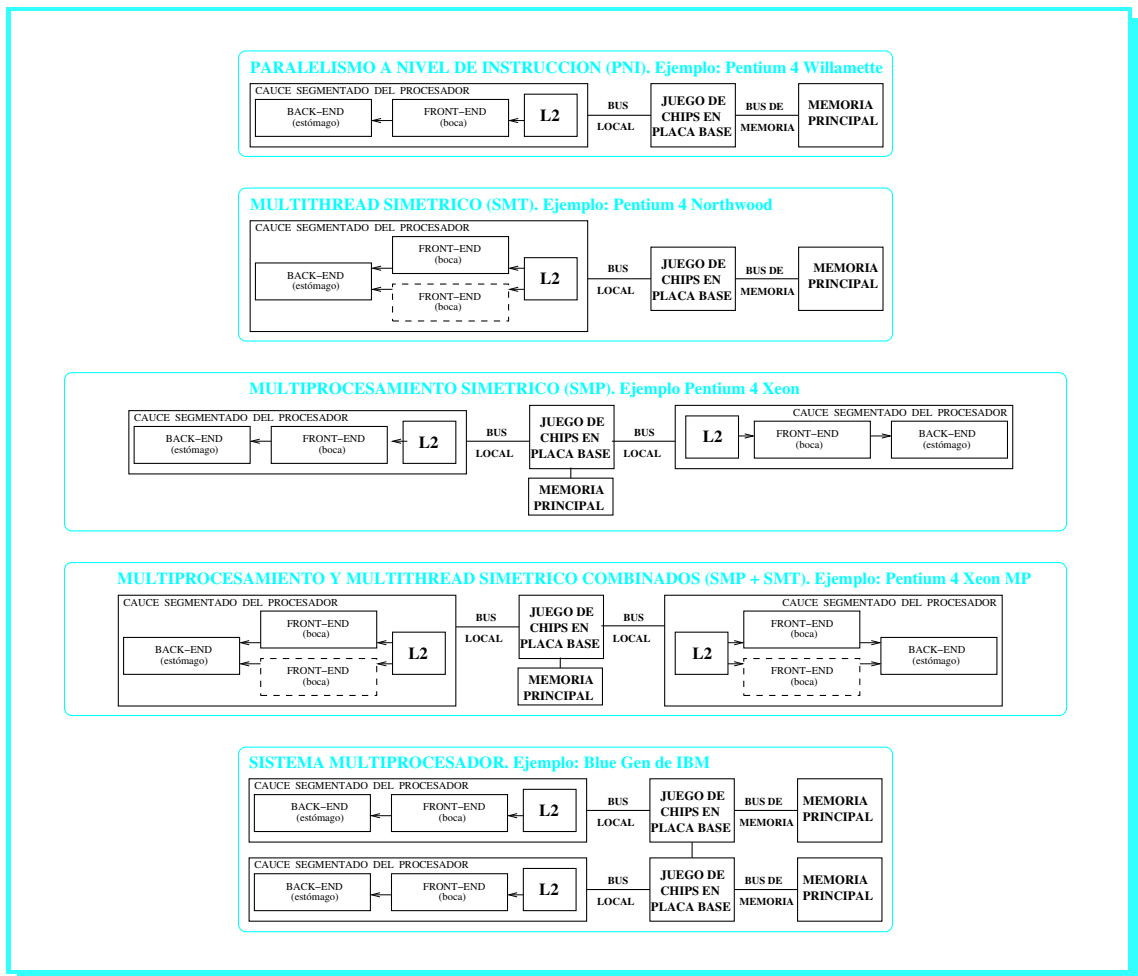


FIGURA 6.12: El paralelismo a distintos niveles de una arquitectura: Multiprocesador, SMP, SMT y PNI.

estrategias

☛ pág. 230

ilusión

La idea SMT se mueve por tanto en un estrato intermedio entre el paralelismo a nivel de instrucción (PNI), que implementan la segmentación y la superescalaridad a nivel monoprocesador, y el paralelismo de varios procesadores que implementan los sistemas multizócalo (SMP). La figura 6.12 ubica cada una de estas estrategias dando ejemplos comerciales.

Dentro ya de SMT, cada procesador lógico dispone de su propio banco de registros, incluyendo su contador de programa aparte y sus registros de control propios, y éstos sí que se encuentran físicamente por duplicado en la circuitería. Incluso parece que se dispone de varias bocas (*front-end*) por las que introducir instrucciones al conjunto, pero esto ya sí es una ilusión, porque lo que en realidad ocurre es que el único *front-end* físico traga instrucciones de forma alterna del conjunto de *threads* que haya en la capa lógica, proporcionando la misma falsa impresión de concurrencia que tiene lugar en los sistemas multiproceso de tiempo compartido, en el que la CPU entera reparte su tiempo entre los diferentes procesos y conmuta entre ellos. Eso sí, el procesador es mucho más ágil en conmutar de thread en un SMT, operación en la que se consume un máximo de 40 ciclos de reloj en el caso del Pentium 4.

analogía

En definitiva, es como si ensanchásemos la boca del procesador para tener comida preparada en una segunda despensa, para que cuando el canal principal de alimentación sufra algún atasco en la digestión (dependencias), pueda activarse este segundo recurso que permita minimizar el tiempo que pasan ociosas las diferentes unidades funcionales del estómago (*back-end*).



Por lo tanto, la idea de SMT es compatible con SMP y PNI, pudiéndose tener cuatro procesadores en una placa base (SMP), cada uno de ellos con dos *threads* concurrentes (SMT), y cada *thread* siendo ejecutado en un procesador segmentado y superescalar (PNI).

compatible

Dentro del Pentium 4 Xeon se juega de hecho con esta configuración, existiendo medidas de rendimiento de Intel para sistemas de dos procesadores físicos y otros dos lógicos que hablan de un incremento en el rendimiento del 30% para el programa Maya (una popular aplicación de renderizado) respecto al Xeon biprocesador convencional. En el Pentium 4 Northwood, Intel estima una mejora en torno al 24% en un SMT de dos procesadores lógicos frente al procesador convencional. Ahora bien, esto sólo es posible en sistemas operativos que sean capaces de explotar este recurso desde el planificador (scheduler) y despachador (dispatcher) de procesos, algo que comercialmente ahora sólo tiene a su alcance Windows XP.

rendimiento

Además, la aplicación tiene que haber sido programada para sacar provecho de este recurso, o de lo contrario, podría verse incluso penalizada al ejecutarse con el recurso SMT activado. Esto es así porque muchos programadores suelen tener por costumbre habilitar un segundo *thread* dedicado exclusivamente a labores de depuración de código cuando se encuentran programando una aplicación: Así aprovechan los recursos de la estación de trabajo en la que normalmente trabajan (que no sólo es SMT, sino también SMP), sabiendo que luego ese segundo *thread* no encuentra oportunidad de activarse porque la plataforma a la que va destinado el producto software, un PC, no es SMP. Con SMT en los PC actuales, este *thread* sí se activa, y puede ralentizar la ejecución del programa convencional.

penalización

De extenderse la implementación del SMT, veremos seguramente una opción en la BIOS del sistema para habilitar o desactivar este recurso. Cuando se disponga de él, nuestra conducta deberá ser:

nuestra conducta

- ❶ Aplicaciones recientes que indiquen que han sido programadas expresamente para aprovechar *HyperThreading* en Pentium 4: Activar SMT.
- ❷ Aplicaciones recientes que no indiquen explícitamente un aprovechamiento del recurso SMT: Se ejecutan igual al margen de que activemos SMT o no (los programadores ya han sido advertidos de sus prácticas perniciosas y reformarán su conducta).
- ❸ Aplicaciones antiguas escritas para PC: Desactivar SMT.

SECCIÓN 6.6

## Comparativa: K7 versus Pentium 4

Le proponemos que nos acompañe a través de un viaje por las entrañas de los dos colosos de séptima generación.

Asumiremos el papel de una instrucción entera idílica que se cuele por la entrada de los dos microprocesadores, un K7 Thunderbird y un Pentium 4 Willamette, y le seguiremos la pista a lo largo de su ejecución en el interior de cada modelo. Seremos así testigos de una frenética carrera por concluir la ejecución antes que el rival.

Vamos a ser generosos en la explicación de cada una de las etapas de procesamiento. Quisiéramos esperar hasta aquí para hacerlo porque creímos que su descripción separada era ciertamente tediosa, y sin embargo, más amena y sobre todo didáctica si se realizaba de forma conjunta. Recomendamos al lector que para un mejor aprovechamiento de esta sección mantenga un ojo en los diagramas de bloques de los respectivos procesadores (figuras 6.2 y 6.4).

descripción de etapas

➔ pág. 196

➔ pág. 210

Terminada la ejecución, repasaremos las principales magnitudes del procesador en base a las variantes de diseño escogidas por cada compañía en sus modelos comerciales, lo que nos permi-

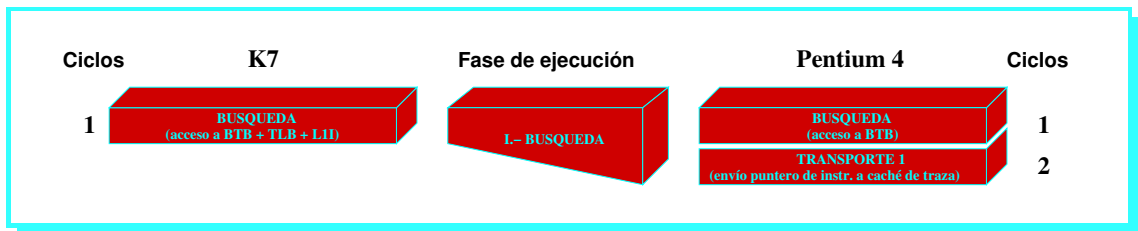


FIGURA 6.13: Los procesadores K7 y Pentium 4 al finalizar su primera fase: Búsqueda de instrucción.

visión de conjunto

tirá relativizar adecuadamente la incidencia de cada una de ellas, sopesar sus ventajas e inconvenientes, y adquirir una valiosa visión de conjunto. Para ello, cuando finalice cada instrucción, corregiremos el número de ciclos consumidos en cada procesador según la frecuencia que asignaremos a cada procesador, contabilizaremos su segmentación y superescalaridad, penalizaremos el impacto de las dependencias de datos y control en cada diseño, y en el tramo final, analizaremos el acceso a memoria para dictaminar el ganador.

Una vez concluya la carrera entre las dos arquitecturas de 0.18 micras, las confrontaremos también para las 0.13 micras en función de los cambios introducidos.

## 6.1 ▶ Ejecución de una instrucción

### 6.1.1 Fase de búsqueda de instrucción

**K7:** El K7 realiza la búsqueda de instrucción utilizando su caché L1I convencional, a la que accede en un solo ciclo a través de su TLB integrada donde se realiza la traducción virtual a física.

**Pentium 4:** El Pentium 4 se sirve del mecanismo de la caché de traza, pero la dirección de acceso a ésta procede de la BTB, cuya latencia consume el primer ciclo. La distancia en el área de integración hasta la caché de traza se lleva el segundo ciclo.

La situación es la que describe la [figura 6.13](#), y el parcial en ciclos queda reflejado en nuestro marcador inaugural (recordar que las victorias parciales van siendo para el procesador que anote **menos** ciclos).

1 - 2

K7 1 - 2 Pentium 4

### 6.1.2 Fase de decodificación de instrucción

**K7:** Para el K7, desglosamos por ciclos:

- **macrodec.** ▶ Se realiza una predecodificación en la que se distingue entre uno de los dos caminos de su cauce: DirectPath o VectorPath.
- **delimitación** ▶ La macroinstrucción de 128 bits obtenida en la búsqueda debe ahora analizarse, para delimitar cuántas instrucciones fuente contiene concretamente. El máximo es de tres instrucciones: Una aritmética, otra de carga de datos de memoria, y eventualmente una tercera de almacenamiento de datos en memoria.
- **empaquetado** ▶ Se agrupa el mayor número de instrucciones que sea posible (máximo tres).
- **microdec.** ▶ Cada instrucción por separado accede a la unidad de control cableada si es de las nuevas RISC o a la ROM de microprograma si es de las viejas CISC. En el primer caso, se obtienen las

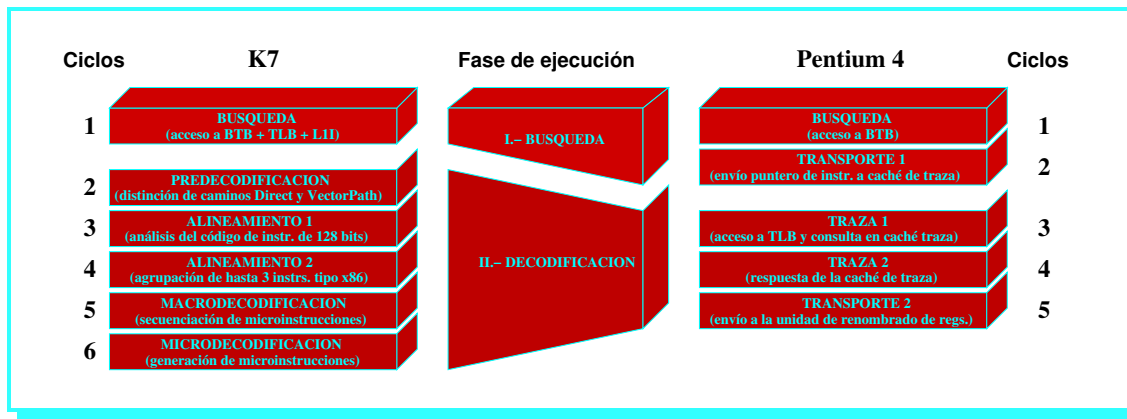


FIGURA 6.14: Los procesadores K7 y Pentium 4 al finalizar su segunda fase: Decodificación de instrucción.

microinstrucciones nativas por hardware; en el segundo, a través de una ROM de proyección y una ROM de microprograma.

► Las tres microinstrucciones se decodifican simultáneamente en los tres decodificadores de que dispone el K7 para no demorar más esta fase. El conjunto de palabras de control (hasta 6 de una anchura aproximada de 120 bits cada una) está ya generado para cada instrucción, y se encamina a la fase de planificación.

Total de ciclos para el K7 en esta fase: 5.

Para el Pentium 4, se accede a la caché de traza, que devuelve ya la instrucción decodificada en forma de un paquete de entre 1 y 4 microinstrucciones (instrucciones de la arquitectura nativa traducida de la x86 de entrada). En el acceso a la caché de traza, se consumen dos ciclos, y un tercero se gasta en conducir a las microinstrucciones hasta la circuitería de planificación de instrucciones.

En este instante, la situación es la que se describe en la figura 6.14. El parcial en ciclos, 5-3 para el Pentium 4, y el subtotal acumulado, 6 - 5, también a favor del Pentium 4.

El Pentium 4 ha adelantado al K7 gracias a la contribución de su caché de traza, que evita el proceso de decodificación convencional. En realidad, los tres ciclos contabilizados aquí podrían considerarse parte de la fase de búsqueda, pues el concepto de caché de traza difumina mucho la frontera entre la búsqueda y la decodificación de instrucción.

Otra observación que queremos hacer es que la elevada frecuencia del Pentium 4 va a corregir después el cómputo de ciclos a su favor frente a la cuenta realizada en el K7, pero también le ha perjudicado ya: Los ciclos 2 y 5 probablemente podrían haber sido absorbidos por alguno(s) de los otros tres si la frecuencia hubiese sido algo inferior, y el marcador estaría ahora decantado bastante más a favor del Pentium 4 (6-3).

Respecto al K7, se ha complicado un poco la vida para tragar tres instrucciones frente a una de su rival, y eso también tendremos que contabilizarlo a su favor cuando corriamos por el factor de superescalaridad. Hasta este momento, la cosa está más igualada de lo que el marcador indica.

K7 6 - 5 Pentium 4

5 - 3

- transporte

K7: 5

Pentium 4

caché traza

frecuencia

superescalar

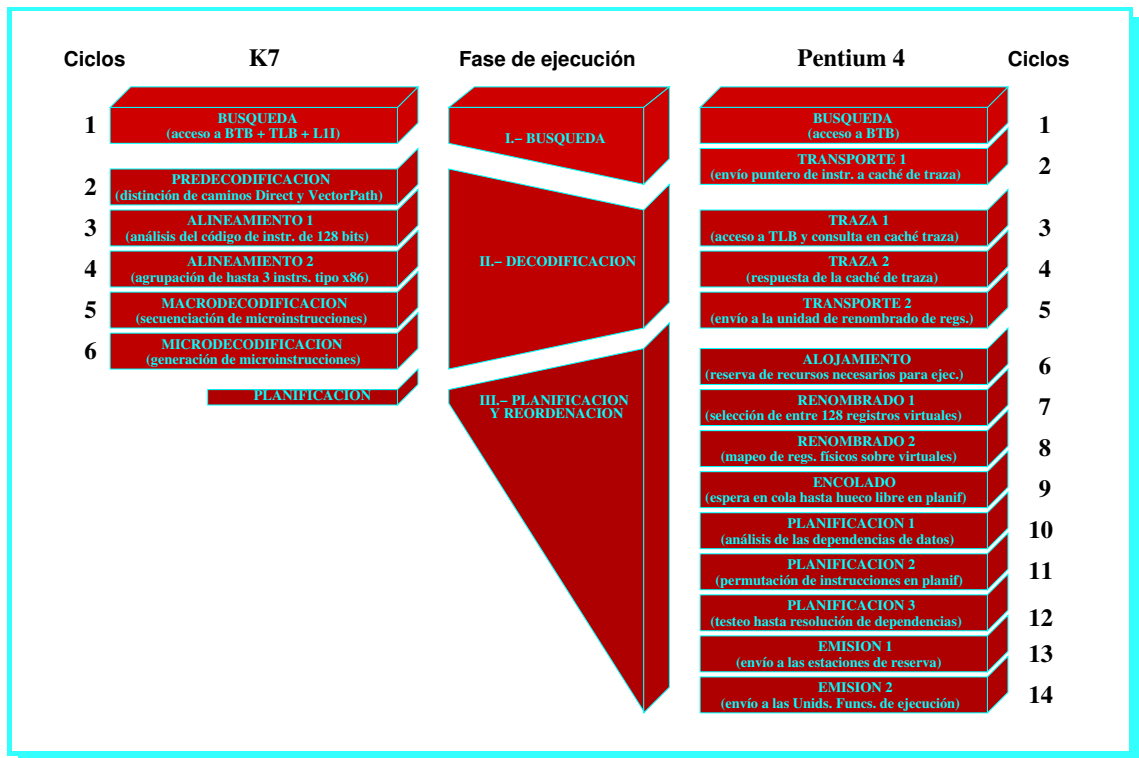


FIGURA 6.15: Los procesadores K7 y Pentium 4 al finalizar su tercera fase: Planificación y reordenación de instrucción.

### 6.1.3 Fase de planificación y reordenación de instrucciones

Para el K7: La planificación de instrucciones está tremendamente simplificada. Apenas existe renombrado de registros para esquivar los falsos riesgos estructurales que de su uso puedan surgir (sólo se adopta en el caso de instrucciones de punto flotante), y la ejecución fuera de orden adopta todo tipo de estrategias conservadoras para concluir todas estas tareas en el increíble tiempo de **medio ciclo de reloj**.

K7: raudo

Para el Pentium 4: Se conoce el impacto que las dependencias tienen sobre un cauce tan largo como el de este procesador, así que se extreman las medidas de precaución. El desglose de sus actividades por ciclos es el siguiente:

P4: previsor

- reserva ➤ Se reservan los recursos necesarios para completar la ejecución.
- renombrado ➤ ➤ Se renombran 128 registros virtuales que se mapean sobre los ocho físicos (2 ciclos).
- colas planif. ➤ Se pasan las microinstrucciones a las colas de planificación.
- dependencias ➤ Se analizan las dependencias de datos.
- reordenación ➤ ➤ Se decide la mejor reordenación de instrucciones posible para tratar de eliminar el mayor número de dependencias del programa original (2 ciclos).
- transporte ➤ ➤ Se envían las microinstrucciones a la correspondiente unidad de ejecución (2 ciclos).

Total de ciclos para el Pentium 4 en esta fase: 9.

En este instante, la situación es la que se describe en la [figura 6.15](#). El Pentium 4 acaba de encajar un parcial en contra de nueve a medio. El subtotal acumulado es de 6.5 - 14 a favor del K7.

La amplia ventana de 126 instrucciones considerada por el Pentium 4 le hacen entreteñerse más de la cuenta en la detección de dependencias y la reordenación de instrucciones. Algo similar le ocurre también en el renombrado de registros. Su miedo al fantasma de las dependencias le hacen ser muy previsor, pero recogerá sus frutos más adelante, cuando evaluemos las dependencias.

ventana de instrucciones

El K7 es justamente el caso opuesto: Opta por no demorarse aquí y se deja mucho trabajo por hacer. Asume el coste de que la ejecución que prosigue va a ser baldía en algunas ocasiones por la mayor incidencia de dependencias sin resolver. Puede permitirse el lujo de afrontar ese coste, pues su cauce de segmentación es la mitad de profundo que el de su rival.

De lo que sí saca rédito el K7 aquí es del mayor tiempo que dedicó a decodificar; de allí obtiene la información que le permite ya discriminar dos flujos bien diferenciados para las instrucciones que lleva dentro: Uno entero, y otro de punto flotante. De esta manera, las instrucciones enteras sacan partido a su extremada sencillez. En el Pentium 4 no se establece tal distinción hasta la siguiente fase de búsqueda de operandos, que ya es la etapa 15, por lo que la planificación y reordenación de instrucciones de punto flotante debe ser también completada en toda su extensión por las instrucciones enteras.

separación de cauces

Se cumple aquí el ecuador de nuestro itinerario. El Pentium 4 parece haberse detenido en el camino, pero a meditar acerca de lo que está por venir. El K7 va por delante, pero camina casi a tientas, precipitado, y le espera algún que otro traspiés que dará más emoción de lo que el marcador refleja hasta este instante.

ecuador

K7 6.5 - 14 Pentium 4

0.5 - 9

### 6.1.4 Fase de búsqueda de operandos

Para el K7: El medio ciclo sobrante de la fase anterior se aprovecha para obtener los valores de los operandos con los que realizar la operación en la fase de ejecución. Percibimos otra ventaja más de no ser tan agresivo en la frecuencia de trabajo: Hay oportunidades para aprovechar los pequeños resquicios que quedan en el tiempo sobrante de un mismo ciclo de reloj.

K7

Para el Pentium 4: Se accede al banco de registros, cuya latencia es de un ciclo. Pero la cosa no acaba ahí, porque la compleja etapa de reordenación de instrucciones y anticipación de datos llevada a cabo en la etapa de planificación anterior le obligan a chequear en esta fase todos los posibles cambios que hay que dar a los valores supuestamente incorrectos que se hayan extraído del banco de registros sin respetar la secuencia de lectura y escritura definida en el programa fuente. Ahí se consume un segundo ciclo, pero había que gastarlo para ser consecuente con las decisiones tomadas en la fase anterior.

Pentium 4

El parcial de esta fase queda en 0.5-2 a favor del K7. Los procesadores alcanzan el estado que se desprende de la [figura 6.16](#), y la distancia entre ambos sigue en aumento. Nuestro marcador queda como sigue:

☛ pág. 236

K7 7 - 16 Pentium 4

0.5 - 2

### 6.1.5 Fase de ejecución

Esta fase apenas tiene trascendencia. Ambos procesadores pueden realizar el cálculo necesario en un solo ciclo de reloj.

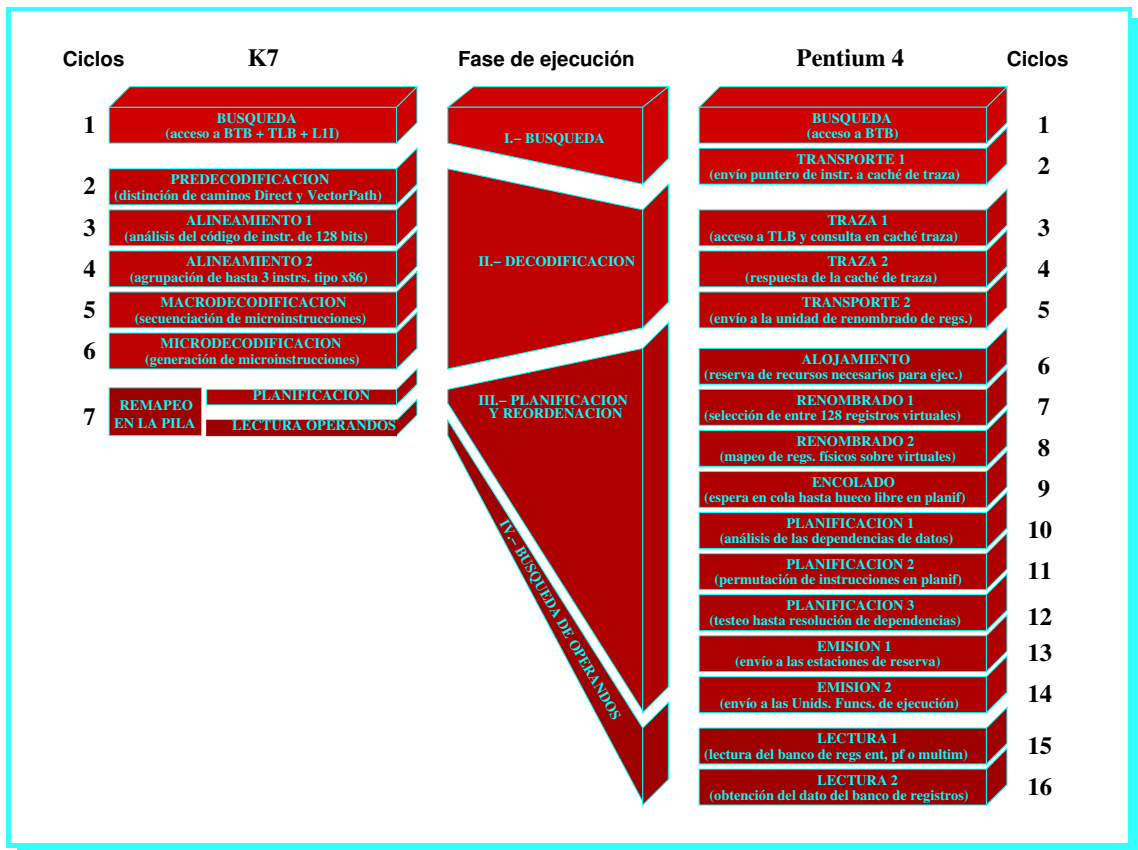


FIGURA 6.16: Los procesadores K7 y Pentium 4 al finalizar su cuarta fase: Búsqueda de operandos.

K7: UF

El K7 dispone de tres ALU y tres sumadores para el cálculo de la dirección efectiva, lo que corrobora el techo de su superescalaridad en seis, factor medido en términos de microinstrucciones nativas de la arquitectura.

Pentium 4: UF

El Pentium 4 sólo dispone de dos ALU, pero al responder éstas en medio ciclo de reloj, cuentan como si se desdoblaran y se dispusiese de cuatro ALU a todos los efectos. Les acompañan una unidad de cálculo de dirección efectiva para operaciones de lectura de memoria, y otra unidad similar para las operaciones homólogas de escritura en memoria. El número de unidades funcionales operativas en paralelo es también de seis, coincidiendo con su rival.

El empate es riguroso a todos los efectos, y nuestro marcador incrementa una unidad a cada casillero.

1 - 1

K7

8 - 17

Pentium 4

### 6.1.6 Fase de escritura

Llegamos a la última fase, que no es mera comparsa a pesar de su sencillez.

K7: BTB

En el K7, una vez concluida la operación en la ALU, se pueden leer los valores de sus *flags* de estado en el mismo ciclo. Estos *flags* sirven para verificar la condición de salto en el ciclo siguiente, y en caso de predicción errónea, notificarlo a la BTB para que traslade a la L1I la lectura del código de la instrucción por la que proseguir la ejecución. Para esta notificación se necesita un segundo

ciclo de reloj, pero era necesario de todas formas para escribir en el banco de registros el resultado de la operación realizada en la fase V, así que ambas operaciones tienen lugar concurrentemente, y el número total de ciclos para esta fase es dos.

Al Pentium 4 se le hace un poco más duro este trabajo. La lectura de los flags tras el ciclo de cómputo debe hacerse ya en el primer ciclo de la fase actual. Esto es consecuencia de seguir la filosofía de diseño de las ALU del Pentium III y anteriores, que tienen la circuitería de los *flags* al final de la etapa de segmentación, y de que el retraso de propagación de la lógica de los *flags* consume alrededor del 25 % de un ciclo de reloj (otra consecuencia de trabajar a un régimen de frecuencia muy elevado: ensanchando un poco el período de reloj, este 25 % hubiera entrado en el ciclo anterior). La verificación de la predicción de salto se lleva el segundo ciclo, en el cual se aprovecha para escribir en el banco de registros el resultado de la operación efectuada en la fase anterior. Finalmente, hay que notificar a la BTB la posibilidad de salto erróneo, y para ello, increíblemente, se habilita un tercer ciclo de reloj en lugar de optimizar mediante las muchas concurrencias que toda esta operativa permite. Como consecuencia de ello, el Pentium 4 pierde el parcial de esta fase por un total de 2-3.

Pentium 4: BTB

Hemos terminado la ejecución de nuestra instrucción entera, reflejándose el conjunto de pasos realizados en la [figura 6.17](#). Pero nuestra comparativa no acaba aquí, pues ahora hay que contabilizar el resto de magnitudes que intervienen en la ejecución. Antes de eso, actualicemos nuestro marcador, claramente decantado en favor del K7.

pág. 238

K7 10 - 20 Pentium 4

2 - 3

## Frecuencia

6.2

Vamos a elegir en nuestra comparativa una frecuencia de 1 GHz para el K7 y de 1.4 GHz para el Pentium 4, valores que representaron la parte central del rango de frecuencias inicial de cada procesador. Si tiene interés en comparar modelos de otra frecuencia, no tendrá más que aplicar el factor de corrección oportuno (por ejemplo, para un K7 a 1.2 GHz, multiplicará por 10 y dividirá por 12 el casillero del marcador para este procesador). La [tabla 6.12](#) muestra más adelante una serie de frecuencias equivalentes para cada procesador.

K7: 1 GHz  
P4: 1.4 GHz

pág. 244

Los diez ciclos obtenidos para el K7 a 1 GHz representan el mismo tiempo que catorce ciclos del Pentium 4 a 1.4 GHz, así que para contabilizar el impacto de la frecuencia, el marcador debe corregirse en esta proporción.

factor de  
corrección

K7 14 - 20 Pentium 4

x1.4 - x1

## Paralelismo a nivel de instrucción

6.3

En el K7, entran hasta tres instrucciones simultáneamente si se consiguen agrupar una aritmética, una de carga en memoria y otra de almacenamiento juntas.

Sin embargo, está bastante reconocido que resulta muy difícil llenar este cupo, pues por ejemplo, las primeras suelen tener un peso mucho mayor en el total de instrucciones de un programa. Incluso habrá ocasiones en las que la conjunción de una aritmética y otra de acceso a memoria no pueda realizarse. Vamos a conceder que por término medio el K7 conseguirá agrupar 1.7 instrucciones en cada ocasión. La posterior microdecodificación a partir de aquí desdobra en una media de 3 microinstrucciones cada instrucción, lo que nos da un total de  $1.7 \times 3 = 5.1$ . Pensamos que es un valor bastante realista, ya que cuando pasamos por la etapa de ejecución vimos allí un

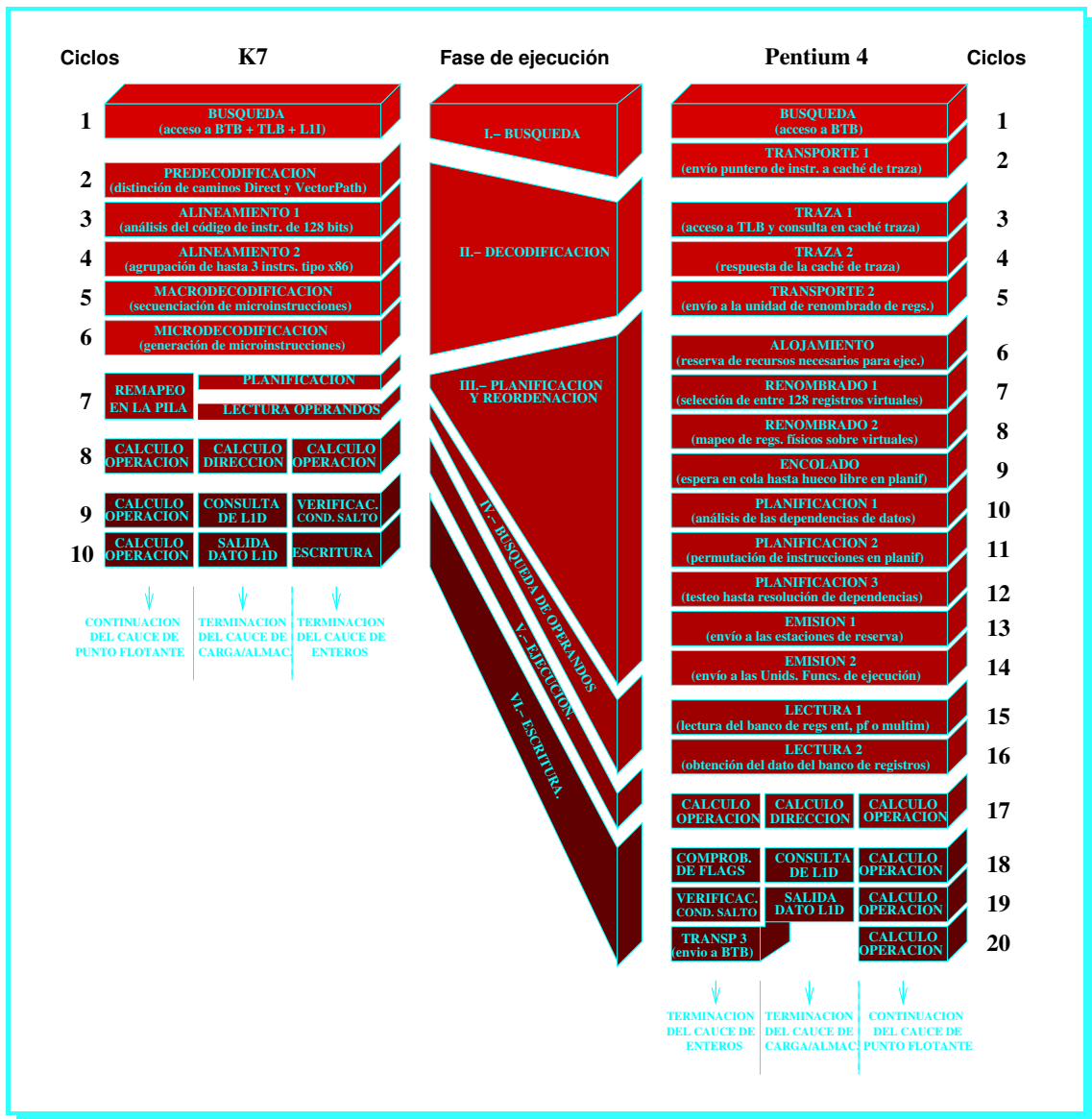


FIGURA 6.17: Los procesadores K7 y Pentium 4 al finalizar la sexta y última fase de escritura que completa la ejecución de una instrucción entera.

máximo de seis unidades funcionales y sólo tres ALU completas, con lo que el techo de seis no podrá conseguirse siempre.

K7: 50 Por otro lado, la segmentación del K7 es de 10 etapas. Ignorando provisionalmente las dependencias, la máxima ejecución concurrente es de 5 por 10, o sea, 50 microinstrucciones.

En el Pentium 4, sólo entra una instrucción en cada ciclo, que luego se desdobra en tres microinstrucciones de media (esta transformación es muy parecida a la microdecodificación que lleva a cabo el K7, y además la analizamos en detalle en la [sección 5.1.1](#)). Su factor superescalar queda así cifrado en 3 microinstrucciones nativas.

pág. 151

Pentium 4: 60 Respecto a la segmentación, presenta 20 etapas, así que la máxima ejecución concurrente es de 60 instrucciones. Eso sí, la concesión que hacemos al Pentium 4 es que todas las unidades funcionales en las que se consumían dos ciclos (caché de traza en los ciclos 3 y 4, renombrado en los ciclos 7 y 8, reordenación en los ciclos 11 y 12 y emisión a las unidades de ejecución en los



Tipo de instrucción origen de la dependencia (A)	Tipo de instrucción destino de la dependencia (B)	K7		Pentium 4	
		B comienza después de	Retraso total	B comienza después de	Retraso total
Aritmética	Aritmética	1 ciclo	2 ciclos	0.5 ciclos	<b>1.5 ciclos</b>
	Lógica	1 ciclo	2 ciclos	1 ciclo	<b>1.5 ciclos</b>
	Rotación/desplaz.	1 ciclo	2 ciclos	2 ciclos	4 ciclos
Lógica	Aritmética	1 ciclo	2 ciclos	0.5 ciclos	<b>1.5 ciclos</b>
	Lógica	1 ciclo	2 ciclos	0.5 ciclos	<b>1.0 ciclos</b>
	Rotación/desplaz.	1 ciclo	2 ciclos	0.5 ciclos	4 ciclos
Rotac/despl	Aritmética	1 ciclo	2 ciclos	2 ciclos	3 ciclos
	Lógica	1 ciclo	2 ciclos	2 ciclos	2.5 ciclos
	Rotación/desplaz.	1 ciclo	2 ciclos	2 ciclos	4 ciclos
Media aritmética		2 ciclos		2.55 ciclos	
Media ponderada aproximada		2 ciclos		2 ciclos	

**TABLA 6.8:** Impacto negativo en ciclos del procesador que tienen las dependencias de datos de un programa sobre la arquitectura del K7 y el Pentium 4. Los medios ciclos que aparecen en las casillas del Pentium 4 son consecuencia de la original implementación de sus ALU al doble de frecuencia que el resto del procesador. La mayor penalización en operaciones de rotación y desplazamiento es consecuencia de su deficiente realización en el Pentium 4, donde una serie de sumas puede ser más rápido que realizar pequeños desplazamientos. En negrita, los cuatro casos más probables, que al ser todos inferiores a 2 en el Pentium 4, provocan un descenso de su media ponderada hasta equipararse de forma aproximada al K7.

ciclos 13 y 14) están a su vez segmentadas en dos etapas para evitar el negativo impacto de las dependencias estructurales. Esto, que presentamos en la [sección 3.3](#) como el concepto de supersegmentación, resulta en la práctica bastante sencillo de llevar a cabo, y dado que si no se hace el rendimiento segmentado cae estrepitosamente, estamos seguros de que ha sido implementado en Intel (si bien no disponemos de información que lo corrobore).

Aunque de manera contrapuesta, ambos procesadores llegan a un potencial de paralelismo muy similar, con ligera ventaja para el Pentium 4 en una proporción de 60 a 50. Esto nos obliga a actualizar nuestro marcador multiplicando los ciclos del K7 por un factor de penalización de 1.2. El marcador se estrecha:

K7 16.8 - 20 Pentium 4

x1.2 - x1

pág. 58 ➔

## Incidencia de las dependencias

◀ 6.4

### 6.4.1 Dependencias de datos

Aquí vamos a simplificar el cálculo de los resultados. La [tabla 6.8](#) resume los ciclos de penalización asociados a cada tipo de dependencia de datos que puede darse en un programa para cada uno de los dos procesadores.

El K7 se comporta de forma fija para todo tipo de dependencias de datos, mientras que el Pentium 4 tiene más problemas con las instrucciones que realizan rotaciones y desplazamientos. Obsérvese que las situaciones que se recogen en la [tabla 6.8](#) no son equiprobables, ya que las aritméticas suelen ser mayoría frente a las lógicas, y éstas a su vez frente a las de rotación y desplazamiento (normalmente, van fusionadas, pero aquí es necesario separar estas últimas). Como el Pentium 4 presenta un comportamiento ligeramente mejor en las situaciones más frecuentes

Concepto	Programa (análisis por cada 100 instrucciones)			
	Dependencias de datos		Dependencias de control	
	K7	Pentium 4	K7	Pentium 4
Instrucciones	60 de tipo ALU		16 de salto	
Dependencias existentes	40 del tipo lectura tras escritura		16, tanto de salto condicional como incond.	
Dependencias resueltas con éxito	8 (20%) Renombre regs: NO. BRC: 1/2 etapa. Ventana reorden: 16 instrucciones.	20 (50%) Renombre regs: 2 et. BRC: 3 etapas. Ventana reorden: 128 instrucciones.	14.33% (90%) (BTB: 2048 posiciones) (Algoritmo de predicción mejor en el Pentium 4)	15.52 (93%) (BTB: 4096 posiciones)
Penalización por cada dep. no resuelta	2 ciclos (valores obtenidos en la <a href="#">tabla 6.8</a> )		10 ciclos (debe realizarse un vaciado completo del cauce de instrs.)	20 ciclos
Penalización:	32 x 2 =	20 x 2 =	(1.66x10)+2+2 = 20.66 ciclos	1.16x20 = 23.33 ciclos
► Total	64 ciclos	40 ciclos	20.66 %	23.33 %
► Porcentual	64 %	40 %		

**TABLA 6.9:** Comparativa del comportamiento de los microprocesadores K7 y Pentium 4 en la resolución de dependencias. ❖ En las dos primeras filas de la tabla caracterizamos nuestro programa típico, compuesto de un 60 % de instrucciones de tipo aritmético/lógico y un 16 % de instrucciones de salto (el 24 % restante serían operaciones de acceso a memoria). De las primeras, consideramos que existen 40 dependencias de datos del tipo lectura tras escritura entre pares de instrucciones adyacentes. ❖ En la tercera fila, cuantificamos el comportamiento de cada procesador en lo referente a la resolución de dependencias. ❖ Las tres últimas filas evalúan la penalización en que se incurre. Como siempre, consideramos una ejecución idílica (CPI = 1), por lo que la cantidad de ciclos de penalización en que se incurre representa realmente un porcentaje de ralentización para cada procesador.

probabilidad y penalización

y bastante peor en las menos probables, los pesos de penalización se contrarrestan con los de probabilidad y resulta una media ponderada prácticamente igual a la del K7.

ventana de instrucciones

Cuando el Pentium 4 atravesaba las fases de reordenación y planificación de instrucciones lo veíamos entretenido durante 9 ciclos. Ahora va a compensar ese esfuerzo realizado: Trabajando sobre su compleja ventana de 128 instrucciones, ha conseguido eliminar alrededor del 50 % de las dependencias de datos del código, penalizando sólo 40 ciclos tal y como hemos mostrado de forma resumida en la [tabla 6.9](#). El K7 pasó muy rápido por esa fase (tan sólo medio ciclo), y pagará las consecuencias de haber infravalorado ese problema: Sólo ha podido deshacer un 20 % de sus dependencias, y el 80 % restante le supone una penalización de 64 ciclos, esto es, un 60 % más que el Pentium 4.

Debemos corregir esta incidencia sobre nuestro marcador: Multiplicamos por 1.6 el casillero del K7, lo que le da la vuelta, colocándolo a favor del Pentium 4. El K7 ha pagado ya su excesiva precipitación. Finalizó la ejecución en la mitad de ciclos, pero no puede evitar que el Pentium 4 se le escape cuando saca fruto de su labor.

x1.6 - x1

K7 26.8 - 20 Pentium 4

### 6.4.2 Dependencias de control

La BTB o unidad de predicción dinámica de salto es el doble de grande en el Pentium 4 que en el K7 (4096 entradas frente a 2048), pero también lo será la penalidad que sufrirá cuando se falle en la predicción de salto (su cauce es exactamente el doble de profundo que el del K7). Parece que una cosa compensa la otra, pero no es exactamente así. Vamos a detallar un poco más:

En general, existe una regla no escrita que se asume en el análisis de los programas: Una de cada seis instrucciones es de salto. El porcentaje de acierto del algoritmo de predicción de salto del Pentium 4 según Intel está en torno al 93%, y puede corroborarse con su patente. Así pues, por cada 100 instrucciones ejecutadas, 16.66 son de salto, y de ellas, el predictor falla en el 7% de las ocasiones, o sea, que se producen  $16.66 \times 0.07 = 1.16$  fallos. Como el error de predicción se descubre en la última etapa de segmentación, cada fallo de predicción supone la pérdida de 20 ciclos, con lo que tenemos un total de  $1.16 \times 20 = 23.33$  ciclos por cada 100 instrucciones del programa.

P4: 93%

En el K7, el algoritmo de predicción de salto es ligeramente inferior, acreditando un 90% de acierto. El fallo, de ocurrir, se descubre también en la última etapa, pero esa vez supone perder tan sólo 10 ciclos, la longitud del cauce segmentado de este procesador. En total,  $16.66 \times 0.10 \times 10 = 16.66$  ciclos de penalidad por cada 100 instrucciones del programa.

K7: 90%

Se ha consultado a la BTB un total de 16.66 veces por cada 100 instrucciones. Dado que la BTB del Pentium 4 es el doble de grande, el K7 no encontrará la condición de salto en la BTB en ciertas ocasiones en las que el Pentium 4 sí lo hará. Suponemos que falla en dos ocasiones más, lo que añade dos ciclos más a su retraso para cambiar la predicción dinámica por otra estática. Tenemos al K7 con  $16.66 + 2 = 18.66$  ciclos, pero no hemos terminado aún.

Ahora debemos contabilizar la mayor probabilidad de error asociada a las dos predicciones que no se realizaron en la BTB, sino mediante predicción estática, que en estos modelos supone la asunción de salto realizado en el código, apuesta certera en el 80% de los casos según está tipificado. Sube un 10% la probabilidad de fallo en dos casos por cada 100 instrucciones ejecutadas, con lo que esta nueva penalización es  $2 \times 0.10 \times 10 = 2$  ciclos. El primer factor de 0.10 se debe al incremento de la probabilidad de fallo en un 10%, y el segundo, a las 10 etapas de segmentación del K7, que es la penalización por cada fallo. El total de ciclos perdidos en el K7 es  $18.66 + 2 = 20.66$ . Como en el Pentium 4 fue de 23.33, el factor de penalización normalizado en contra del Pentium 4 es de 1.13. Actualizamos los marcadores y los guarismos quedan muy ajustados. Nos queda tan sólo una faceta por contabilizar para concluir nuestro análisis.

predicción de salto

K7 26.8 - 22.6 Pentium 4

$x1-x1.13$

## Acceso a memoria

## ◀ 6.5

Esta es la vertiente que más trabajo cuesta analizar de forma analítica. Ya nos hemos perdido suficientemente en cálculos intrincados como para continuar por esta línea, así que proseguiremos con un análisis algo más fácil de digerir. En el acceso a memoria, el K7 cuenta con tres ventajas:

K7: 3 ventajas

- ❶ La caché L1 del K7 es bastante más grande: 64 Kbytes para memoria de datos frente a 8 Kbytes en el Pentium 4, y 64 Kbytes para memoria de instrucciones frente a 32 Kbytes equivalentes de la caché de traza en el Pentium 4. La organización de estas cachés es mejor en el Pentium 4 (4 líneas por conjunto frente a sólo 2 en el K7), pero no es suficiente para contrarrestar tanta diferencia de tamaño. Esta diferencia es consecuencia del menor área de integración requerido por el núcleo arquitectural del K7 frente al del Pentium 4, lo que deja espacio para colocar aditivos como éste.

L1 más grande

precarga mejor

- ② En las operaciones de precarga de datos de memoria como las que habilitan las cachés de primer nivel, la latencia de la memoria se oculta mejor en el K7, pues puede comenzar la precarga de datos ocho ciclos después del acceso a la caché de instrucciones (el acceso a la caché de instrucciones está en la etapa 1 y el puerto de acceso a caché de datos en la etapa 9). En el Pentium 4, se deben esperar quince ciclos (la distancia que existe en la etapa de acceso a la caché de traza y la etapa 18 de acceso a la caché de datos).

doble puerto

- ③ En las operaciones de acceso a memoria, sea cual sea el nivel de la jerarquía en el que se termine encontrando el dato, el K7 cuenta con la ventaja de que dispone de doble puerto para lectura/escritura, mientras que en el Pentium 4 se dispone de uno sólo.

A favor del Pentium 4 vemos ventajas en los buses, tanto en la comunicación interna con caché L1, que es de 256 bits frente a 64 en el K7, como en la conexión por el bus local. Esto es muy importante para el sistema como ente global, pero no es éso lo que se evalúa aquí. Por lo tanto, barriendo la jerarquía de memoria de dentro a afuera, tenemos:

p. 13/Vol. 2

p. 84/Vol. 2

- ① **Memoria principal.** Tanto ella como el tráfico con los periféricos es un aspecto externo, que analizaremos a lo largo del [capítulo 10](#) para completar la pieza que le falta al puzzle (recomendamos especialmente la lectura de la [sección 10.13.7](#), donde se contrastan DDRAM y RDRAM).
- ② **Caché L2.** Aquí tenemos un claro empate entre los dos procesadores, puesto que ambos disponen de 256 Kbytes, integradas, y responden a la velocidad del procesador contando con similar infraestructura.
- ③ **Caché L1.** El bus más ancho de L1 del Pentium 4 no es suficiente para compensar su escuálida capacidad de 8 Kbytes. En contraste, la terna de mejoras que el K7 presenta aquí es bastante significativa, permitiéndole enjugar la desventaja que traía del apartado anterior. Ya tenemos veredicto final.

veredicto

K7

Tablas

Pentium 4

## 6.6 Valoración final

pág. 244

El K7 a 1 GHz presenta a nuestros ojos un rendimiento para aplicaciones enteras muy similar al Pentium 4 de 1.4 GHz. Extrapolando frecuencias a un rango más elevado, la equivalencia estaría entre un K7 de 1.4 GHz y un Pentium 4 de 2 GHz. La [tabla 6.12](#) resume estas equivalencias.

pág. 243

No se sienta decepcionado si no hemos podido proclamar un ganador. No se trataba de eso. Nuestro análisis tiene ciertas deficiencias, y realiza bastantes suposiciones como para que pueda ser tomado con rigurosidad, pero lo más importante es que refleja las virtudes y carencias de cada modelo. Nuestras aportaciones principales aparecen resumidas en la [tabla 6.10](#).

dos escuelas  
divergentes

Si existiera una fórmula mágica para desarrollar el procesador perfecto, la computación sería una aburrida dictadura. Lo bonito es presenciar formas de hacer lo mismo a través de caminos que divergen desde sus inicios; ya en la inmadura fase de decodificación donde el procesador ni siquiera conoce a lo que se enfrenta, el tratamiento de las instrucciones es casi contrapuesto en los dos modelos analizados.

Creemos que el verdadero ganador de un análisis como el nuestro no es ésta o aquella compañía, sino sus clientes. Ellas sólo quieren que su procesador quede por encima del rival; ellos, saber cómo se consigue eso. Y para ello, sólo hay que echar un vistazo a qué características del programa fuente elegido hicieron más daño a cada procesador, algo que se aprecia con claridad meridiana en nuestro análisis, y alterar los porcentajes a conveniencia.

Rasgo de la configuración del microprocesador	Marcador		Comparativa/ Conclusiones
	K7	P4	
Fases de una instrucción entera			
❖ Búsqueda	1	2	Penaliza 1 ciclo la mayor frecuencia del Pentium 4.
❖ Decodificación	6	5	Caché de traza P4 elimina camino crítico decod. x86. K7 apuesta superescalar.
❖ Planificación/Reorden	6.5	14	Prevención riesgos dependencias superior en el P4.
❖ Búsqueda de operandos	7	16	Penaliza 1.5 ciclos la mayor frecuencia del P4.
❖ Ejecución	8	17	Empate técnico.
❖ Escritura	10	20	Penaliza 1 ciclo la mayor frecuencia del P4.
Frecuencia			
1.4 GHz en el Pentium 4 frente a 1.0 GHz en K7	14	20	40% ▲ P4 P4 recupera sus ciclos.
Paralelismo a nivel de instrucción			Factor
❖ Superescalaridad			66% ▲ K7 5 vs 3
❖ Segmentación			100% ▲ P4 20 vs 10
❖ Conjunción de ambas	16.8	20	20% ▲ P4 60 vs 50
Análisis de dependencias			Ciclos
❖ Dependencias datos	26.8	20	60% ▲ P4 40 vs 64
❖ Dependencias control	26.8	22.6	13% ▲ K7 20.6 vs 23.3
Acceso a memoria			
Cachés integradas	Tablas		L1 más grande en el K7

TABLA 6.10: Evolución de nuestro marcador y resumen de nuestra comparativa entre los microprocesadores K7 y Pentium 4.

Sensibilidad ante los desafíos planteados por cada tipo de instrucción		
Característica del programa	K7	Pentium 4
Instrucciones aritméticas con dependencias	Muy alta	Media
Instrucciones de salto	Baja	Media
Instrucciones de acceso a memoria	Baja	Media

TABLA 6.11: Sensibilidad del rendimiento de cada microprocesador a las características de un benchmark o programa de evaluación utilizado en su comparación. El K7 renuncia a un exhaustivo análisis de las dependencias de datos, dado que no cursa una apuesta tan agresiva por la segmentación. El resultado es que termina viéndose afectado más que el Pentium 4, sobre el papel bastante más sensible a este tipo de conflictos dada la profundidad de su cauce segmentado.

Hemos escogido los parámetros de la forma más objetiva posible, aunque ésa no fue nuestra principal preocupación. Si ahora llega Intel o AMD y le pide preparar unas pruebas de rendimiento basado en un conjunto de aplicaciones software (benchmark) en el que éste o aquel procesador se muestre un 30% mejor que su rival, estamos seguros de que sabrá cómo hacerlo. Esa es realmente nuestra contribución: En la [tabla 6.11](#) presentamos la sensibilidad de cada procesador a las características de un benchmark con el que se quiera analizar. A mayor sensibilidad, mayor ralentización sufrirá el procesador ante la elevada presencia de operaciones del tipo mostrado.

selección de  
parámetros

pág. 243

Frecuencias equivalentes para los modelos de séptima generación y 0.18 micras	
AMD	Intel
Athlon Thunderbird 1 GHz	Pentium 4 Willamette 1.4 GHz
Athlon Thunderbird 1.2 GHz	Pentium 4 Willamette 1.7 GHz
Athlon Thunderbird 1.4 GHz	Pentium 4 Willamette 2.0 GHz
Bajo fabricación de 0.13 micras	
Athlon XP 1800+ (1.533 GHz)	Pentium 4 Northwood 2.1 GHz
Athlon XP 2000+ (1.666 GHz)	Pentium 4 Northwood 2.3 GHz
Athlon XP 2400+ (2 GHz)	Pentium 4 Northwood 2.8 GHz
Athlon XP 2700+ (2.16 GHz)	Pentium 4 Northwood 3.06 GHz

TABLA 6.12: Frecuencias equivalentes para los modelos de Intel y AMD de séptima generación.

cauce de punto  
flotante

La confrontación en aplicaciones de punto flotante resulta bastante más complicada que su homóloga entera, y pensamos que es mucho menos didáctica porque exige un mayor detalle que dificulta la visión de conjunto que buscábamos. Eso sí, avisamos de que es una historia completamente diferente a la narrada aquí, y que muy pocas cosas son extrapolables.

cauce  
multimedia

Respecto a las aplicaciones multimedia, el Pentium 4 cuenta con la clara ventaja de su extensión SSE2, bastante más potente que la Enhanced 3DNow!. No obstante, en las versiones de 0.13 micras, AMD ha mejorado este aspecto e Intel no, por lo que ya se ha diluido buena parte la diferencia entre ambas.

## 6.7 ▶ Comparativa a 0.13 micras

Los parámetros del K7 y el Pentium 4 que son el armazón de su arquitectura no han cambiado desde su llegada, ni van a cambiar en los próximos años. Nos estamos refiriendo a la segmentación, la superescalaridad, el tratamiento de dependencias, y las unidades funcionales de ejecución, ...

La llegada de las 0.13 micras a estos procesadores trastoca los parámetros más ajenos a la arquitectura interna, como son la frecuencia, las conexiones por el bus y la jerarquía de cachés. Repasemos cada uno de ellos por separado:

- ❶ **Frecuencia.** Ha aumentado en proporciones similares en estos dos últimos años. Si en 2001 las frecuencias más usuales eran 1 GHz para el K7 y 1.4 GHz para el Pentium 4, a comienzos de 2003 esas frecuencias se han trasladado al Athlon XP 2700+ (2.25 GHz) y al Pentium 4 Northwood de 3.06 GHz. Como puede comprobarse, el porcentaje del 40 % a favor de Intel que hemos utilizado a lo largo de nuestro análisis ha envejecido con una salud de hierro.
- ❷ **El bus local.** AMD ha mejorado el bus desde 2x 133 MHz hasta 2x 166 MHz, mientras que Intel lo ha hecho desde 4x 100 MHz hasta los 4x133 MHz. El porcentaje de mejora en ambos casos ronda el 30 %.
- ❸ **Las cachés.** Ninguno introduce mejoras en el primer nivel, y ambos se ponen de acuerdo para ampliar la L2 hasta los 512 Kbytes (modelos Barton y Northwood).

persiste  
el empate

En consecuencia, el empate técnico sigue persistiendo tres años después de haberse introducido en el mercado ambas arquitecturas.

A partir de aquí, AMD emprende su singladura con el K8, e Intel ha jugado la baza del *HyperThreading* para entretener al mercado entretanto pone en el mercado su próxima arquitectura de octava generación.

## Otros fabricantes

El mercado de fabricantes de microprocesadores para PC experimentó un giro radical a mediados de 1999. Casi al unísono, y durante los meses de Mayo y Junio, las tres alternativas que existían a Intel y AMD cerraron sus puertas: Cyrix, de la que hemos hablado en el contexto de la sexta generación, Centaur y Rise. Los productos de estas dos últimas siempre tuvieron una jurisdicción muy limitada, acotándose en el mercado norteamericano sin más proyección internacional. Eran casi diseños de culto.

tres cierres:  
Cyrix  
Centaur  
Rise

Cuesta tanto trabajo abrirse camino en un mercado tan exigente y competitivo como el de los microprocesadores, que si no se dispone del auspicio de uno de los pesos pesados del sector, las cabalgadas en solitario terminan más temprano que tarde, así que la supervivencia de este tipo de iniciativas se ampara bajo un nombre clave: Diversificación, o el arte de diseminar la actividad empresarial en distintos frentes. Cyrix, Centaur y Rise adolecieron de esta clase de políticas, y así les fue.

diversi-  
ficación

El único que quizá merece una sucinta explicación sobre su final es Cyrix, quizá por haberlo mencionado en todas las generaciones hasta ahora, o quizá porque llegó VIA para salvarlo y mantenerlo aún vigente.

### Cyrix

#### ◀ 7.1

Tras el lanzamiento del MediaGX de Cyrix, National Semiconductor, el que entonces era su propietario, comenzó a acusar graves problemas financieros. Pero aún antes de cerrar sus puertas, Cyrix dió algunos coletazos más. Primero, ensamblaron el computador WebPad utilizando el MediaGX como microprocesador, y seguidamente, rediseñaron por completo la arquitectura del procesador incluyendo un nuevo cauce para punto flotante. El código de referencia fue Cayenne, siguiendo la cadena de alimentación mejicana abierta con el Jalapeño de sexta generación.

Cayenne

Fue entonces cuando VIA decidió adquirió Cyrix. Muchos pensaron entonces que VIA pretendía fortalecer su posición como fabricante de juegos de chips para placa base, utilizando toda la logística que Cyrix poseía para desarrollar e integrar sus chips con el solo propósito de arrebatar a Intel cuota de mercado en *chipsets*.

Pero las intenciones de VIA eran bien distintas: Diversificar actividades para diluir riesgos financieros. Tomando el núcleo arquitectural del Cayenne como semilla, VIA finalizó el diseño de un nuevo procesador, de código de referencia Joshua (a éstos les ha dado por nombres místicos y bíblicos, qué le vamos a hacer) y denominación comercial Cyrix III.

Joshua

El nuevo procesador vió la luz el 22 de Febrero de 2000, iniciando su andadura a 466 MHz y tres velocidades de bus: 66, 100 y 133 MHz. Disponía de caché de segundo nivel de hasta 256 Kbytes y estaba descaradamente dirigido al segmento de gama baja, donde competiría con el Celeron de Intel y el Duron de AMD. Empleaba tecnología de integración de 0.18 micras, el mismo empaquetado Socket370 de los Pentium III y Celeron de Intel con los que aprovechaba toda su infraestructura de placas base, y el juego de instrucciones multimedia 3DNow! de AMD. Su coste: Tan sólo \$80, una vez más lo más atractivo de este tipo de productos.

coste

Desde entonces, VIA se ha dedicado a mejorar la competitividad de este microprocesador. La familia de código de referencia Samuel traspasó la frontera del Gigahercio manteniendo unos precios que atraen más que otros, pero lo cierto es que desde un prisma comercial, en el contexto europeo siguen siendo unos completos desconocidos.

Samuel



## Resumen



La séptima generación traslada las innovaciones a la entrada del procesador, donde el controlador de bus local desaparece por fin como claro cuello de botella del sistema. Esto revierte a su vez notables cambios sobre los chips de la memoria principal y la placa base.

**fechas** La variable temporal, que había resultado decisiva para Intel en confrontaciones anteriores, juega por primera vez en su contra en la séptima generación: El K7 se anuncia en Junio de 1999, año y medio antes de que vea la luz el Pentium 4.

**K7** El K7 es un buen diseño, y durante todo el tiempo en que se mide con el Pentium III le saca los colores. Parece el momento de AMD, pero enseguida demostrará que el liderazgo le queda grande: Su primer paso adelante, el K7 Thunderbird, calca el proceder de Intel dos años antes con su Pentium III Coppermine (vuelta al Socket e integración L2 de 256 Kbytes). El segundo paso, el Duron, calca al Celeron de Intel de cuatro años atrás. Y el tercer paso, el Athlon XP, calca las mejoras del Pentium 4 Northwood y, de paso, copia de Intel lo que menos nos gusta de él: Las especificaciones rimbombantes en torno a la frecuencia y respaldadas por una fuerte campaña de marketing.

**Pentium 4** El Pentium 4 es una arquitectura muy dependiente de una frecuencia elevada, pero nadie mejor que Intel sabe que de todas las magnitudes del procesador, es la gran seductora de los clientes neófitos. Sin grandes alardes, Intel ha vuelto a marcar distancias en ventas respecto a sus competidores. Primero se ha sacado de la chistera el concepto de caché de traza, algo costosa para el rendimiento que produce; luego apuesta todo a la memoria RDRAM, para salir escaldado y terminar abrazando la DDRAM; finalmente, ha introducido el *HyperThreading* como último bastión para mantener vigente su arquitectura de séptima generación entretanto nos presenta al competidor del K8.

No sabemos si la próxima generación acabará como ésta, pero sí podemos adelantar que ha empezado igual. AMD ya tiene en la calle su diseño, y a Intel, en las previsiones más optimistas que conocemos, aún le faltan 18 meses para replicar con el suyo. ¿Será capaz una vez más de enjugar esta desventaja inicial? Se lo contaremos en nuestra edición de 2005. Como aperitivo, le invitamos al [capítulo 7](#) para que conozca las credenciales del K8.

pág. 255



## La anécdota: El curso cíclico de la historia



En 1582, el Papa Gregorio XIII, asesorado por el jesuita Christopher Clavius, decidió ajustar el calendario al año solar real para evitar el deslizamiento del equinoccio de primavera, decretando para ello que en octubre de ese año, el día siguiente al 4 fuera el 15 y que, en adelante, sólo 97 de cada 400 años fueran bisiestos en lugar de 100. Muchas personas protestaron por ello, al considerar que se les estaban usurpando diez días de vida. Y es que hay gente que no entiende que el método utilizado para contar no altera lo que se mide.

Cinco siglos más tarde, los dirigentes de AMD decidieron escamotear 33 MHz de cada 100 MHz a la frecuencia encubierta de sus procesadores Athlon-XP. No se recuerdan protestas populares por aquel agravio, pues al parecer la comunidad tecnológica consideró lícito el empleo de unas especificaciones adulteradas que nosotros reprobamos. Y es que hay dirigentes que no quieren entender que si se establece un método para contar, debe ser el mismo para todos.



Bastante peor que no entender las cosas es no querer entenderlas. ¿Cuándo se juega más con la ignorancia del pueblo, en el siglo XVI o en pleno siglo XXI?

## Cuestionario de evaluación

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

**1** 🗣️ Nos encontramos en la fase inaugural de una nueva generación de microprocesadores, con los primeros modelos comerciales en la calle. Atendiendo a como se ha comportado el mercado hasta ahora, podemos vaticinar que...

- a** No es un buen momento para encarar la compra de un PC.
- b** Pronto cambiaremos a un nuevo proceso de fabricación basado en distancias de integración más pequeñas.
- c** La frecuencia más baja del nuevo producto es superior a la frecuencia más alta del último de la generación anterior para una misma marca.
- d** En vista de que el ciclo completo de desarrollo de un procesador se estima en unos seis años, ha transcurrido al menos un lustro desde que inauguramos la generación anterior, y pasará al menos otro hasta que inauguremos la siguiente.

**2** 🗣️ ¿Qué es una caché de traza?

- a** La caché de instrucciones de primer nivel del Pentium 4.
- b** Una caché que almacena las instrucciones ya decodificadas.
- c** Las dos respuestas anteriores son válidas.
- d** La pregunta está mal formulada, pues sólo existe el concepto de *traza de caché*.

**3** 🗣️ ¿Cómo ha evolucionado la familia del procesador K7?

- a** En tecnología de integración: 0.25, 0.18, 0.13 micras.
- b** En caché L2: Interna de 512 Kbytes, integrada de 256 Kbytes, integrada de 128 Kbytes.
- c** En paralelismo a nivel de instrucción: Factor de superescalaridad 3, 4 y 5.
- d** En el conjunto de instrucciones: 3DNow!, Enhanced 3DNow!, Hyper-Enhanced 3DNow!.

**4** 🗣️ Tenemos ante nosotros a un K7 de 800 MHz. Podemos asegurar con plena certeza que se trata de un modelo fabricado mediante:

- a** CMOS de 0.25 micras e interconexiones de aluminio.
- b** CMOS de 0.18 micras e interconexiones de aluminio.
- c** CMOS de 0.18 micras e interconexiones de cobre.
- d** No podemos asegurar nada a priori.

**5** 🗣️ ¿Qué rasgo comparten todos los modelos de microprocesador K7 del mercado?

- a** El zócalo Socket A y la caché L2 integrada.
- b** La integración CMOS de 0.18 micras e interconexiones de cobre.
- c** El conjunto de instrucciones Enhanced 3DNow!.
- d** La caché L1 de 128 Kbytes y la L2 de 256 Kbytes.

**6** 🗣️ ¿Qué fase del procesador K7 consume un mayor número de ciclos?

- a** Búsqueda.
- b** Decodificación.
- c** Reordenación.
- d** Ejecución.

**7** 🗣️ ¿Qué es un Duron?

- a** Un K7 con 1/4 de su tamaño de caché L2.
- b** Un K7 de bajo coste.
- c** Las dos respuestas anteriores son correctas.
- d** Un Device-Under-Rambus-Or-Network.

**8** 🗣️ Disponemos de un modelo de procesador Duron. Sabemos que es...

- a** Una arquitectura intermedia entre el K6 y el K7.

**b** Un K7 Thunderbird con el bus local más lento: 200 MHz frente a 266 MHz.

**c** Un K7 Thunderbird con la cuarta parte de caché: 64 Kbytes frente a 256 Kbytes.

**d** Las dos respuestas anteriores son correctas.

**9** ¿Qué multiplicador fue el más utilizado por AMD para la caché L2 interna de su primer Athlon de 0.25 micras?

**a** 1/3 de la frecuencia del procesador.

**b** 1/2 de la frecuencia del procesador.

**c** 1 (la misma frecuencia del procesador).

**d** La L2 del primer Athlon es externa.

**10** A la hora de discriminar el peregrinar de una instrucción entera y otra de punto flotante por el cauce segmentado del procesador K7, las únicas etapas que son compartidas por ambas son

**a** Las de la fase de búsqueda, como en cualquier otro procesador.

**b** Las de la fase de búsqueda y decodificación, como en cualquier otro procesador.

**c** Las de la fase de búsqueda, decodificación y búsqueda de operandos.

**d** Todas hasta la entrada en la ALU en el primer caso y en la FPU en el segundo, como en el Pentium 4.

**11** Hay procesadores cuyas mejoras son marginales respecto a su predecesor, otros que introducen cambios sustanciales en el rendimiento, y otros que representan ya una nueva arquitectura. En el mundo del software, si el producto original constituye la versión 1.0, los primeros estarían etiquetados como la versión 1.01, los segundos como la 1.1 y los terceros como la 2.0. Si adjudicamos al Athlon original la versión 7.0 de AMD, ¿cómo catalogarías a los procesadores Athlon Thunderbird, Athlon XP y Clawhammer, respectivamente? (considerar en todos ellos la primera versión aparecida)

**a** 7.1, 7.2 y 8.0.

**b** 7.01, 7.02 y 7.1.

**c** 7.1, 7.11 y 8.0.

**d** 7.01, 7.11 y 8.0

**12** AMD nos ha encargado el diseño de un nuevo procesador basado en la arquitectura K7, pero que

rompe la compatibilidad con el código x86, aceptando el conjunto de instrucciones máquina nativo desde un principio. Sin embargo, se quieren aprovechar todas las unidades funcionales que sea posible de la vieja arquitectura. ¿Cuál de las siguientes es imposible reutilizar?

**a** El acelerador para las predicciones de salto ó BTB.

**b** El búfer de reordenación circular ó BRC.

**c** El acelerador para las traducciones de memoria virtual a física ó TLB.

**d** Los bancos de registros y las unidades aritméticas que se nutren de ellos.

**13** El primer microprocesador Pentium 4 lanzado al mercado en Noviembre de 2000 disponía de

**a** 1.4 GHz, 0.18 micras y 256 Kbytes de caché L2 integrada.

**b** 1.4 GHz en su núcleo, 400 MHz en su bus local y 100 MHz en su caché L2.

**c** Una caché L1 de mayor tamaño que su L2.

**d** Todas las respuestas anteriores son correctas.

**14** En un Pentium 4 de 2 GHz

**a** Todas sus unidades funcionales se encuentran sincronizadas a un periodo de reloj que emite dos mil millones de pulsos por segundo.

**b** Es imposible sincronizar 42 millones de transistores a una frecuencia de reloj tan elevada debido a los retrasos en el transporte de la señal por un área de integración tan grande.

**c** La mayoría de sus unidades funcionales están sincronizadas por esa señal de reloj, pero el controlador de bus responde cinco veces más lento, y las ALU, dos veces más rápido.

**d** Lo más que pueden hacer la placa base y la memoria principal es trabajar a la vigésima parte de esa velocidad.

**15** El sistema de memoria con el que dialoga un Pentium 4 Willamette se compone de

**a** Un nivel de caché integrada y memoria principal RDRAM.

**b** Dos niveles de caché integrada y memoria principal RDRAM.

**c** Dos niveles de caché integrada y memoria principal RDRAM ó DDRAM.

**d** Dos niveles de caché integrada y memoria principal RDRAM ó DDRAM dependiendo del puente norte del juego de chips que le acompañe en placa base.

- 16 ¿Cuál era la configuración de caché predominante en el mercado en el momento del lanzamiento del Pentium 4?
- a 256 Kbytes de caché L3 interna.
  - b 256 Kbytes de caché L2 integrada.
  - c 256 Kbytes de caché L1 integrada.
  - d 32 Kbytes de caché L1 integrada.
- 17 ¿Cuál es la principal contribución del Pentium 4 desde el punto de vista de su arquitectura interna?
- a La tecnología de integración: 0.13 micras.
  - b El paralelismo a nivel de instrucción: Puede ejecutar hasta 100 instrucciones de forma simultánea.
  - c El conjunto de instrucciones: Se deshace por fin del tormentoso legado x86.
  - d La presencia de una caché de traza.
- 18 ¿Qué aspecto altera una caché de traza con respecto al rendimiento de una caché convencional?
- a A igual número de transistores, aumenta el índice de aciertos, pero reduce el beneficio que se obtiene en cada uno de ellos.
  - b A igual número de Kbytes, el número de instrucciones máquina que caben en ella es inferior.
  - c A igual velocidad, ralentiza la computación de la instrucción que sale de ella, ya que éstas ocupan más espacio, y ya se sabe que en microelectrónica, "más grande, más lento".
  - d Altera los tres aspectos mencionados en las opciones anteriores, pero precisamente en el sentido contrario al que éstas estipulan.
- 19 ¿Pueden montarse cachés de traza a diferentes niveles de una jerarquía de memoria como ocurre con las cachés convencionales? Asumir que se prescinde tan sólo de las cachés de instrucciones, manteniendo íntegramente la L1D y la parte proporcional de datos en cualquier caché unificada (L2, L3, ...).
- a Sí, la extensión del concepto es inmediata, colocando progresivamente niveles más grandes y lentos.
  - b No, porque cada instrucción se decodifica una sola vez.
  - c Sólo sería posible en arquitecturas como la del K7, donde se efectúa una decodificación también a diferentes niveles: Macrodecodificación y microdecodificación.
- d Sólo sería posible en arquitecturas como la del Pentium 4, donde el número de etapas de segmentación es muy elevado.
- 20 ¿Cómo se cubre el Pentium 4 de la vulnerabilidad que supone un cauce de ejecución segmentado en 20 etapas?
- a Ampliando la frecuencia de trabajo del bus local hasta los 400 MHz e incorporando el interfaz de conexión dual con memoria RDRAM y DDRAM.
  - b Incorporando nuevos bancos de registros e incrementando el tamaño de caché.
  - c Aumentando sobremanera el tamaño de la BTB y el BRC en relación a su predecesor, el Pentium III.
  - d Haciendo trabajar a las ALU internas al doble de velocidad que el resto del procesador.
- 21 Intel nos encarga el diseño de un procesador lo más parecido posible en su hardware al Pentium 4 y que conserva su mismo patillaje y zócalo, pero que acepta un repertorio de instrucciones propio, incompatible con el código x86. Indica cuál de las cuatro unidades funcionales siguientes es más ajena a la selección del conjunto de instrucciones, y por lo tanto, puede reutilizarse para nuestro diseño.
- a El controlador de bus local.
  - b La caché de traza.
  - c Los diferentes bancos de registros.
  - d Las ALU al doble de frecuencia de reloj.
- 22 Elige la sentencia que mejor resume la arquitectura del Pentium 4
- a Gran número de transistores para construir unas amplias cachés en sus dos niveles integrados dentro del procesador.
  - b Elevada frecuencia para un profundo cauce de segmentación.
  - c Dependencia del compilador dado su marcado diseño RISC.
  - d Gran número de transistores y elevada frecuencia que conducen a un marcado diseño superescalador.
- 23 Entre la quinta y la séptima generación, el número de etapas de segmentación ha crecido más en las arquitecturas
- a De AMD.
  - b De Intel.
  - c De Intel, porque en AMD ha venido descendiendo.

**d** En ambas marcas la evolución ha venido siendo muy similar.

**24** ¿Cuál de las siguientes es la característica más sobresaliente del procesador K7 en comparación a los modelos equivalentes provenientes de Intel?

- a** Frecuencia de reloj.
- b** Tecnología de integración.
- c** El tamaño de la caché L1.
- d** El tamaño de la caché L2.

**25** ¿Qué microprocesadores hemos visto fabricados tanto en formato Slot como en Socket?

- a** El Pentium II, III y 4 de Intel y el K6-2, K6-III y K7 de AMD.
- b** El Pentium III y su versión Celeron por parte de Intel y el K7 y su versión Duron por parte de AMD.
- c** El Pentium III de Intel y el K7 de AMD, cambiando en ambos casos de 512 Kbytes de caché interna bajo el formato Slot, a 256 Kbytes de caché integrada bajo el formato Socket.
- d** Ninguno. Cada microprocesador se ha fabricado exclusivamente en un tipo de formato.

**26** ¿Qué procesador es más rápido a la hora de ejecutar un programa, el K7 de 1 GHz o el Pentium 4 de 1.4 GHz?

- a** Siempre el K7.
- b** Siempre el Pentium 4.
- c** Empatán en todos los casos.
- d** Están muy igualados, pero dependerá de las características del programa a ejecutar el que la balanza se decante a favor de uno u otro.

**27** ¿En qué facetas saca ventaja el K7 frente al Pentium 4?

- a** El grado de superescalaridad.
- b** El tamaño de las cachés de primer nivel.
- c** Las dos respuestas anteriores son correctas.
- d** Ninguna de las respuestas anteriores es correcta.

**28** ¿En qué facetas saca ventaja el Pentium 4 frente al K7?

- a** La segmentación en el caso ideal.

**b** La resolución de las dependencias de datos.

**c** Las dos respuestas anteriores son correctas.

**d** Ninguna de las respuestas anteriores es correcta.

**29** Pretendemos idear una fórmula para relacionar el número cardinal de una generación de microprocesadores,  $G$ , con el número de etapas de segmentación en los modelos de Intel para esa generación,  $I$ . La expresión matemática que acierta con error máximo de una etapa para  $G$  restringida a los valores 5, 6 y 7 (esto es, entre quinta y séptima generación) es:

**a**  $I = 5 \times 2^{(G-5)}$

**b**  $I = 5 \times (G - 4)$

**c**  $I = 3 \times (G - 3)$

**d**  $I = G + (G - 5) \times 5$

**30** Sabemos que el Pentium 4 tiene una arquitectura diferente de la del Pentium III porque

**a** Su ordinal se escribe en números arábigos, frente a la numeración romana de todos los modelos predecesores en los que se comparte la misma arquitectura base del Pentium.

**b** El tiempo que separa sus fechas de lanzamiento es muy superior al que separa las de los modelos que comparten arquitectura, como el Pentium II y III, por ejemplo.

**c** La tecnología de integración del primer Pentium 4 es diferente de la del último Pentium III.

**d** Ninguno de los razonamientos anteriores nos permite inferir que la arquitectura del Pentium 4 es diferente de la del Pentium III.

**31** ¿Cuál de los siguientes microprocesadores cuenta con un marcado diseño VLIW?

**a** El K7.

**b** El Pentium 4.

**c** El Pentium 4 con HyperThreading.

**d** Ninguno de los anteriores.

**32** ¿Qué conjunto de instrucciones es más parecido al del microprocesador Pentium 4?

**a** El del Pentium III.

**b** El del K7.

**c** El del Itanium.

- d** El del Pentium 4 Celeron.
- 33** ¿Qué rasgo ha caracterizado unívocamente al microprocesador Pentium II, Pentium III y Pentium 4, respectivamente?
- a** Tecnología de integración: 0.25, 0.18 y 0.13 micras.
- b** Caché: L2 interna, L2 integrada, L3 interna.
- c** Instrucciones multimedia: MMX, SSE y SSE2.
- d** Frecuencia de reloj: 500 MHz, 1000 MHz y 1500 MHz.
- 34** Hemos desarrollado un microprocesador de ciertas carencias y lentitud en su fase de búsqueda/decodificación, y de extremada agilidad y rapidez en su fase de reordenación/ejecución. ¿Con qué arquitectura le encuentras un mayor parecido?
- a** El Pentium Pro (sexta generación en Intel).
- b** El K6 (sexta generación en AMD).
- c** El Pentium 4 (séptima generación en Intel).
- d** El K7 (séptima generación en AMD).
- 35** ¿Qué pareja de microprocesadores recorta su tamaño de caché L2 cuando ésta pasa de ser interna a ser integrada en el mismo chip de la CPU?
- a** El K6-III y el Pentium II.
- b** El K6-2 y el Pentium III.
- c** El Pentium II y el Pentium III.
- d** El Pentium III y el K7.
- 36** ¿Qué característica destacarías de Intel como fabricante de microprocesadores?
- a** Su capacidad de adelanto a las exigencias del mercado.
- b** La rapidez de procesamiento bruto en sus ALU y FPU.
- c** La integración de enormes cachés dentro del chip.
- d** El bajo precio de sus productos.
- 37** El aspecto que resume la metamorfosis producida desde el K6 al K7 es
- a** La superior frecuencia de reloj.
- b** El diferente formato de instrucción.
- c** La mayor anchura en los buses.
- d** Son de generaciones diferentes.
- 38** ¿Qué microprocesador representa la transición entre dos generaciones?
- a** El Katmai.
- b** El Willamette.
- c** El Celeron.
- d** El Deschutes.
- 39** ¿Qué compañía partió con ventaja en el lanzamiento de su primer microprocesador de séptima generación?
- a** Intel, porque desarrollaba productos más baratos.
- b** Intel, porque su arquitectura era más potente.
- c** AMD, porque su K7 vio la luz un año antes.
- d** AMD, gracias a la generosa cobertura proporcionada por los fabricantes de placa base.
- 40** ¿Qué compañía llegó primera al hito del Gigahercio en la frecuencia de reloj de un microprocesador?
- a** Intel y AMD. Tanto el Pentium III como el K7 lo consiguieron, y casi al alimón.
- b** Compaq, a través del Alpha 21364 de Digital.
- c** Seguro que fue algún diseño RISC orientado al segmento de servidores.
- d** AMD, con su K8.
- 41** ¿Qué antiguo inconveniente de los microprocesadores de AMD es un parámetro que juega ahora a su favor?
- a** La frecuencia de reloj.
- b** La fecha de comercialización de sus diseños.
- c** El precio.
- d** Los tres anteriores.
- 42** ¿Qué innovación tecnológica apoyada por Intel se volvió en su contra al iniciarse la andadura de la séptima generación de microprocesadores?
- a** El bus AST.
- b** La memoria RDRAM.

- c** El formato ATX de las placas base.
- d** Las tres anteriores.
- 43** ¿Qué razón estuvo detrás de la vuelta al formato Socket con la llegada del microprocesador K7 de 0.18 micras (Thunderbird)?
- a** AMD se pasa a Slot cuando las grandes cachés de segundo nivel (L2) le obligan, y regresa al Socket cuando éstas L2 reducen su capacidad y aumentan su velocidad, imitando el proceder de Intel con su Pentium III.
- b** Que el voltaje sufrió una drástica reducción.
- c** Siempre que Intel pase de Slot a Socket, AMD recorrerá el camino inverso y viceversa.
- d** Las tres respuestas anteriores son correctas.

Un grupo de viejos alumnos de nuestra escuela ha fundado Sol Valley, una novel empresa dedicada al diseño de procesadores para PC. Su primer producto es Decaomnium, un revolucionario procesador de séptima generación (2003). Ellos lo denominan "la arquitectura diez", porque aseguran haber alineado sus principales magnitudes en torno a ese número. Por el momento, conocemos que cuenta con 10 millones de transistores, fabricados a 0.10 micras y dispuestos en 10 etapas de segmentación.

- 44** Corren rumores de que Decaomnium logra alcanzar una frecuencia de 10 GHz. ¿Qué argumentarías al respecto?
- a** Atendiendo exclusivamente al número de transistores y su integración a 0.10 micras, tiene al alcance dicho valor.
- b** Atendiendo exclusivamente a la integración a 0.10 micras y a las 10 etapas de segmentación, tiene al alcance dicho valor.
- c** Atendiendo exclusivamente a los 10 millones de transistores y las 10 etapas de segmentación, tiene al alcance dicho valor.
- d** Es un farol de sus creadores.
- 45** Dicen de Decaomnium que dispone de un factor 10 de superescalaridad. ¿Qué tienes que decir al respecto?
- a** Atendiendo al número de transistores, apuesto a que es un farol.
- b** Atendiendo a la integración a 0.10 micras, apuesto a que es un farol.
- c** Atendiendo al número de etapas de segmentación, apuesto a que es un farol.

- d** Atendiendo exclusivamente a las tres magnitudes anteriores, lo considero un valor factible.

- 46** ¿Cuenta Decaomnium con ingredientes suficientes para apostar por su supersegmentación?

- a** Sí.
- b** No. Hacen falta más transistores.
- c** No. Sería necesario contar de partida con una segmentación más profunda.
- d** No. No conocemos la frecuencia del diseño, magnitud clave para poder supersegmentar.

- 47** Se especula con la posibilidad de que Decaomnium disponga de un único nivel de caché integrada, con una L1D de 10 Kbytes y una L1I también de 10 Kbytes. ¿Es una alternativa razonable?

- a** No. Por las magnitudes que conocemos, esas capacidades se quedan muy por debajo de lo que el diseño puede dar de sí.
- b** Sí, pero colocando ambas cachés en un chip aparte.
- c** No es lo mejor desde el punto de vista del rendimiento, pero con 10 millones de transistores en total, pocas alegrías nos podemos conceder en este sentido.
- d** Esa elección descompensa el diseño desde el punto de vista de la segmentación y superescalaridad consideradas.

- 48** ¿Cuál es el punto fuerte de Decaomnium para ser competitivo frente al K7 de AMD?

- a** La elevada frecuencia.
- b** El bajo coste del producto (integración y materiales).
- c** El paralelismo a nivel de instrucción.
- d** Sus cachés integradas.

- 49** ¿Cuál es el punto fuerte de Decaomnium frente al Pentium 4 de Intel?

- a** La tecnología de integración.
- b** El número de etapas de segmentación.
- c** El factor de superescalaridad.
- d** El conjunto de instrucciones.

- 50** De las tres magnitudes que se dieron como conocidas para Decaomnium, ¿Cuál ha sido la peor elección?

**a** El bajo número de transistores, que no permitirá cubrirse apropiadamente de las dependencias para un cauce de segmentación bastante largo, ni construir rápidas ALU que estén a la altura de la frecuencia a que aspira el diseño.

**b** Disponer de pocos transistores no es un gran problema si se compensa adecuadamente con las múltiples ventajas que reúne un diseño de gran simplicidad. El problema son las micras: Apostar por 0.10 cuando los gigantes del sector acaban de llegar a las 0.13 obliga cuando menos a dudar de que Sol Valley conseguirá dar con alguien que pueda integrarle el diseño.

**c** Las 0.10 micras serían un problema para 100 millones de transistores, pero no para 10. El problema son las etapas de segmentación: Dado que es el mejor mecanismo para explotar el paralelismo a nivel de instrucción cuando la infraestructura hardware es limitada, debiera haberse apostado por una veintena de etapas.

**d** Las tres respuestas anteriores son correctas.

**51** ¿Satisface Decaomnium las premisas básicas para considerarla una arquitectura RISC?

**a** Sí. Apuesta por la simplicidad y la frecuencia elevada.

**b** Sí, aunque 10 etapas de segmentación son muchas para un RISC.

**c** Sí, aunque 10 millones de transistores son pocos para un RISC.

**d** No.

**52** Asumiendo los tres siguientes rasgos de Decaomnium: (A) Pocos transistores, (B) Estrecha distancia de integración, y (C) Elevada frecuencia, ¿Cuáles juegan a favor y cuáles en contra de adjudicar al procesador un voltaje de alimentación lo más bajo posible?

**a** A y B a favor; C en contra.

**b** Todos a favor.

**c** Todos en contra.

**d** A a favor; B y C en contra.

**53** ¿Dispone Decaomnium de una gran facilidad para la incorporación de un conjunto de instrucciones multimedia?

**a** Sí.

**b** No. Un diseño RISC como se presupone que abandera Decaomnium es incompatible con la idea de las instrucciones multimedia.

**c** No. 10 etapas de segmentación son pocas para la complejidad que introducen las instrucciones multimedia.

**d** Tendríamos que conocer algunos detalles sobre el conjunto de instrucciones máquina del procesador, pero en cualquier caso haría falta tener una mayor holgura en la disponibilidad de transistores.

Establecemos un baremo para puntuar la memoria caché de una serie de procesadores en sus dos primeros niveles (L1 y L2), otorgando un punto por cada una de las premisas siguientes:

- Su presencia externa, sin importar el tamaño, en caso de existir configuraciones que dispongan de ella en placa base.
- Cada 64 Kbytes disponibles en otro chip adjunto (caché interna).
- Cada 16 Kbytes integrados en el chip procesador (caché integrada).
- Cada 2 Kbytes de espacio equivalente en la caché de traza respecto al tamaño convencional de una L1I.

Los puntos conseguidos en el primer nivel cuentan el triple que los del segundo. Por lo tanto, puntuaremos el primer y el segundo nivel por separado, y luego multiplicaremos por tres el primero antes de sumarlo con el segundo para obtener la puntuación total de cada procesador. Las cantidades residuales no puntúan (por ejemplo, 8 Kbytes de caché integrada no obtienen punto, y 3 Kbytes de caché de traza puntúan lo mismo que 2 Kbytes).

El ranking se establece entre los siguientes 12 procesadores: Por parte de Intel, el primer Celeron, el Klamath, el Deschutes, el Katmai, el Coppermine, el Willamette y el Tualatin. Por parte de AMD, el primer K6, el Sharptooth, el primer Athlon, el Duron y el Thunderbird. Ambas series se han enumerado respetando la evolución temporal seguida por cada firma, y de la serie de doce ordenada según el baremo de puntuación que hemos establecido, obtenemos 7, 13, 14, 14, 14, 22, 28, 28, 32, 37, 38 y 40 puntos.

Se recomienda construir una tabla con doce filas (una para cada procesador) y cinco columnas (L1 integrada, L1 de traza, L2 externa, L2 interna y L2 integrada) donde registrar la puntuación de cada procesador en cada caché, acumular su puntuación global, identificarla con alguna de las doce cantidades indicadas anteriormente, y finalmente proceder a responder a las siguientes cuestiones:

**54** Hay un procesador tan especial que puntúa en una casilla que ningún otro lo hace. ¿Cuál es?

**a** El Tualatin.

**b** El Willamette.

**c** El primer K6.

**d** El primer Celeron.

**55** 🗡️ Hay un procesador que no puntúa en una casilla en la que sí lo hacen todos los demás. ¿Cuál es?

- a** El Celeron.
- b** El K6.
- c** El Duron.
- d** El Willamette.

**56** 🗡️ Existe un empate a 28 puntos entre dos procesadores a pesar de que presentan configuraciones muy dispares. ¿Cuáles son?

- a** El primer Celeron y el primer K6.
- b** El Sharptooth y el Duron.
- c** El Athlon y el Thunderbird.
- d** El Coppermine y el Tualatin.

**57** 🗡️ Existe un triple empate a 14 puntos entre tres procesadores que dadas sus similitudes no podrían desempatar por mucho que cambiáramos el baremo de puntuación. ¿Cuáles son?

- a** Primer K6, primer Celeron y primer Pentium II (Klamath).
- b** Klamath, Deschutes y Katmai.
- c** Coppermine, Tualatin y Willamette.
- d** Sharptooth, Athlon y Thunderbird.

**58** 🗡️ ¿Qué procesador brilla tanto en uno de sus niveles de caché que la puntuación aislada de este nivel supera ya los 30 puntos?

- a** El Willamette con su caché de traza.
- b** El Tualatin con su caché L2.
- c** El Athlon con su caché L1.
- d** Los tres anteriores.

**59** 🗡️ ¿Qué procesador queda en último lugar en nuestra clasificación?

- a** El primer K6.
- b** El primer Celeron.
- c** El Klamath.
- d** El Duron.

**60** 🗡️ El baremo anterior no favorece a las configuraciones equilibradas. Si otorgáramos un plus de 10 puntos al procesador que menos diferencias presenta entre la puntuación de su L1 y la de su L2, ¿cambiaría el ganador del ránking?

- a** Sí, sería el que ahora queda segundo.
- b** Sí, sería el que ahora queda tercero.
- c** Sí, sería el que ahora queda cuarto.
- d** No.

**61** 🗡️ Si introdujésemos en nuestro ránking al último procesador Pentium (esto es, la versión MMX), ¿tendría la oportunidad de superar a alguno de los últimos clasificados?

- a** Sí, al primer Celeron.
- b** Sí, al primer K6.
- c** No.
- d** No, aunque quedaría empatado con el primer Celeron.

**62** 🗡️ Si introdujésemos en nuestro ránking al Athlon XP, ¿Con qué procesador quedaría siempre empatado al margen de nuestro baremo?

- a** Con el Athlon.
- b** Con el Thunderbird.
- c** Con el Sharptooth.
- d** Con ninguno de los tres anteriores.

**63** 🗡️ Para evaluar la caché, más fácil que usando el baremo anterior lo tenemos fijándonos en la distancia de integración de los transistores de cada modelo, ya que de las cinco magnitudes estudiadas para el procesador, es la que más facilita la consecución de grandes cachés integradas. ¿Corrobora nuestro ránking esta percepción, al menos atendiendo al primer y último clasificado?

- a** Sí, porque el líder obedece al único modelo de 0.18 micras, y el último clasificado, al único de 0.35 micras.
- b** No, el que gana a todos no es de 0.18 micras, y el que pierde ante todos mejora luego sus cachés manteniéndose su fabricación de 0.35 micras.
- c** No, el que gana a todos no es de 0.18 micras, y el que es de 0.18 micras no gana.
- d** No, no hay ningún modelo de 0.18 micras y sí varios de 0.35 micras en la lista de 12 modelos evaluados.



# La octava generación

## Sumario

<b>7.1. El K8 de AMD</b> . . . . .	<b>256</b>
7.1.1. Etimología . . . . .	256
7.1.2. Frecuencia . . . . .	257
7.1.3. Tecnología de integración . . . . .	257
7.1.4. Paralelismo a nivel de instrucción . . . . .	258
7.1.5. Memoria interna . . . . .	260
7.1.6. Conjunto de instrucciones . . . . .	260
7.1.6.1. Arquitectura de 64 bits . . . . .	260
7.1.6.2. x86-64 e IA-64: Dos aproximaciones diferentes . . . . .	261
7.1.6.3. La implementación del x86-64 . . . . .	263
7.1.6.4. Extensiones multimedia . . . . .	265
7.1.7. El interfaz del procesador . . . . .	266
7.1.7.1. Lo que queda del bus local . . . . .	267
7.1.7.2. HyperTransport . . . . .	267
7.1.7.3. La nueva jerarquía de memoria . . . . .	268
7.1.7.4. El doble puerto de acceso a memoria principal . . . . .	269
7.1.8. Las diferencias entre Athlon 64 y Opteron . . . . .	271
<b>Resumen</b> . . . . .	<b>271</b>
<b>La anécdota: Alpha 21464, o el primer procesador arácnido</b> . . . . .	<b>273</b>
<b>Cuestionario de evaluación</b> . . . . .	<b>273</b>

Como ya advertimos en el capítulo anterior, las generaciones de microprocesadores han dejado de ser un marco que sincronice en el tiempo los modelos de las distintas compañías, y este hecho se refleja en mayor medida aún en la octava generación, donde hasta la fecha hay una única arquitectura: El K8 de AMD.

En la edición de 2001 de Arquitectura del PC desvelamos las intenciones de Intel de inaugurar esta generación con el Pentium 8. Desde entonces, ha trascendido que el procesador, con código de referencia Nehalem, disfrutará del nuevo proceso de integración de 65 nm. para concentrar más de 100 millones de transistores en la mitad de área de silicio que ocupa el actual Northwood. El procesador dispondrá de un bus en torno a los 1200 MHz, y de respetarse la pauta temporal que tradicionalmente ha separado dos generaciones, deberá ver la luz antes de que finalice 2004.

Entretanto ocurre todo eso, analizaremos la realidad que es ya el K8. Los modelos comerciales que se derivan de esta arquitectura se adjuntan en la [tabla 7.1](#). De allí puede desprenderse un pro-

Nombre comercial	Fecha lanz.	T.I. (nm.)	Socket (pines)	SMP	Controlador de memoria	Caché L2	Código de referencia
Athlon 64	09/03	130	Socket(754)	1	1 x DDRAM	1024	Clawhammer
Athlon 64	09/03	130	Socket(754)	1	1 x DDRAM	256	París
Athlon 64	2004	90	Socket(754)	1	1 x DDR-II	1024	San Diego
Opteron UP	04/03	130	Socket(940)	1	2 x DDRAM	1024	Sledgehammer
Opteron DP	04/03	130	Socket(940)	2	2 x DDRAM	1024	Sledgehammer
Opteron MP	06/03	130	Socket(940)	4/8	2 x DDRAM	1024	Sledgehammer
Opteron UP	2004	90	Socket(940)	1	2 x DDR-II	1024	Venus
Opteron DP	2004	90	Socket(940)	2	2 x DDR-II	1024	Troy
Opteron MP	2004	90	Socket(940)	4/8	2 x DDR-II	1024	Athens
Pentium 8?	2004?	90/65	N/D	N/D	DDR-II	N/D	Nehalem

TABLA 7.1: Nombre comercial y código de referencia para los procesadores de octava generación anunciados hasta la fecha. La casilla T.I. indica la distancia de integración en nanómetros. La casilla SMP indica la capacidad del procesador para el multiprocesamiento simétrico: 1 significa monoprocesador, y por consiguiente orientación al mercado doméstico del PC; 2 es biprocesador para estaciones de trabajo y pequeños servidores, y 4/8 es la gama más alta.

ceder de AMD muy similar al exhibido por Intel en generaciones pasadas, donde permanece un núcleo arquitectural común para dos segmentos del mercado: el PC doméstico y las plataformas de tipo servidor. Y aunque ése haya sido el resultado final, conviene distinguir la forma en que se ha llegado a él en cada caso:

- 5 En quinta generación, la arquitectura base es la gama baja (el Pentium), y sobre él se colocan tímidas mejoras en placa base para lograr algún sistema servidor.
- 6 En sexta generación, el embrión es de gama alta (el Pentium Pro). Año y medio más tarde se reconvierte al mercado doméstico (Pentium II y III), y otro año y medio después se le incorporan mejoras en los buses y la caché para hacerlo regresar a la gama alta (Xeon).
- 7 En séptima generación, Intel pone en liza dos arquitecturas en paralelo: El Pentium 4 para el segmento doméstico y el Itanium para estaciones de trabajo y servidores. El primero usurpa luego el terreno del segundo con nuevos modelos Xeon, pero el segundo ha permanecido fiel a su jurisdicción, y tiene su propio nicho de mercado, por lo que no tiene cabida en un libro dedicado al PC como éste.
- 8 En la octava generación, el guión de Intel que más se parece al que ha seguido AMD es el del Pentium Pro: El K8 se gestó como un ambicioso proyecto orientado a gama alta, y por el camino se fue pensando en cómo recortarlo para que fuera también viable para perfiles domésticos, hecho que justifica nuestra cobertura del mismo a continuación.

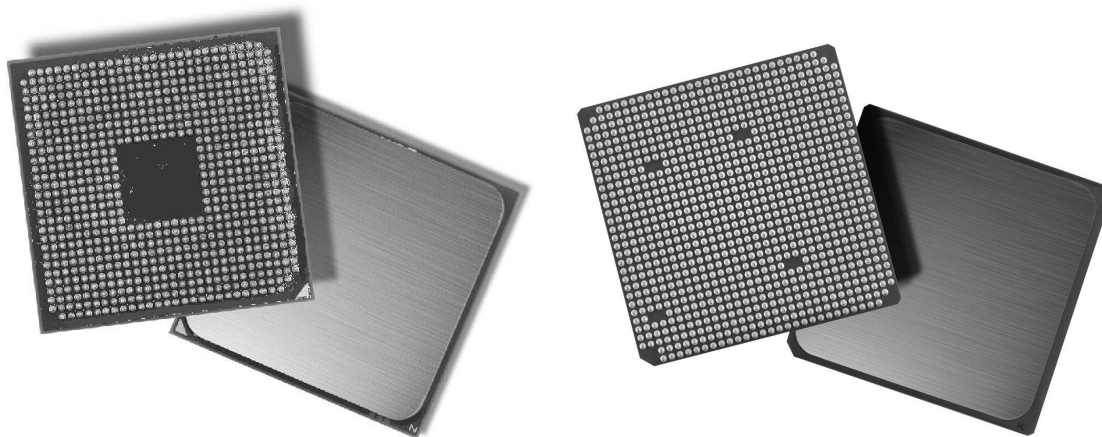
## SECCIÓN 7.1

### El K8 de AMD

#### 1.1 ► Etimología

códigos de referencia

Los códigos de referencia del procesador K8 fueron en sus prolegómenos un tanto belicosos



(a) Athlon 64 (Clawhammer).

(b) Opteron (Sledgehammer)

Fotos cortesía de AMD

FOTO 7.1: Empaquetado y patillaje de los dos modelos de K8. (a) El Athlon 64 con 754 pines. (b) El Opteron, similar en dimensiones, pero con 940 pines.

(Clawhammer significa *martillo de carpintero*, mientras que Sledgehammer es aún más atroz: *Maza rompepiedras*), así que nos alegra la rectificación de AMD aprovechando los modelos de 90 nanómetros, en favor de bonitas ciudades como San Diego, París y Atenas.

Pero el procesador no se conocerá comercialmente por ninguna de estas denominaciones. Ni siquiera por la de K8. El nombre comercial elegido para la versión doméstica es Athlon 64, que por un lado vive de las rentas de la popularidad alcanzada por el nombre de la generación anterior (estrategia que es copiada de la que nos ha enseñado Intel con su saga Pentium), y por el otro, pretende dejar bien claro que se trata de la primera arquitectura de 64 bits de la compañía.

Athlon 64

Para la versión de gama alta (servidores), el nombre comercial es Opteron, procedente del latín, *Optimus*, que significa *el óptimo*. El nombre tiene menos gracia que una multa. Le sobra vanidad, le falta originalidad (todos sabemos quién usó por primera vez el latín para bautizar procesadores), y para rematar el despropósito, se cambia el sufijo por -ron, que llevaba seis años reservado en exclusiva como insignia de la gama baja (Celeron, Duron).

Opteron

## Frecuencia

### ◀ 1.2

La frecuencia inicial del K8 ha sido un secreto bastante bien guardado. Finalmente, parece que el procesador comienza su andadura con versiones a 1.4 GHz, 1.6 GHz, 1.8 GHz y 2 GHz, para continuar subiendo de 200 en 200 MHz hasta superar los 3 GHz a finales de 2003.

inicial

Sea como fuere, llamamos la atención sobre las cifras múltiplo de 200 MHz, sobre las que apostamos a ciegas al margen de especulaciones, ya que la especificación HyperTransport sobre la que se basa la implementación del bus (ver [sección 7.1.7.2](#)) contempla todas sus frecuencias como múltiplos de esa base, y el procesador obtiene a su vez su propia frecuencia como múltiplo de la que rige en su bus.

del bus

◀ pág. 267

## Tecnología de integración

### ◀ 1.3

La versión más modesta del K8 (Athlon 64 con L2 de 256 Kbytes) cuenta con 67 millones de transistores alimentados a 1.55 voltios, y las versiones con L2 de 1 Mbyte superan ya los cien millones de transistores.

número de transistores

Paralelismo a nivel de instrucción	Arquitectura	
	K7	K8
Segmentación		
Etapas de búsqueda y decodificación (front-end)	6	8
Etapas de ejecución de instrucciones enteras (back-end)	4	4
Total cauce entero	6 + 4 = 10	8 + 4 = 12
Ejecución de instrucciones de punto flotante	9	9
Total cauce de punto flotante	6 + 9 = 15	8 + 9 = 17
Superescalaridad		
Factor sobre microinstrucciones nativas	Entre 3 y 6	Entre 3 y 6
Factor medio	5	5

TABLA 7.2: El paralelismo a nivel de instrucción del K8 guarda una extraordinaria similitud la arquitectura K7 de séptima generación.

En la fabricación del chip se utiliza en primera instancia una tecnología de integración de 130 nanómetros, pero que dista mucho de seguir el procedimiento convencional. La capa de polisilicio sobre la que se asientan los transistores se sustituye por un cristal mediante la técnica **SOI** (Silicon-On-Insulator), que minimiza las fugas de electrones (goteo del transistor - ver [figura 3.3](#)) manteniendo una elevada densidad en los portadores de carga de las regiones dopadas a su paso por el canal del transistor. Al apenas incidir estas fugas, ya no necesitan ser contrarrestadas con subidas de tensión en la puerta, lo que redonda en menor consumo, menor calentamiento y frecuencia más elevada (entre un 20% y un 30% respecto al chip fabricado convencionalmente según los investigadores de IBM responsables del hallazgo en 1999, y que ya lo emplearon para la fabricación del procesador Power 4 a finales de 2000 en alianza con Motorola).

SOI

pág. 51

menos calor y  
más frecuencia

patillaje

pág. 257

área

zócalo

A pesar de su complejidad estructural y de su patillaje cercano al millar de pines, las versiones prototipo que conocemos del K8 sorprenden por su miniatura (ver [foto 7.1](#)). Apenas ocupan un área de integración de 100 mm<sup>2</sup> (el último Athlon XP que comparte con él fabricación en 9 niveles de metalización ocupa 80 mm<sup>2</sup> para sus 37 millones de transistores, y los 42 millones del Pentium 4 de 180 nanómetros ocupan 247 mm<sup>2</sup>). En las versiones con caché L2 de 1 Mbyte, el área de integración se expande hasta los 180 mm<sup>2</sup>.

El zócalo del procesador es de tipo Socket, y el patillaje se distribuye en él como ya es tradicional en los encapsulados PGA (Pin Grid Array) del K7, esto es, formando un cuadrado inscrito. La diferencia está en que para pasar de los 462 pines del Thunderbird a los 752 del Athlon 64, el cuadrado tiene 10 filas de pines en lugar de 8, y cada una de ellas está más poblada. Para alcanzar los 940 pines del Opteron en un espacio similar se opta por recubrir de patillaje la zona central que antes quedaba diáfana.

## 1.4 ► Paralelismo a nivel de instrucción

En esta faceta del procesador no nos esperan grandes novedades, pero esto no significa que vengan malas noticias. Si algo bueno tenía la arquitectura K7 era el haber encontrado un equilibrio entre frecuencia y paralelismo a nivel de instrucción, otorgando casi un 50% de peso a cada una en la mágica ecuación que tipifica el rendimiento del sistema ( $T_{CPU} = NI \times CPI \times T$ , ver [sección 3.5.1](#)).

equilibrio

pág. 95

El logro no es baladí, pues no tiene precedentes en la extensa andadura de Arquitectura del PC, donde habremos analizado ya una veintena de diseños. Por tanto, parece muy sensato que AMD no haya querido arriesgar un ápice en el corazón de su nueva arquitectura, aplicando un dicho que es muy informático: "Cuando algo funciona, no lo arregles".

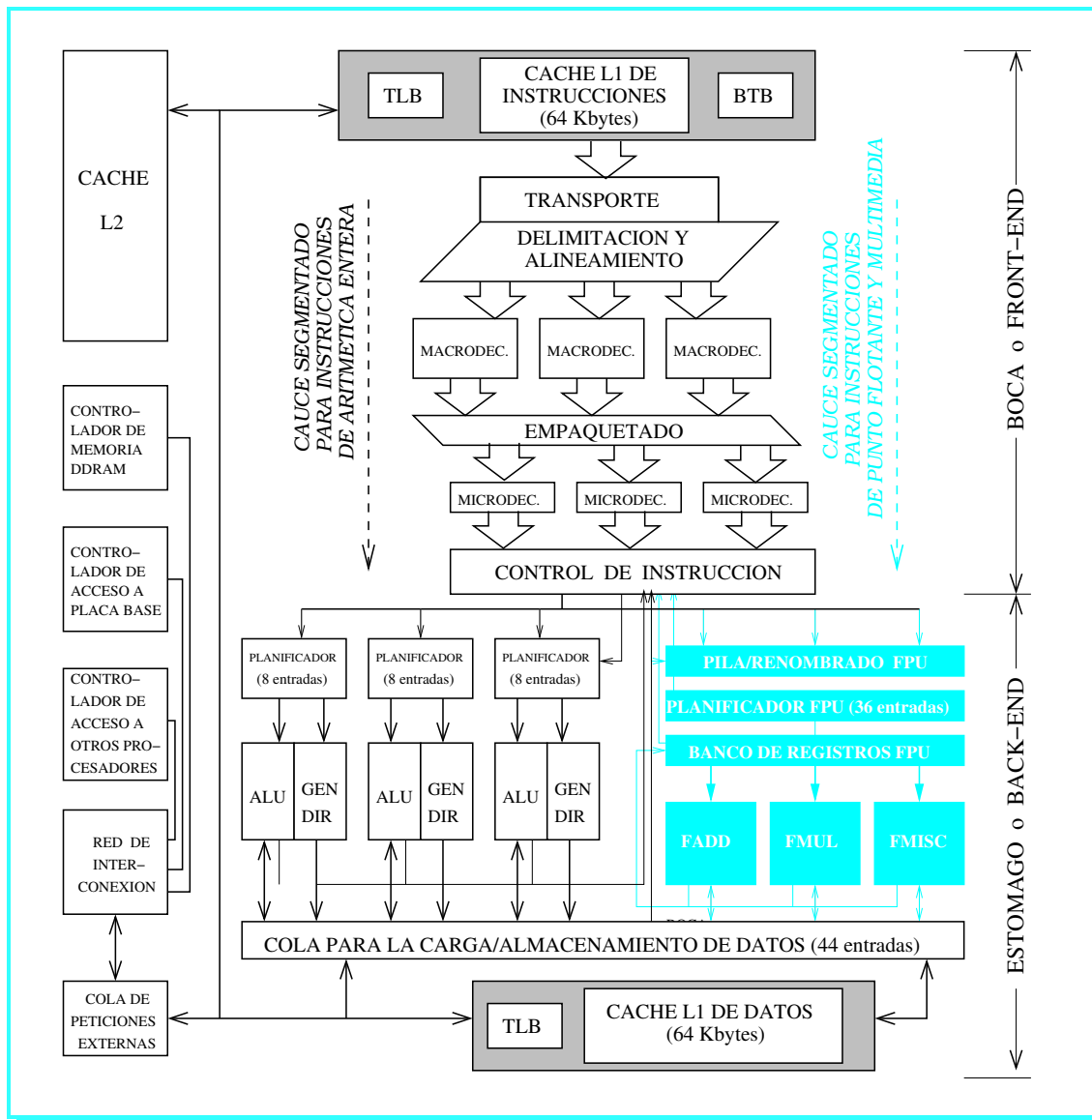


FIGURA 7.1: Diagrama de bloques de la arquitectura K8 de AMD, herencia del K7 en el complejo sistema de decodificación y los dos cauces de ejecución (entero y de punto flotante/multimedia).

La microarquitectura del K8, que adjuntamos en la [figura 7.1](#), es sospechosamente parecida a la de su generación anterior, con una *front-end* o boca del procesador que tiene el mismo sistema de decodificación de instrucciones en dos niveles que vimos para el K7, y un *back-end* o estómago en el que tienen cabida el mismo número de unidades funcionales.

En la boca (*front-end*), la principal complicación nos llega porque conviven las instrucciones viejas del K7 (que se sigue tragando el procesador para garantizar la compatibilidad con las viejas aplicaciones), con las instrucciones del nuevo conjunto x86-64. Además, se incluye una mayor sofisticación en el empaquetado y la delimitación de microinstrucciones, con objeto de aprovechar mejor el paralelismo de que se dispone en el estómago.

Esta mayor complejidad de la boca se traduce en el incremento de dos etapas más de segmentación para las fases de búsqueda y decodificación. A partir de ahí, los dos cauces de segmentación, entero y de punto flotante/multimedia, tienen una descomposición análoga a la del K7,

microarquitect.

boca  
estómago

compatibilidad

segmentación

◀ pág. 258  
superes-  
calaridad  
pág. 197 ▶

resultando cauces de 12 y 17 etapas de segmentación, respectivamente (ver [tabla 7.2](#)).

En el estómago (*back-end*), sigue vigente la entrada de un mínimo de tres y un máximo de seis microinstrucciones nativas (tres enteras y tres de punto flotante), otorgando para el factor de superescalaridad la misma media realista de cinco que ya concedimos al K7 (ver [sección 6.1.10](#)).

## 1.5 ▶ Memoria interna

banco de  
registros  
pág. 265 ▶

El banco de registros de propósito general amplía su anchura, desde los 32 bits de la generación anterior hasta los 64 bits de la actual (ver [figura 7.3](#)), algo lógico sabiendo que hemos migrado desde una arquitectura de 32 bits a otra de 64 bits.

memoria caché

La confianza en la arquitectura del K7 se traslada hasta el nivel de la memoria interna. Así, se respetan los dos niveles de memoria caché que ya se habían introducido en el modelo Thunderbird del K7, y que el Athlon XP se encargó de consolidar: Dos cachés de primer nivel separadas para datos e instrucciones, gemelas de 64 Kbytes, y una caché L2 conjunta de 256 Kbytes.

ampliación

Tan sólo se introduce como novedad la ampliación a 1 Mbyte de la L2 en algunas versiones de Athlon 64 y en todas las del modelo Opteron, algo natural por su clara vocación al segmento servidor en el que se tiene por costumbre engrandecer la jerarquía de caché.

## 1.6 ▶ Conjunto de instrucciones

x86-64

Un denominador común a todos los modelos de la arquitectura K8, y que a buen seguro lo será también para los futuros diseños que se vayan posicionando en esta generación, es el procesamiento de datos de 64 bits en sus unidades funcionales internas, lo que conlleva la adopción de un nuevo conjunto de instrucciones: El x86-64.

novedades

Llegamos a la clave para la explotación de los recursos hardware desde la capa software. Después de tres generaciones en las que las novedades del conjunto de instrucciones se reducían a los leves retoques introducidos por las instrucciones multimedia, se acomete un cambio en el formato de instrucción nativo del procesador, aunque eso sí, dejando la puerta abierta a la compatibilidad con las viejas aplicaciones x86.

### 1.6.1 Arquitectura de 64 bits

arquitectura  
de 64 bits

Antes de proseguir, conviene clarificar el concepto de **arquitectura de 64 bits**, pues vamos a utilizarlo profusamente en lo sucesivo.

para las  
compañías

- 1 Para las compañías fabricantes, representa una buena oportunidad para confundir a través de la publicidad. Así quisieron ponerle al Pentium la vitola de arquitectura de 64 bits. Pero tener un bus de datos de 64 bits no es suficiente, porque el talante de los 64 bits nos lo dan aspectos más internos del chip.

entre los  
usuarios

- 2 Entre los usuarios menos avezados, la creencia más extendida se asemeja a la siguiente definición: “un microprocesador que dispone de al menos un camino de datos o un subconjunto de registros de 64 bits”. Pero tampoco es una visión correcta, porque el camino de datos de punto flotante del Pentium es de 80 bits, y sus registros de punto flotante también, y sin embargo, se trata de una arquitectura de 32 bits.

en los libros

- 3 En libros o *journals* algo más técnicos, la referencia más común alude a “la cantidad de datos que un procesador puede manipular en una sola instrucción”, igualmente ambigua si tomamos una instrucción multimedia del conjunto SSE2 del Pentium 4, donde se procesan datos de 128 bits en una arquitectura de 32 bits.

- ④ En nuestra opinión, una arquitectura de 64 bits es *aquella que cuenta con un banco de registros de propósito general de 64 bits, dispone de unidades aritmético-lógicas que actúan sobre datos de tipo entero de 64 bits, y presenta la habilidad para calcular direccionamientos a memoria con punteros de 64 bits*. O dicho de forma más simple: Que los 64 bits constituyen su forma **natural** de procesar los operandos y las direcciones de un programa.

nuestra  
definición

Obsérvese cómo esta versión reducida de la definición puede extrapolarse al conjunto de instrucciones sin problemas. Y es que a veces la frontera entre una arquitectura y su conjunto de instrucciones es bastante difusa, en tanto en cuanto ambos van madurando de la mano a través de un proceso iterativo de sucesivas realimentaciones que tiene lugar durante la concepción de un nuevo microprocesador, según describimos en la [sección 3.5.2.1](#).

☛ pág. 98

### Ejemplo 7.1: EVOLUCIÓN DE UNA ARQUITECTURA DE N BITS

El concepto de arquitectura de N bits puede afianzarse repasando la trayectoria de Intel: Cómo extendió su microprocesador 8088 de 8 bits para engendrar sus primeras arquitecturas y conjuntos de instrucciones de 16 bits (IA-16), o cómo amplió estas últimas hasta llegar a la primera arquitectura de 32 bits (IA-32), el 80386, cuyo armazón ha sido reforzado durante 15 años de historia con sucesivos modelos que llegan casi hasta nuestros días.

El 8088 disponía de registros de propósito general de 8 bits (ej: A) que se extendieron a 16 bits (ej: AX, con submitades superior e inferior de 8 bits, AH y AL, que podían accederse de forma independiente para garantizar compatibilidad hacia atrás). A la llegada del 80386, estos registros pasaron a ser de 32 bits (ej: EAX, de nuevo con la submitad inferior accesible independientemente como AX). Evidentemente, las unidades de procesamiento de enteros permitían trabajar con 16 y 8 bits desde el 8086 en adelante, y con 32, 16 u 8 bits desde el 80386 hasta nuestros días.

Respecto al direccionamiento, atravesó una época un tanto confusa en los 8086 y 80286, porque estos modelos componían direcciones de 20 y 24 bits respectivamente a partir de una pareja de direcciones de 16 bits conocida como segmento y desplazamiento <sup>a</sup>. Pero con la llegada del 80386, esta idea quedó definitivamente abolida para dar paso a los punteros de 32 bits, con los que se direccionaban 4 Gbytes de memoria. Esta ingente cantidad, que aún hoy sigue quedándonos grande, es una de las razones que explican la prolongada longevidad de las arquitecturas de 32 bits. IA-64, el nuevo conjunto de instrucciones de 64 bits para el Itanium de Intel, aborda la extensión a 64 bits a través de direcciones de memoria de esta longitud, tal y como el 80386 hizo en su día con los 32 bits.

<sup>a</sup>Una de las chapuzas históricas de Intel, en este caso para poder abarcar un mapa de memoria de 1 Mbyte y 16 Mbytes respectivamente a partir de direcciones de 16 bits, y que fue la pesadilla de tantos programadores de ensamblador.

La [figura 7.2](#) muestra la existencia de direcciones de memoria de 64 bits tomando como referencia la arquitectura del procesador K8 de AMD. Respecto a la anchura de los datos, allí también resumimos la evolución seguida por los bancos de registros de propósito general para computación entera en las diferentes arquitecturas en función de su número de bits.

☛ pág. 262

## 1.6.2 x86-64 e IA-64: Dos aproximaciones diferentes

Según indicamos en la [sección 3.5.1](#), el diseño de un conjunto de instrucciones va inevitable-

☛ pág. 95

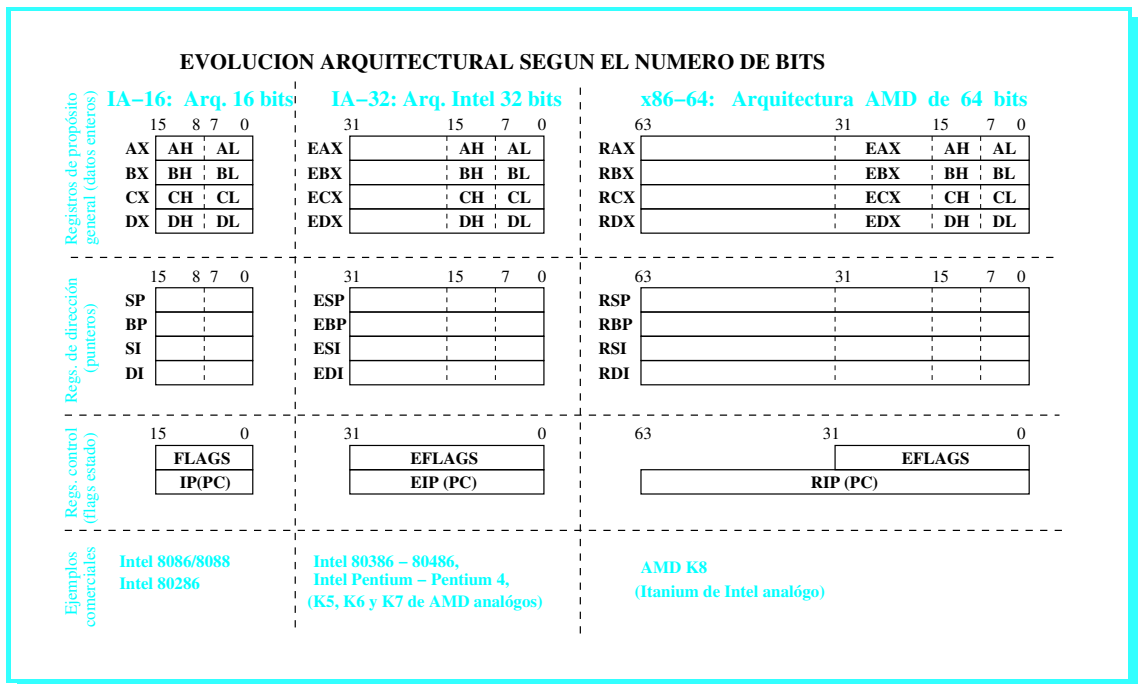


FIGURA 7.2: Evolución de los bancos de registros de propósito general, registros de dirección y registros de estado en las arquitecturas PC de 16, 32 y 64 bits.

mente ligado al diseño de la arquitectura de un microprocesador, y de todos sus elementos, el que más influye es su carácter: RISC, CISC, o esa tercera solución que puede ubicarse entre ambas: VLIW (ver [sección 3.5.3](#)).

☛ pág. 104

IA-64

Por otro lado, resulta inevitable hablar del nuevo conjunto de instrucciones de 64 bits de AMD, el x86-64, sin contrastarlo con su referente más cercano e importante, el IA-64 de Intel para su arquitectura Itanium.

El Itanium ha basado su diseño en un procesador de corte VLIW, mientras que AMD lo ha hecho en un CISC. ¿Por qué no se ha contado esta vez con los RISC? Ya dijimos en la [sección 3.5](#) que la historia ha sido pendular a este respecto, oscilando entre RISC y CISC por décadas:

☛ pág. 94

VLIW en Intel  
CISC en AMD

En los 70 abundaban los RISC, en los 80 primaron los CISC, en los 90 volvieron los RISC, y ahora parece vivirse un resurgir de los CISC: Según el Microprocessor Report, en el lustro 1995-2000, las arquitecturas CISC, representadas desde el Pentium de 133 MHz en 1995 hasta el K7 de 1 GHz en el 2000, han recortado un 50 % de ganancia a sus homólogas RISC para el conjunto de programas del benchmark SPEC. Y según AMD, las mejoras en rendimiento están cada vez menos relacionadas con el diseño de conjunto de instrucciones en sí, y más con su implementación. Así que más que plantear una revolución conceptual, AMD ha dirigido sus esfuerzos a los recursos hardware que ejecutan su nuevo conjunto de instrucciones.

rendimiento  
en IA-64

En el Itanium se ha considerado un conjunto de instrucciones completamente nuevo, anteponiendo de forma decidida el rendimiento a la compatibilidad. El procesador puede ejecutar programas escritos para arquitecturas IA-32 de 32 bits porque se le ha habilitado un modo de emulación para ello, pero la nueva arquitectura no ha cuidado lo más mínimo el tratamiento de este tipo de códigos: Se ejecutan como buenamente se pueda sobre una arquitectura en la que hasta el último transistor está pensado para acelerar la ejecución de aplicaciones de 64 bits nativas del conjunto IA-64.

El K8 contrasta con la visión anterior. Ejecuta código nativo x86 porque es compatible con las arquitecturas de 32 bits, pero a su vez incorpora novedades de 64 bits para ejecutar programas



Aspecto	x86-64 de AMD	IA-64 de Intel
Compatibilidad con código antiguo x86	Plena. Las aplicaciones de 16, 32 y 64 bits pueden ejecutarse concurrentemente	Limitada. Las aplicaciones de 32 bits corren bajo emulación y sin concurrencia
Rendimiento aplicaciones de 32 bits	Pleno	Penalizado
Rendimiento aplicaciones de 64 bits	Penalizado	Pleno
Portabilidad de aplcs. de 32 bits sobre 64 bits	Se ejecutan sin modificación alguna	Se debe reescribir la aplicación para obtener un rendimiento satisfactorio
Programación de nuevas aplicaciones	Más alejado de la circuitería, aunque permite utilizar el método y las herram. de programación antiguas	Más cercano a los recursos hardware existentes, aunque requiere entrenamiento previo
Diseño de la arquitectura subyacente	CISC	VLIW

TABLA 7.3: Comparativa entre los conjuntos de instrucciones x86-64 de AMD para K8 e IA-64 de Intel para Itanium.

de estas características. Por lo tanto, se trata de un paso bastante más conservador, quizá porque el enorme esfuerzo que supone migrar a un nuevo conjunto de instrucciones no pueda ser un objetivo realista para AMD después de ver cómo las ha pasado el gigante Intel.

compatibilidad en x86-64

Asegurando plena compatibilidad hacia atrás, el riesgo que corre AMD es bastante más bajo, ya que el nuevo producto se encuentra desde el principio arropado por el software antiguo, pero al contrario que Intel, lo que prima es lo viejo, y lo que pueda conseguirse de mejora con los retoques de 64 bits, bienvenido sea. Con todo esto, el Itanium es mejor plataforma si se quiere ejecutar código puro de 64 bits, pero mientras la capa software no se encuentre a la altura en este aspecto, la aproximación del K8 resulta una alternativa muy válida.

comparativa

Teniendo asumido que el K8 quedará por debajo del Itanium en rendimiento, la cuestión real es saber a cuánta distancia. Porque si es poca, la derrota será muy dulce para AMD: Tendrá abierta la puerta al succulento mercado de servidores y estaciones de trabajo sin haber realizado la ingente inversión de Intel, y esto seguramente le dotará también de un precio más competitivo que puede compensar su menor velocidad.

precio

En resumen, Intel aporta más cosas al conjunto de instrucciones del procesador con la llegada de los 64 bits, pero AMD puede obtener pingües beneficios habiendo asumido un riesgo muy inferior. La tabla 7.3 sintetiza nuestra comparativa entre ambas aproximaciones.

conclusión

### 1.6.3 La implementación del x86-64

La estrategia de AMD para extender el código de instrucción x86 a 64 bits es muy sencilla, y está basada en la misma idea que ya se utilizó en su día para extender las arquitecturas de 16 a 32 bits: Se habilita un nuevo modo para el procesador denominado modo extenso (*large mode*), que se activa mediante un bit global LMA (*Long Mode Active*) para habilitar los recursos con que cuenta el hardware para la computación de 64 bits. Dentro de este modo, hay dos submodos: El modo de 64 bits puro, y el modo de compatibilidad con las viejas aplicaciones bajo sistema operativo de 64 bits. Adicionalmente, se dispone de un tercer modo para conmutar a aplicaciones

modo extenso

submodos

Modo	Sistema Operativo requerido	Necesario recompilar aplicación	Usa HW nuevo	Anchura por defecto		
				Punteros (bits)	Operandos (bits)	Regs. prop. general
Extenso de 64 bits	Nuevos de 64 bits	Sí	Sí	64	32	64
Extenso compatible		No	No	32	32	32
Tradicional	De 32 bits	No	No	32	32	32

TABLA 7.4: Principales características de los tres modos de ejecución x86-64. La anchura por defecto puede cambiarse anteponiendo un prefijo a la correspondiente instrucción.

y sistemas operativos de 32 bits, esto es, tener la máquina trabajando bajo el paradigma actual.

#### ❑ Modo extenso de 64 bits

nuevos recursos

Este modo permite aprovechar todos los nuevos recursos de que dispone la arquitectura K8 para computación de 64 bits, y que se resumen en los siguientes elementos:

- Direcciones virtuales de 64 bits.
- Registros extendidos de propósito general:
  - De 64 bits. Los ocho tradicionales de propósito general para aritmética entera (AX, BX, ... ) se han ensanchado desde 32 a 64 bits, incorporándose además otros ocho nuevos (R8-R15).
  - De 128 bits. Los ocho SSE (XMM0-7) ya existentes para computación de punto flotante y multimedia 3DNow! Professional se amplían con otros ocho de nueva incorporación (XMM8-15).
- Un puntero de instrucción de 64 bits (el tradicional contador de programa o PC, que toma ahora el mortuario nombre de RIP), junto a los modos de direccionamiento relativos a éste para poder manejarse en espacios de direcciones superiores al Terabyte.

pág. 265

La [figura 7.3](#) muestra los nuevos registros de la arquitectura K8 extendidos sobre la del K7.

anchura por defecto prefijos operaciones parciales

En el modo extenso, se asumen por defecto punteros de 64 bits y operandos de 32 bits, pudiendo cualquiera de ellos cambiarse explícitamente anteponiendo un prefijo al código de instrucción. Cuando se computan operaciones de 32 bits y el resultado se almacena en un registro de propósito general de 64 bits, los bits superiores se completan con ceros, mientras que cuando se realizan operaciones de 8 o 16 bits, la parte superior del registro permanece intacta para preservar la semántica antigua ya definida para estas operaciones en los códigos de instrucción x86.

#### ❑ Modo extenso de compatibilidad con 32 bits

para la transición

Este modo permite a los sistemas operativos garantizar la compatibilidad binaria con las aplicaciones existentes de 16 y 32 bits, pudiendo éstas ser ejecutadas, sin necesidad de ser recompiladas, bajo un sistema operativo de 64 bits trabajando en modo extenso.

Eso sí, se deberá renunciar a las nuevas prestaciones, esto es, el espacio de direcciones virtuales deberá limitarse a 4 Gbytes, y los operandos pueden conmutar entre 16 y 32 bits de anchura, pero nunca alcanzar los 64 bits.

#### ❑ Modo tradicional de 32 bits

En este modo no está disponible ninguna prestación nueva del K8, y el código se ejecutará en

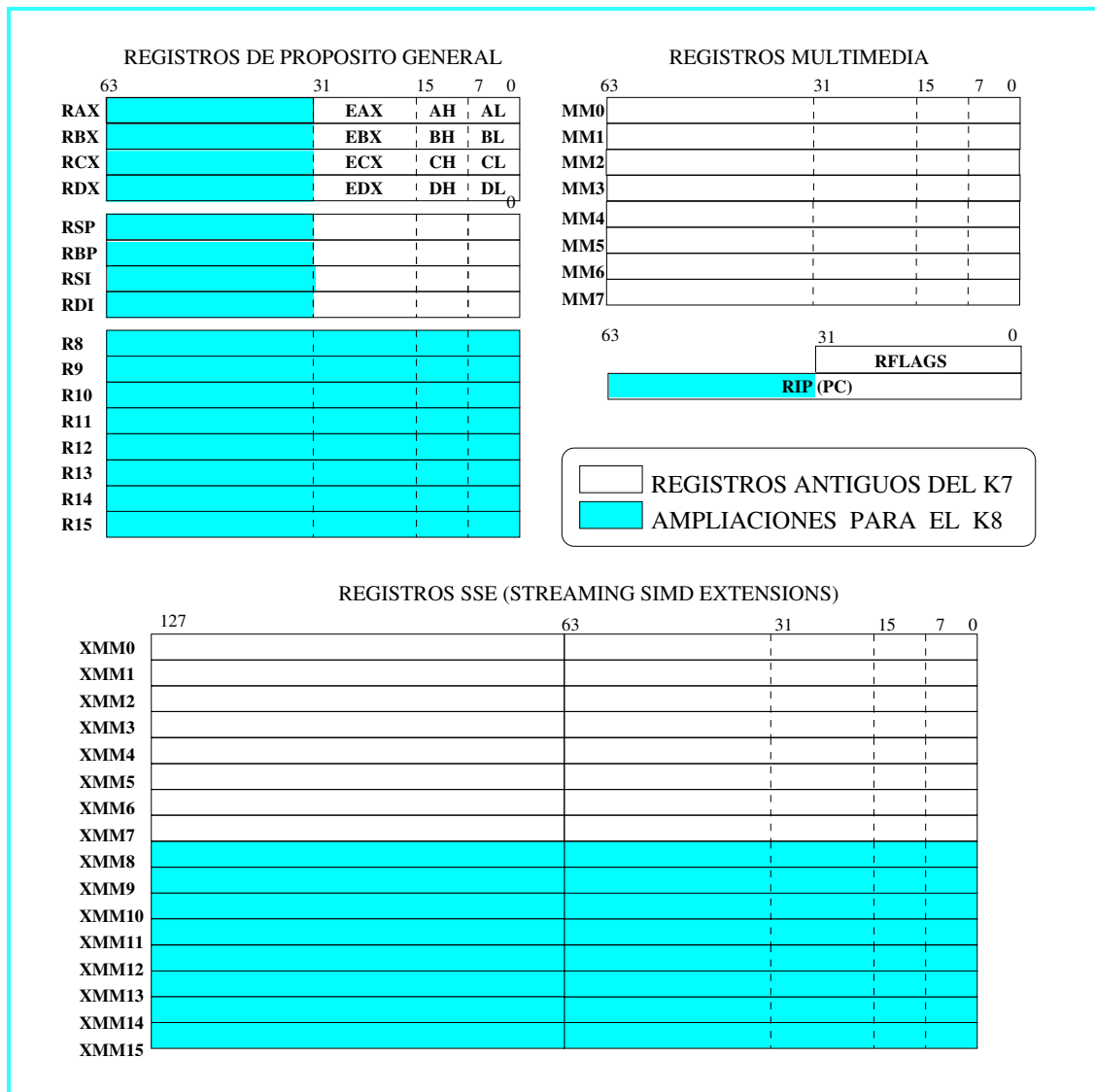


FIGURA 7.3: El banco de registros de la arquitectura K8 extendido sobre el existente en el K7.

todo momento como si de un procesador x86 estándar se tratase, garantizándose plena compatibilidad y coexistencia con cualquier aplicación que funcione en los modelos de microprocesadores de generaciones anteriores.

La tabla 7.4 resume las principales características de cada modo.

➔ pág. 264

### 1.6.4 Extensiones multimedia

La familia K8 al completo dispone de una extensa cobertura para instrucciones multimedia, incluyendo soporte para MMX, 3DNow!, 3DNow! Professional (el incluido en el Athlon XP y que combina el subconjunto Enhanced 3DNow! y las extensiones SSE del Pentium III), y la garantía de la compatibilidad con el conjunto de instrucciones SSE2 de Intel gracias a la incorporación de los nuevos registros XMM8-15.

completa  
cobertura

De esta manera, en el K8 termina de hacerse realidad el movimiento que comenzó con el Ath-

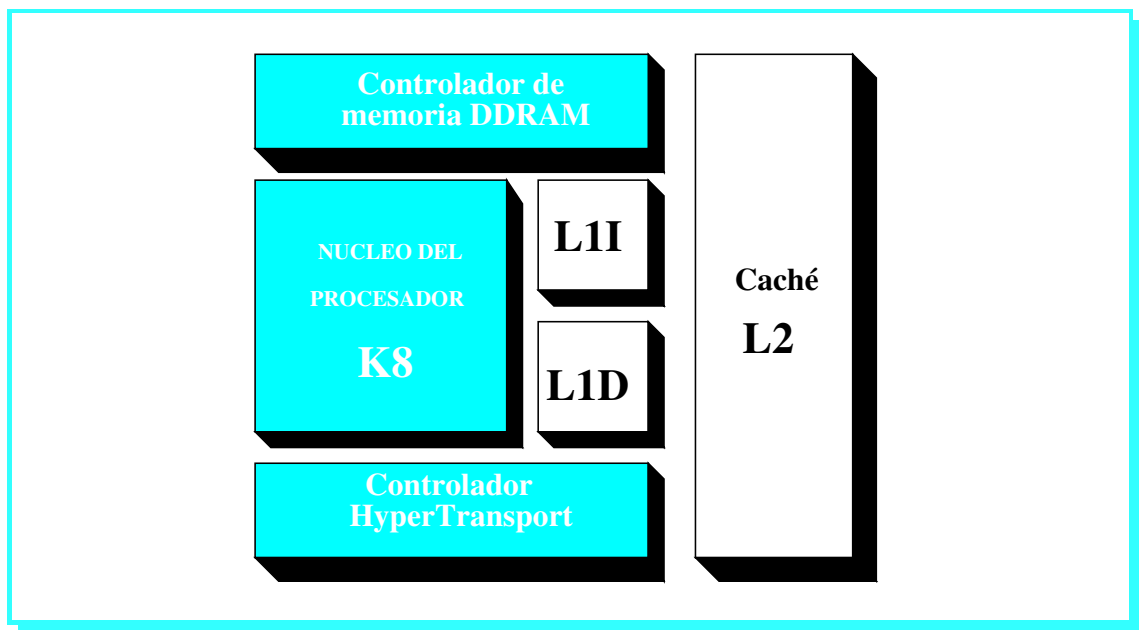


FIGURA 7.4: Los principales bloques funcionales de la arquitectura K8.

☛ pág. 119  
convergencia

lon XP y que ya vislumbrábamos en la [figura 3.20](#): Los universos SSE de Intel y 3DNow! de AMD, divergentes en su fase de gestación, convergen a un punto común en su fase de maduración.

## 1.7 ▶ El interfaz del procesador

Si el núcleo interno del procesador apenas presenta retoques, preparémonos para todo lo que llega en su frontera con el exterior, porque las novedades van a salpicar a todo el sistema: El bus local, el bus de memoria principal y la arquitectura del juego de chips de la placa base sufren una profunda metamorfosis, y habrá para ellos un antes y un después de la octava generación.

p.161/Vol.1.2

A lo largo de la [sección 13.2.3](#) desvelamos el panorama futuro de la jerarquía de memoria del PC que creemos más factible para los próximos años, donde apostamos por un aumento de la L2 antes que por la llegada de una L3, dando precisamente a la arquitectura K7 como ejemplo de dicha trayectoria (ver [ejemplo 13.2](#)).

p.162/Vol.1.2

La apuesta de AMD ahora en el K8 no se centra en el nivel de memoria que queda por delante de la L3 en la jerarquía, sino en el que queda por detrás: Memoria principal.

pág. 194  
jugarretas  
pasadas

Ya dijimos en la [sección 6.1.5](#) que el mercado jugó dos malas pasadas al K7 en sus inicios:

- ❶ No dotarlo de la logística necesaria en el juego de chips de la placa base para poder aprovechar el ancho de banda de 1.6 Gbytes/sg. que atesoraba su bus de 200 MHz en 1999.
- ❷ No acompañarlo de la memoria RDRAM, lista por entonces a 2x400 MHz (ancho de banda exacto de 1.6 Gbytes), por haber caído Rambus en el entramado financiero de Intel.

AMD trató entonces de paliar esta rémora lanzando sus propios juegos de chips de placas base para arropar al K7, singladura en la que no tuvo excesiva fortuna.

escarmiento

La situación que se produce ahora con el K8 es mimética: Una nueva arquitectura, transgresora del universo Intel, que necesita del concurso de ese otro mundo que conforma, al margen del poder establecido, VIA y otras compañías. La disyuntiva para AMD estaba en asumir el riesgo de volver a tropezar en la misma piedra o aprender del error. Con una baza ya consumida en este

sentido, la carta que se juega ahora con el K8 es bien distinta: Si nadie va a darme fuera lo que necesito, lo voy a fabricar yo, pero esta vez, en lugar de ofrecerlo a Asus o QDI para que no pongan el debido interés porque Intel es más seductora, lo voy a traer a mis dominios, integrándolo como una unidad funcional más del procesador, que para eso llegan los 90 nm. y tenemos licencia para embutir más cosas en el chip.

Así llegamos a una de las principales novedades del diagrama de bloques del procesador: La presencia de un controlador interno para memoria DDRAM (ver [figura 7.4](#)). De las muchas versiones ya existentes para la DDRAM (consultar [tabla 10.13](#)), se han incluido en la versión inicial del procesador las especificaciones PC-1600, PC-2100 y PC-2700 para módulos DIMM Unbuffered o Registered, conceptos que explicamos en la [sección 10.5.2.3](#).

Puesto que hay que sincronizar la velocidad de la memoria con la del bus y las tres especificaciones anteriores de memoria DDRAM funcionan a 2x100, 2x133 y 2x166 MHz, ya tenemos la frecuencia del bus que accede directamente a memoria. Estas frecuencias irán mejorando a medida que el consorcio DDR-II vaya lanzando las sucesivas versiones que tiene anunciadas en agenda (ver [sección 13.3.3](#)). El K8 se encontrará con este consorcio en sus versiones del procesador a 90 nanómetros, donde incluirá ya pleno soporte para DDR-II, inicialmente a 2x400 MHz según las intenciones ya hechas públicas por AMD.

controlador  
de DDRAM  
integrado

☛ [pág. 266](#)  
☛ [p. 72/Vol. 2](#)  
☛ [p. 25/Vol. 2](#)

frecuencia  
del bus

☛ [p. 165/Vol. 2](#)

### 1.7.1 Lo que queda del bus local

Cada generación se cobra su víctima en la arquitectura PC, y si el Pentium 4 con su caché de traza se llevó por delante el concepto de caché de instrucciones, con el K8 salta por los aires la noción de bus local tal y como lo conocemos desde la cuarta generación: Esa arteria que dialoga de forma compartida con todo lo que se encuentra fuera del procesador. Ahora esos diálogos tienen tres partidas diferentes (memoria, periféricos y otros procesadores), que se escinden por hasta cuatro arterias de comunicación disjuntas:

tres partidas,  
cuatro caminos

- ① La vía que parte del controlador de memoria para el acceso exclusivo a los zócalos DIMM que son jurisdicción de cada procesador.
- ② El bus independiente para el acceso al resto de zócalos y tarjetas ubicados en placa base.
- ③ El canal de comunicaciones que entra en juego en sistemas biprocesador para comunicar los dos procesadores vecinos.
- ④ Otro canal similar, habilitado sólo en sistemas dotados de cuatro procesadores, para comunicarse con el otro vecino que tiene cada procesador una vez éstos han sido dispuestos sobre una red con topología malla.

Los cuatro buses pueden observarse en el diagrama de bloques de la arquitectura global para un sistema dotado de cuatro procesadores, que adjuntamos en la [figura 7.5](#), donde ocupan los cuatro puntos cardinales de cada procesador, responsables en última instancia de su patillaje cercano al millar de pines.

☛ [pág. 269](#)

Los tres últimos buses enumerados se implementan bajo una especificación común desarrollada por el consorcio HyperTransport, y que pasamos a ver a continuación.


### 1.7.2 HyperTransport

En su fase inicial de gestación del K8, AMD denominó al bus local del procesador bajo el código de referencia LDT (*Lightning Data Transport*). Entonces la concepción de la arquitectura no era tan pretenciosa en su interfaz externo, teniéndose en mente una mera adaptación de la tecnología EV6 del procesador Alpha, que tan buen resultado había dado en los K7 monoprocesador.

LDT

no EV6 SMP

Sin embargo, cuando comenzaron a desarrollarse las primeras plataformas multiprocesador, EV6 dió numerosos problemas, y AMD desechó esta posibilidad, poniéndose a trabajar en una nueva especificación.

<p>orígenes</p> <p>p. 137/Vol. 2 ➔</p> <p>la anticipación de nVidia</p> <p>p. 144/Vol. 2 ➔</p>	<p>Así nació HyperTransport ( <a href="http://www.hypertransport.org">http://www.hypertransport.org</a>), un consorcio de especificación para buses de altas prestaciones del que también forman parte firmas tan conocidas en el mundo del PC como nVidia. Una vez más, el mercado de las tarjetas gráficas (ver <a href="#">capítulo 12</a>) se adelanta para enseñarnos lo que está por llegar: Las GeForce FX para 2003 basadas en el chip NV30 (ver <a href="#">tabla 12.3</a>) disponen ya de controlador de memoria integrado para memoria DDR-II, y la siguiente generación nForce basada en el chip con código de referencia Crush incorpora ya toda la tecnología HyperTransport.</p>
<p>etimología</p> <p>pág. 229 ➔</p> <p>equivocos consorcio</p>	<p>El nombre HyperTransport es otro bautismo que no nos gusta, ya que se parece demasiado al HyperThreading de Intel para su Pentium 4 Northwood (ver <a href="#">sección 6.5.3</a>). No obstante, el hecho de que la idea de HyperTransport provenga de un consorcio, mientras que la de Intel sea algo unilateral, nos hace sospechar que AMD fue primero esta vez. Sea como fuere, debe quedar claro que son dos nombres muy similares para dos ideas que no se asemejan un ápice conceptualmente.</p>
<p>objetivos</p>	<p>HyperTransport es un bus de interconexión punto a punto pensado para reemplazar a la extensa jerarquía de buses multinivel del PC, aunque está dirigida a todo tipo de plataformas, incluyendo servidores y sistemas empujados. El objetivo de la especificación a nivel software es mantener la compatibilidad con la especificación PCI, dado que ésta se encuentra consolidada como estándar en todas esas plataformas.</p>
<p>propiedades</p>	<p>Las propiedades del bus pueden resumirse en el empleo de voltaje reducido de 1.2 voltios, un multiplicador de 2x sobre la frecuencia nominal, un protocolo basado en paquetes de datos para eliminar numerosas líneas de control y comandos, y el soporte de diálogos asimétricos y de diferente anchura y velocidad.</p>
<p>ancho de banda</p> <p>p. 221/Vol. 2 ➔</p> <p>p. 243/Vol. 2 ➔</p>	<p>El ancho de banda del bus en sus primeras versiones es muy variado, ya que se consideran anchuras de 2, 4, 8, 16 y 32 bits, y frecuencias de 2x200, 2x400 y 2x800 MHz. La versión que utiliza el K8 en sus diálogos interprocesador es la de 16 bits y 2x800 MHz, para un total de 3.2 Gbytes/sg., aunque el bus permite la comunicación bidireccional simultáneamente (<i>full-dúplex</i>), con lo que realmente pueden circular 6.4 Gbytes/sg, la mitad de ellos en cada dirección.</p> <p>En el enlace del K8 con los zócalos de los periféricos, la versión a utilizar dependerá evidentemente de la que utilice el juego de chips de la placa base, puesto que debe sincronizarse con ella. Por ejemplo, si es una placa con bus AGP 3.0, que dispone de una anchura de 32 bits y una frecuencia de 8x66 MHz (ver <a href="#">tabla 16.1</a>), puede tener sentido incorporar la versión de 32 bits y 2x400 MHz de HyperTransport, que aunque proporciona el mismo ancho de banda que la que lleva el K8 para dialogar con sus homólogos, se ajusta mejor a la sincronización procedente de AGP. De forma similar, si lo que tenemos enfrente es un controlador SCSI (ver <a href="#">tabla 16.12</a>), se torna como más propicia la versión de 16 bits y 2x200 MHz por coincidir con la máxima anchura de SCSI.</p>

### 1.7.3 La nueva jerarquía de memoria

NUMA La arquitectura K8 utiliza el concepto de **NUMA (Non-Uniform Memory Architecture)**, una antigua idea ampliamente utilizada en sistemas multiprocesador, más devota en el pasado a arquitecturas masivamente paralelas, dotadas incluso de miles de procesadores.

El precedente más reciente en el que AMD se ha basado para llevar a cabo su implementación es el bus Sequent, cuya especificación data de 1996. De forma paralela, Silicon Graphics también desarrolló una variante muy parecida en su familia de multiprocesadores Origin bajo un paradigma de memoria compartida-distribuida, utilizando para ello un microprocesador propio: El R10000.

En esencia, el modelo NUMA utilizado en el K8 combina acertadamente la simplicidad del

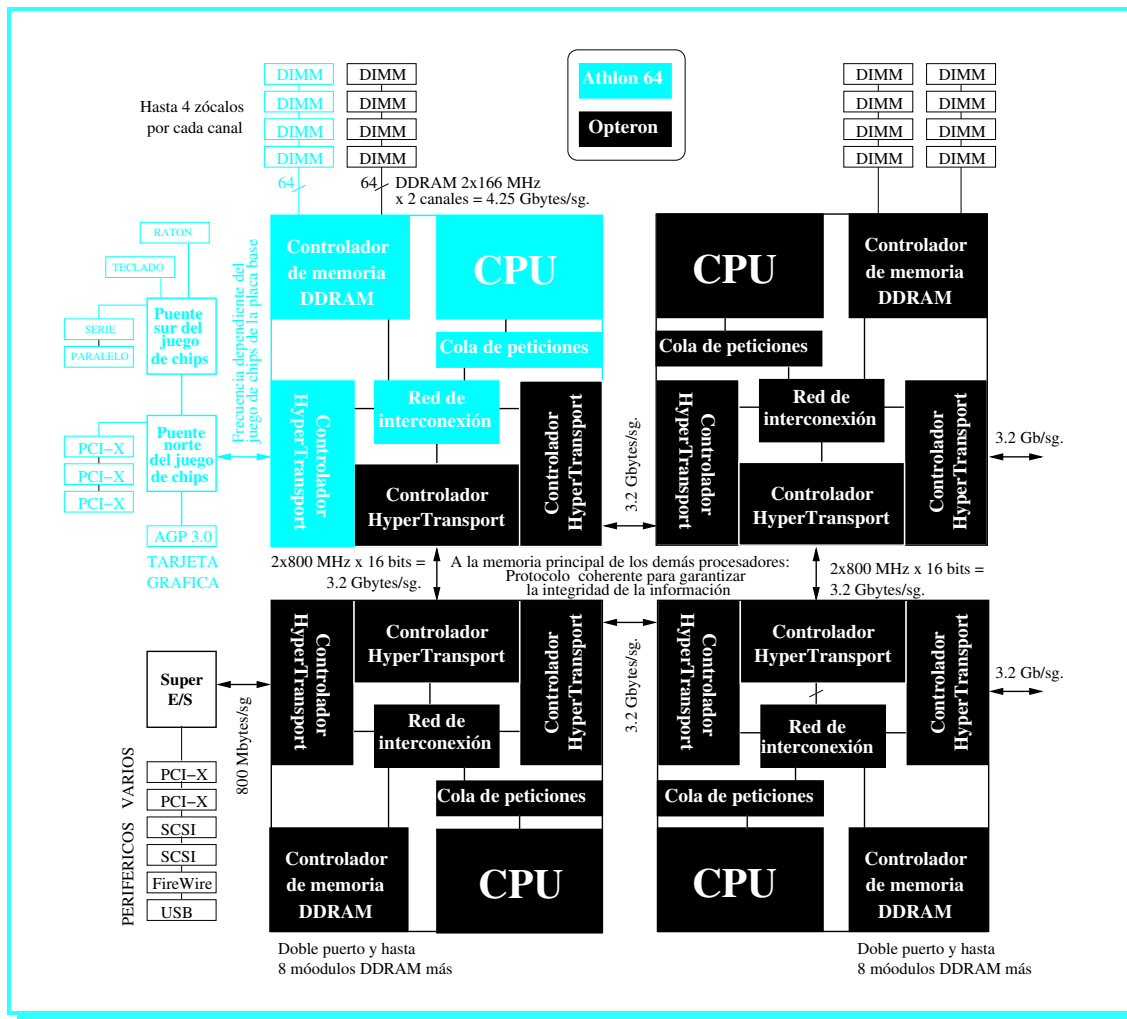


FIGURA 7.5: La arquitectura K8 en su variante más general: Un multiprocesador dotado de cuatro procesadores Opteron funcionando bajo multiprocesamiento simétrico. Se ha delimitado en un segundo color lo que corresponde a la arquitectura monoprocesador bajo Athlon 64.

multiprocesamiento simétrico SMP (*Symmetric Multiprocessing*) con la escalabilidad del multiprocesador masivamente paralelo. La memoria principal se estructura en dos niveles: Uno local a cada procesador, conformado por hasta ocho zócalos DIMM, y otro global, donde se sitúa la memoria perteneciente al resto de procesadores (en caso de existir), y al cual se accede mediante controladores HyperTransport adicionales al que ya existe en la arquitectura monoprocesador para acceder al conjunto de periféricos del sistema (AGP, PCI, etc). En este sentido, la implementación del K8 ha adoptado la versión 3.0 del bus AGP con multiplicador de hasta 8x (ver tabla 16.1) y la variante más sofisticada del bus PCI, PCI-X, con multiplicador de hasta 4x (ver tabla 15.2).

dos niveles

• p. 221/Vol. 2  
 • p. 207/Vol. 2

### 1.7.4 El doble puerto de acceso a memoria principal

Si desde el punto de vista arquitectural, la novedad del K8 es su controlador de memoria integrado, de cara al rendimiento resulta mucho más efectivo el hecho de haber dotado a éste de un doble puerto, puesto que esto multiplica por dos el ancho de banda de la DDRAM que conectemos al procesador.

doble puerto

No es que se esté apostando por una memoria que se entrelace nuevamente en anchura al

← p.48/Vol.2

nivel de módulo como ya lo hicieran en su día los módulos DIMM (ver [sección 10.11](#)), sino que ese entrelazado se implementa externamente, considerando 128 líneas de datos de entrada al procesador, lo que redunda una vez más en su numerosísimo patillaje.

La estrategia es buena, pero hay que decir que se la vimos antes a Intel con su juego de chips i850 para Pentium 4, en el que el doble puerto permitía una comunicación perfecta entre el bus de datos de 64 bits del procesador y la versión de 32 bits de la memoria RDRAM (ver [figura 6.11](#)). La forma de trabajar del K8 con DDRAM es exactamente igual a la mostrada en aquella Figura, con dos salvedades:

pág. 229  
copia de Intel

- ❶ En el caso del K8 no hay tramo de comunicaciones entre la CPU y el juego de chips de la placa base, sino directamente con memoria principal.
- ❷ Tanto el procesador como la memoria son el doble de anchos de lo que allí se muestra: 128 bits para el procesador y 64 bits para memoria principal.

Conviene no confundir este doble puerto del controlador de memoria DDRAM con el multiplicador 2x intrínseco a la memoria DDRAM. Ambos duplican ancho de banda, pero el primero lo hace doblando la anchura del bus, y el segundo, la frecuencia efectiva (recordemos que el ancho de banda es el producto de estas dos magnitudes). Es un buen momento para afianzar conceptos con un ejemplo.



### Ejemplo 7.2: EL COLOSAL ANCHO DE BANDA DEL K8 DE 90NM

En el plan de trabajo del consorcio DDR-II se dibuja un techo de frecuencia de 4x667 MHz, pero vamos a ser más conservadores. Tomaremos la frecuencia de 4x400 MHz como el valor más realista que sale al mercado a finales de 2003. Incluso puede retrasarse unos meses, porque la cita es a mediados de 2004, cuando según AMD, está prevista la fabricación del K8 a 90 nm, la versión que incluye el controlador de memoria DDR-II.

Juntando el ancho de datos de la memoria DDR-II (64 bits), su frecuencia de 400 MHz, su multiplicador interno 4x, y el desdoble 2x gracias al doble puerto del K8, el ancho de banda con memoria principal asciende a 25.6 Gbytes/sg., esto es, diez veces más de lo que el Athlon XP más avanzado (2x166MHz de bus sobre memoria DDR) proporciona a mediados de 2003.

Por si el ejemplo no es suficientemente contundente, un par de años más tarde debemos tener disponible (siempre según la agenda del consorcio DDR-II) la DDRAM de 4x667 MHz, con la que el ancho de banda asciende hasta 42.48 Gbytes/sg. Y recordemos que aún tiene que dar sus frutos la iniciativa DDR-III (ver [tabla 13.3](#)), por lo que a este ritmo acabaremos la década sobrepasando los 100 Gbytes/sg.

Claro que si extrapolamos a esa fecha la frecuencia del procesador de Enero de 2003 (3 GHz) siguiendo la Ley de Moore, nos salen 100 GHz de reloj, y volvemos a una situación comparativa muy similar a la actual: No tenemos mas que recordar que cuando el procesador llegó a 1 GHz, el ancho de banda de la memoria se encontraba en torno a 1 Gbyte/sg.

pág. 271  
ancho de banda

La [figura 7.6](#) resume todos los anchos de banda que tiene a su alcance el procesador K8 combinando sus diferentes versiones para DDR, DDR-II, DDR-III, monoprocesador y multiprocesador. No hay mas que echar un vistazo a los valores que se manejan para darse cuenta de que difícilmente el acceso a memoria será una rémora para el extraordinario potencial que presenta esta nueva arquitectura.



Frec. nominal (denominación)	Año previsto	Anchura del bus	Multip. del bus	Ancho de banda sobre Athlon 64 (Gbytes/sg)	Ancho de banda sobre Opteron de doble puerto	Ancho de banda sobre multiprocesador SMP dotado de cuatro Opteron
DDR	133 MHz (PC-2100)	2003	2x	2.12	4.25	17.00
	166 MHz (PC-2700)	2003	2x	2.66	5.32	21.31
DDR-II	400 MHz	2005	4x	12.80	25.60	102.40
	667 MHz	2006	4x	21.34	42.68	170.07
DDR-III	800 MHz	2008	4x	25.60	51.20	204.80
	1.5 GHz	2010	4x	48.00	96.00	384.00

FIGURA 7.6: Evolución prevista del ancho de banda en el controlador de memoria DDRAM integrado del K8. El valor más previsible para 2005 en el Athlon 64 es cinco veces superior al que disfruta el Athlon XP en 2003 (apenas 2.7 Gbytes/sg.).

## Las diferencias entre Athlon 64 y Opteron

◀ 1.8

Una vez expuestos todos los buses de la arquitectura K8, enumeramos como colofón las dos grandes diferencias entre Athlon 64 y Opteron, aunque pudieron ya extraerse gráficamente a partir de la [figura 7.5](#):

➡ [pág. 269](#)

- 1 El Athlon 64 lleva un único controlador de memoria DDRAM, que no puede sacar provecho del doble puerto de memoria, mientras que el Opteron sí lo lleva por duplicado.
- 2 El Athlon 64 lleva sólo un controlador HyperTransport, mientras que el Opteron lo incluye por triplicado.

La consecuencia más directa de todo esto es que el zócalo del Opteron tiene bastante más patillaje, y aunque ambos disponen de formato Socket, debemos tener claro que las placas base de uno y otro son completamente diferentes e incompatibles entre sí. La [tabla 7.5](#) sintetiza las principales características de uno y otro.

zócalo

➡ [pág. 272](#)



## Resumen



La octava generación de procesadores se encuentra en un estado incompleto. Ha llegado la arquitectura K8 de AMD, pero falta aún la de Intel.

Las principales novedades del K8 se sitúan en los estratos de más alto nivel del procesador, y es que tomando como referencia sus cinco magnitudes internas, los cambios se concentran en la parte superior de nuestra pirámide (ver [figura 3.1](#)):

➡ [pág. 49](#)

- 1 Sin cambios en la frecuencia del procesador, donde no veremos sino incrementos en línea con lo que ha venido acostumbrándonos el mercado. La frecuencia de bus de 2x800 MHz garantiza la utilización de multiplicadores en mínimos históricos.

frecuencia

Característica	Athlon 64	Opteron
Frecuencia estimada	1.8, 2, 2.2, 2.4, 2.6, 2.8 GHz	1.4, 1.6, 1.8, 2, 2.2, 2.4 GHz
Tecnología integr.	130 nm. y SOI	130 nm y SOI
Millones de transistores	67 (256 Kb. L2) >100 (1 Mb. L2)	>100
Area de integración (mm <sup>2</sup> )	104 (256 Kb. L2) 180 (1 Mb. L2)	180 en todas sus versiones
Zócalo	Socket de 754 pines	Socket de 940 pines
Segmentación	12 (enteros), 17 (punto flotante)	12 (enteros), 17 (p. f.)
Superescalaridad	5	5
Memoria caché L1	2 x 64 Kbytes	2 x 64 Kbytes
Memoria caché L2	256 Kbytes - 1 Mbyte	1 Mbytes
Memoria principal	1 x (DDRAM, DDR-II)	2 x (DDRAM, DDR-II)
Conjunto instrucciones	x86-64, compatible x86	x86-64, compatible x86
Buses HyperTransport	1 x (16 bits, 2x800 MHz)	3 x (16 bits, 2x800 MHz)
Multiprocesador SMP	1	1, 2, 4, 8

TABLA 7.5: Resumen de las características del Athlon 64 y el Opteron.

- tecnología** ② Al nivel de integración, contabilizamos sólo SOI como novedad. Más adelante, como ya es costumbre, bajaremos desde 130 nm., en esta ocasión hasta los 90 nm.
- paralelismo** ③ Al nivel del paralelismo a nivel de instrucción, un *front-end* o boca del procesador con el mismo sistema de decodificación en dos niveles que tenía el K7, y un *back-end* o estómago en el que tienen cabida el mismo número de unidades funcionales, con una segmentación y superescalaridad similares a las del K7.
- memoria** ④ En la memoria interna, dos episodios bien distintos: Por un lado, el banco de registros ampliado a 64 bits y la memoria caché conservando los mismos dos niveles del K7 Thunderbird y XP. Por otro lado, una memoria principal revolucionaria, con un controlador de memoria DDRAM integrado, y ese doble puerto de acceso interno que permite duplicar el ancho de banda de la memoria DDRAM que incorporaremos.
- conjunto de instrucciones** ⑤ En el conjunto de instrucciones, la novedad del x86-64, aunque eso sí, garantizando siempre la compatibilidad con aplicaciones de 32 bits anteriores.
- bus local** Las novedades introducidas en la memoria salpican al bus local, que ahora se escinde en dos carriles: El bus DDRAM para memoria principal y el bus HyperTransport para el resto de zócalos y tarjetas del sistema. Si el equipo es un multiprocesador SMP, este segundo carril se incluye por triplicado en la versión Opteron, donde las dos vías adicionales atienden la comunicación con la memoria de los demás procesadores, integrada en el sistema como un segundo nivel de memoria
- NUMA** principal bajo una arquitectura NUMA.



## La anécdota: Alpha 21464, o el primer procesador arácnido



En esa moda cada vez más establecida de bautizar el diseño de un microprocesador con nombres esotéricos en la etapa previa a la de su denominación comercial, algunas elecciones parecen inverosímiles al margen de los subjetivos gustos que cada cual pueda tener.

Aunque la moda fue iniciada por Intel y secundada por AMD, otros fabricantes tradicionalmente más austeros como Compaq terminaron sumándose a esta moda colorista.

Antaño, Compaq era bastante metódica en los códigos de referencia asignados a sus procesadores. Así, en la familia Alpha 21x64, se utilizaron las denominaciones EV4, EV5, EV6 y EV7 para sus microprocesadores 21064, 21164, 21264 y 21364, en clara alusión a los 4, 5, 6 y 7 niveles de pistas de metal utilizados en la integración de los respectivos diseños. Pero esta tendencia se rompió con el Alpha 21464, apodado con el código de referencia Araña.

EV#

Desde los tiempos de Cyrix con su viejo Jalapeño para denominar al sucesor del MediaGX ya bajo el auspicio de VIA (último trimestre de 1999) no veíamos a un procesador bautizado con nombre español, e incluyendo además la letra eñe en su denominación. Ambos vocablos tienen su origen en las preferencias de cierto personal mejicano de las plantillas de esas compañías. En concreto, lo de araña proviene de un símil de las ocho patas del arácnido con los ocho threads simultáneos que se pretendía que ejecutara el procesador, pues el rasgo SMT (Simultaneous Multi-Threading) era lo más novedoso de su arquitectura interna.

Jalapeño

El personal de Compaq trabaja en Estados Unidos en su mayoría, y puesto que allí no se dispone de eñe en el teclado, la plantilla dejó casi establecido el nombre de *Arana* para aquella creación. Imagino que desconocían que ese vocablo tiene un sórdido significado en castellano que todo buen amante de los crucigramas conoce: *Trampa, engaño*.

sin eñe

Tampoco sabían de lo premonitorio de aquel gazapo: El 21464 comenzó a gestarse antes incluso del año 2000, pero en 2001, Intel adquirió a golpe de talonario la división de arquitectura Alpha de Compaq con objeto de utilizar sus ideas para reforzar la línea de servidores abanderada por el Itanium, precisamente temiendo lo que podría suponerle la llegada del K8 que acabamos de analizar. Incluso más tarde, utilizaría esas mismas ideas en el segmento doméstico (véase HyperThreading - [sección 6.5.3](#)).

HyperThreading

[pág. 229](#)

Y del *Arana*, nunca más se supo ...

## Cuestionario de evaluación

En las cuestiones que presentan varias respuestas válidas, deberá quedarse con la que considere más exacta y/o completa. Las soluciones a todas las cuestiones se encuentran al final de este volumen.

1 En un procesador de 64 bits,

- a Los buses de datos y direcciones son de 64 bits.
- b El bus de datos es de 64 bits; el de direcciones no tiene por qué.
- c La forma normal de procesar datos y direcciones es de 64 bits.

d El bus de direcciones es de 64 bits; el de datos no tiene por qué.

2 Una arquitectura de 64 bits es aquella

a Sobre la que se monta un conjunto de instrucciones de 64 bits.

- b** Que dispone de al menos un camino de datos de 64 bits.
- c** Que dispone de al menos un banco de registros de 64 bits.
- d** Que dispone de un banco de registros de propósito general de 64 bits sobre el que actúan unidades aritmético-lógicas para procesar tanto datos como direcciones de esa misma longitud.

**3** 🗣️ ¿A qué arquitectura se parece más el Athlon 64?

- a** Al Athlon original.
- b** Al Athlon Thunderbird.
- c** Al Athlon XP.
- d** Al Opteron.

**4** 🗣️ ¿Qué tipo de memoria principal puede aprovechar la arquitectura K8 a pleno rendimiento?

- a** La memoria DDRAM de 128 bits.
- b** La memoria DDRAM de 64 bits.
- c** La memoria RDRAM de 64 bits.
- d** La memoria RDRAM de 32 bits.

**5** 🗣️ ¿Qué bus de una arquitectura K8 dispone de un multiplicador de reloj 2x?

- a** El bus de acceso a memoria principal.
- b** El bus de acceso al juego de chips en placa base.
- c** El bus de acceso a otros procesadores.
- d** Los tres anteriores.

**6** 🗣️ ¿Qué anchura tiene el bus de datos de una arquitectura K8?

- a** 16 bits.
- b** 32 bits.
- c** 64 bits.
- d** Escogería la opción **a** si se refiriera al tramo que conecta con otros procesadores, la opción **b** si se refiriera al tramo que conecta con la tarjeta gráfica AGP, y la opción **c** si se refiriera al tramo que conecta con memoria principal.

**7** 🗣️ Ponemos a ejecutar una aplicación A sobre un sistema operativo B y dos plataformas monoprocesador diferentes: una bajo Athlon 64 y otra bajo Opteron. El programa finaliza su ejecución antes en el Athlon 64. ¿Qué ha ocurrido?

- a** El Athlon 64 trabaja a mayor frecuencia que el Opteron.
- b** La aplicación es de 32 bits.
- c** El sistema operativo es de 32 bits.
- d** La aplicación que se lanzó en el Athlon 64 no se ha recompilado antes de lanzarla en el Opteron.

**8** 🗣️ ¿Qué procesador con arquitectura K8 es compatible con las aplicaciones software de 32 bits provenientes de un procesador de séptima generación?

- a** El Athlon 64.
- b** El Opteron.
- c** Los dos.
- d** Ninguno.

**9** 🗣️ ¿En qué aspecto se parece más la arquitectura del K8 a la del K7?

- a** En el rango de frecuencia inicial.
- b** En el paralelismo a nivel de instrucción (segmentación y superescalaridad).
- c** En la composición de sus niveles de memoria interna (banco de registros, L1 y L2).
- d** Apenas se asemeja en ninguna de las características anteriores, pues para eso es una nueva arquitectura de otra generación.

**10** 🗣️ ¿Cuál es la principal razón de que el patillaje del Opteron ascienda a casi mil pines (descontando que la mitad son para repartir la alimentación)?

- a** La enorme caché L2 de que dispone internamente.
- b** El doble puerto de acceso a memoria principal con controlador integrado en el chip.
- c** Los tres enlaces HyperTransport de que dispone.
- d** Las tres respuestas anteriores incrementan el patillaje, y de menos a más en el orden listado.

# El futuro del procesador

## Sumario

<b>8.1. Memoria histórica de la última década . . . . .</b>	<b>276</b>
8.1.1. Lección 1. 1993-94: Frecuencia y algo más . . . . .	276
8.1.2. Lección 2. 1995-96: Apuesta prematura por las arquitecturas de 32 bits . . . . .	276
8.1.3. Lección 3. 1997-98: El falso reclamo de las instrucciones multimedia . . . . .	276
8.1.4. Lección 4. 1999-2000: El hardware esquivo los condicionamientos software . . . . .	277
8.1.5. Lección 5. 2001-2002: Ansiedad por los 64 bits . . . . .	277
<b>8.2. Perspectivas para la próxima década . . . . .</b>	<b>277</b>
8.2.1. Frecuencia . . . . .	278
8.2.1.1. Interna . . . . .	278
8.2.1.2. Externa . . . . .	278
8.2.2. Tecnología de integración . . . . .	279
8.2.3. Paralelismo a nivel de instrucción . . . . .	280
8.2.3.1. Superescalaridad . . . . .	280
8.2.3.2. Segmentación . . . . .	281
8.2.3.3. Arquitectura multinúcleo . . . . .	281
8.2.3.4. Paralelismo en las comunicaciones . . . . .	281
8.2.4. Memoria caché . . . . .	282
8.2.5. Conjunto de instrucciones . . . . .	282
8.2.6. El microprocesador del año 2010 . . . . .	282
8.2.7. El microprocesador del año 2015 . . . . .	282
<b>8.3. El futuro de la computación . . . . .</b>	<b>283</b>
8.3.1. Aritmética entera . . . . .	283
8.3.2. Aritmética de punto flotante . . . . .	284
8.3.3. El supercomputador más potente . . . . .	285
<b>La anécdota: La ley de Moore . . . . .</b>	<b>286</b>

Algunos visionarios ya postularon la desaparición del PC a corto plazo. Larry Ellison, presidente de Oracle, fue uno de ellos. En el año 2000, ya espetó la conocida frase: “El PC es el único electrodoméstico cuyo manual de instrucciones es más grande que el propio aparato”. Esperemos, Larry, que la extensión de nuestra obra termine dándote la razón en breve, en cuyo caso, te haremos llegar gustosamente un ejemplar para tu regocijo personal.

Afirmaciones como ésta han estado siempre lejos de nuestra percepción, aunque al mismo tiempo debemos confesar que fueron tremendamente enriquecedoras. Por ello, nuestra aspiración es también la de aportar cosas nuevas. En primer lugar, vamos a tratar de coger algo de inercia del pasado, para posteriormente tratar de proyectarnos hacia el futuro.

## SECCIÓN 8.1

## Memoria histórica de la última década

El decenio 1993-2002 ha sido muy revelador en la arquitectura del PC, y permite extraer jugosas lecciones que poco nos dirán acerca de cómo debe ser el procesador del año 2010, pero mucho aportarán sobre cómo NO debería ser. Esta sección pretende así estrechar el cerco de posibilidades que se barajan en el camino desde el presente, enseñándonos las sendas por las que no hay que transitar para llegar a nuestra cita con el 2010.

## 1.1 ▶ Lección 1. 1993-94: Frecuencia y algo más

simplicidad  
irreal

El nacimiento del Pentium en 1993 trajo consigo un inusitado interés en la frecuencia del procesador, e instauró una costumbre. Era el parámetro estrella en el que todos se fijaban a la hora de comprar un equipo. El tiempo se ha encargado de demostrar que tal simplicidad roza lo grotesco. Por su extraordinaria complejidad, un procesador sólo puede estudiarse en su conjunto. De hecho, los análisis que nosotros hemos realizado en capítulos precedentes ya incluyen un cierto exceso de simplificación.

una sola  
magnitud

Sin embargo, en determinados círculos, como los adolescentes ensimismados en los juegos o los serios usuarios de aplicaciones software, se sigue pidiendo una mayor concreción. Se busca una magnitud de medida, una sola, que nos ponga a salvo de conocer medianamente la complejidad de un procesador. Misión imposible: El que quiera comprar mejor, tendrá que saber mejor lo que compra. Y al que quiera simplificar, más le vale ir con cuidado, puesto que el recorrido por esta década está jalonado de casos en los que las publicaciones periódicas y el acervo popular sólo han proporcionado el señuelo equivocado.

## 1.2 ▶ Lección 2. 1995-96: Apuesta prematura por las arquitecturas de 32 bits

paralelismo

El Pentium (1993) aceptaba dos instrucciones por ciclo de reloj, pero con severas restricciones. El Pentium Pro (1995) ya era capaz de manejar tres instrucciones simultáneas sin apenas restricciones, duplicando además el número de etapas de segmentación. Esto cuadruplica el rendimiento sobre el papel, siendo la clave para entender que su arquitectura haya permanecido vigente hasta nuestros días, rentabilizándose mediante secuelas como los Pentium II y III.

prematura  
optimización

Con tanta ventaja sobre un Pentium que arrasó en ventas, muchos se preguntarán cómo pudo el Pentium Pro pasar tan desapercibido al mercado. El culpable no fue otro que una prematura optimización para un software de 32 bits que no existía. A pesar de sus exiguos 200 MHz, el Pentium Pro volaba seis años más tarde cuando se enfrentaba al Windows NT y Windows 2000, los sistemas operativos de 32 bits. Pero ya era demasiado tarde, porque sus hermanos mayores lo habían retirado del mercado.

## 1.3 ▶ Lección 3. 1997-98: El falso reclamo de las instrucciones multimedia

falso reclamo

El fallo de previsión del Pentium Pro se pagó muy caro, pero sus secuelas, los Pentium II y III, colocaron un parche certero y justificaron así su éxito.

En las instrucciones multimedia que aparecieron más tarde, ni el embrión MMX justificaba su buena estrella, ni sus numerosas derivaciones (3DNow!, SSE, ...) merecieron mejor suerte.

infrauti-  
lización

A estos conjuntos de instrucciones le sacaríamos partido contando con un compilador, sistema operativo o programador que hiciera un buen uso de ellas. Había una mayor diversificación de

riesgos porque se instrumentaron numerosos agentes para aprovecharlas, pero ninguno de ellos quiso saber nada de ellas. Empezando por los programadores de juegos, como John Carmack (creador de los juegos Quake y Doom entre otros) que las desecharon en sus creaciones de la época al descubrir una pérdida de 200 ciclos máquina para conmutar al modo MMX.

programadores

#### Lección 4. 1999-2000: El hardware esquivo los condicionamientos software ◀ 1.4

El siguiente episodio ya no tuvo la osadía de confiarse a la capa software. En 1999, AMD lanzó una nueva arquitectura con rasgos en los que se tenía plena certeza que serían convenientemente aprovechados por los programas. El K7 se convirtió así en una hermosa lección: La frecuencia avanza, pero las optimizaciones sobre la jerarquía de memoria son más notables, y la simultaneidad en la ejecución de instrucciones alcanza el sobresaliente.

equilibrio

Las mejoras se dirigen a todas las instrucciones, sin distinción alguna. El resultado es un gran rendimiento, rematado con un precio muy competitivo. El mercado apadrina el producto a pesar del corporativismo que maniató a todo lo que no fuera Intel, y cae por fin su dictadura.

rendimiento a bajo coste

#### Lección 5. 2001-2002: Ansiedad por los 64 bits ◀ 1.5

El último episodio nos lo ofrece el número de bits de una arquitectura, con la transición desde los 32 hacia los 64 bits. Se vendió la falsa impresión de que el número de bits era sinónimo de potencia, cuando lo único cierto es que de nuevo vuelve a cometerse la pifia de dejar que el hardware vaya por delante del software en las cosas que depende manifiestamente de éste.

Una arquitectura tiene vías alternativas a la mejora en frecuencia para aumentar su rendimiento, pero esta vía no será los 64 bits hasta que el software no haya demostrado su capacidad para afrontar el reto. Las nuevas arquitecturas Itanium y K8 son el mejor ejemplo de que el hardware está disponible bastante antes de que el software pueda sacarle todo el potencial que esconde.

HW por delante

Esto no sería un problema muy grave si el hardware no fuera un producto tan perecedero: En no pocas ocasiones, la siguiente generación de productos hardware llega al mercado antes de que estén disponibles los sistemas operativos y las aplicaciones que permiten aprovechar todo el potencial de la generación presente.

perecedero

### SECCIÓN 8.2

## Perspectivas para la próxima década

A la hora de atisbar lo que nos deparará el futuro, nada mejor que asomarse a los informes emitidos por la ITRS (International Technology Roadmap for Semiconductors), un consorcio que agrupa a los mejores especialistas de los principales fabricantes de procesadores (Intel, AMD, Motorola, Sun, y un largo etcétera), cuyo objetivo básico consiste en proporcionar una guía de referencia sobre la evolución futura de las principales magnitudes del procesador.

ITRS

objetivo

El comité ITRS nació en 1992 bajo el nombre NTRS, entonces bajo un ámbito norteamericano (la N inicial proviene de National). Enseguida la iniciativa atrajo a empresas europeas y japonesas, y hoy en día es un conglomerado de expertos de la industria del semiconductor a escala mundial que ha adquirido una gran relevancia con la llegada del límite de la integración en silicio y los desafíos tecnológicos que esto plantea.

cosmopolita

ITRS realiza estimaciones a 15 años vista, por lo que su último informe disponible, emitido a finales de 2001, nos llevará de una forma científica y más o menos realista a lo que se espera del procesador del año 2016. La [tabla 8.1](#) nos muestra los datos más relevantes que hemos extraído de este estudio, y que pasamos a comentar.

previsiones a 15 años

▶ pág. 278

Evolución prevista para el microprocesador del futuro							
Magnitud del procesador	A medio plazo				A largo plazo		
	2004	2005	2006	2007	2010	2013	2016
<b>Variables eléctricas</b>							
Frecuencia (GHz)	4	5.17	5.63	6.73	11.50	19.35	28.75
Tensión de alimentación (voltios)	1.1	1.0	1.0	0.9	0.8	0.7	0.6
Potencia disipada (vatios)	85	92	98	104	120	138	158
<b>Tecnología de integración</b>							
Distancia de integración (nm.)	90	80	70	65	45	32	22
Anchura puerta del transistor (nm.)	54	45	40	35	25	18	13
Densidad de integración (Mt./cm <sup>2</sup> )	138	174	219	276	552	1104	2209
<b>Fabricación</b>							
Diámetro de la oblea (cm.)	30	30	30	30	30	45	45
Número de máscaras utilizadas	25	25	27	27	27	29	29
Máximo número de pines	1600	1760	1936	2140	2782	3616	4702
<b>Constitución interna</b>							
Millones de transistores (Mt.)	193	243	307	386	773	1546	3092
Area de integración (mm <sup>2</sup> )	140	140	140	140	140	140	140
Niveles metálicos de interconexión	8	9	9	9	10	10	10
<b>Funciones de coste</b>							
Coste/pin (máx, en céntimos €)	1.17	1.06	1.03	1.00	0.98	0.93	0.78
Coste (€/100 mill. transistores)	38	27	19	13.3	4.71	1.66	0.59

Fuente: International Technology Roadmap for Semiconductors. Edición 2001

TABLA 8.1: Estimaciones de ITRS para el microprocesador del futuro.

## 2.1 ► Frecuencia

### 2.1.1 Interna

Lo primero que llama la atención de las estimaciones de ITRS es el valor conservador otorgado a la **frecuencia** interna del procesador. La Ley de Moore se atenúa hasta reducir su velocidad a más de la mitad: si la progresión histórica ha sido duplicar cada 18 meses, el tiempo para esa mejora pasa a estar en torno a los cuatro años.

Reconocemos no dar mucha credibilidad a ITRS aquí, pues es la variable en la que históricamente ha cometido su mayor error estimativo. Por ejemplo, en la edición anterior del informe ITRS (1999), se daba al procesador una frecuencia de 1.1 GHz para 2005, y de 2.2 GHz para el 2014. Así que al igual que ya hicimos en aquella ocasión, apostamos a que los valores reales serán superiores a éstos. En nuestra contra tenemos esta vez un argumento más: Tanto Intel como AMD lanzaron por su cuenta previsiones anteriores al informe ITRS de 2001, en las que que hablaban de 10 GHz en 2010

### 2.1.2 Externa

Respecto a la frecuencia del bus local, ITRS le asigna el mismo valor que para el chip en su conjunto, pero con una particularidad: Distingue entre la elevada velocidad para un selecto y reducido grupo de pines, y la relativa baja frecuencia para la gran mayoría del patillaje que con-



forma el resto del bus. En el montante más numeroso estarán las partidas dedicadas a direccionamiento, control y alimentación, mientras que los datos se multiplexarán por un bus estrecho muy rápido, que responderá a razón de elevados multiplicadores internos sobre el reloj general, de forma similar a como muestra el cronograma de la [figura 10.35](#) para el caso del bus de memoria RDRAM.

selección de pines

p.81/Vol.2

A nivel eléctrico, las frecuencias tan elevadas que se apuntan se encuentran con la dificultad en la propagación de señales externas debido a dos causas: El incremento de la inductancia y capacitancia eléctricas, y la degradación de la señal: en los puntos de conexión entre el patillaje y su zócalo por un lado, y en los vínculos al interior del chip por el otro. Para mitigar todo esto, se apunta la posibilidad de prescindir del tradicional zócalo del procesador, fijando éste de forma directa sobre las pistas de la placa base.

fijación directa

## Tecnología de integración

### 2.2

La otra pifia de ITRS en su edición de 1999 no creemos que vuelva a repetirse ahora. Se estimó que la resolución litográfica del transistor del procesador, que siempre ha ido por detrás de la utilizada para la memoria DRAM, alcanzaría por fin a ésta en el 2015: Los resultados experimentales han demostrado que esto ocurrirá a mediados de 2004. A partir de ahí, la fabricación del procesador requerirá una maquinaria de litografía más sofisticada que la de las memorias DRAM.

litografía

Completando la terna de variables eléctricas, voltaje de alimentación y potencia disipada condicionan su evolución al que es el objetivo número uno del chip en la actualidad: Mitigar los excesos de calor que padece. La reducción de voltaje es una estrategia bastante saturada, ya que se encuentra atezada desde hace un tiempo por el ruido eléctrico, y de hecho, ha venido reduciéndose en sus tres últimos avances al 85 % del valor de referencia obtenido en el proceso de fabricación anterior, cuando teóricamente se le presupone un 70 %.

excesos térmicos

Puesto que el número de transistores del chip sigue creciendo de forma importante, no queda más remedio que distribuir la corriente dispersa entre un mayor número de pines. Esa es la principal razón de que estemos rondando los 5000 pines para el año 2016 en las estimaciones del consorcio ITRS.

patillaje

Aunque parezca superlativo, este patillaje representa una continuación de tendencia, puesto que si tomamos el presente de los casi 1000 pines del Opteron en 2003, hace trece años (1990), lo que teníamos era un 80486 con patillaje rondando los doscientos pines, cinco veces menos que el valor actual.

tendencia

La distribución de corriente también vuelve a apoyarse sobre el área del chip mucho antes que sobre sus niveles apilados en altura, ya que sus capas de metalización se cifran en diez, tres más que los siete con que cuenta ahora Intel para su Pentium 4, un incremento de similar magnitud respecto a los tres niveles de que disponía el 80486 de Intel en 1990.

capas de metalización

A pesar de todo, esos 158 vatios representan mucho calor, y para que sea posible su funcionamiento veremos todo un desfile de nuevas técnicas de disipación de calor y refrigeración que tendrán como difícil misión que el chip no sobrepase en su interior los noventa grados Celsius, la frontera aproximada donde se certifica su defunción.

disipación

Avisamos, una vez más, del papel protagonista que está llamado a tener la variable térmica a la hora de adquirir un chip. Piénsese que si calculamos el calor que genera el procesador con las propiedades eléctricas que ITRS apunta para 2016, nos sale un valor por encima de  $1000 \text{ W/cm}^2$ , unas 25 veces superior al máximo tolerado por el chip, que viene a ser de  $50 \text{ W/cm}^2$ . Por lo tanto, a partir de ya, cuando vayamos a comprar un procesador, deberemos interesarnos casi por igual en sus muchos gigahercios y en sus pocos vatios de potencia disipada. Y además, estaremos obligados a adquirir sistemas de disipación y refrigeración cada vez más sofisticados.

temperatura

El área de integración y su coste asociado en la fabricación de chips han sido tradicionalmente

las variables de comportamiento más predecible:

- |                     |  |
|---------------------|--|
| área de integración | <p>❶ El área de integración, porque ha permanecido prácticamente constante durante un largo período de tiempo, anclado en torno a los <math>140 \text{ mm}^2</math> en los microprocesadores para PC, y en <math>310 \text{ mm}^2</math> en el segmento de servidores, valores por cuya continuidad se apuesta de forma decidida.</p>  |
| coste               | <p>❷ Respecto al coste, su comportamiento es casi lineal, tan sólo con un leve salto en la depreciación que coincide con la puesta en marcha de la fabricación con obleas de 45 cm de diámetro (ya sabemos que a mayor oblea, menor coste por chip). En sucesivos procesos de fabricación cobra especial relevancia el coste asociado a la verificación del chip. Si en el Pentium hubo muchos más ingenieros dedicados al diseño que a la verificación del chip, en los diseños actuales el equipo de verificación está compuesto por centenares de personas. Esta labor adquiere tintes desafiantes en los 65 nm., donde las dimensiones de la anchura de puerta y sus posibles defectos se encuentran ya por debajo de la longitud de onda utilizada por la instrumentación de microscopía que se ha venido empleando hasta la fecha.</p> |

proceso de fabricación	<p>Lo más llamativo de todas las variables de integración lo hemos reservado para el final. La saturación de la tecnología de silicio comienza a notarse en su progreso, e ITRS apunta a que a partir de ahora no inauguraremos un nuevo proceso de fabricación (reducción de la distancia de integración) cada dos años, tal y como venía siendo costumbre en las dos últimas décadas, sino cada tres. Así, si ITRS había situado anteriormente las 0.35 micras en 1995, las 0.25 micras en 1997 y las 0.18 micras en 1999, desplaza las 0.13 micras (130 nm.) desde 2001 hasta 2002. A partir de ahí, deja la puerta abierta a cualquiera de las dos evoluciones, puesto que el siguiente proceso, 90 nm., lo sitúa en 2004, y el próximo, ya de sólo 65 nm., en 2007 (recordar que cada valor representa el 70 % de la distancia anterior).</p>
saturación	

Intel	<p>Probablemente ITRS haya decidido no discutir unilateralmente con Intel, uno de sus miembros más influyentes, quien sigue apostando por mantener el ritmo de innovación tecnológica de dos años en sus plantas de fabricación, según indicamos en la <a href="#">tabla 3.1</a>. Lo cierto es que sus modelos Tualatin y Northwood cumplieron ambos la cita de 0.13 micras en 2001, y al menos habrá que concederle el beneficio de la duda en lo que resta de año para presenciar la llegada de las 0.09 micras.</p>
pág. 52	

## 2.3 ► Paralelismo a nivel de instrucción

### 2.3.1 Superescalaridad

desdoble	<p>Las estrategias de superescalaridad que existen en un microprocesador actual parecen haberse acomodado al conjunto de instrucciones. Se impone así un grado de superescalaridad que aprovecha el desdoble que produce la conversión de código 80x86 a código nativo de cada arquitectura en la fase de decodificación, para alimentar simultáneamente a un número de unidades funcionales en la fase de ejecución que oscila entre tres (Pentium 4) y cinco (K7 y K8).</p>
----------	---

La tendencia de habilitar un código de instrucción del tipo VLIW (*Very Long Instruction Word*), en el que el compilador ha detectado las ligaduras entre las instrucciones agrupándolas en paquetes de ejecución lo más independientes posibles (y donde los posibles conflictos han sido ya analizados e incorporados al macrocódigo), parece que no termina de cuajar al nivel del PC por el miedo escénico a sacrificar la compatibilidad con las viejas aplicaciones.

Entretanto esto no ocurra, el paralelismo a nivel de instrucción de un microprocesador para PC descansará más sobre la segmentación que sobre la superescalaridad.

### 2.3.2 Segmentación

Estamos viendo la llegada de cauces de segmentación cada vez más profundos, pero nuestra impresión al respecto es que volveremos sobre nuestros pasos, como ya ha ocurrido repetidas veces entre las filosofías RISC y CISC.

Un cauce de segmentación cercano a las veinte etapas sólo puede sustentarse sobre las bases de un programa excesivamente benévolo. A estas alturas, no vamos a pedirle a los programadores que se fijen en los detalles de la arquitectura hardware para que saquen al software del que es su estado natural y nos lo disfracen del ente que a nosotros nos gustaría que fuese.

cauce extenso

Tampoco esperamos que el compilador nos saque las castañas del fuego reconvirtiendo un código *espagueti* en otro limpiamente secuencial: Acumulamos sangrantes experiencias a este respecto, y va siendo hora de afrontar que esta tarea es un poco indigesta.

vías saturadas

Por último, tampoco esperamos que los predictores de salto, la BTB y las estrategias de reordenación de instrucciones den más de sí de lo que han demostrado ya. Todo eso es tierra demasiado trillada; ha habido tanta gente trabajándola de forma intensiva durante tanto tiempo, que todo el petróleo que había se ha debido encontrar ya.

BTB

### 2.3.3 Arquitectura multinúcleo

Así las cosas, las novedades al nivel de la arquitectura del chip han llegado explotando el paralelismo a un nivel superior al de instrucción: Los threads o hebras del sistema operativo. La ventaja de subir un nivel es que se construye sobre lo que ya tenemos, es decir, disponemos de una nueva forma de paralelismo que se suma a las anteriores y es compatible con ellas.

suma y sigue

Para lograrlo, hay que replicar la estructura segmentada con que se cuenta, obteniendo un procesador con diversos núcleos de ejecución. Tanto la idea del Chip MultiProcessing (CMP) de Sun un par de años atrás como la más reciente y popular HyperThreading de Intel apuntan en esta dirección, explotando paralelismo a nivel de thread (SMT - Simultaneous MultiThreading).

CMP

HyperThreading

Aunque en su fase de rodaje inicial estas ideas se han puesto en marcha con valores conservadores de tan sólo dos núcleos replicados, no creemos que la cosa se vaya a quedar ahí en vista del ingente patrimonio de transistores que ITRS divisa en el horizonte del procesador para dentro de una década.

### 2.3.4 Paralelismo en las comunicaciones

Si la idea anterior replica el interior del chip, ésta le complementa replicando sus vías de comunicación externas. Y si lo anterior llevaba el sello de la casa Intel, ésto lo lleva de AMD, porque ha sido el nuevo K8 el que ha dado el paso firme hacia adelante en esta dirección.

replicar vías

La idea de sustituir un bus local común por enlaces punto a punto específicos con cada uno de los elementos con que el procesador quiera conectar se enmarca en la tendencia actual de hacer vías más rápidas de comunicación limitando su anchura (Fire Wire es un ejemplo actual muy ilustrativo - ver [sección 16.3.2](#)). Si además se dispone de área de integración para embutir en el interior del chip procesador los controladores para todas estas comunicaciones, tendremos una doble mejora en el rendimiento: Menos diálogos y más concurrencia.

☛ p. 255/Vol. 2  
coms. serie

Pero al contrario que en el caso anterior, esta vez sí creemos que desde un primer momento se apuesta fuerte por la idea, ya que el diseño del K8 contempla hasta cuatro arterias de comunicación con controlador integrado, destacando una conexión de doble puerto e increíblemente veloz con los chips de memoria principal DRAM.

apuesta fuerte

## 2.4 ► Memoria caché

p.157/Vol.2 ➔

p.161/Vol.2 ➔

p.166/Vol.2 ➔

más grandes

No L3

El futuro de la caché lo dibujamos en el que es su sitio natural dentro del volumen II, donde dedicamos el [capítulo 13](#) a dibujar el futuro de la memoria del PC en general (ver [sección 13.2.3](#)) y de la caché en particular (ver [sección 13.4](#)).

No obstante, podemos apuntar brevemente tres caminos básicos: La L1 y L2, que continuarán aprovechando las mejoras de integración para crecer en tamaño (1 Mbyte es el siguiente paso, en este caso, presente en varias configuraciones de Opteron para 2003). La L3, que representa el camino cerrado que cada vez tiene menos posibilidades de abrirse. Y memoria principal, la nueva apertura que ha usurpado a las tres cachés el puerto de datos más rápido del procesador para acercar los chips de memoria principal a sus confines.

## 2.5 ► Conjunto de instrucciones

80x86

Teníamos muchas esperanzas puestas en la octava generación para que saldase una cuenta pendiente en esta faceta del procesador: La conversión de código 80x86 a código nativo de una arquitectura, tarea que le sustrae en torno al 20% de su rendimiento. Sin embargo, la iniciativa más agresiva en este sentido (Itanium), se ha instalado fuera del mercado PC, y la que abarca su jurisdicción dentro de él (K8), se ha limitado a seguir en la línea de siempre.

multimedia

Entre lo lejos que queda la próxima oportunidad y el pesimismo que nos invade, creemos que el futuro difícilmente cambiará esta forma de hacer las cosas, así que aquí no vislumbramos a medio plazo otra cosa que no sean aditivos encasillados en el mundo de las aplicaciones multimedia.

## 2.6 ► El microprocesador del año 2010

rendimiento

Si las previsiones de ITRS son correctas, en el año 2010 tendremos a nuestra disposición un procesador de 11.5 GHz y 770 millones de transistores. Tomando como referencia actual el Pentium 4 Northwood de 3 GHz y 55 millones de transistores, nos sale un procesador que proporciona un rendimiento en torno a sesenta veces superior por un coste total de unos 60 €, 35 € para pagar el área de silicio y 25 € destinados a financiar el patillaje.

refrigeración

Considerando que el chip disipará en torno a 120 W, nuestra percepción es que puede costarnos más el collar que el perro, es decir, terminaremos invirtiendo más dinero en el sistema de refrigeración del PC que en sus chips constituyentes.

## 2.7 ► El microprocesador del año 2015

modularidad

Unos años más adelante, en el 2015-2016, tendremos a la frecuencia flirteando con un reloj de 30 GHz para sincronizar 3000 millones de transistores que se comunicarán con el exterior por un patillaje cercano a los 5000 pines. No imaginamos un procesador de semejante frecuencia y densidad de patillaje sin estar soldado directamente a la placa base, así que aquí las mejoras del rendimiento nos obligarán a hipotecar la modularidad a que el PC nos tiene acostumbrados, regresando paulatinamente a sus orígenes de máquina compacta.

coste

A pesar de su sofisticación, el área de integración del procesador costará unos 18 €, mientras que sólo en patillaje el coste total ascenderá a 36 €. Es decir, pagaremos el doble de dinero por 5000 pines que por 3000 millones de transistores.

memoria

Pero la cifra más grande para nuestro PC del 2016 no se la llevan los transistores. Según recogemos en nuestro plan de desarrollo estimado para las futuras nanotecnologías de memoria (ver [tabla 13.10](#)), para ese mismo año se espera que esté disponible la memoria por sonda nanométrica

p.180/Vol.2 ➔

Año	Microprocesador	Ciclos de ejecución	Frecuencia de reloj inaugural	Max. ejecución simultánea de instrucciones	MIPS
1979	Intel 8088	2	4.77 MHz	1	0.2
1989	Intel 80486	1	25 MHz	5	20
1999	AMD K7	1	500 MHz	72	2.000
2009	Previsiones de Intel	1	10 GHz	?	100.000

TABLA 8.2: Breve evolución de la potencia bruta de procesamiento de instrucciones aritméticas de números enteros con el procesador más significativo del mercado de PC al inaugurarse cada una de las tres últimas décadas. Los números de la tabla delatan un ritmo de crecimiento de dos órdenes de magnitud por década en la métrica de rendimiento MIPS (Millones de Instrucciones Por Segundo), que prevé mantenerse en el futuro de cumplirse las previsiones de los fabricantes.

o PRAM (ver [sección 13.7.5](#)), y con ella, la capacidad de almacenamiento de nuestro PC superará el listón del Terabyte ( $2^{40}$  bytes), o lo que es lo mismo, más de un billón de bytes.

◀ p. 177/Vol. 2

SECCIÓN 8.3

## El futuro de la computación

### Aritmética entera

◀ 3.1

Los cálculos de tipo entero han sido tradicionalmente los más veloces en términos del número de ciclos consumidos, frecuencia de reloj y número de operaciones simultáneas realizadas.

Los primeros 8088 ya realizaban una operación de éstas por cada dos ciclos de reloj de 4.77 MHz en 1979. Diez años más tarde (1989), la frecuencia rondaba ya los 50 MHz. Y otros diez años después, el ritmo evolutivo nos llevaba a frecuencias de 1 GHz mientras se trabajaba afanosamente en el grado de superescalaridad: Un microprocesador puntero de 1999 como el K7, era capaz de procesar 72 instrucciones enteras de forma simultánea. La [tabla 8.2](#) sintetiza todas estas cifras.

ritmo evolutivo

Para el 2009, recogemos las previsiones de Intel, quien a finales de 2000 indicó que para entonces estaría disponible un microprocesador a 10 GHz trabajando a menos de un voltio y superando con creces los cuatrocientos millones de transistores. Según estas estimaciones, este procesador desarrollaría una potencia de computación entera de 20.000 millones de cálculos por segundo y 100.000 MIPS (*Million of Instructions Per Second*). Como a 10 GHz tenemos 10.000 millones de ciclos de reloj por segundo, las cuentas que nos salen son que el paralelismo a nivel de instrucción de ese año conseguirá ejecutar diez instrucciones por cada ciclo de reloj de forma sostenida.

estimaciones

Esto pone de manifiesto la futura saturación del paralelismo a nivel de instrucción, ya que la arquitectura del PC tendría que llegar a ejecutar hasta veinte instrucciones por ciclo para intersectar de forma exacta a la línea de tendencia marcada en los últimos treinta años, donde la computación entera ha seguido un ritmo de progreso de dos órdenes de magnitud (factor 100) por cada década transcurrida (ver la última columna de la [tabla 8.2](#)).

saturación del paralelismo

El dato muestra una vez más los desequilibrios de la computación en favor de la frecuencia, buscando antes el aspecto cuantitativo de hacer las cosas en el período de reloj más corto que el aspecto cualitativo de realmente aprovechar al máximo cada uno de esos ciclos.

desequilibrios

La conclusión, desde otra perspectiva, es que la arquitectura del PC confía mucho más su suerte a la ingeniería microelectrónica que a la ingeniería del software.

conclusión

Década	Logros conseguidos por la FPU
Años 60	Circuitería para operaciones de punto flotante en simple y doble precisión para dar soporte a las aplicaciones de tipo científico.
Años 70	Migración de esta circuitería a los minicomputadores ampliando su rango de disponibilidad a todo tipo de disciplinas científicas
Años 80	Estandarización de la circuitería empleada en la Unidad de Punto Flotante e incorporación de la misma a la arquitectura del computador personal
Años 90	Ampliación de las unidades de ejecución del procesador al mundo multimedia siguiendo un modelo SIMD y reutilizando el banco de registros de la FPU

TABLA 8.3: Evolución histórica de las unidades de procesamiento de punto flotante (FPU) desde su nacimiento.

### 3.2 ► Aritmética de punto flotante

**expectativas** Se esperaba algo más de este tipo de computación, pues la forma en que se procesan los datos de punto flotante lleva asociado inherentemente un elevado grado de paralelismo a nivel de instrucción. Sobre el papel, esto era suficiente para apostar por un ritmo de desarrollo para la computación de punto flotante similar al demostrado por su homóloga entera.

**logros** Sin embargo, la experiencia demuestra que no ha sido así. Si echamos un vistazo al diagrama de bloques de un microprocesador actual, descubriremos múltiples ALU de cierta riqueza computacional, mientras la FPU (*Floating-Point Unit*), en contraste, continúa su cabalgada en solitario. La [tabla 8.3](#) resume sus principales logros e innovaciones a lo largo y ancho de las cuatro últimas décadas.

Si nos circunscribimos al segmento doméstico más modesto, la FPU ha sido soberanamente maltratada por los fabricantes:

**pila** ■ Primero, con esa enrevesada estrategia de direccionar a los operandos de punto flotante a través de una pila, rémora histórica de los diseños de Intel que posteriormente AMD también adoptaría.

**registros** ■ Segundo, con el aprovechamiento del banco de registros de punto flotante, generoso en anchura (80 bits), para albergar allí los formatos de instrucción de 64 bits con múltiples operandos que caracterizan a la computación multimedia o MMX. Otra tropelía que ambas firmas enquistarían en todos sus procesadores con la extensión MMX o superior.

Esperábamos que semejantes agravios fueran corregidos con presteza, pero lo cierto es que no se subsanaron hasta la llegada del Pentium 4 y el K7, y al mantenerse durante tanto tiempo entre nosotros, el daño ya estaba hecho.

**ocio y juegos** La realidad es que un equipo doméstico realiza una parte muy elevada de su trabajo sirviéndose de la computación entera y multimedia. Y para esto, son las ALU y las unidades MMX las que deben estar optimizadas. El que quiere algo más, como potencia en la ejecución de las complejas operaciones de *blending* (*mezclado*) y *renderizado* (conversión del espacio 3D al 2D de nuestro monitor) que abandera la industria de los juegos para PC, se compra una aceleradora gráfica o en su defecto una buena tarjeta gráfica y el problema está resuelto poniendo sobre la mesa apenas 200 euros de más. Para eso el PC es un sistema modular.

Potencia de cálculo	Monoprocesador de gama baja	Multiprocesador más potente	
		Equipo	Microprocesador
KiloFLOPS (10 <sup>3</sup> FLOPS)	8088/80286 1980-1985	Connection Machine Años 80	Sparc (Sun)
MegaFLOPS (10 <sup>6</sup> FLOPS)	80386s, 80486s y Pentium (1985-2000)	Intel Paragon	i860 (Intel)
		Cray T3D/T3E	Alpha (Digital)
		IBM SP-2	PowerPC (Motorola)
		Origin 2000	R10000 (SGI)
GigaFLOPS (10 <sup>9</sup> FLOPS)	Servidores IA-64 (2001 - Itanium)	IBM Deep Blue Dic'96 - 1 GFLOPS	PowerPC (Motorola)
TeraFLOPS (10 <sup>12</sup> FLOPS)	Aún por llegar	IBM ASCI White Jun'00 - 7 TFLOPS	PowerPC G3 (Motorola)
		Earth Simulator Mar'02 - 35 TFLOPS	SX (NEC)

TABLA 8.4: Evolución de la potencia de cálculo para números reales en los microprocesadores para PC y comparativa con sus homólogos, los supercomputadores más potentes del mercado de su misma época (en FLOPS u operaciones de punto flotante por segundo). En el mercado doméstico, situamos la llegada de la computación Megaflop con la aparición de las primeras FPU integradas en el chip microprocesador, y la llegada de la computación Gigaflop, con los primeros equipos bajo arquitectura IA-64 (Itanium) y frecuencias superiores a 1 GHz. El ritmo evolutivo a que progresa la potencia de cálculo en el mercado doméstico es en este caso de tres órdenes de magnitud por cada quince-veinte años transcurridos.

## El supercomputador más potente

### 3.3

Si en lugar de observar la evolución de la computación en punto flotante desde el peldaño más bajo nos encaramamos al más alto, divisaremos desde allí a los supercomputadores más rápidos del momento, y cambiará tanto nuestra percepción del entorno como el horizonte que oteamos.

supercom-  
putadores

Las aplicaciones que requieren supercomputación son eminentemente científicas, y se caracterizan por tener una cuantiosa componente de procesamiento de números reales. No es de extrañar, pues, que su ritmo evolutivo haya sido muy superior al que hemos descrito para el PC, tal y como hemos querido reflejar en la [tabla 8.4](#), donde se comparan ambas trayectorias.

ritmo evolutivo  
de la supercom-  
putación

Allí podemos comprobar que en el instante presente nos encontramos en plena era del Teraflop, o de los supercomputadores con una potencia de cálculo de 10<sup>12</sup> FLOPS (operaciones de punto flotante por segundo). Para conocer cuánto nos queda para inaugurar la era del GigaFLOP, lo primero que debemos saber es qué capacidad de cálculo atesora el computador más potente del mundo en la actualidad.

era del  
Teraflop

En 2001, el techo de la computación lo ostentaba el ASCI White de IBM instalado en el Lawrence Livermore National Laboratory y utilizado por el Ministerio de Defensa Norteamericano. Con un total de 8192 procesadores PowerPC, la potencia bruta del ASCI White asciende a 5 TeraFLOPS.

ASCI White

En 2002, ese techo ascendió hasta los 35 TeraFLOPS, propulsado por el Earth Simulator, un supercomputador japonés ubicado en Yokohama.

Earth  
Simulator

El Earth Simulator está basado en la arquitectura SX de NEC, una familia de supercomputadores con 15 años de vigencia y seis generaciones anteriores, SX-1 a SX-6. Dispone de 640 nodos, cada uno de ellos compuesto de 8 procesadores vectoriales con un rendimiento pico de 8 GFLOPS y una memoria principal de 16 Gbytes, para un montante total de 10 Terabytes.

NEC

La memoria es de tipo compartido, utilizándose para llegar a ella desde los procesadores una red *cross-bar* de interconexión de una sola etapa con 83.000 cables de cobre, para un total de

memoria

- coste** 2.900 kilómetros de cableado. El coste del conjunto supera los 350 millones de euros y ocupa el espacio de cuatro pistas de tenis. Su misión consiste en crear un modelo computacional del planeta Tierra y simular en él todo tipo de fenómenos atmosféricos y meteorológicos, de tal forma que por ejemplo podamos adelantarnos al futuro para conocer con mucha antelación el efecto que el calentamiento global del planeta producirá a finales de siglo.
- misión**
- TOP 500** Según Hans Meuer, el padre del ranking TOP 500 que cada año enumera los quinientos computadores más potentes del planeta, el Earth Simulator no sólo es cinco veces más rápido que el segundo de la lista en las pruebas de rendimiento Linpack, sino que proporciona una potencia superior a la que alcanzarían conjuntamente las 19 máquinas que le siguen en el ranking.



## La anécdota: La ley de Moore



El mundo de la computación destila sobre el papel el glamour de una ciencia exacta, quizá porque para muchos de sus usuarios, los computadores denotan precisión antes incluso que velocidad.

**orden o caos** La realidad indica, muy al contrario, que en este mundo está más presente la anarquía que el orden. En la fabricación de chips rigen las normas empíricas, tan sólo acomodadas por la regularidad estadística a gran escala; en la arquitectura, la selección de los bloques funcionales de un procesador se hace más por meras hipótesis o intuiciones que por la aplicación de reglas analíticas; y qué decir de la velocidad de un chip, determinada a posteriori en base a *prueba a ver qué ha salido y dictamina a grosso modo*.

Estando así las cosas, a pocos puede extrañar que muchos ingenieros en computación hayamos echado en falta una formulación acaso más científica para nuestra disciplina. Y a pesar de todo, una de las contadas contribuciones que hemos aportado al respecto ha sido la Ley de Moore.

Vaya en nuestro descargo que el vertiginoso ritmo evolutivo de la tecnología apenas sí nos ha concedido un respiro para establecer unas bases más sólidas y una metodología algo más elegante. Pero si nuestra Ley es la de Moore, apaga y vámonos.

**ley o conjetura** Leyes son las de Newton en física o las de Maxwell en electromagnetismo. Nuestra *ley* no tiene rigor alguno, ni demostración conocida. Fue una mera conjetura de un señor que, privilegiadamente situado, observó la evolución del número de transistores por chip en sus inicios, planteando la posibilidad de que ese ritmo se mantuviese en el futuro. Lo gracioso del asunto es que todos nos hemos encargado de hacer realidad esas proyecciones futuras para poder apadrinar aquel hecho como nuestra *ley*.

Al ingeniero jefe de Intel que le tocó decidir la frecuencia del primer Pentium III, seleccionó los 450 MHz porque había transcurrido año y medio desde el primer Pentium II a 233 MHz, y claro, *tocaba* doblar el valor para así seguir la misma progresión enunciada por Gordon Moore, aunque ésta fuera formulada sobre otra variable tan distinta como el número de transistores. Pudieron ser perfectamente 350 MHz, o 600 MHz, pero entonces se quebraba la tendencia que sus predecesores en la empresa habían respetado con exquisito mimo durante tantos años por provenir del fundador de la compañía.

**vaticinio o manipulación** Ahora sabemos que no ha sido la ley la que ha predicho los acontecimientos, sino más bien al contrario: éstos han sido artificialmente manipulados para que así haya ocurrido. Es como el que vaticina que va a levantarse todos los días a las ocho en punto de la mañana durante los próximos diez años y luego se dedica a actuar con disciplina espartana para hacer realidad su vaticinio. Eso sí, por algún motivo le interesará también levantarse a esa hora para que no le tomen por loco, de la misma manera que la industria de los chips ha descubierto que se encuentra cómoda siguiendo el ritmo marcado por la Ley de Moore.

Las previsiones de Intel siguen diciéndonos que es posible apurar la tecnología de silicio hasta las 0.03 micras manteniendo la evolución marcada por la Ley de Moore durante los próximos diez años. Muy bien, Intel, seguiremos levantándonos a las ocho en punto de la mañana hasta nueva orden.



# Ejercicios suplementarios

---

Las soluciones a estos ejercicios se encuentran al final del libro, a continuación de las soluciones de los cuestionarios de los capítulos.

## Relacionar la dimensión RISC/CISC de un microprocesador

De entre las características que se listan a continuación, algunas son más propias de una arquitectura RISC y otras de una CISC. Clasifícalas colocando R o C en el hueco de la parte izquierda según corresponda.

- \_\_\_\_\_ Elevado número de registros de propósito general.
- \_\_\_\_\_ El programa que se está ejecutando tiene pocas instrucciones máquina.
- \_\_\_\_\_ El programa en ejecución tiene muchas instrucciones máquina y muy repetidas.
- \_\_\_\_\_ Integración de la caché de segundo nivel y la TLB dentro del propio chip.
- \_\_\_\_\_ Poca dependencia del compilador en términos de rendimiento.
- \_\_\_\_\_ Ejecución de una instrucción por ciclo de reloj.
- \_\_\_\_\_ Diseño segmentado y superescalar.
- \_\_\_\_\_ Instrucciones de alto nivel para el procesamiento multimedia.
- \_\_\_\_\_ El programa gasta pocos ciclos de reloj en procesar cada instrucción máquina.
- \_\_\_\_\_ Unidad de Control Microprogramada.

## Clasificar las mejoras del procesador Pentium

Relaciona cada una de las siguientes mejoras del microprocesador Pentium (a la izquierda) con el **primer** modelo comercial que la implementó (a la derecha). La correspondencia es biyectiva.

- |  |              |
|--|--------------|
| a.- Caché L1 gemelas de 16 Kbytes cada una.        | 1.- MMX.     |
| b.- Frecuencia de reloj por encima de los 500 MHz. | 2.- Pro.     |
| c.- Frecuencia de reloj por encima de los 2 GHz.   | 3.- Klamath. |

- d.- Optimizado para sistemas operativos de 32 bits.
  - e.- Formato Slot.
  - f.- Vuelta al formato Socket tras haber pasado por el Slot.
  - g.- Multiplicador de 12x respecto a su bus local.
  - h.- Fabricado a 0.25 micras.
  - i.- Dotado de 512 Kbytes de caché L2.
  - j.- Dotado de caché de traza.
- 4.- Deschutes.
  - 5.- Celeron.
  - 6.- Katmai.
  - 7.- Coppermine.
  - 8.- Tualatin.
  - 9.- Willamette.
  - 10.- Northwood.

# Para saber más

---

## Acerca de la historia del PC

### ► Libros

Una semblanza histórica de la microelectrónica a lo largo de los primeros setenta años del siglo XX:

⇒ J. Millman. *Microelectrónica. Circuitos y Sistemas Analógicos y Digitales*. Cuarta Edición. Editorial Hispano-Europea, Barcelona, 1988 (páginas 5-21).

La historia del computador narrada por uno de sus principales protagonistas:

⇒ M. Wilkes. *Computing Perspectives*. Morgan Kaufmann Publishers, 1995.

Los orígenes del computador y los principios básicos de su funcionamiento:

⇒ D. A. Patterson, John L. Hennessy. *Estructura y diseño de computadores*. 3 volúmenes. Ed. Reverté, 2000. (1.400 páginas en total). Coste total aproximado: 70€. Desde el punto de vista histórico, recomendamos ante todo la lectura del primero de los tres volúmenes.

El nacimiento del PC desde la perspectiva de los ingenieros de Xerox que diseñaron el primer computador personal:

⇒ Douglas Smith, Robert Alexander. *Fumbling the Future: How Xerox Invented, Then Ignored, the First Personal Computer*. 1988.

Multitud de historias, unas verídicas y otras verosímiles, para impregnarse del ambiente que rodea al Silicon Valley:

⇒ Po Bronson. *El nudista del turno de noche y otras historias de Silicon Valley*. Tusquets Editores. 2000. (336 páginas).

Para aprender más sobre la familia de microprocesadores 68000 de Motorola:

⇒ Ronald J. Tocci. *Microprocessors and Microcomputers: The 68000 Family*. Ed. Prentice Hall. Enero, 1986. Coste aproximado: 67€.

Para conocer más sobre los primeros modelos de la familia de microprocesadores 80x86 de Intel:

⇒ James L. Antonakos. *An introduction to the Intel Family of Microprocessors: A Hands-On Approach utilizing the 8088 Microprocessor*. Segunda edición. Ed. Prentice Hall. Junio, 1996. 675 páginas. Coste aproximado: 85€.

Todos los detalles sobre la familia 80x86 de Intel hasta alcanzar la sexta generación con el Pentium Pro:

⇒ Brey Barry. *Intel Microprocessors 8086/8088, 80186/80188, 80286, 80386, 80486, Pentium and Pentium Pro Processor*. Cuarta Edición. Ed. Prentice Hall. Noviembre, 1996. 896 páginas. Coste aproximado: 100 €.

### ► Artículos

El artículo en el que Wilkes acuñó el término *Microprogramación*:

◆ M. Wilkes. *The Best Way to Design an Automated Calculating Machine*. Conferencia Inaugural de Computación de la Universidad de Manchester, 1951, páginas 16-18.

El primer artículo en el que Wilkes describió la técnica de la microprogramación:

◆ M. Wilkes, J.B. Stringer. *Microprogramming and the Design of the Control Circuits in an Electronic Digital Computer*. Proceedings Cambridge Philosophy Society, número 49, páginas 230-238, 1953.

Los principios básicos que guiaron el desarrollo del Alto, uno de los principales precursores del PC, explicados por su propio creador:

◆ A. Kay. *Microelectrónica y ordenador de uso personal*. Investigación y Ciencia (edición en español de la revista Scientific American), Noviembre de 1997. Pags. 149-160.

El nacimiento del microprocesador 4004 narrado por sus protagonistas 21 y 25 años después, respectivamente:

◆ F. Faggin. *The Birth of the Microprocessor*. Revista Byte, vol. 17, número 3, páginas 145-150, 1992.

◆ F. Faggin, M. Hoff, S. Mazor y M. Shima. *The History of the 4004*. Revista IEEE Micro. Diciembre 1996.

La convención utilizada por Intel en la denominación de sus primeros productos y algunos datos curiosos sobre su fabricación:

◆ Andrew M. Volk, Peter A. Stoll, Paul Metrovich. *Recollections of Early Chip Development at Intel*. Intel Technology Journal Q1, 2001.

### ► Patentes

Novedades arquitecturales del 4004 a través de la patente que registra su innovación:

• M. Hoff, S. Mazor y F. Faggin. *Memory System for Multi-Chip Digital Computer*. Patente norteamericana número 3,821,715 dada a Intel Corporation el 28 de Junio de 1974.

Novedades arquitecturales del 8080 a través de la patente que registra su innovación:

• F. Faggin, M. Shima y S. Mazor. *Computer Employing a Plurality of Separate Chips*. Patente norteamericana número 4,010,499 dada a Intel Corporation el 1 de Marzo de 1977.

### ► Páginas Web

Una cronología histórica de los hitos más importantes en la historia del computador, desde 1966 hasta nuestros días:

🔗 <http://www.islandnet.com/~kpolsson/comphist>

## Sobre componentes en general

### ► Libros

Un completo repaso a todos los componentes, aunque su información se va quedando ya un poco obsoleta:

⇒ H.P. Messmer. *The Indispensable PC Hardware Book*. Tercera Edición. Editorial Addison-

Wesley, 1997. 1384 páginas. ISBN: 0-201-40399-4.

### ► Páginas Web

Una recopilación de los principales sitios del World Wide Web dedicados al hardware del computador:

☞ <http://users.erols.com/chare/hardware.htm>

Noticias sobre el mundo del PC en tiempo real, más desde un punto de vista mercantilista que técnico:

☞ <http://www.theregister.co.uk>

Noticias sobre el mundo de las tecnologías de información en general, muchas de ellas relacionadas con los componentes de un computador:

☞ <http://www.ctnews3d.com>

Completa lista de pruebas de rendimiento comparativas sobre equipos PC y sus componentes (puesta al día):

☞ <http://www.cpuscorecard.com>

Información general sobre los componentes del PC:

☞ <http://www.pcguide.com>

## Sobre microprocesadores

### ► Libros

Los aspectos de diseño de un microprocesador y los principios básicos de su funcionamiento:

☞ D. A. Patterson, John L. Hennessy. *Estructura y diseño de computadores*. 3 volúmenes. Ed. Reverté, 2000. Aprox. 1.400 páginas en total. Coste total aproximado: 70 €.

☞ W. Stallings. *Computer Organization and Architecture*. Fifth Edition. Ed. Prentice Hall. 2000. 748 páginas. Coste aproximado: 60 €.

Los aspectos de más bajo nivel relacionados con la circuitería de los microprocesadores en:

☞ William Kleitz. *Digital and Microprocessors Fundamentals: Theory and Applications*. Ed. Prentice Hall. Septiembre, 1999. 434 páginas. Coste aproximado: 100 €.

Para aprender más en profundidad la filosofía de diseño RISC:

☞ Patrick H. Stakem. *A Practitioner's Guide to RISC Microprocessor Architecture*. Ed. John Wiley & Sons. Abril, 1996. 600 páginas. Coste aproximado: 95 €.

☞ Clemente Rodríguez, Gonzalo Alvarez, Olatz Arregi, Jesús M. Pérez. *Microprocesadores RISC. Evolución y Tendencias*. Editorial Ra-ma, 1999. 219 páginas. Coste aproximado: 20 €.

Para conocer más sobre las técnicas de renombrado de registros, ejecución fuera de orden y otras optimizaciones relacionadas con el diseño de los Pentium Pro, II y III en el contexto de la sexta generación:

☞ D. Papworth. *Tuning the Pentium Pro Microarchitecture*. IEEE Micro, Abril, 1996.

☞ T. Shanley. *Pentium Pro and Pentium II System Architecture*. Reading, Massachussets. Ed. Addison-Wesley, 1998.

El algoritmo de predicción dinámica implementado en la BTB del Pentium Pro, II y III explicado al detalle en un artículo de sus creadores:

☞ Yeh, T., Patt, Y. *Two-level Adaptive Training Branch Prediction*. Proceedings 24th Annual Int'l

Symposium on Microarchitecture, 1991.

Más detalles sobre el procesador K6 de AMD:

⇒ Howard Kalish, Jerry Isaac. *The AMD K6-3D Processor: Revolutionary Multimedia Performance*. Ed. Abacus Publishing Company. Junio, 1998. 534 páginas. Coste aproximado: 30€.

Para saber más sobre la arquitectura de los procesadores RISC PA de Hewlett-Packard:

⇒ Gerry Kane. *PA-RISC 2.0 Architecture*. Ed. Prentice Hall. Diciembre, 1995. Coste aproximado: 43€.

La arquitectura de los Alpha de Digital, en:

⇒ Richard L. Sites, Richard T. Witek. *Alpha Architecture Reference Manual*. Tercera Edición. Ed. Digital Press. Abril, 1998. 420 páginas. Coste aproximado: 60€.

La arquitectura de los microprocesadores Power PC de Motorola:

⇒ Cathy May, Ed. Sikha. *The Power PC Architecture: A Specification for a New Family of RISC Processors*. Ed. Morgan Kaufmann Publishers. Junio, 1994. 518 páginas. Coste aproximado: 70€.

### ► Patentes

Las patentes más sobresalientes del Pentium 4:

Caché de Trazas:

☞ <http://www.patents.ibm.com/details?&pn=US06018786>

Algoritmo de predicción de salto:

☞ <http://www.patents.ibm.com/details?&pn=US06055630>

### ► Páginas Web

Los principales fabricantes de microprocesadores:

Intel: ☞ <http://www.intel.com>

Advanced Micro Devices (AMD): ☞ <http://www.amd.com>

Cyrix: ☞ <http://www.cyrix.com>

Hewlett-Packard: ☞ <http://welcome.hp.com/country/es/spa/welcome.htm>

Compaq Computer: ☞ <http://www5.compaq.com>

Motorola: ☞ <http://www.mot.com>

Transmeta: ☞ <http://www.transmeta.com>

Dos buenos documentos técnicos en PDF sobre la arquitectura del K7:

☞ [http://chipanalyst.com/pc\\_processors/articles/v12/121301.pdf](http://chipanalyst.com/pc_processors/articles/v12/121301.pdf)

☞ <http://www.amd.com/products/cpg/athlon/techdocs/pdf/22007.pdf>

Los detalles sobre el algoritmo de predicción de salto del K7:

☞ <http://www.s.netic.de/ak/k7doc.pdf>

Los manuales con todos los detalles acerca del Pentium 4 Willamette editados por Intel:

☞ <http://developer.intel.com/design/processor/future/manuals>

La página Web del procesador Crusoe mantenida por Transmeta con todo tipo de información al respecto:

☞ <http://www.crusoe.com>

Los manuales técnicos de procesadores consolidados junto a artículos sobre arquitecturas más

recientes:

🔗 <http://www.x86.org>

Las magnitudes principales de cada microprocesador, bien sintetizadas en tablas con completa y actualizada información:

🔗 <http://www.sandpile.org>

Las páginas Web que analizan los productos más innovadores del mundo del PC a medida que van apareciendo en el mercado. Aparecen listados por orden de recomendación:

En inglés:

🔗 <http://www.tomshardware.com>

🔗 <http://www.anandtech.com>

🔗 <http://www.sharkyextreme.com>

🔗 <http://sysdoc.pair.com>

🔗 <http://hothardware.com>

🔗 <http://www.xbitlabs.com>

🔗 <http://www.hardware-unlimited.com>

🔗 <http://www.extremeoverclocking.com>

🔗 <http://www.hardwaresecrets.com>

🔗 <http://www.pchardware.ro>

🔗 <http://www.cpu-central.com>

🔗 <http://www.ocworkbench.com>

En castellano:

Muy completo (precios, artículos, productos), aunque le cuesta estar al día:

🔗 <http://www.duiops.net/hardware/>

Menos completo pero más actualizado:

🔗 <http://www.meristation.com/HARD/elhard.htm>

La versión html de la popular revista *PC Actual* y poca cantidad de información técnica útil:

🔗 <http://www.vnunet.es/pcactual/>

Más un grupo de noticias que una Web que suministre información:

🔗 <http://www.mundopc.com/index.php>

## Diccionarios de terminología informática

### ▶ Libros

☞ Arturo Moreno Martín. *Diccionario de informática y telecomunicaciones*. Editorial Ariel/Practicum, Barcelona 2001. 842 páginas. 29 €. ISBN: 84-344-2886-5.

### ▶ Páginas Web

En inglés. Posibilidad de consulta por término, orden alfabético o categorías:

🔗 <http://what-is.techtarget.com>

En inglés. Relativo a terminología del computador e Internet:

✎ <http://www.webopedia.com>

En castellano, como parte de unas páginas más pretenciosas:

✎ <http://www.conozcasuhardware.com>

Extenso vocabulario informático de Juan Antonio Millán, escritor técnico y columnista del diario El País. Muy completo, no sólo por las definiciones, sino por sus etimologías (en castellano):

✎ [http://jamillan.com/v\\_index.htm](http://jamillan.com/v_index.htm)



# Soluciones a las cuestiones

---

## Capítulo 1

1. **b** Jack Kilby no participó en la invención del transistor. Gordon Moore vaticinó el potencial del silicio, pero no jugó un papel tan importante en su consecución. Por último, la contribución de Faggin se limita a diversos desarrollos de una misma idea: La integración de un microprocesador en silicio.
2. **a** En concreto, comenzaron a desarrollarse para el computador Alto (1972). Para los nostálgicos, recomendamos la lectura de su artículo *Microelectrónica y ordenador de uso personal* publicado en el *Investigación y Ciencia* de Noviembre de 1977, en cuya entrada (pag. 149) Kay escribe "El ritmo de progreso de la microelectrónica hace pensar que dentro de una década mucha gente poseerá un ordenador del tamaño de un cuaderno y la capacidad de una gran ordenador actual. ¿En qué podrá ayudarles este sistema?". Y dos páginas más adelante, al pie de una ilustración con las primeras ventanas e iconos de la historia, puede leerse "Las VENTANAS, marcos del monitor dentro de la pantalla, capacitan al usuario para organizar y editar información a diferentes niveles de claridad. Una vez creadas, se solapan como hojas de papel...".
3. **d** Data de 1945, y fue ideado por J.W. Mauchly y J.P. Eckert.
4. **b** Busicom, calculadora encargada a Intel por la empresa del mismo nombre en 1970.
5. **c**
6. **b**

## Capítulo 2

1. **b** Aunque la L2 del Pentium Pro era interna, no integrada.
2. **b** El 8086 disponía de 29.000 transistores y el Pentium de 3.100.000.
3. **c** Motorola desaparece en la cuarta y AMD comienza a resurgir con fuerza a partir de la sexta.
4. **c** Primero era de IBM, luego de National Semiconductor y finalmente pasó a manos de VIA, el fabricante de placas base, a finales de 1999. Ninguna de las otras compañías sufrió cambios importantes en su accionariado

durante ese período.

## Capítulo 3

1.  a
2.  c
3.  d
4.  d La ejecución de una instrucción en el tiempo que delimita un solo ciclo de reloj del procesador puede ser parcial (como en las arquitecturas CISC) o total (el objetivo último que se persigue en las arquitecturas RISC). Respecto al matiz de ejecutar **una** o **varias** instrucciones en ese tiempo, debemos inclinarnos por la pluralidad puesto que todos los microprocesadores actuales tienen implementadas estrategias de paralelismo a nivel de instrucción como la segmentación o la superescalaridad.
5.  a Es la única afirmación que es siempre cierta, puesto que la mejora consiste en acelerar un parámetro y dejar el resto inalterables (se trata en ambos casos del mismo equipo).
  - La afirmación **b** se refiere a dos computadores distintos, y en ese caso, el que uno tenga una frecuencia de reloj mayor para su microprocesador no significa necesariamente que vaya a ser más rápido, ya que esto dependerá de las características (tamaño, velocidad, ...) del resto de componentes (placa base, memoria caché, memoria principal, buses, ...).
  - La afirmación **c** es ambigua, puesto que son cinco las magnitudes que reflejan en conjunto la presteza con que se ejecuta un programa, y de ellas la que mejor refleje la rapidez en la ejecución será la que se encuentre más potenciada en cada diseño en concreto. Por ejemplo, si se opta por un diseño de frecuencia elevada y un bajo paralelismo a nivel de instrucción, entonces la afirmación es correcta; en cambio, si se prima éste último frente a la primera, la afirmación pasa a ser falsa.
  - La afirmación **d** es falsa, pues la única incidencia destacable que ocurre a mitad de frecuencia es que el programa tarda el doble de tiempo en ejecutarse, sin producirse alteración alguna en la forma en que se lleva a cabo su ejecución.
6.  a
  - Las opciones **b** y **c** son más convenientes cuanto más rápido es el microprocesador, por lo que quedan directamente descartadas.
  - La opción **d** es válida tanto para procesadores rápidos como para lentos, mientras que la **a** suele producir la ralentización de la frecuencia que nos indica el enunciado de la pregunta. Por lo tanto, se presenta como la mejor opción para aprovechar lo poco positivo que tiene el condicionamiento de partida.
7.  b
8.  c La correspondencia entre **tecnología de integración** y **distancia de integración** estuvo dándose hasta las 0.35 micras. A partir de ahí, las tecnologías de 0.25, 0.18 y 0.13 micras utilizan transistores con anchura de puerta de 0.20, 0.13 y 0.07 micras, respectivamente. Esto invalida las opciones **a** y **b**, y en consecuencia, también la opción **d**.
9.  b Las otras tres opciones también mejoran el microprocesador, pero encarecen su coste.
10.  b La frecuencia de reloj dobla su valor cada 18 meses en un crecimiento claramente exponencial que no puede compararse en su evolución con ninguno de los otros tres parámetros. Precisando un poco más: La tecnología de integración en micras es la variable que presenta un mayor grado de estancamiento en términos porcentuales (30% de mejora cada dos años). El tamaño de la caché L1 está creciendo paulatinamente, pero aunque la dependencia del sistema con respecto a

este elemento es elevada en términos de rendimiento, se está solventando más por medio del incremento de niveles en la jerarquía de cachés (aparición de cachés L2 y L3) que por el aumento de tamaño del primero de esos niveles. El número de transistores está aumentando, pero al ser una variable directamente ligada al precio del chip, el incremento se está produciendo de una forma algo más lenta y ordenada.

11.  d Todas son correctas excepto la opción b . La más completa es obviamente la opción d .
12.  b
13.  d
14.  d La opción a , que es la que menos puede entenderse de forma directa, es cierta porque en una misma oblea de silicio caben más chips si se fabrican bajo un proceso de 0.13 micras (los transistores se encuentran más apilados).
15.  d Es la única variable, de las que aparecen como respuesta, sobre la que se produce una incidencia cuadrática (sobre las otras tres, la influencia es lineal).
16.  a
17.  a Si la frecuencia no presenta un valor muy ajustado, las posibilidades de aumentar la ejecución simultánea de operaciones crecen de forma considerable.
18.  c
19.  a Efectivamente, resulta una alternativa más económica.
20.  c
21.  c
22.  a
23.  c
24.  d
25.  b
26.  b
27.  a A mayor número de etapas de segmentación, mayor número de instrucciones se encuentran simultáneamente dentro del cauce segmentado, y recordemos que son éstas las que se ven afectadas por las dependencias de datos. Por otro lado, no existe una relación directa entre el grado de segmentación y de superescalaridad de un código.
28.  d
29.  c
30.  c El microprocesador que conjunta las dos estrategias es al menos tan eficiente como el que implementa cualquiera de ellas por separado.
31.  b Las opciones a y c parecen también candidatas. sin embargo, debemos hacer dos observaciones: Una BTB reduce los conflictos por dependencias de control, pero no puede eliminarlos por completo, ya que se basa en una predicción sobre un evento futuro que como tal puede fallar.

La anulación de la secuencia de instrucciones ejecutada tras una predicción errónea de salto y el correspondiente restablecimiento de la secuencia correcta de instrucciones a ejecutar tras el salto es más compleja cuando tenemos una BTB, ya que además de todo lo anterior debe actualizarse la BTB para que registre este error y cambie la predicción para poder acertar en un futuro.

32.  **c** Descartamos el resto por los siguientes motivos:  
 Opciones **a** y **b**: La ejecución fuera de orden establece el orden en que se realiza la operación (aritmética, lógica, ...) asociada a cada instrucción máquina. La TLB, en cambio, no es más que un recurso para acelerar la traducción de direcciones de memoria virtuales a físicas, y por lo tanto, se encuentra emparentada con las unidades de almacenamiento de datos e instrucciones (cachés de datos e instrucciones en su fase de búsqueda).  
 Opción **d**: Cuando la arquitectura dispone de una caché de traza, la fase de búsqueda se encuentra más cercana a las fases de ejecución, pero aún le queda un largo trecho por recorrer hasta llegar a ellas dentro del cauce segmentado del procesador, ya que si el diseño incluye ejecución fuera de orden, antes debe pasar por el BRC y las fases de renombrado de registros, reordenación y planificación de instrucciones.
33.  **b**  
 Opción **a**: Descartada porque realizando un control segmentado de las instrucciones (esto es, haciendo que las señales de control viajen por el cauce junto con el código de la instrucción y sus datos), la Unidad de Control no es sustancialmente diferente de la de un microprocesador multiciclo no segmentado.  
 Opción **c**: Eliminada porque el tamaño del banco de registros está relacionado con el volumen de datos que maneja un programa, no con su ejecución segmentada o no.
34.  **d** Y únicamente nos resta descartar la opción **a**, ya que el tamaño de la memoria requerida por un programa y el factor de superescalaridad que mejor aprovecha son dos aspectos que no guardan relación alguna.
35.  **b** Para la segmentación, el techo sí es una instrucción por ciclo de reloj, pero con la superescalaridad podemos ejecutar múltiples instrucciones en cada ciclo de reloj.
36.  **c** El número de etapas de segmentación y el factor de superescalaridad apenas guardan una relación, salvo para determinar el paralelismo a nivel de instrucción con que cuenta el procesador en su conjunto.
37.  **a**  
 Opción **b**: Si por algo se caracteriza la segmentación es por ser una alternativa barata de implementar que no requiere la replicación de hardware.  
 Opción **c**: El número de etapas de segmentación y el factor de superescalaridad apenas si guardan relación.  
 Opción **d**: Un conjunto de instrucciones RISC aboga por la simplicidad de las instrucciones, y difícilmente así podrán constar de 50 etapas.
38.  **b** Es necesario replicar la circuitería diez veces en la unidad de proceso e implementar una sofisticada unidad de control, lo que disparará el número de transistores de que consta el diseño de una forma exagerada.
39.  **b** Aumentar el factor de superescalaridad supone la replicación (copia) de toda la circuitería involucrada en esa computación, mientras que la segmentación sólo complica un poco el diseño de la unidad de control de procesador, resultando mucho más barata de implementar. Del resto de opciones posibles:  
 Opción **a**: No existen tales dependencias entre la puesta en marcha de la segmentación y la superescalaridad.  
 Opción **c**: La incidencia de las dependencias es similar en cualquier forma de paralelismo a nivel de instrucción.  
 Opción **d**: La presencia de caché integrada no está relacionada con ninguna forma de paralelismo a nivel de instrucción. Sólo la distinción entre caché de datos e instrucciones en el primer nivel de la jerarquía reduce las dependencias estructurales en procesadores segmentados.
40.  **a**
41.  **d**
42.  **b** La transferencia se dirige siempre hacia la caché de primer nivel ó L1.
43.  **d**

44. **a**
45. **d**
46. **d** En concreto, se utilizan entre 4 y 6 transistores por cada bit de datos en cualquiera de las modalidades de caché vistas, ya sea externa, interna o integrada. Puntualizaremos que el interfaz con el controlador de caché (al que NO está referido la pregunta) sí es diferente en cada caso, pero reduce la complejidad del controlador en cada una de las transiciones expuestas, ya que sus necesidades de diálogo se acrecientan conforme bascula hacia ubicaciones más externas al microprocesador. Esto le hace requerir un **menor** número de transistores en su evolución desde la modalidad externa a la interna o desde ésta hacia la integrada.
47. **c**
48. **b** Ya no hay configuraciones en las que el procesador y el bus local funcionen a la misma velocidad, por lo que las opciones **a** y **c** quedan descartadas. Entre las otras dos, la opción **b** es mas equilibrada.
49. **d**
50. **d** De hecho, este bus desaparece cuando la caché se integra en el conjunto.
51. **c** La opción **a** no es correcta ni por sí mismo ni con la enmienda de la opción **b**, ya que sólo en L2, el Pentium 4 se va a los 256 Kbytes. Respecto a la opción **d**, es todo un disparate, ya que cada transistor cumple idéntica función de almacenamiento al margen de su tamaño.
52. **c**
53. **a**
54. **d**
55. **c** Tanto los bancos de registros como las cachés de primer nivel son mayores en un RISC que en un CISC, y por eso a un RISC le hace menos falta una caché de segundo nivel. No obstante, hay diseños RISC donde la opción **d** es cierta (como en el Alpha), pero hay otros en los que no lo es (como en el UltraSparc).
56. **a** La elegimos por eliminación. Veamos con un poco más de detalle por qué hemos de rechazar las otras alternativas:
- ⊗ El bus local no merece la pena mejorarlo porque en este caso disponemos de una caché L2 bastante grande que reduce mucho la dependencia del sistema con respecto al bus. Además, desde los años ochenta no se ve una arquitectura en la que no haya un multiplicador de reloj para diferenciar la frecuencia del microprocesador de la del bus local.
  - ⊗ El factor superescalar no tiene sentido mejorarlo hasta 20, ya que debido a las dependencias que presentan las instrucciones de los programas, resulta complicado aprovechar más de un factor 6. Además, en el ejemplo resulta comparativamente muy caro.
  - ⊗ Aumentar más la caché L2 sólo se amortizaría en una arquitectura con un bus local suficientemente lento, y tampoco es el caso, ya que en el ejemplo disponemos de un bus a 200 MHz.
57. **b** El PowerPC de Motorola es la única familia que ha llegado al factor seis de superescalaridad en sus modelos comerciales.
58. **c** La respuesta está entre el PA8000 y el R10000, si bien Hewlett-Packard desplazó en aquel marco temporal parte de sus esfuerzos en el desarrollo de sus propios procesadores hacia la iniciativa del Itanium en conjunción con Intel.
59. **c**
60. **d**
61. **b**

62. **d**
63. **d** CISC es una filosofía de diseño para el conjunto de instrucciones de un procesador.  
 La opción **b** muestra tres tipos de buses y/o zócalos de la placa base, y la opción **c** materiales utilizados en la implementación del microprocesador; ninguna de las dos tiene nada que ver con lo que se pregunta.  
 Respecto a la opción **a**, muestra tres modelos de Pentium II con modificaciones en las características de la caché L2, pero idénticos en cuanto a su conjunto de instrucciones.
64. **a**
65. **d** Las razones para descartar el resto de opciones son las siguientes:  
 Opción **a**: El controlador de caché no tenemos por qué cambiarlo, puesto que las etiquetas de consulta que se utilizan para localizar los datos que alberga corresponden a direcciones de memoria física, no virtual.  
 Opción **c**: La traducción virtual a física es realizada por la TLB, en la que sí debemos cambiar la dirección de entrada que se le proporciona, que ahora reflejaría un interfaz de 64 líneas.  
 Opción **b**: El controlador de bus local también se vería alterado, ampliándose el patillaje correspondiente a las líneas de dirección de 32 a 64 con objeto de direccionar un espacio de direcciones de memoria virtuales de hasta  $2^{64}$  palabras.
66. **d** Esas cantidades no indican nada acerca del ancho del bus de datos del procesador. Habría que conocer el tamaño de la palabra de memoria principal, algo que tampoco puede determinarse a partir del tamaño de memoria principal.
67. **b** Un aumento lineal en cualquiera de las tres magnitudes indicadas (distancia de integración, frecuencia y número de transistores) provoca un aumento lineal sobre la potencia disipada del microprocesador. Evaluándolas por ese orden, y siempre en relación a PK1, la potencia de cada procesador se incrementa según los siguientes factores multiplicativos:  
 - PK2. 0.75x (180 nm. frente a 250 nm. en PK1), 1.3x (2.6 GHz frente a 2 GHz en PK1), 1x (los mismos 50 Mt. que PK1). En conjunto: 1x (misma potencia que PK1).  
 - PK3. 0.5x (130 nm. frente a 250 nm. en PK1), 1x (misma frecuencia en ambos procesadores) y 4x (200 Mt. frente a 50 Mt. en PK1). En conjunto: 2x (el doble que PK1).  
 - PK4. 0.5x, 2x, 1x. En conjunto: 1x.

La siguiente tabla resume nuestro análisis:

Magnitud analizada: Mayor potencia disipada	Las tres magnitudes conocidas			Factor de incremento total respecto a PK1
	Distancia de integr.	Frecuencia	Núm. transistores	
Efecto de un incremento lineal de cada magnitud sobre la potencia	Aumento lineal	Aumento lineal	Aumento lineal	
Cambios respecto a PK1 en la potencia disipada por cada procesador:				
- PK2 (180 nm, 2.6 GHz, 50 Mt)	0.75x	1.3x	1x	=
- PK3 (130 nm, 2 GHz, 200 Mt)	0.5x	1x	4x	2x
- PK4 (130 nm, 4 GHz, 50 Mt)	0.5x	2x	1x	=

68. **c** De cara a evaluar la repercusión sobre el área de integración, la distancia de integración incide cuadráticamente, la frecuencia no entra en juego y el número de transistores influye linealmente. Nuestra tabla evaluadora queda en este caso como sigue:

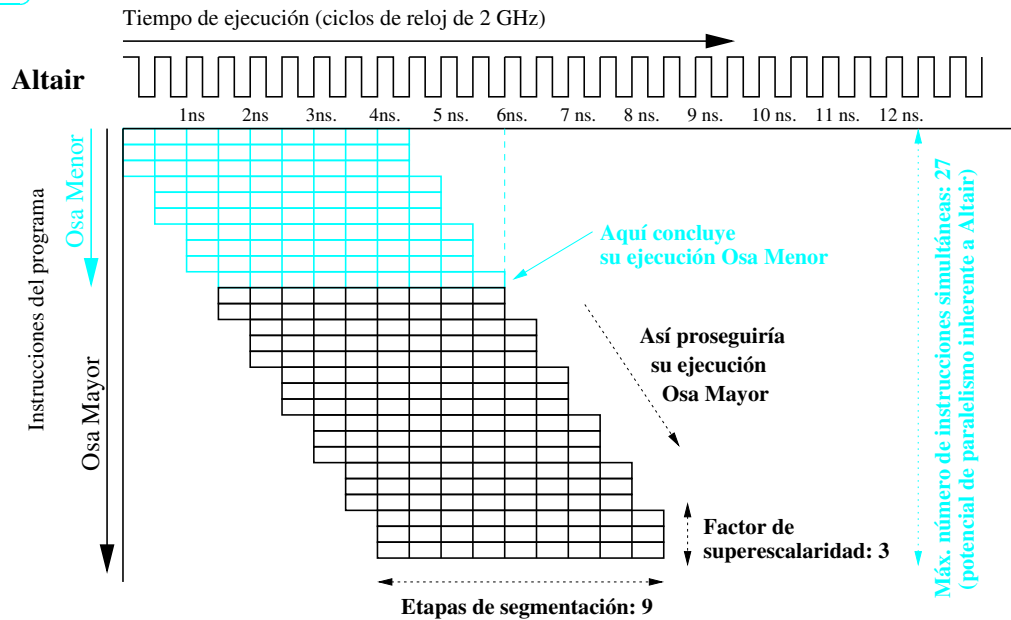
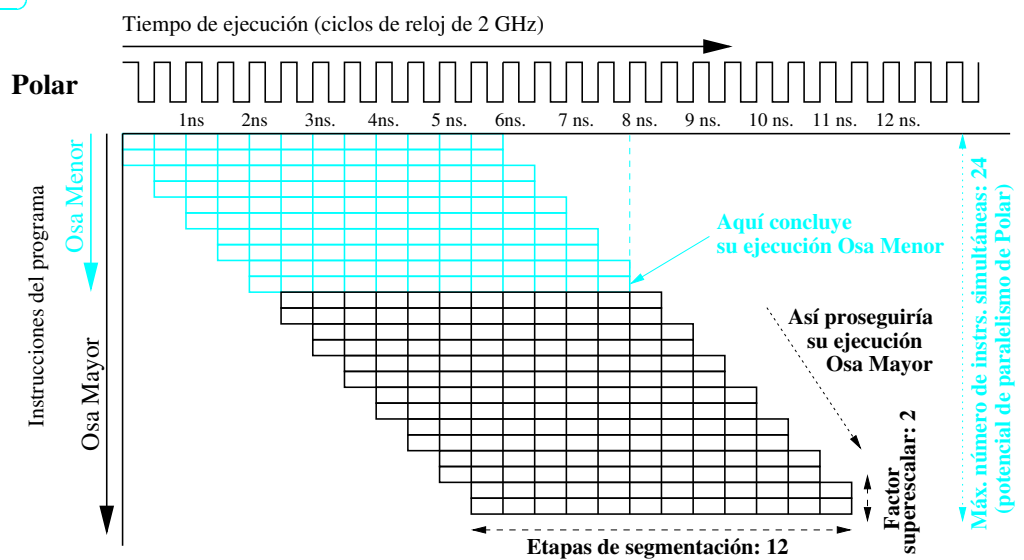
Magnitud analizada: Menor área de integración	Las tres magnitudes conocidas			Factor de incremento total respecto a PK1
	Distancia de integr.	Frecuencia	Núm. transistores	
Efecto de un incremento lineal de cada magnitud sobre el área de int.	Aumento cuadrático	No entra en juego	Aumento lineal	
Cambios respecto a PK1 en el área de integración de cada procesador:				
- PK2 (180 nm, 2.6 GHz, 50 Mt)	$(0.75x)^2$	1x	1x	0.6x
- PK3 (130 nm, 2 GHz, 200 Mt)	$(0.5x)^2$	1x	4x	1x
- PK4 (130 nm, 4 GHz, 50 Mt)	$(0.5x)^2$	1x	1x	0.25x

69. **b** Sobre el rendimiento bruto, la distancia de integración no incide, mientras que la frecuencia y el número de transistores lo hacen de forma lineal. En consecuencia, nuestra evaluación queda como sigue:

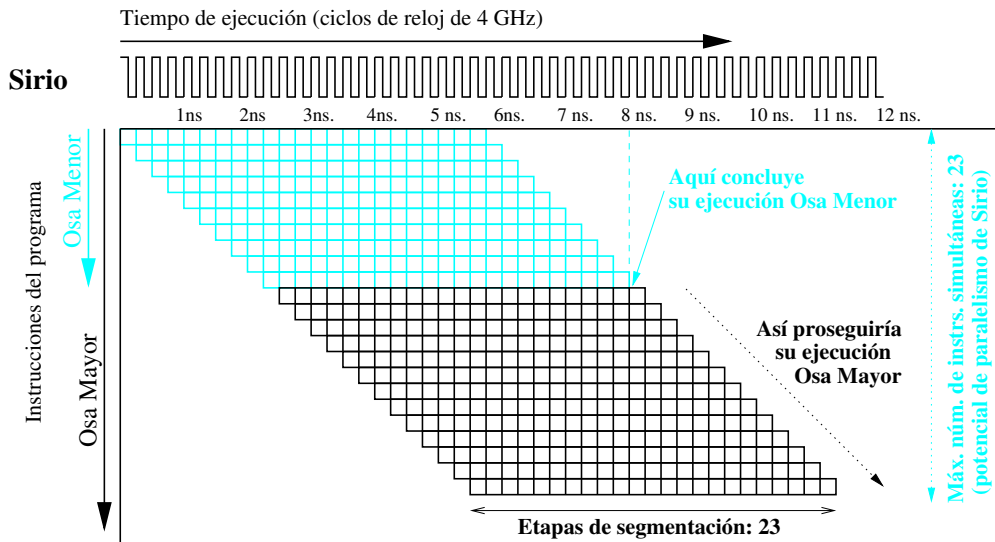
Magnitud analizada: Mayor rendimiento bruto	Las tres magnitudes conocidas			Factor de incremento total respecto a PK1
	Distancia de integr.	Frecuencia	Núm. transistores	
Efecto de un incremento lineal de cada magnitud sobre el rendim.	No entra en juego	Aumento lineal	Aumento lineal	
Cambios respecto a PK1 en el rendimiento bruto de cada procesador:				
- PK2 (180 nm, 2.6 GHz, 50 Mt)	1x	1.3x	1x	1.3x
- PK3 (130 nm, 2 GHz, 200 Mt)	1x	1x	4x	4x
- PK4 (130 nm, 4 GHz, 50 Mt)	1x	2x	1x	2x

70. **c** Si la maquinaria de fabricación se supone amortizada, los costes del procesador están relacionados con la materia prima utilizada, y ésta es directamente proporcional al área de integración del procesador, aspecto que evaluamos anteriormente, resultando ganador PK4.
71. **b** A la hora de debatirse por el procesador que más fácilmente implementa la superescalaridad, y asumiendo la simplicidad que supone la sola intervención de las magnitudes propuestas, nos quedamos con el número de transistores como la más representativa de todas ellas. Recordemos que el concepto de superescalaridad consiste en la replicación de unidades funcionales, y que por lo tanto, cada incremento en el factor de superescalaridad supone un incremento similar en el número de transistores del diseño. Nuestro ganador aquí es por lo tanto PK3 con sus 200 millones de transistores.
72. **c** La supersegmentación lleva implícita una subdivisión del ciclo de reloj del procesador, por lo tanto, el modelo de más alta frecuencia debe ser el elegido, en este caso, PK4.
73. **b** Las cachés son las grandes consumidoras de transistores en la integración de un procesador. Por lo tanto, el modelo de 200 millones de transistores (Mt.) es el más proclive a disponer de una gran caché L2 integrada. Si los modelos comerciales de 256 Kbytes de L2 integrada se llevan en torno a 15-20 Mt., para un tamaño ocho veces superior está claro que hablamos de un modelo de más de 100 Mt., cantidad sólo superada por PK3.
74. **b** Para ampliar la funcionalidad del procesador, se necesita disponer de un mayor patrimonio de transistores con el que construirlas. Sin embargo, nuevas operaciones no incrementan el número de transistores de una forma tan apreciable como lo hacen otros elementos como las celdas de caché. Por ello, aquí el número de transistores no es un elemento tan discriminante como en el caso anterior, y la distancia de integración también entra en juego. Al evaluar los dos parámetros conjuntamente, es PK3 el que sale mejor parado.
75. **c** El diseño RISC aboga por una arquitectura simple que conjugue perfectamente la antítesis del principio microelectrónico "más grande, más lento". En consecuencia, no se condiciona el diseño a la disponibilidad de muchos transistores, sino a la velocidad de éstos. Para ello, se necesita una elevada frecuencia y una ínfima distancia de integración que lo haga posible, y el procesador con mejores cualidades en el conjunto de ambas es PK4.
76. **d** En principio, cuantos más transistores haya disponibles, más fácil será la construcción de defensas frente a las dependencias, como un gran búfer de reordenación circular (BRC) o un gran búfer de destino del salto (BTB).

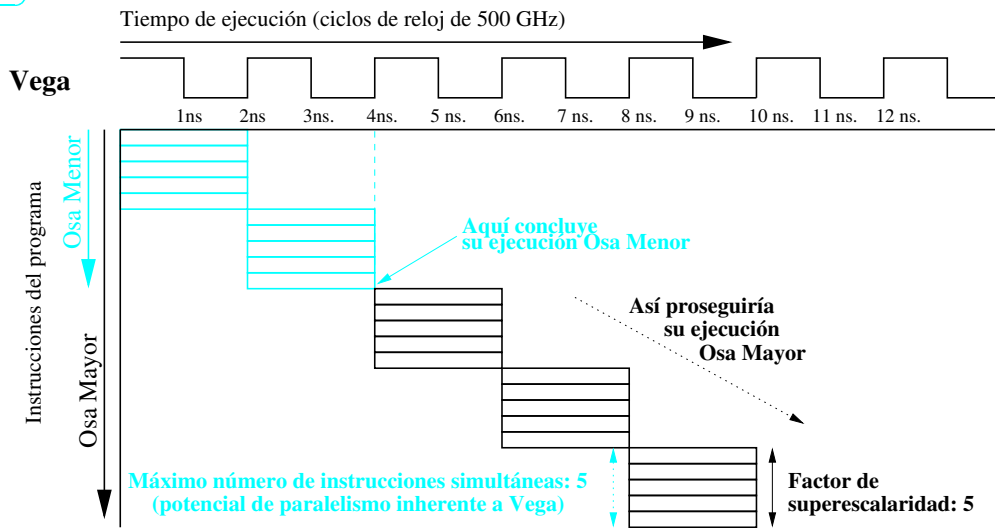
Pero la experiencia nos indica cierta indefinición al respecto, en vista de que existen modelos dotados de ingentes cantidades de transistores que prefieren dedicarlos a otros menesteres. Por lo tanto, si bien PK2 y PK4 son los candidatos más proclives y PK3 queda en un segundo plano, debemos conocer en qué medida el diseñador ha decidido engordar sus BRC y BTB para protegerse frente a las dependencias antes de pronunciarnos sobre un procesador en concreto.

77. **b**78. **c**79. **c**





80. a



81. a Altair puede ejecutar hasta 27 instrucciones simultáneamente, según se aprecia en su diagrama de ejecución anterior.

82. d Vega sólo puede ejecutar 5 instrucciones simultáneamente (consultar su diagrama de ejecución), ya que sus cinco carriles superescalares no se encuentran segmentados.

83. a Para conocer el tiempo de ejecución de Osa Mayor en cada procesador, la mejor forma de proceder si nos fijamos en los diagramas consiste en calcular el número de ciclos (NC) necesarios para ejecutar el programa, y luego multiplicarlo por el periodo de reloj (T).

En Altair, por ejemplo, cada ciclo comienzan su ejecución 3 nuevas instrucciones, luego al cabo de  $10^6/3$  ciclos ya han comenzado todas, y ahora hay que sumar los ciclos que tardan en concluir las últimas que comenzaron, esto es, el número de etapas de segmentación menos uno. Por lo tanto, en la expresión para NC, el factor de superescalaridad se coloca dividiendo, y el de segmentación sumando. La expresión para T es simplemente la inversa de la frecuencia:

$$\text{Altair: } T_{CPU} = NC \times T = \left(\frac{10^6}{3} + (9 - 1)\right) \times \left(\frac{1}{2GHz}\right) = 333342 \times 0,5ns. = 166671 \text{ nanosegundos}$$

$$\text{Polar: } T_{CPU} = NC \times T = \left(\frac{10^6}{2} + (12 - 1)\right) \times \left(\frac{1}{2GHz}\right) = 500011 \times 0,5ns. = 250005,5 \text{ ns.}$$

$$\text{Sirio: } T_{CPU} = NC \times T = \left(10^6 + (23 - 1)\right) \times \left(\frac{1}{4GHz}\right) = 1000022 \times 0,25ns. = 250005,5 \text{ ns.}$$

$$\text{Vega: } T_{CPU} = NC \times T = \frac{10^6}{5} \times \left(\frac{1}{0,5GHz}\right) = 200000 \times 2ns. = 400000 \text{ ns.}$$

Por lo tanto, Altair es el que menos tarda y Vega el que más, quedando Polar y Sirio empatados en posiciones intermedias.

84. **d** (ver la respuesta anterior)

85. **c** La media de ciclos por instrucción ejecutada se obtiene dividiendo el número total de ciclos (NC) de Osa Mayor, que ya calculamos anteriormente, por el número de instrucciones ejecutadas (NI), que son  $10^6$  para el caso de Osa Mayor. Desglosando para cada procesador:

$$\text{Altair: CPI} = \frac{NC}{NI} = \frac{333342}{10^6} = 0,333$$

$$\text{Polar: CPI} = \frac{NC}{NI} = \frac{500011}{10^6} = 0,500$$

$$\text{Sirio: CPI} = \frac{NC}{NI} = \frac{1000022}{10^6} = 1,000$$

$$\text{Vega: CPI} = \frac{NC}{NI} = \frac{200000}{10^6} = 0,200$$

El mayor CPI corresponde a Sirio, pero esta rémora la compensa con una frecuencia de reloj muy superior a la de Vega, el procesador con menor CPI de todos.

86. **d** (ver la respuesta anterior)

## Capítulo 4

1. **b** El K5 apenas se vendió y además es una remodelación de una arquitectura de NexGen, el MMX no es una arquitectura en sí, sino un leve refinamiento de otra, y el K6 no es de quinta generación, sino de sexta.
2. **d**
3. **c** En efecto, para que la instrucción de tipo flotante pueda disfrutar de ejecución superescalar tiene que ser extremadamente sencilla.
4. **b**
5. **a**
6. **c** Cuando el K5 fue lanzado al mercado, el Pentium era casi un estándar en la compra de un PC, y eso resulta muy difícil de enjugar. Respecto al descarte de las opciones que se acompañan:
  - No es la opción **a** porque el K5 incluía aspectos como la superescalaridad de 4 y la ejecución fuera de orden que ya hubiera querido para sí el Pentium.
  - Tampoco es la opción **b**, pues el K5 siempre fue una alternativa más barata que el Pentium.
  - Finalmente, descartamos **d** porque la incompatibilidad fue siempre un bulo que no se correspondía con la realidad.
7. **b** Los dos tienen igual número de patillas y son iguales externamente. El MMX tiene el doble de tamaño en su caché L1, pero eso no le impide compartir el zócalo. La razón es que los Pentiums MMX tienen un área externa que funciona a 3.3 voltios y un núcleo interno que lo hace a 2.7 voltios, y los zócalos para Pentium clásico no suelen venir preparados para esta última eventualidad.
8. **a** Las opciones **c** y **d** constituyen la materia prima en la fabricación de los transistores que componen el microprocesador, pero el hecho de que se denominen semiconductores delata que sus propiedades físicas no son las mejores para la transmisión de los electrones.
9. **c** El formato indicado en la opción **d** es una mera invención del autor, que lo ha tomado prestado del nombre más formal dado a la especificación del bus FireWire.

## Capítulo 5

1. **d**
2. **d** El Pentium es un contraejemplo para la opción **c**, y el K6-III, para la opción **b**.
3. **c** En concreto, fueron las instrucciones MMX, que también se incluyeron en el Pentium II, opción que no se encuentra entre las respuestas de la pregunta. Respecto a la opción **b**, decir que el Pentium MMX no es un diseño de sexta generación como reza el enunciado de la pregunta. Y respecto a la opción **d**, el K6-2 fue el primero en incluir las instrucciones 3DNow!, pero las MMX llegaron mucho antes al contexto de la sexta generación.
4. **b** En efecto, tanto el Pentium III Coppermine como el K6-III original la incorporan. La opción **a** señala un aspecto de séptima generación (se da en el Pentium 4 por primera vez) y las opciones **c** y **d** se dan ya en la quinta generación.
5. **b** La opción **a** no es correcta porque los zócalos de tipo Socket son un diseño que vuelve a estar de actualidad. La opción **c** se corresponde precisamente con uno de los aspectos negativos del zócalo tipo Slot 1, no del Socket 7.
6. **c**
7. **d** Intel siempre ha incluido buenas ALUs y FPUs en sus microprocesadores, fué la primera en incluir cachés de segundo nivel dentro del microprocesador (en el Pentium Pro), y además es fabricante de placas base, lo que unido a su gran cuota de mercado hace que la mayoría de fabricantes siguieran sus especificaciones en el lanzamiento de productos durante la sexta generación.
8. **b** Lo resumíamos con el símil de “modelo de boca estrecha y estómago glotón”.
9. **c**
10. **c** Fue Digital con su Alpha 21264 en el año 1996.
11. **d**
12. **a**
13. **b** El Celeron tiene un coste inferior, pero también es inferior su rendimiento. La caché se ha extinguido sólo en su segundo nivel, y además el propósito de su cambio no está en querer reducir el tamaño del microprocesador, sino en disminuir su coste de fabricación, ya que se perseguía a todo coste retirar del mercado a sus antecesores de zócalo Socket 7.
14. **b** El Celeron carece de memoria caché L2, mientras que el Celeron A dispone de una caché L2 de 256 Kbytes sincronizada a la velocidad del procesador.
  - ⊗ La frecuencia del procesador (opción **a**) se descarta porque existen frecuencias, como los 300 MHz, en las que Intel fabricó ambos modelos (de hecho, fue eso lo que provocó que Intel colocara la A en la versión más reciente para evitar confusiones).
  - ⊗ La frecuencia del bus local (opción **c**) se descarta porque es la misma en ambos modelos (66 MHz).
  - ⊗ Respecto al zócalo (opción **d**), el Celeron sin caché se presenta en formato SEC y SEP para acoplarse en ambos casos a un zócalo Slot 1, mientras que el Celeron A coexiste inicialmente en SEP para compartir el Slot 1, y posteriormente evoluciona a PPGA y FC-PGA, formatos que se acoplan a un zócalo Socket 370.
15. **d** Rechazamos las otras opciones por los siguientes motivos:
  - ⊗ Opciones **a** y **b**: El tamaño de caché se ha recortado siempre, ya sea prescindiendo de ella por completo (primeros modelos bajo Pentium II) o reduciéndolo a la mitad (últimos modelos bajo Pentium III).
  - ⊗ Opción **c**: La velocidad del bus se congeló en los 66 MHz hasta la llegada del modelo Celeron a 800 MHz (el primero con bus a 100 MHz, pero ya derivado del Pentium III Coppermine con bus de 133 MHz).

16.  d
17.  c Esta pregunta conlleva un problema de índole temporal, ya que el precio de los microprocesadores cambia cada trimestre y en el enunciado de la pregunta no se indica ninguna fecha de referencia sobre la que apoyarnos. La cuestión se formuló en un examen de Junio de 1998. Por aquel entonces, la respuesta correcta era la opción **c**, aunque la cosa está en todo caso entre esta opción y la **d**, dado que las otras dos respuestas pueden descartarse independientemente del marco temporal en el que nos situemos.
- La opción de subir la tecnología de integración no abarata el coste de fabricación del chip, sino que lo encarece por requerir una mayor área de silicio. Sólo podría abarataarse si el hecho de aumentar la distancia entre las pistas y el volumen de integración obliga a prescindir de algunas de sus unidades funcionales, pero esto queda descartado en tanto en cuanto el chip Pentium II es el mismo en ambos casos.
  - La opción **b** es siempre falsa, ya que ningún Pentium II posee una caché L3 interna (llega sólo hasta la L2).
  - Puestos a elegir entre las dos últimas opciones, a fecha Junio de 1998, bajar la frecuencia de 300MHz a 233 MHz reducía el coste de 60.000 a 30.000 ptas (esto es, un 50%), mientras que esperar seis meses no suele abaratar tanto el precio, salvo en excepciones muy puntuales y siempre que se adquiriera muy próximo a su lanzamiento.
18.  a El primero disponía de un bus a 66 MHz y un multiplicador de 5, mientras que el segundo integraba un nuevo controlador para bus a 100 MHz y un multiplicador de 3.5.
19.  b Para el descarte del resto de opciones:
- Hay Pentium II con bus local a 100 MHz que no son Xeon, lo que elimina la opción **a**.
  - También hay Pentium II que igualan en el tamaño de caché L2 con respecto a algunos Xeon. Por tanto, la opción **c** es verdadera parcialmente (sólo en algunos modelos del microprocesador).
  - Con respecto a la tecnología de integración, es de 0.25 micras en ambos casos, con lo que la opción **d** no es un elemento diferenciador.
20.  c
21.  d
22.  b El número de transistores es el mismo (9.5 millones en este caso) y su distancia de integración también (0.25 micras en este caso), pues ambos pasan por la misma maquinaria de fabricación. Respecto al voltaje, aunque la versión de 500 MHz podría obtenerse a partir de la de 450 MHz subiendo un poco su voltaje, la respuesta hace referencia al voltaje nominal (esto es, el valor especificado por el fabricante, no el manipulado por el usuario).
23.  d La versión de 0.18 micras a que hace alusión la opción **a** presenta una caché L2 integrada y formato FC-PGA, mientras que la de 0.25 micras a que hace alusión la opción **b** presenta una caché L2 interna y formato SECC2. La mejor elección es por tanto la conjunción de ambas.
24.  d Efectivamente, la pregunta hace referencia al coste de la materia prima por unidad fabricada, y con 0.18 micras gastamos menos área de silicio porque los transistores son más pequeños. La maquinaria de litografía de 0.18 micras es más cara, pero eso no lleva a que la opción **a** pueda ser cierta, ya que la pregunta presupone la existencia de dicha maquinaria, y por tanto, no entra en juego el coste de amortización de la planta de fabricación.
25.  c El matiz que refleja la opción **d** podría provocar diferencias de criterio en aquellas cachés que se encuentran fuera del chip procesador pero dentro de su placa de circuito impreso, esto es, las que el libro clasifica como cachés *internas*. Sin embargo, la L1 del K6-2 está integrada dentro del chip de la CPU y por tanto es propia del procesador, mientras que su L2 se encuentra siempre en placa base y es externa al mismo.
26.  d
27.  b
28.  d En preguntas como ésta, se trata de elegir la menos falsa de las cuatro opciones que se dan.
29.  d El Pentium II no innova en el conjunto de instrucciones, puesto que copia las extensiones MMX a secas. El K6-2 sí innova, con sus extensiones 3DNow!, y el primer Pentium III también, con sus extensiones SSE. Respecto a

las mejoras en caché, el primer Pentium III es un calco del último Pentium II, y sólo el de 0.18 micras (Coppermine) incluye la caché L2 integrada de 256 Kbytes.

30.  a El Pentium II ya incluye las 57 del MMX y por tanto, aumenta el conjunto respecto al Pentium clásico. Por otra parte, la aclaración entre paréntesis de la pregunta descarta que el modelo Pentium pueda referirse de forma implícita a la versión MMX de este procesador.
31.  c Veamos las diferencias opción por opción.
- En a, la diferencia es la incorporación del conjunto de instrucciones MMX.
  - En b, las diferencias son dos: (1) La ampliación del conjunto de instrucciones con las 24 instrucciones del conjunto 3DNow!, y (2) En algunos modelos, el bus de 100 MHz.
  - En d, estamos ante dos arquitecturas completamente diferentes.
- Todas las diferencias enumeradas introducen cambios en el sistema de decodificación de instrucciones de un procesador, pero no así los cambios realizados en la caché L2 (opción c), que no trascienden sobre el decodificador de instrucción del procesador.
32.  d Es la única pareja de microprocesadores dada en la que cada elemento pertenece a una generación diferente.
33.  d
34.  b En efecto, la memoria caché es con diferencia la unidad funcional que más transistores requiere en su implementación (más de diez millones por cada 256 Kbytes de caché).
35.  c Descartamos el resto por los siguientes motivos:
- Opción a: El K6 sigue teniendo la caché L2 externa.
  - Opciones b y d: Tanto el primer Pentium II como el primer Athlon parten ya de caché L2 interna, no externa.

## Capítulo 6

1.  b Un nuevo proceso de fabricación permite acometer grandes cambios arquitecturales para aprovechar el mayor número de transistores disponibles y la mayor velocidad a que éstos conmutan. Las compañías suelen sacar los nuevos modelos al final de un proceso de fabricación para que la inminente llegada del nuevo proceso de fabricación pueda ser empleado en la subsanación de las numerosas vías de agua que se le descubren a toda arquitectura novel en su fase inicial de rodaje. Respecto al resto de opciones, las descartamos por las siguientes razones:
- Opción a: Sí es un buen momento para comprar, pero no el producto recién estrenado, sino el último de la generación anterior, que ha sufrido una enorme depreciación en su precio final, y que tras sus sucesivas mejoras ha alcanzado un elevado grado de madurez en su estado de optimización. Los productos en su fase inicial de comercialización tienen un precio sobredimensionado y cierta probabilidad de pagar su inmadurez con errores o derivando hacia callejones sin salida donde quedaríamos atrapados.
  - Opción c: Lo normal es que el último producto de la generación precedente termine su ciclo de vida bastante por encima de la frecuencia a que se estrenó el nuevo, ya que según lo expuesto en la opción b, pronto llega un nuevo proceso de fabricación que permite aumentar la frecuencia, y éste se aprovecha no sólo en los modelos nuevos, sino en las últimas remesas de los viejos.
  - Opción d: Es correcto que el ciclo de desarrollo se estima en unos seis años, pero el tiempo real que transcurre entre generaciones se ha acortado a la mitad porque las compañías como Intel o AMD ponen a trabajar a sus ingenieros en arquitecturas de dos generaciones diferentes de forma solapada en el tiempo.
2.  c
3.  a
- En las opciones b y d es incorrecto el último valor de la secuencia.

- ⊗ En la opción **c**, es incorrecta la secuencia completa, ya que no ha habido evolución alguna al respecto de la superescalaridad en el contexto de la séptima generación.
4. **d** En efecto, ese modelo en concreto ha sido muy heterogéneo habiendo sido fabricado tanto con tecnología de cobre como de aluminio, y tanto a 0.25 micras como a 0.18 micras.
5. **c**
6. **b** De los 10 ciclos que tarda en ejecutarse una instrucción entera, 5 de ellos se dedican a la decodificación en un complejo proceso en el que se suceden tareas de predecodificación, alineamiento, macrodecodificación y microdecodificación.
7. **c**
8. **c** La arquitectura del Duron es la misma que la del K7, y tiene muy poco que ver con la del K6. Entre el K7 y el Duron existe una diferencia en la frecuencia de bus en algunos modelos y una diferencia en el tamaño de caché L2 en *todos* los modelos.
9. **a**
10. **b** Hasta después de la fase de decodificación es imposible separar el tratamiento de las instrucciones en el cauce de segmentación, ya que se desconoce el tipo de instrucción de que se trata. Por otra parte, la búsqueda de operandos requiere ya la separación de las instrucciones, puesto que éstas se alimentan de bancos de registros diferentes.
11. **c** El Thunderbird modifica la caché L2 de interna a integrada, con un efecto mayúsculo en el rendimiento. El Athlon XP sólo produce leves retoques en la prebúsqueda en cachés y TLB y la inclusión de un nuevo conjunto de instrucciones multimedia, 3DNow! Professional (la caché L2 de 512 Kbytes llega en el último modelo, Barton, y se pide comparar el primero de ellos). El K8 es ya una nueva arquitectura, y por lo tanto, debemos catalogarlo como de octava generación.
12. **a** La BTB contiene en cada una de sus entradas: El código de instrucción de salto del programa fuente (compatible x86), su comportamiento pasado (si saltó o no las últimas N veces), y la dirección de salto correspondiente. Es el primero de estos campos de información el único que está obligatoriamente expresado con códigos de instrucción x86, y que dada su heterogeneidad en el formato de instrucción sería imposible respetar en nuevos códigos de instrucción con un mínimo de elegancia.  
Las demás opciones se descartan por los siguientes motivos:  
⊗ Opción **b**: El BRC aloja instrucciones nativas, puesto que éstas ya han pasado por las fases de búsqueda y decodificación, y por lo tanto, sirve para la nueva arquitectura.  
⊗ Opción **c**: El TLB sólo contiene direcciones de memoria, y en la medida en que el bus de direcciones se mantenga en 32 bits, puede reaprovecharse en la nueva arquitectura.  
⊗ Opción **d**: El banco de registros y las unidades aritmético-lógicas pueden reutilizarse si se utilizan formatos similares para los operandos de las nuevas instrucciones. Además, puede extenderse el formato a otros nuevos que vayan implementados aparte sin renunciar a la utilización de las viejas unidades funcionales.
13. **a**  
⊗ La opción **b** es falsa porque la caché L2 no va a 100 MHz.  
⊗ La opción **c** refleja algo que le ocurrió al Duron de AMD, no al Pentium 4 de Intel.
14. **c** Existen técnicas para tener en cuenta los retardos en el transporte de la señal que permiten sincronizar cantidades ingentes de transistores a frecuencias muy elevadas. Esto descarta la opción **b** y a la vez da cierta vigencia a la opción **a**, que podría ser cierta en algunos procesadores, aunque no lo es en el caso particular que nos ocupa.
15. **d** En concreto, con el juego de chips i850 (alias Tehama) se proporcionó inicialmente el interfaz con RDRAM, mientras que con el posterior Brookdale se le dotó de soporte para DDRAM.
16. **b** Esa configuración se repitió tanto en los Pentium III y K7 que había en el mercado cuando salió el Pentium 4, sino también en este último.

- La presencia de L3 relatada en la opción **a** sólo se da en el modelo Xeon, y su tamaño es muy superior al indicado.
- La opción **c** no es correcta porque la L1 apenas ha seguido una uniformidad entre las diferentes casas comerciales, cambiando tanto su capacidad como su organización.
- Respecto a la opción **d**, ningún modelo llegó a alcanzar tal capacidad para su caché L1.
17.  **d**
18.  **d**
19.  **c** Efectivamente, con una decodificación progresiva es posible implementar un esquema de traza múltiple que se encargue de decodificar cada uno de sus pasos intermedios.
20.  **c** La opción **b** mitiga la presencia de dependencias de datos, pero es una solución menos eficaz que la que damos por válida.
21.  **a** Si se mantiene el patillaje es porque los buses de datos y direcciones y las líneas de control son idénticas al Pentium 4 que conocemos, por lo que lo más lógico es mantener las mismas formas de diálogo. Ahora bien, el formato de instrucción que recibiremos por el bus de datos será diferente y deberá decodificarse de forma distinta, lo que repercute en cambios sobre la caché de traza, descartando la opción **b**. Los bancos de registros o las ALU del procesador sólo requerirían cambios si en el nuevo repertorio de instrucciones se definieran nuevos operandos u operaciones, respectivamente, descartando en ese caso las opciones **c** o **d**.
22.  **b** Las razones para eliminar el resto de opciones son las siguientes:
- Opción **a**: Aunque la caché de traza produce un buen rendimiento y consume muchos transistores, no podemos decir lo mismo de una escasa caché L1D de tan sólo 8 Kbytes y de una caché L2 de 256 Kbytes.
- Opción **c**: Con unas instrucciones escindidas en 20 etapas, difícilmente podemos pensar en su simplicidad.
- Opción **d**: El diseño superescalar no es algo que caracterice al Pentium 4, pues conserva el mismo factor tres de sus hermanos menores de sexta generación.
23.  **b**
24.  **c** En Intel, la L1 se queda en 32 Kbytes en total (a lo sumo), mientras que en AMD es cuatro veces superior.
25.  **c**
26.  **d**
27.  **c**
28.  **c**
29.  **a** Con esta fórmula resultan los valores 5, 10 y 20, muy similares a los correctos: 5 (para el Pentium, 6 si es la versión MMX), 11 (para los Pentium II y III de sexta generación) y 20 (para los Pentium 4 de séptima generación). Del resto de fórmulas resultan los siguientes valores:
- Opción **b**: 5, 10 y 15.
- Opción **c**: 6, 9 y 12.
- Opción **d**: 5, 11 y 17.
30.  **d**
- Opción **a**: Rechazada porque el cambio de tipo de numeración viene propiciado por motivos de marketing y no refleja metamorfosis arquitectural alguna.
- Opción **b**: El espacio de tiempo que separa dos modelos de diferente arquitectura no es significativo, puesto que aunque un refinamiento de una arquitectura existente tarda mucho menos en implementarse que una arquitectura nueva, los sucesivos refinamientos de una arquitectura vieja se van escalando en el tiempo de forma estratégica por las compañías para entretener al mercado mientras se desarrolla la arquitectura nueva, proporcionando así la ilusión de una progresión tecnológica continua en el tiempo.
- Opción **c**: No es correcta, pues en los dos casos (Coppermine y Willamette) se trataba de las 0.18 micras.

31.  d
32.  a
33.  c
- En la opción **a**, los Pentium II y Pentium III se solapan a 0.25 micras (Deschutes y Katmai, respectivamente), y los Pentium III y 4 lo hacen también a 0.18 micras (Coppermine y Willamette, respectivamente).
  - En la opción **b**, Deschutes y Katmai también comparten los 512 Kbytes de caché interna, y Coppermine y Willamette los 256 Kbytes de caché integrada, quedando inédita la L3 salvo en los modelos Xeon de la gama servidora.
  - Respecto a la opción **d**, la frecuencia de reloj del Pentium II tan sólo alcanza los 450 MHz (aunque los Pentium III y 4 sí superaron los 1000 y 1500 MHz, respectivamente).
34.  b Los rasgos que se indican en el enunciado de la pregunta llevan el sello distintivo de AMD. Nos decantamos por su sexta generación porque aunque en su séptima generación la fase de reordenación/ejecución sigue siendo más rápida que la de búsqueda/decodificación, la incorporación de una gran caché de primer nivel acelera ésta última frente a la arquitectura del K6.
35.  d El Pentium III con la transición del Katmai al Coppermine, y el K7 con la transición del Athlon al Thunderbird.
36.  b
- Las opciones **a** y **c** no se cumplen en algunos de sus modelos como el Celeron.
  - La opción **d** no se cumple en casi ninguno (Intel da más calidad que cantidad en sus cachés).
  - Además, en los aspectos que recogen las opciones **a** y **d**, Intel no es precisamente una empresa modelo).
37.  d
38.  b
39.  c
40.  a En efecto, ambas compañías lo lograron en un mismo marco temporal (Febrero de 2000), pero al tratarse en ambos casos de diseños experimentales, resulta difícil establecer objetivamente un ganador para esa carrera.
41.  b En el aspecto de la frecuencia, AMD ha ido casi siempre por detrás de competidores como Intel. En el precio, ha ido casi siempre por delante, al margen de que nunca ha sido una compañía con productos caros. En cambio, en la fecha de comercialización, entre el K5, que por retrasarse en su salida apenas se vende, y el K7, que por adelantarse supone un rotundo éxito, el cambio es muy sustancial.
42.  b
- El bus AST es una mera invención del autor.
  - La memoria RDRAM perjudicó a Intel en los inicios del Pentium 4 porque se apostó exclusivamente por ella y encarecidísimo el coste de los equipos, dificultando las ventas del microprocesador. Además, el primer procesador que realmente aprovechaba las prestaciones de la RDRAM era el K7, y en sus inicios, rivalizaba con un Pentium III en el que su lento bus local no sacaba provecho del elevado ancho de banda de la RDRAM.
43.  a
44.  a Pocos transistores y una buena integración son los dos principales ingredientes para lograr un calentamiento exiguo, y con éste, ya tenemos los tres puntales necesarios para conseguir una frecuencia de ensueño.
45.  a El factor de superescalaridad de 10 supone disponer de 10 copias de buena parte de las unidades funcionales de la arquitectura base. Eso nos deja apenas un millón de transistores para ella, un valor ridículo en séptima generación.



46. **a** Todo apunta a una frecuencia muy elevada, y con un escaso número de transistores, el paralelismo debe explotarse por la vía de la segmentación frente a la de la superescalaridad. Y con diez etapas, aún queda margen para aumentar ese valor pensando que estamos ubicados en séptima generación, donde suele simultanearse la ejecución de 50-60 instrucciones tomando K7 y Pentium 4 como referencia.
47. **c** Descartamos el resto por los siguientes motivos:
- ⊗ Opción **a**: Con 10 millones de transistores, poco más puede dar de sí la arquitectura, sabiendo además que las cachés se llevan un gran número de ellos.
  - ⊗ Opción **b**: Las cachés de primer nivel han sido siempre integradas y resulta incongruente ubicarlas fuera del chip procesador.
  - ⊗ Opción **d**: La segmentación y la superescalaridad se encuentran ligadas muy débilmente al tamaño de las cachés del sistema.
48. **a** Teniendo a su alcance una frecuencia de 10 GHz, tenemos una mejora cinco veces superior a los valores medios del rango de frecuencias del K7. Ninguno de los otros tres aspectos presentados aquí pueden lograr semejante grado de mejora, por los siguientes motivos:
- ⊗ Opción **b**: 10 millones de transistores son pocos desde el punto de vista del coste de la materia prima, pero la maquinaria necesaria para su integración es tremendamente cara, lo que dispararía el coste unitario para una PIME (pequeña empresa) que no tiene a su alcance la fabricación de un gran volumen de unidades.
  - ⊗ Opción **c**: El paralelismo a nivel de instrucción tiene también una cara positiva, la segmentación, y otra negativa, la superescalaridad. Con ésta última limitada por la paupérrima disponibilidad de transistores en el Decaomnium, y teniendo el K7 en la superescalaridad una de sus mejores virtudes, resulta inviable crear un diseño compensado que en el cómputo global del paralelismo pueda ganar de forma clara al K7.
  - ⊗ Opción **d**: Las cachés de primer nivel del K7 son muy superiores a las propuestas para el Decaomnium en función de la limitación a 10 millones de transistores.
49. **a** El Pentium 4 dispone de modelos comerciales de 0.18 y 0.13 micras, mientras que el Decaomnium se va a las 0.10 micras de salida. Además, en las otras tres opciones apenas presenta ventajas frente al Pentium 4 por las siguientes razones:
- ⊗ Opción **b**: El Pentium 4 dispone de partida del doble de etapas de segmentación que el Decaomnium, y aunque éste aún dispone de la baza de una eventual supersegmentación que es factible, difícilmente sobrepasaría al Pentium 4 en la que es una de sus apuestas más claras.
  - ⊗ Opción **c**: En el factor de superescalaridad el Pentium 4 es deficiente, pero el Decaomnium lo es más por contar con un número de transistores tan bajo.
  - ⊗ Opción **d**: Con la información de que disponemos, no es posible inferir apenas elementos de juicio como para discernir si el conjunto de instrucciones del Decaomnium es mejor o peor que el del Pentium 4.
50. **b** No podemos aumentar alegremente el número de etapas de segmentación precisamente porque, según reza la opción **a**, las dependencias arruinan la eficiencia al no poder disponer de mecanismos de cobertura. Diez millones de transistores y diez etapas de segmentación son valores consecuentes el uno con el otro, y la elección mala son las micras por ser excesivamente pretenciosa para una compañía novel. Además, las opciones **a** y **c** incluyen coletillas falsas:
- ⊗ Opción **a**: La construcción de rápidas ALU es una de las cosas a las que menos entorpece el que contemos con pocos transistores, primero porque suele ser la unidad funcional que más fácilmente alcanza elevadas velocidades y segundo porque es la que menos transistores demanda.
  - ⊗ Opción **c**: Las 0.10 micras plantean el mismo problema tecnológico para 10 millones de transistores que para 100.
51. **a** Aunque no nos indican nada acerca del conjunto de instrucciones, con 10 millones de transistores no tenemos mucho margen para diseñar instrucciones complejas, y si para considerar la simplicidad como el abanderado de nuestro diseño en todos los niveles del procesador.
52. **a** Pocos transistores requieren poca alimentación para el conjunto. Distancias de integración cortas demandan una menor energía para que la excitación de los electrones los haga atravesar la puerta del transistor, que es el elemento que controla su conmutación. Por último, una frecuencia elevada aumenta el número de conmutaciones de cada transistor, y cada una de estas operaciones consume una fracción de energía del procesador (es, en definitiva, lo que hace que el chip se caliente).
53. **d** Descartamos el resto por los siguientes motivos:
- ⊗ Opción **a**: No se puede responder afirmativamente de manera incondicional, ya que todo está mediatizado por

cómo se haya diseñado el conjunto de instrucciones actual.

⊗ Opción **b** : Existen abundantes ejemplos en el mercado de procesadores RISC con aditivos multimedia en su conjunto de instrucciones (familias como los R# de SGI, los PA# de HP y los PowerPC de Motorola), que demuestran que ambas ideas son perfectamente conjugables.

⊗ Opción **c** : El cauce de segmentación multimedia suele desligarse del cauce de segmentación entero y de punto flotante.

54. **b** Efectivamente, es el único que puntúa en la casilla correspondiente a la L1 de traza. La siguiente Tabla resume la puntuación obtenida por cada procesador en cada uno de los apartados analizados:

Nombre del procesador	Caché L1				Caché L2						Puntuación	
	Integrada		Traza		Integrada		Interna		Externa		Suma	TOTAL
	Kb.	Punt.	Kb.	Punt.	Kb.	Punt.	Kb.	Punt.	Kb.	Punt.		
Thunderbird	128	8			256	16					$8 \times 3 + 16$	40
Tualatin	32	2			512	32					$2 \times 3 + 32$	38
Willamette	8	0	14	7	256	16					$7 \times 3 + 16$	37
Primer Athlon	128	8					512	8			$8 \times 3 + 8$	32
Duron	128	8			64	4					$8 \times 3 + 4$	28
Sharptooth	64	4			256	16					$4 \times 3 + 16$	28
Coppermine	32	2			256	16					$2 \times 3 + 16$	22
Katmai	32	2					512	8			$2 \times 3 + 8$	14
Deschutes	32	2					512	8			$2 \times 3 + 8$	14
Klamath	32	2					512	8			$2 \times 3 + 8$	14
Primer K6	64	4							SI	1	$4 \times 3 + 1$	13
Celeron	32	2							SI	1	$2 \times 3 + 1$	7

55. **d** Sólo cuenta con 8 Kbytes de L1 integrada que no llegan al mínimo requerido para puntuar, que está establecido en 16 Kbytes, y que sí alcanzan todos los demás.
56. **b** Uno es de sexta y otro de séptima generación.
57. **b** Los tres presentan la misma arquitectura interna, incluyendo las cachés.
58. **b** Se trató del primer procesador para la gama PC en llegar a los 512 Kbytes de caché L2 integrada.
59. **b**
60. **d** De los cuatro que superan los 30 puntos y que tienen en la mano superar los 40 puntos del ganador, es nuevamente el líder el más equilibrado.
61. **d** Ambos tienen una L1 integrada de 32 Kbytes en total y una L2 externa.
62. **b** Ambos presentan la misma configuración de caché, tanto en L1 como en L2.
63. **c** El Tualatin es el único modelo de 0.13 micras y no le sirve para ganar. Respecto a los modelos de 0.35 micras, tenemos a los primeros Celeron, K6 y Pentium II en la lista de nuestro ránking.

1. **c**
  - ⊗ La opción **b** elimina la imprecisión de la opción **a**, pero tampoco es válida porque ése es precisamente el perfil del Pentium, y éste no es de 64 bits, sino de 32 bits.
  - ⊗ La opción **d** tampoco puede ser, puesto que no existe en el mercado una configuración de tales características, y en cualquier caso, el bus de datos debería disponer de al menos 64 líneas.
2. **d** Atendiendo a la definición que acuñamos en el libro.
  - ⊗ La opción **a** es una solución recurrente que no podemos dar como válida.
  - ⊗ La opción **b** no es correcta, pues entonces el Pentium MMX sería una arquitectura de 80 bits solo porque sus operandos lo son.
  - ⊗ La opción **c** tampoco sirve, ya que en ese caso el Pentium 4 sería una arquitectura de 128 bits simplemente por gozar de la existencia del banco de registros XMM de esa longitud (el que utiliza el conjunto de instrucciones multimedia SSE2).
3. **d** Les diferencia fundamentalmente el número de enlaces HyperTransport (sólo uno en el Athlon 64 por tres en el Opteron), pero el núcleo de la arquitectura es muy parecido en ambos casos. Los otros tres procesadores son de una generación anterior, diferenciándoles cambios más que notables.
4. **b** Aunque el bus de datos tiene 128 bits, se montan dos carriles de 64 bits, ya que no hay memoria DDRAM de 128 bits.
5. **d** En efecto, 2x166 MHz en el caso de la memoria, 2x800 en el acceso a otros procesadores, y 2x200, 2x400 o 2x800 en el acceso a placa base, siempre dependiendo de la versión del juego de chips de que se disponga.
6. **d** Las opciones **a** y **c** son ambas correctas, y lo normal es que la anchura del enlace HyperTransport que conecta con AGP tenga una anchura de 32 bits, en vista de que esa anchura coincide con todas las versiones de AGP entre la 1.0 y la 3.0 actual.
7. **a** Los dos procesadores se comportan igual bajo código de 32 y 64 bits.
8. **c** Es la estrategia de AMD para oponerse al miedo al cambio de los usuarios después de 25 años acostumbrados a la compatibilidad hacia atrás.
9. **b** Coincide en la superescalaridad, y sólo alarga los cauces de segmentación en dos etapas. Respecto a los descartes, la frecuencia inicial del K7 fue de 500 MHz, mientras que en la jerarquía de memoria hay notables diferencias en el banco de registros.
10. **c** Efectivamente, las respuestas están listadas en orden de menor a mayor influencia en el patillaje, con la salvedad de que el tamaño de caché L2 integrada no afecta en nada al patillaje funcional del chip (ya hemos descontado la partida correspondiente al suministro de corriente, según reza el enunciado). Entre las opciones **b** y **c**, ésta última es más correcta porque influye mucho más.

## Ejercicios suplementarios

### ☆ Relacionar la dimensión RISC/CISC de un microprocesador

Las diez letras que llenan los diez casilleros son, respectivamente: R, C, R, R, C, R, R, C, R, C.

### ☆ Clasificar las mejoras del procesador Pentium

La correspondencia sería:

- a con 1.
- b con 6.
- c con 10.
- d con 2.
- e con 3.
- f con 7.
- g con 5.
- h con 4.
- i con 8.
- j con 9.

# Glosario conceptual

Un número de página subrayado localiza una definición, en *cursiva*, una explicación, y normal, una mención.

Además, hemos utilizado la siguiente simbología en los términos del glosario:

- I** : Precede a todo vocablo en inglés.
- III** : Serigrafía en placa base, chip, ...
- ≈: Ente software (instrucción, programa).
- Ⓢ: Ente hardware (pin, chip, computador).

## Alfanuméricos

área de integración, 279

**I**3DNow

y sus variantes, véase glosario comercial

## A

abstracción (del cto. de instrs.), 94

**I**AFR66 **III**, 177

amigabilidad (del cto. de instrs.), 95

**I**API ≈, 113

multimedia más relevantes, 113

área de integración, 51

**arquitectura de 64 bits**, 260

en los libros, 260

entre los usuarios, 260

para las compañías, 260

asimilación (del cto. de instrs.), 95

Athlon, véase K7 en glosario comercial

Athlon 64, véase K8 en glosario comercial

atto-, 121

## B

búfer de prebúsqueda de instrucciones/datos, 73

**I**backside bus, véase bus trasero

**banco de registros**, 73

anchura, 73

tamaño, 73

**I**BASIC ≈, 27

bipolar (transistor), 52

**I**BRC (**B**uffer de **R**eordenación **C**ircular), 156

en el K7, 197

en el Pentium 4, 222

**I**BTB (**B**ranch **T**arget **B**uffer), 69

acierto en su acceso, 69

en el K7, 198

en el Pentium 4, 240

fallo en su acceso, 69

funcionamiento, 69

información que contiene, 69

bus

frontal, 75

trasero, 75

## C

caché, véase memoria caché

camino crítico, 57

carga retrasada, 99

cauce-u (del Pentium), 137

cauce-v (del Pentium), 137

Celeron, véase glosario comercial

chip (nacimiento), 10

**I**Chip MultiProcessing (CMP), 281

**I**CISC, 96

versus RISC, 97

clónicos, 28

**I**CMOS, 52

**I**code inlining, 103

**compilador**

soporte para instrs. multimedia, 112

**conjunto**

organización en BTB y TLB, 158

**conjunto de instrucciones**, 48

amigabilidad del interfaz, 95

construcción del, 98

historia pendular, 95

nivel de abstracción, 94

previsiones futuras, 282

rapidez en su asimilación, 95

riqueza del lenguaje, 95

conjunto de trabajo, 111

**I**CPU, véase microprocesador

**I**critical word first, 79

## D

**dependencias**

anticipación, 69

de control, 65

de datos, 65

estructurales, 66

penalización, 66

soluciones, 66

**I**DIB (Dual Independent Bus), 77

directorio caché, 85

**I**DirectPath, 196

distancia de integración, 52

Duron, véase glosario comercial

## E

**I**early restart, 79

ejecución especulativa, 68

**ejecución fuera de orden**

complejidad, 67

concepto, 67

riesgos, 67

verificaciones, 68

**ensamblador**

soporte para instrs. multimedia, 112

**I**EPROM (nacimiento), 17

estación de reserva, 156

exa-, 121

exbi-, 121

extensiones multimedia, véase instrucciones multimedia

## F

familia 68000, 37

familia 80x86, 37

fempto-, 121

**I**FPU, 284

**frecuencia de reloj**

errores estimativos, 278

estereotipos, 50

forma, 49

magnitud inversa, 49

magnitudes, 49

origen, [49](#)  
 previsiones futuras, [278](#)  
 ubicación del oscilador, [49](#)  
**I**frontside bus, [véase](#) bus frontal

## G

gibi-, [121](#)  
 giga-, [121](#)

## H

**H**high end (gama de usuario PC), [166](#)  
**H**yper-pipelined, [219](#)  
**H**yperThreading, [281](#)

## I

instrucciones de relleno, [99](#)  
**I**nstrucciones multimedia  
 estandarización, [118](#)  
 falso reclamo, [276](#)  
 infrautilización, [119](#)  
 operandos, [106](#)  
 rasgos hardware, [111](#)  
 rasgos software, [111](#)  
 relación con el API DirectX, [118](#)  
 resumen, [118](#)  
 invariantes, [103](#)

## J

juego de chips, [85](#)

## K

K5, [véase](#) glosario comercial  
 K6, [véase](#) glosario comercial  
 K7, [véase](#) glosario comercial  
 K8, [véase](#) glosario comercial  
 kibi-, [121](#)  
 kilo-, [121](#)  
**I**KNI, [véase](#) SSE en glosario comercial

## L

**L**ey  
 de Moore, [12](#), [278](#), [286](#)  
 litografía, [279](#)  
**L**ow end (gama de usuario PC), [166](#)

## M

macroinstrucción, [195](#)

**M**ainframes, [24](#)  
 mebi-, [121](#)  
 mega-, [121](#)  
**m**emoria caché  
 área de control, [85](#)  
 con lectura anticipada, [79](#)  
 de primer nivel ó L1, [73](#)  
 de segundo nivel ó L2, [73](#)  
 de tercer nivel (L3), [80](#)  
 en los años 80, [72](#)  
 en los años 90, [72](#)  
 espacio en silicio, [75](#)  
 externa, [82](#)  
 integrada, [82](#)  
 interna, [82](#)  
 jerarquía, [75](#)  
 no bloqueante, [77](#)  
 nomenclatura, [82](#)  
 origen, [72](#)  
 previsiones futuras, [282](#)  
 relación con los buses, [84](#)  
 segmentada, [78](#)  
 víctima, [79](#)  
 zócalo opcional, [85](#)  
**m**emoria de microprograma, [97](#)  
 ralentización, [97](#)  
 versatilidad, [97](#)  
 memoria virtual, [39](#)  
 micra, [52](#)  
 micro-, [121](#)  
 microinstrucción, [195](#)  
**m**icroprocesador  
 definición, [48](#)  
 influjo empresarial, [35](#)  
 Unidad de Control, [véase](#) en la U  
 Unidad de Proceso, [véase](#) en la U  
**M**id end (gama de usuario PC), [166](#)  
 mili-, [121](#)  
 minicomputadores, [24](#)  
**M**IPS, [283](#)  
**M**MMU (Memory Management Unit),  
[39](#)  
**M**MX y sus variantes, [véase](#) glosario  
 comercial  
**M**MOS, [51](#)  
**M**ountain View, [8](#), [13](#)  
 multimedia  
 instrucciones, [véase](#) en la I  
 multiplicador de frecuencia, [133](#)


## N

nano-, [121](#)  
**n**ivel de caché  
 omisión, [88](#)  
 relevancia, [91](#)  
 nueva economía, [14](#)  
**N**UMA (Non-Uniform Memory Ar-  
 chitecture), [268](#)

## O

Opteron, [véase](#) K8 en glosario comercial

## P

Palo Alto, [14](#)  
**p**aralelismo a nivel de instrucción  
 frente a paralelismo de datos, [108](#)  
 idea, [58](#)  
 previsiones futuras, [280](#)  
**P**C  
 origen, [7](#)  
 pebi-, [121](#)  
 Pentium, [véase](#) glosario comercial  
 Pentium 4, [véase](#) glosario comercial  
 Pentium II, [véase](#) glosario comercial  
 Pentium III, [véase](#) glosario comercial  
 período de reloj, [49](#)  
 peta-, [121](#)  
 pico-, [121](#)  
**P**ipelining, [véase](#) segmentación  
 placa base, [48](#)  
**P**R (Pentium Rated), [203](#)  
**P**PR-XXX , [183](#)  
**p**redicción de salto  
 comportamiento adaptativo, [69](#)  
 concepto, [68](#)  
 dinámica, [69](#)  
 estática, [68](#)  
 saltos incondicionales, [69](#)

## Q

**q**uinta generación  
 amigable interfaz, [134](#)  
 bajo coste, [134](#)  
 frecuencia, [133](#)  
 funcionalidad, [134](#)

## R

**r**eloj  
 oscilador, [49](#)  
 renombrado de registros, [156](#)  
 repertorio de instrucciones, [véase](#) con-  
 junto de instrucciones  
 riqueza (del cto. de instrs.), [95](#)  
**R**ISC, [96](#)  
 inconvenientes, [97](#)  
 ventajas, [97](#)  
 versus CISC, [97](#)  
**r**itmo evolutivo

de la computación de punto flotante, [284](#)

de la computación entera, [283](#)

de la supercomputación, [285](#)

■ **ROM** de microprograma, [196](#)

## S

saltos retrasados, [100](#)

■ **SECC**, [167](#), [174](#)

■ **SECC2**, [165](#), [174](#)

**segmentación**

concepto, [59](#)

coste, [59](#)

etimología, [59](#)

futuro de la, [281](#)

rendimiento, [59](#)

■ **SEPP**, [167](#), [174](#)

**séptima generación**

asincronismo temporal, [191](#)

conductas previas, [191](#)

enfrentando a AMD e Intel, [191](#)

ese alumno aventajado, [191](#)

un estandarte por bando, [192](#)

visión de conjunto, [192](#)

servidor, [150](#)

**sexta generación**

grandes cachés en el interior del chip, [150](#)

uso extensivo del PNI, [150](#)

■ **Silicon Valley**, [8](#), [14](#)

■ **SIMD**, [108](#)

funcionamiento, [108](#)

monoprocesador, [108](#)

origen, [108](#)

paralelismo de datos, [108](#)

sincronismo, [108](#)

sistemas empotrados, [24](#)

■ **Slot 1**, [162](#), [168](#), [174](#), [193](#)

■ **Slot 2**, [166](#)

■ **Slot A**, [162](#), [193](#)

■ **SMP (Symmetric MultiProcessing)**, [42](#)

en los años 70, [24](#)

en séptima generación, [229](#)

en sexta generación, [150](#)

■ **SMPT**, [229](#)

concepto, [229](#)

estrategias, [230](#)

ilusión, [230](#)

penalización, [231](#)

rendimiento, [231](#)

■ **Socket 370**, [168](#)

■ **Socket 5, 7, 8, 9 y 370**, [174](#)

■ **SOI (Silicon-On-Insulator)**, [258](#)

■ **SSE** y sus variantes, véase glosario comercial

■ **Sunnyvale**, [13](#)

supercomputador, [285](#)

**superescalaridad**

concepto, [60](#)

coste, [61](#)

futuro de la, [280](#)

rendimiento, [60](#)

**supersegmentación**

definición, [63](#)

en el análisis K7 vs. P4, [239](#)

relación con

frecuencias elevadas, [63](#)

superescalaridad, [63](#)

■ **Symbol** ☞, [15](#)

## T

tebi-, [121](#)

**tecnología de integración**

definición, [50](#)

efectos laterales, [56](#)

importancia, [50](#)

influencia sobre

coste de fabricación, [55](#)

frecuencia, [55](#)

número de transistores, [54](#)

potencia disipada, [55](#)

temperatura, [55](#)

voltaje, [55](#)

plantas de fabricación, [53](#)

previsiones futuras, [279](#)

tera-, [121](#)

■ **THERMTRIP** ☞, [205](#)

Thunderbird, véase glosario comercial

**tipo de caché**, [91](#)

omisión, [88](#)

■ **TLB**, [89](#), [103](#)

en el Athlon XP, [206](#)

en el K7, [195](#)

**transistor**

de germanio, [8](#)

de silicio, [8](#)

MOS, [50](#)

## U

Unidad de Control, [48](#), [97](#)

Unidad de Proceso, [48](#), [97](#)

## V

vectores de interrupción, [104](#)

■ **VectorPath**, [196](#)

ventana de relleno, [99](#)

■ **VLIW (Very Long Instruction Word)**, [104](#)

cualidades, [106](#)

idea, [104](#)

misión del compilador, [104](#)

origen, [104](#)

## X

XMM8-15 ☞, [265](#)

## Y

yocto-, [121](#)

yotta-, [121](#)

## Z

zepto-, [121](#)

zetta-, [121](#)

■ **ZIF (Zero Insertion Force)**, [174](#)

zócalo del microprocesador, [133](#)





# Glosario comercial

Se ha empleado la siguiente simbología para los términos incluidos en este glosario:

- ☉® : Un hardware comercial (procesador, chipset, tarjeta, ...)
- ≈® : Un software comercial (programa, sistema operativo, ...)
- ☉™ : Una marca hardware registrada.
- ≈™ : Una marca software registrada.
- ⌠ : Una casa o firma comercial.
- ☒ : Código de referencia interno de un procesador ó juego de chips.

## Alfanuméricos

- 3DNow ≈™, 115
  - API para, 113
  - base, 115
  - contexto histórico, 95, 114
  - en el K6-2, 178
  - en procesadores Cyrix, 245
  - falso reclamo, 276
  - funcionalidad, 116
  - implementación, 115
  - rango de operandos, 115
  - sus 21 instrucciones, 116
- 3DNow Professional ≈™, 117, 264
  - 72 instrucciones, 117
  - API para, 113
  - en el Athlon XP, 207
  - en el K8, 265
- 3Dfx Interactive ⌠
  - API para MMX de, 113
- 5x86 ☉®, 146
- 6x86 ☉®, 146
- 6x86 de Cyrix ☉®, 183
- 6x86L ☉®, 183
- 6x86MX ☉®, 183
- 80386DX ☉®, 36
- 80386SX ☉®, 36
- 80486DX ☉®, 36
- 80486SX ☉®, 36

## Números

- 3708 ☉®, 16
- 4001 ☉®, 15
- 4002 ☉®, 15
- 4003 ☉®, 15
- 4004 ☉®, 36, 38
  - características, 17
  - inicio, 15
  - la serie, 15
- 6800 ☉®, 24
- 8008 ☉®, 21, 36, 38
  - definición, 19
  - origen, 19
- 8080 ☉®, 21, 36
- 8086 ☉®, 26, 36–38
  - nacimiento, 25
- 8088 ☉®, 36
  - nacimiento, 25
- 68000 ☉®, 21, 38
- 68020 ☉®, 37
- 68030 ☉®, 37
- 68040 ☉®, 37, 72
- 80286 ☉®, 28, 36, 37
- 80386 ☉®, 28, 37
- 80486 ☉®, 37, 72
  - tipo de caché, 81

## A

- AGTL+ ☉™, 193, 211
- Alpha ☉®, 267
- Alpha 21164 ☉®, 78
- Alpha 21264 ☉®, 57, 63, 72, 78, 118
- Alpha 21364 ☉®, 63
- Altair ☉®, 23, 27
- AltiVec ≈™, 118
- Alto ☉®, 14, 21
- AMD ⌠, 12, 14, 43, 115, 118, 134, 191, 257, 258, 266–268, 284
  - fundación, 13
- Apple ⌠, 21, 24, 38, 40
- Apple II ☉®, 24

Araña ☼, 273  
 Argon ☼, 191, 192  
 Arquitectura del PC ≈<sup>TM</sup>, 192, 255, 258  
 ASCII White ☉®, 285  
 ASML ⤴, 54  
 Asus ⤴, 267  
 Athens ☼, 256, 257  
 Athlon, véase K7  
 Athlon 64, véase K8  
 Athlon XP ☉®, 202, 260  
   API para, 118  
   encapsulado, 205  
   extensiones multimedia, 207  
   fabricación, 205  
   frecuencia, 202  
   frecuencia de bus, 204  
   instrucciones multimedia, 117  
   integración, 205  
   mejoras térmicas, 205  
   memoria caché, 206  
   metalización, 205  
   segmentación, 206  
   superescalaridad, 206

## B

Barton ☼, 191, 207, 211, 244  
 Brookdale ☼, 227  
 Burroughs ⤴, 21  
 Busicom ⤴, 15  
 Busicom ☉®, 15

## C

Caruso ☼, 21  
 Cascades ☼, 175  
 Cayenne ☼, 183, 245  
 Celera Genomics ⤴, 63  
 CeleronA ☉®  
   tipo de caché, 81  
 Centaur ⤴, 115, 245  
 Centrallab ⤴, 8  
 Clawhammer ☼, 256, 257  
 Coca-Cola ⤴, 209  
 Commodore ⤴, 24  
 Commodore-64 ☉®, 24  
 Compaq ⤴, 28, 63, 118  
 Control Data ⤴, 21  
 Coppermine ☼, 173  
 Covington ☼, 166  
 CP/M ≈®, 23, 27  
 Crush ☼, 268  
 Crusoe ☉®, 104, 106  
 CTC ⤴, 19  
 Cyrix ⤴, 43, 115, 145, 245  
 Cyrix III ☉®, 245

## D

Data General ⤴, 24  
 Datamation (revista) ≈®, 17  
 Datapoint 2200 ☉®, 19  
 Deschutes ☼, 161  
 Digital Research ⤴, 21, 24, 27, 57, 63, 118  
 Direct3D ≈<sup>TM</sup>, 115  
   empleo en K6-2, 178  
 DirectX ≈<sup>TM</sup>, 118  
   6.0, 113  
   7.0, 113  
   8.0, 113  
   empleo en K6-2, 178  
   funciones, 113  
   versiones, 113  
 Doom ≈®, 277  
 Duron ☉®, 200  
 Duron con DDR integrado ☉®, 201  
 Dynabook ☉®, 20

## E

Earth Simulator ☉®, 285, 286  
 EDSAC ☉®, 17  
 Electro-Technical Industries ⤴, 15  
 Electronic News (revista) ≈®, 8, 17  
 Enhanced 3DNow ≈<sup>TM</sup>  
   API para, 113  
   contexto histórico, 95, 114  
   en el K7, 198  
   recursos hardware, 112, 117  
   sus 24 instrucciones, 117  
 ENIAC ☉®, 8  
 EV6 ☉<sup>TM</sup>, 193, 267

## F

Fairchild Semiconductor ⤴, 8, 12  
 Foster ☼, 191

## G

Gallatin ☼, 191  
 GeForce FX ☉®, 268  
 Gimp ≈®, 2  
 Glide ≈<sup>TM</sup>, 113

## H

Hewlett-Packard ⤴, 24, 38, 104, 118  
 Honeywell ⤴, 21  
 HyperThreading ☉<sup>TM</sup>, 268  
   concepto, 229  
   origen, 273  
 HyperTransport ☉<sup>TM</sup>, 268, 269  
   ancho de banda, 268

etimología, 268  
objetivos, 268  
orígenes, 268  
propiedades, 268

## I

i432 Ⓒ, 21, 24, 26  
i845 Ⓒ, 227  
i850 Ⓒ, 227, 270  
i860 Ⓒ, 21  
i960 Ⓒ, 61  
IA-64 ≈<sup>TM</sup>, 104, 192, 262  
comparativa versus x86-64, 263  
compatibilidad, 262  
rendimiento, 262  
IBM Ⓐ, 24, 25, 27, 28, 40, 43, 146, 258, 285  
IBM-PC Ⓒ, 26  
nacimiento, 25  
IDT Ⓐ, 43  
IEEE Computer (revista) ≈, 17  
Intel Ⓐ, 14, 17, 21, 23, 24, 25, 26–28, 35, 37–40, 43, 61,  
104, 118, 134, 191, 192, 266–268, 270, 280,  
283, 284  
fundación, 13  
relación con Rambus, 194  
Investrónica Ⓐ, 28  
Itanium Ⓒ, 21, 106, 192, 261, 262  
tipo de caché, 81

## J

Jalapeño ⚡, 245, 273  
Joshua ⚡, 245

## K

K5 Ⓒ, 42, 144  
segmentación, 145  
tipo de caché, 81  
versiones, 145  
K6 Ⓒ, 43  
análisis de su configuración de caché, 89  
caché L1, 177  
conjunto de instrucciones, 177  
fecha de salida, 176  
frecuencia, 176  
millones de transistores, 176  
segmentación, 176  
superescalaridad, 177  
tipo de caché, 81  
unidades funcionales, 177  
voltaje, 176  
K6-2 Ⓒ, 114  
éxito comercial, 178  
conjunto de instrucciones, 178  
frecuencia, 177

instrucciones multimedia, 115  
novedades, 177  
rendimiento, 178  
tipo de caché, 81  
K6-2+ Ⓒ, 179  
K6-III Ⓒ, 179  
tipo de caché, 81  
K6-III+ Ⓒ, 179  
K7 Ⓒ, 62, 114  
área de integración, 193  
análisis de su configuración de caché, 90, 91  
ancho de banda del bus local, 194  
bus de 2x100 MHz, 192  
código de instrucción, 195  
caché L1, 194  
caché L2, 194  
como opción de compra, 192  
competidores, 191  
concurrencia máxima, 193  
controlador de caché, 87  
cuota de mercado, 192  
decodificación, 195  
dependencias  
de control, 198  
de datos, 197  
frecuencia, 192  
incompatibilidades, 194  
instrucciones multimedia, 117  
intereses corporativos, 194  
interfaz  
con caché, 194  
con memoria principal, 195  
lo mejor, 199  
lo peor, 199  
mejoras en el bus, 192  
mejoras en la FPU del, 284  
multiplicadores, 192  
planificador de instrucción, 197  
predicción de salto, 198  
protocolo de bus, 193  
relación con memoria principal, 194  
segmentación, 197  
superescalaridad, 197  
tamaño de línea, 195  
tipo de caché, 81  
TLB, 195  
unidades funcionales  
de punto flotante, 198  
enteras, 198  
versus Pentium 4, 231  
zócalo, 193  
K8 Ⓒ, 267  
área de integración, 258  
ampliación de caché, 260  
ancho de banda, 270  
banco de registros, 260  
boca o front-end, 259  
bus local, 267  
códigos de referencia, 256

controlador de DDRAM integrado, 267  
diferencias entre Athlon 64 y Opteron, 271  
doble puerto en el bus, 269  
dos niveles de memoria principal, 269  
estómago o back-end, 259  
etimología, 257  
frecuencia del bus, 257  
frecuencia del bus de memoria, 267  
frecuencia inicial, 257  
memoria caché, 260  
menos calor y más frecuencia, 258  
microarquitectura, 259  
número de transistores, 257  
patillaje, 258  
segmentación, 259  
superescalaridad, 260  
utilización de SOI, 258  
voltaje, 257  
zócalo, 258  
zócalo del procesador, 271

Katmai ⚡, 116, 168

Klamath ⚡, 161

## L

LaTeX ≈®, 2

LDT ⚡, 267

Lectrice Ⓞ®, 21

Linpack ≈™, 286

Linux ≈®, 2

## M

M1 de Cyrix Ⓞ®, 183

M2 Ⓞ®, 183

Macintosh Ⓞ®, 21, 38

MAX ≈™, 118

Maya ≈®, 231

Mc-Donalds Ⓜ, 209

MDMX ≈™, 118

MediaGX Ⓞ®, 245

Mendocino ⚡, 167

Merced ⚡, 106

Microsoft Ⓜ, 21, 25, 25, 27, 118

APIs para MMX y sucesoras de, 113

MIPS Ⓞ®, 59

MIPS-V Ⓞ®, 118

MMX ≈™, 284

antecedentes, 109

API para, 113

cómo se ejecutan, 112

coexistencia de instrucciones, 112

compatibilidad

hardware, 112

software, 112

contexto histórico, 95

falso reclamo, 276

independencia de la arquitectura, 110

mejoras en velocidad, 110

origen, 109

recursos hardware, 111

sus 57 instrucciones, 110

tipo de caché, 81

Morgan ⚡, 191, 201

Motorola Ⓜ, 21, 24, 26, 37–40, 57, 118, 191

MS-DOS ≈®, 25, 28

MSI Ⓜ, 168

MSX Ⓞ®, 24

Multiflow Computers Ⓜ, 104

MVI ≈™, 118

MXi Ⓞ®, 183

## N

National Semiconductor Ⓜ, 43, 245

NCR Ⓜ, 21

NEC Ⓜ, 285

Nehalem ⚡, 255, 256

NexGen Ⓜ, 176

nForce Ⓞ®, 268

Nippon Calculating Machines Company Ⓜ, 15

Nocona ⚡, 191

Northwood ⚡, 191, 207, 211, 226, 244, 255, 268

NV30 ⚡, 268

nVidia Ⓜ, 268

Nx586 Ⓞ®, 144

## O

Olivetti Ⓜ, 15

OpenGL ≈™, 113, 114

Opteron Ⓞ®, 260, 272

Opteron, véase también K8

Origin Ⓞ®, 268

## P

P6 ⚡, 151, 208

P67 ⚡, 208

PA7x00 Ⓞ®, 118

PA8x00 Ⓞ®, 118

Palomino ⚡, 191

París ⚡, 256, 257

PC-AT Ⓞ®, 28

Pentium Ⓞ®, 36, 41, 57, 134

BTB, 138

diseño terminal, 135

distancia de integración, 136

formato de instrucción, 139

frecuencia, 135




memoria caché, 72












MMX, 140

multiplicador, 135




rasgos CISC, 138

segmentación, 59, 63, 137








superescalaridad, 62, 137  
 tipo de caché, 81  
 TLB, 138  
 unidades funcionales, 139  
 Pentium 4 , 36, 59, 62, 115, 267, 268  
 área de integración, 211  
 bancos de registros, 216  
 BRC, 222  
 BTB, 222  
 bus de datos, 213  
 bus de direcciones, 212  
 caché  
   de datos, 219  
   de segundo nivel (L2), 219  
   de traza, 216  
 competidores, 191  
 conjunto de instrucciones, 215  
 conversión a instrucciones nativas, 215  
 decodificación de instrucciones, 215  
 devaluación, 211  
 estructura básica, 209  
 etimología, 209  
 fiabilidad, 210  
 frecuencia, 210  
 frecuencia de bus, 211  
 grupo de trabajo, 209  
 instrucciones multimedia, 117, 222  
 interfaz  
   con DDRAM, 213  
   con RDRAM, 215  
   con SDRAM, 213  
 mejoras en la FPU del, 284  
 memoria interna, 216  
 número de transistores, 211  
 Northwood, 226  
   área de integración, 227  
   bus de 4x133 MHz, 226  
   frecuencias más elevadas, 226  
   L2 de 512 Kbytes, 226  
   zócalo mPGA478, 226  
 popularidad, 209  
 repliegue a 32 bits, 208  
 resolución de dependencias, 221  
 segmentación, 219  
 superescalaridad, 221  
 tipo de caché, 81  
 unidades de ejecución, 224  
 ventana de instrucciones, 222  
 versiones, 227  
 versus K7, 231  
 zócalo, 211  
 Pentium 8 , 255  
 Pentium II , 36, 43  
   Deschutes, 164  
   disipador de calor, 162  
   frecuencia, 161  
   memoria caché, 73, 162  
   segmentación, 162  
   sus dos versiones, 161


tecnología de integración, 161  
 tipo de caché, 81  
 zócalo, 162  
 Pentium II Celeron , 166  
 Pentium II Xeon , 165  
 Pentium III , 36, 114  
   caché L2, 174  
   competidores, 191  
   CPU\_ID, 169, 172  
   frecuencia, 168, 173  
   instrucciones multimedia, 116  
   jerarquía de caché, 168  
   memoria caché, 73  
   SSE, 170  
   tipo de caché, 81  
   zócalo, 173, 174  
 Pentium III Celeron , 167  
 Pentium III Xeon , 175  
 Pentium Pro , 36, 42  
   BTB, 158  
   caché L1, 159  
   caché L2, 159  
   chip caché L2, 152  
   conjunto de instrucciones, 159  
   decodificación, 153  
   frecuencia, 152  
   memoria caché, 72  
   memoria de microprograma, 153  
   multizócalo, 159  
   para servidores, 151  
   segmentación, 152  
   superescalaridad, 152  
   supersegmentación, 152  
   sus secuelas, 151  
   tecnología de integración, 152  
   tipo de caché, 81  
   transistores, 152  
 PL/M , 21  
 Power 4 , 258  
 Power PC , 40, 57, 61, 118  
 Prescott , 191  
 Prestonia , 191

## Q


QDI , 267  
 Quake , 277  
 QuantiSpeed , 206


## R


R10000 , 86, 118, 268  
 R2000 , 118  
 R4000 , 86  
 R5000 , 86  
 R8000 , 86  
 Rambus , 194, 266  
 Rank Xerox , 14


Rise , 43, 245


## S


Samuel , 245

San Diego , 256, 257


Seattle Computer Products , 27


Sequent , 268


Sharpooth , 179

Silicon Graphics , 86, 118, 268


API para MMX de, 113


Sinclair Research , 23

Sledgehammer , 191, 256, 257

Spectrum , 23

Spitfire , 191, 201

SRAM 3101 , 13

SSE , <sup>TM</sup>

API para, 113

contexto histórico, 95, 114, 115


en el Pentium III, 171

etimología, 116

falso reclamo, 276

recursos hardware, 112, 117

sus 70 instrucciones, 116

SSE2 , <sup>TM</sup>, 265

API para, 113


en el Pentium 4, 222

en P4 vs. K7, 244

recursos hardware, 117


sus 144 instrucciones, 117


Sun Microsystems , 38, 118

Super 7 , 176


## T

Tablet PC , 21


Tandon , 25


Tanner , 175


Tejas , 191


Texas Instruments , 8, 9, 17, 19, 24


Thoroughbred A , 191


Thoroughbred B , 191, 205


Thunderbird , 191, 193, 199, 260


Timna , 202

TMS-1000 , 24

TOP 500 , <sup>TM</sup>, 286


Transmeta , 21, 104, 106, 217

Troy , 256


Tualatin , 173, 207, 226


## U


UltraSparc , 118

Univac , 21


## V


Venus , 256

VIA , 43, 266

VIS , <sup>TM</sup>, 118

## W

WebPad , 245


Willamette , 191, 208

Winchip C6 , 43

Windows , 21

Windows 98 , <sup>TM</sup>


API para 3DNow, 115

Windows XP , <sup>TM</sup>, 202, 207

cobertura para HyperThreading, 231

## X

X-Box , 114

x86-64 , <sup>TM</sup>, 260, 262

anchura por defecto, 264

comparativa versus IA-64, 263

compatibilidad, 263

implementación, 263

modo

compatible con 32 bits, 264


extenso, 263

tradicional, 265


nuevos recursos, 264


operaciones parciales, 264

prefijo, 264

Xeon , 39


tipo de caché, 81


Xerox , 14


Xfig , <sup>TM</sup>, 2

## Z


Z8 , 23

Z80 , 23

Z8000 , 23

Zilog , 23, 24, 26

ZX80 , 23

ZX81 , 23

# Glosario

## institucional

Se han empleado los siguientes prefijos clave para los términos incluidos en este glosario:

(COM) : Comité o consorcio de especificación.

(CON) : Congreso o conferencia.

(FER) : Feria o certamen de informática.

(INS) : Instituto u organismo de investigación.

(LAB) : Laboratorio de investigación.

(NOR) : Normativa de estandarización.

(PER) : Persona.

(PRO) : Proyecto de investigación.

(UNI) : Universidad.

(INS) ARPA (Advanced Research Projects Agency), 14

(PER) Bardeen, 8

(UNI) Berkeley, 14

(PER) Brattain, 8

(PRO) C5, 23

(PER) Carmack, John, 277

(PER) Carmean, Doug, 209

(UNI) Carnegie-Mellon, 14

(PRO) Chess, 25

(PER) Eckert, J. Presper, 8

(PRO) ELI, 104

(PER) Ellison, Larry, 275

(PER) Faggin, Federico, 15, 23

(PER) Fairchild, Sherman Mills, 8

(PER) Feeney, Hal, 19

(PER) Fisher, Josh, 104

(PER) Frohman-Bentchkoswky, Dov, 17

(PER) Gates, Bill, 27

(PER) Gelbach, Eduard, 17

(PER) Grove, Andy, 13, 208

(PER) Hoerni, Jean, 10

(PER) Hoff, Marcian E. (Ted), 15

(COM) HyperTransport, 267, 268

(INS) IEC (International Electrotechnical Commission), 121

(INS) ISO (International Standardization Organization), 121

(COM) ITRS (International Technology Roadmap for Semiconductors), 277

(PER) Jobs, Steve, 24

(PER) Kay, Alan, 20

(PER) Kelly, M.J., 8

(PER) Kilby, Jack Sinclair, 8

(PER) Kildall, Gary, 21, 27

(PER) Klein, Tom, 15

(PER) Lampson, Butler, 20

(LAB) Lawrence Livermore National Laboratory, 285

(PER) Mauchly, John W., 8

(PER) Maxwell, James Clerk, 286

(PER) Mazor, Stanley, 15

(PER) Meuer, Hans, 286

(PER) Meyer, Dirk, 193

(UNI) MIT, 14

(PER) Moore, Gordon, 8, 286

(PER) Morse, Stephen, 25

(PER) Newton, Isaac, 286

(PER) Noyce, Robert Norton, 8

(PER) Otellini, Paul, 208

(PER) Pake, George, 14

(LAB) PARC, 20

(LAB) PARC (Palo Alto Research Center), 14

(PER) Paterson, Tim, 27

(PER) Paul Allen, 27

(PRO) QL (Quantum Leap), 23

(PER) Ravenel, Bruce, 25

(PER) Rock, Art, 13

(PER) Shima, Masatoshi, 15, 23

(COM) SIA (Semiconductor Industry Association),  
52

(PER) Sinclair, Clive, 23

(UNI) Stanford, 14, 59

(PER) tío Sam, 14

(PER) Thacker, Charles, 20

(UNI) UCLA, 14

(UNI) Utah, 14

(PER) Wozniak, Steve, 24

(UNI) Yale, 104