
Evolución histórica de los computadores

En los sólo 50 años de vida de los computadores, los avances en su arquitectura y en la tecnología usada para implementarlos han permitido conseguir una evolución en su rendimiento sin precedentes en ningún otro campo de la ingeniería. Dentro de este progreso la tecnología ha mantenido un ritmo de crecimiento constante, mientras que la contribución de la arquitectura ha sido más variable.

En los primeros años de los computadores (desde el 45 hasta el 70) la mejora provenía tanto de los avances tecnológicos como de innovaciones en el diseño. En una segunda etapa (aproximadamente de los 70 a mediados de los 80) el desarrollo de los computadores se debió principalmente al progreso en la tecnología de semiconductores, que obtuvo mejoras impresionantes en densidad, velocidad y disipación de potencia. Gracias a estos avances el número de transistores y la frecuencia de reloj se incrementaron en un orden de magnitud en la década de los 70 y en otro en la de los 80.

Posteriormente tanto la tecnología como la arquitectura tuvieron una influencia fundamental en dicha evolución, cuyo ritmo se ha acelerado actualmente. En la década de los 90 el número de transistores y la frecuencia de reloj se han multiplicado por 20.

Muchos anuncian que este proceso comenzará a hacerse más lento a medida que nos aproximemos a los límites físicos de la tecnología de semiconductores. Según F. Faggin [Fagg96a], a partir de entonces las innovaciones en la arquitectura de los procesadores serán el motor fundamental de su progreso.

Para estudiar este proceso con mayor detalle usaremos una clasificación de los computadores en generaciones. Estas se dividen habitualmente basándose en la tecnología empleada, aunque los límites entre una y otra son más bien difusos. Cada nueva generación se caracteriza por una mayor velocidad, mayor capacidad de memoria, menor consumo y menor tamaño que la generación anterior. Existen algunas diferencias a la hora de realizar la clasificación en generaciones tecnológicas pero en general podemos decir que la Tabla 1 presenta la clasificación más ampliamente aceptada. En ella se destacan los principales avances tecnológicos y arquitectónicos que tienen lugar en cada una de las etapas.

Generación	Fechas	Característica tecnológica básica	Otros avances tecnológicos	Avances arquitectura
Primera	1946-1957	Válvula de vacío	Memoria ferritas Cintas magnéticas Disco magnético	Modelo Von Neumann
Segunda	1958-1963	Transistor	Incremento capacidad memorias	Memoria virtual Interrupciones Segmentación
Tercera	1964-1971	Circuito integrado SSI-MSI	Disco Winchester	Microprogramación memoria cache
Cuarta	1972-1980	LSI Microprocesadores	Memorias de semiconductores	
Quinta	1981-	VLSI	Incremento capacidad memorias y discos	Computadores RISC Superescalares Supersegmentados

ANTECEDENTES: LAS MÁQUINAS DE CALCULAR MECÁNICAS

Los computadores surgen por la necesidad de las personas de realizar cálculos, que llevó a la invención de distintas máquinas para calcular. Ya en el 3.000 a. C. se inventó el ábaco, que puede considerarse el primer antecedente. Pero el primer paso importante en la historia de computadores lo constituyen las primeras máquinas de calcular mecánicas, que se remontan al siglo XVII, construidas por B. Pascal (1642-43) y G. Leibnitz (1674).

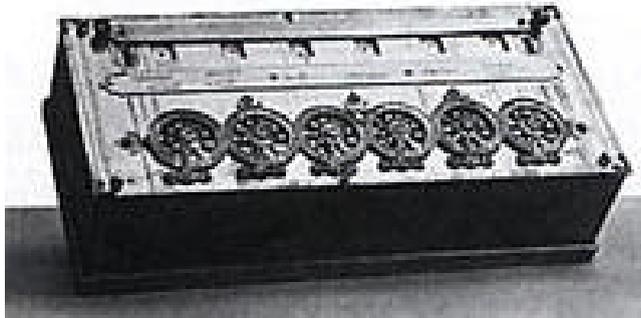


Figura 1: Máquina de Pascal

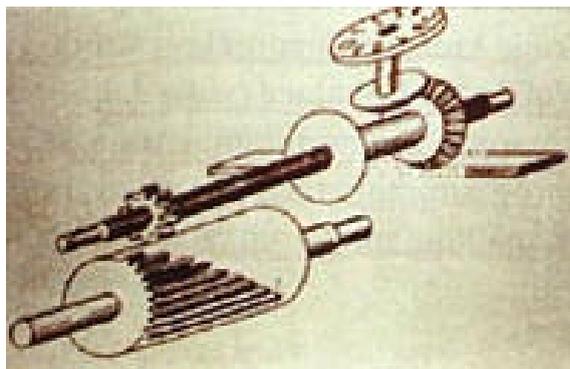


Figura 2: Máquina de Leibnitz

Posteriormente, en el siglo XIX, C. Babbage ideó dos máquinas: diferencial (1822-32) y analítica (1834-35). La primera, que únicamente realizaba un algoritmo, tenía una estructura muy simple. Su principal novedad la constituyó la salida de resultados por medio de perforaciones sobre una placa. La máquina

analítica, por su parte, constituye la primera máquina de propósito general. Era capaz de realizar cualquier operación matemática automáticamente. Tenía una unidad de almacenamiento, una unidad de procesamiento, una unidad de entrada de datos y otra de salida, y en cierto modo, su estructura se mantiene aún en los computadores modernos. No llegó a implementarse por dificultades tecnológicas, pero los computadores electromecánicos Harvard Mark I y Mark II, basados en relés, fueron diseñados por H. Aiken en 1944 y 1947 respectivamente, a partir de estas ideas de Babbage [Coel02].

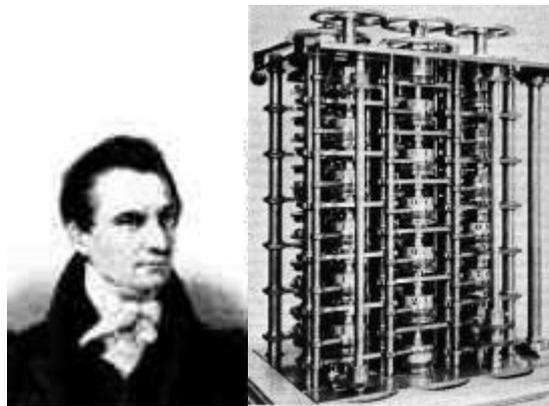


Figura 3: Charles Babbage y su máquina diferencial

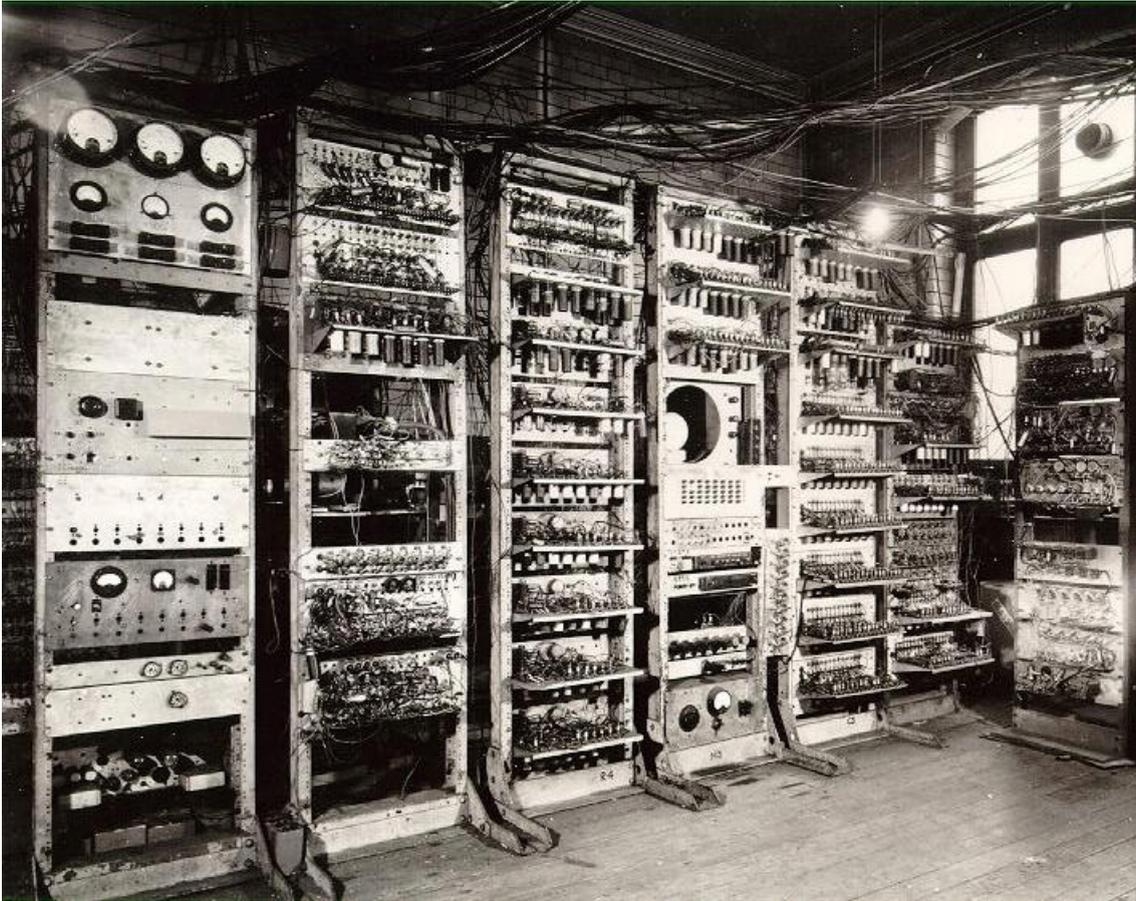


Figura 4: Imagen de la Mark I basada en las ideas de Babbage

Cabe destacar también el papel de Ada Lovelace, (hija de Lord Byron), en el desarrollo histórico de las computadoras modernas que fue casi totalmente ignorado hasta 1979, cuando el Departamento de Defensa de los Estados Unidos decidió utilizar su nombre para el nuevo lenguaje de programación que utilizarían como estándar para desarrollar su propio software interno. Desde entonces, parece que una nueva luz se ha producido sobre la vida de esta matemática tan talentosa, que fue una de las pocas personas que conoció y comprendió detalladamente el trabajo de Charles Babbage, además de haber escrito el primer programa para la inexistente Máquina Analítica.

PRIMERA GENERACIÓN: LAS VÁLVULAS DE VACÍO (1946-1957)

Tecnología básica

En 1904, Fleming patentó la válvula de vacío diodo, con idea de utilizarla para mejorar las comunicaciones de radio. En 1906, Forest añade un tercer electrodo al flujo de corriente de control del diodo de Fleming, para crear la válvula de vacío de tres electrodos.

Los computadores mecánicos tenían grandes dificultades para conseguir aumentar su velocidad de cálculo, debido a la inercia de los elementos móviles. Por ello el uso de válvulas de vacío supuso un gran paso adelante en el desarrollo de los computadores, tanto en velocidad como en fiabilidad, y dio lugar a lo que se conoce como Primera Generación de computadores.

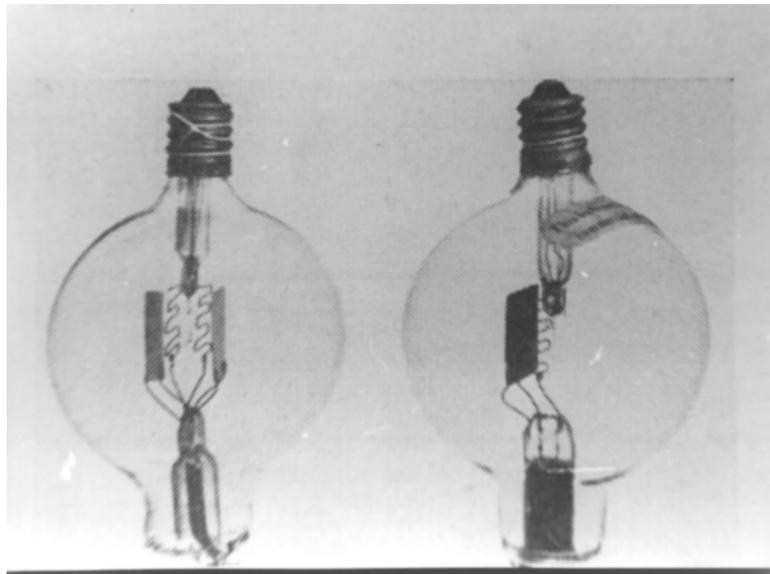


Figura 5: Imagen de las primeras válvulas de vacío

Avances en arquitectura

Habitualmente se considera que los computadores comenzaron con el ENIAC en 1946 y, de acuerdo con esto, la IEEE Computer Society celebró en

1996 los primeros 50 años de los computadores modernos. Sin embargo, J. V. Atanasoff había construido en 1939 un prototipo de computador digital electrónico que usaba aritmética binaria. Por eso desde 1973 se le reconoce como creador del computador moderno [CaBM96].

Si la bomba atómica fue el secreto mejor guardado por los norteamericanos durante la Segunda Guerra Mundial, su equivalente en Inglaterra fue el Colossus, la primera computadora completamente electrónica del mundo que se diseñó explícitamente para poder descifrar los mensajes secretos de los nazis y que A. Turing, T. Flowers y M.H.A. Newman presentaron en Diciembre de 1943 e hicieron operacional en Bletchley Park [Dani96]. . Esto marcó el inicio de la escuela inglesa de cómputo electrónico que le dio al mundo la primera computadora con programa almacenado de la historia, la primera unidad de control microprogramada y muchas otras valiosas contribuciones a la computación moderna.

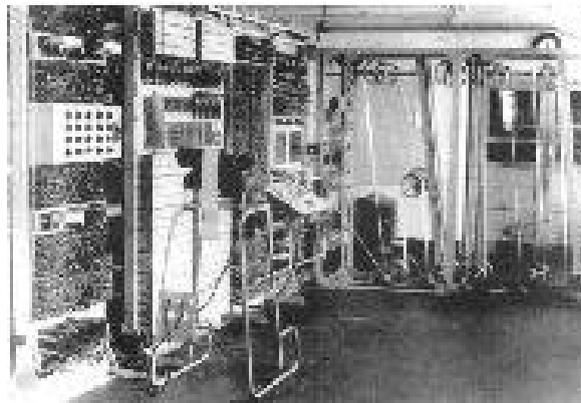


Figura 6: El Colossus, primera computadora totalmente electrónica

Pero es en 1946 cuando se considera que comienza la historia de los computadores. En la Universidad de Pennsylvania, J.P. Eckert y J.W. Mauchly mostraron al mundo el primer computador electrónico de propósito general: el ENIAC (Electronic Numerical Integrator and Calculator). Pesaba 30 toneladas y consumía 150 KW. Estaba construido con 18.000 válvulas de vacío y permitía realizar cálculos a una velocidad 100 veces mayor que una persona. Era programable y la programación se efectuaba mediante cables y conmutadores.

Los datos se introducían mediante tarjetas perforadas. Sus principales inconvenientes eran la tediosa tarea de programación, y la limitada capacidad de almacenamiento.

Para eliminar la necesidad de programar manualmente el computador J. Von Neumann propone un computador de programa almacenado denominado EDVAC (Electronic Discrete Variable Automatic Computer). Su diseño se denomina hoy "modelo Von Neumann", y se sigue manteniendo en la mayoría de computadores actuales, con unidad aritmético-lógica, unidad de control, unidades de entrada/salida, y memoria.

Basado en las ideas del EDVAC, M. Wilkes, de la Universidad de Cambridge, construye en 1949 el EDSAC (Electronic Delay Storage Automatic Calculator), que utiliza la noción de memoria jerárquica y una arquitectura basada en acumulador.



Figura 7: EDVAC

También von Neumann junto con H. Goldstine y A. Burks comenzó a construir un nuevo computador de programa almacenado, denominado IAS (Institute for Advanced Study) cuyo diseño no terminó hasta 1952. El IAS constaba de una memoria principal para almacenar datos e instrucciones, una unidad aritmético-lógica, una unidad de control que interpreta las instrucciones

y provoca su ejecución, y una unidad de entrada/salida dirigida por la unidad de control.

En 1951 Wilkes introduce la idea de la microprogramación para el diseño ordenado de la unidad de control. Esta idea no fue realizable ya que para almacenar los microprogramas se requería una memoria muy rápida, que no estaba disponible en ese momento. Por esta razón, la idea quedó como una mera conjetura académica durante una década. Una vez más, como le sucedió a Babbage, una innovación arquitectónica tuvo que esperar hasta que la tecnología avanzara para permitir su implementación[Wilk51][Wilk53] .

Otras tecnologías

Al mismo tiempo, en el MIT el equipo de J. Forrester trabaja en un computador de propósito especial para tratamiento de señales en tiempo real, el proyecto Whirwind (1949), cuya principal aportación es la utilización de la memoria de ferritas. Esta ha sido la principal tecnología de memoria durante varias décadas [deMi90]. Cada punto de memoria es un toro o anillo de ferrita, que puede presentar dos direcciones de magnetización. Las primeras ferritas fabricadas tenían un diámetro exterior de 3 mm, tenían una capacidad de 2 Kbytes y el tiempo de acceso era de unos 30ms. La conexión de los anillos de ferrita a los transductores se hacía mediante hilos de cobre barnizados que debían hacerse pasar por el interior de las ferritas. Este proceso era de difícil automatización, por lo que debía hacerse a mano.

Por otra parte se extiende el uso de cintas magnéticas para el almacenamiento masivo. También aparece el primer disco magnético en el año 1956, que se utilizó en la máquina RAMAC (Random Access Method of Accounting and Control) de IBM, con una capacidad de 5Mbytes y un tiempo de aproximación de 1s [HePa02].



Figura 8: Usuario utilizando una de las primeras versiones del RAMAC de IBM

Los primeros computadores comerciales de esta generación, que aparecieron en la década de los 50, fueron el UNIVAC I y II (Universal Automatic Computer), fabricados por Eckert y Mauchly y la serie 700 de IBM.

En esta primera generación de computadores, las instrucciones se procesaban en serie: se buscaba la instrucción, se decodificaba y luego se ejecutaba. La velocidad típica de procesamiento que alcanzaron los computadores era aproximadamente 40.000 operaciones por segundo. Eran equipos de gran tamaño, escasa capacidad y difícil mantenimiento, que disipaban mucho calor. Los trabajos se realizaban en monoprogramación y no existía sistema operativo, por lo que los periféricos de entrada/salida dependían directamente del procesador. Se programaba en lenguaje máquina, lo que exigía programadores muy especializados.

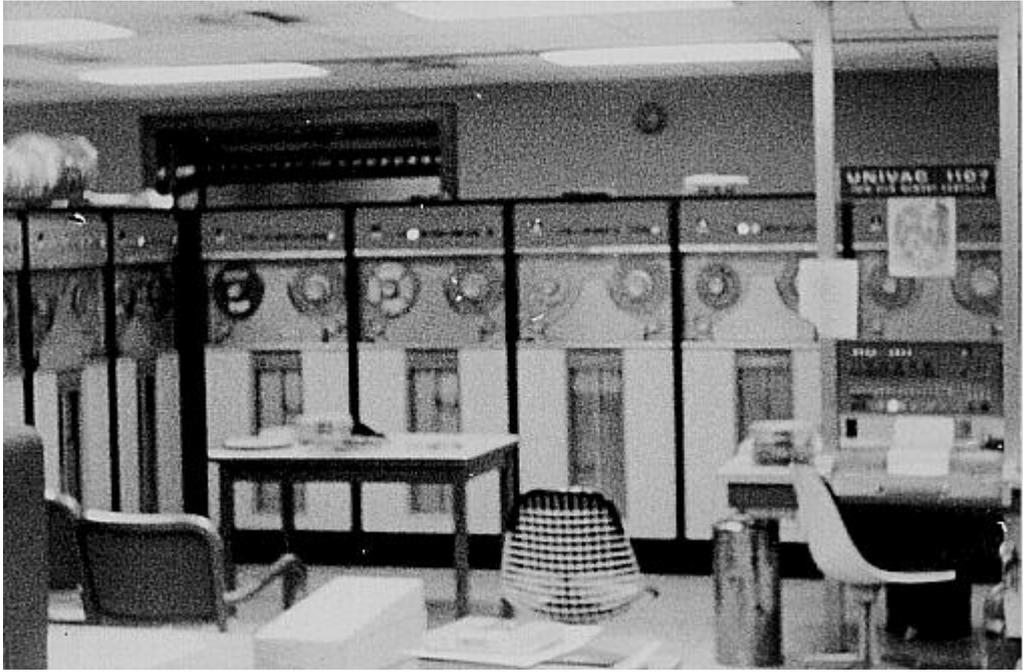


Figura 9: Imagen del UNIVAC II

SEGUNDA GENERACIÓN: LOS TRANSISTORES (1958-1963)

Tecnología

La invención del transistor tuvo lugar en 1948 en los laboratorios Bell por W.B. Shockley, J. Bardeen y W.H. Brattain. Poco a poco la industria de semiconductores fue creciendo y los productos industriales y comerciales sustituían los dispositivos de válvulas de vacío por implementaciones basadas en semiconductores.

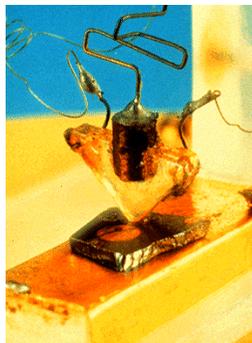


Figura 10: El transistor

La nueva tecnología permite aumentar el rendimiento y la fiabilidad, y reducir de forma drástica el tamaño de los computadores, dando lugar a la Segunda Generación de computadores. La velocidad de ejecución de la CPU se incrementó enormemente, hasta alcanzar 200.000 operaciones por segundo. La disminución de tamaño de los módulos permitió introducir unidades lógicas y aritméticas y unidades de control más complejas.

Por otra parte, el tamaño de la memoria principal de ferritas creció de 2 Kpalabras a 32 Kpalabras, y el tiempo de aproximación cayó de 30 ms a 1,4 ms.

Avances en arquitectura

El incremento de la complejidad de las unidades de control, permitió introducir una de las innovaciones arquitectónicas que posteriormente se ha utilizado en gran escala: la segmentación de operaciones. Con esta técnica, la decodificación de una instrucción se solapa con la búsqueda de la instrucción siguiente y con la ejecución de la anterior.



Figura 11: IBM 7030

En 1961 aparece el IBM 7030 o Stretch, el primer computador que usa segmentación. También tiene memoria entrelazada y predicción de saltos. No tuvo éxito comercial debido, entre otras causas, a que no alcanzó el rendimiento esperado porque el tiempo para recuperarse de un error de predicción era muy largo.

Entre las innovaciones arquitectónicas más importantes de esta generación puede destacarse la utilización de memoria virtual, para facilitar la tarea del programador a la hora de escribir programas demasiado largos para residir completamente en memoria principal. Estos programas debían estar formados por varios segmentos que se cargaban alternativamente desde la memoria secundaria hasta la memoria principal, bajo la supervisión del programa principal. La memoria virtual intentaba aliviar a los programadores de este peso, gestionando automáticamente los dos niveles de la jerarquía de memoria, formada por la memoria principal y la secundaria [HePa02]. La memoria virtual,

y la utilización de interrupciones para la E/S se utilizaron por primera vez en el sistema ATLAS (1962), desarrollado por Ferranti en la Universidad de Manchester, que también usaba segmentación.

El CDC 6600 diseñado por S. Cray de Control Data Corp. en 1964 fue el primer supercomputador comercial de éxito. Tenía arquitectura de carga-almacenamiento y empleaba técnicas de segmentación, además de paralelismo a nivel de unidades funcionales, lo cual le permitía un rendimiento de 9 MFLOPs, superior en un orden de magnitud al del 7094 de IBM.



Figura 12: CDC 6600 y CDC 7600 diseñados por Seymour Cray

Además de la serie 7000 de IBM, otra máquina comercial de esta generación fue el PDP-1, lanzado en 1960 por DEC. Una de las innovaciones interesantes del PDP-1 fue el empleo de un terminal de vídeo, así como cierto grado de capacidades gráficas sobre la pantalla de 512x512 pixels.

Con estos computadores comenzó la utilización de lenguajes de programación de alto nivel [Stal97], como FORTRAN -cuyo primer compilador desarrolló IBM en 1957-, COBOL y LISP (ambos en 1959), y apareció el procesamiento por lotes, que será el germen de los sistemas operativos.

TERCERA GENERACIÓN: LOS CIRCUITOS INTEGRADOS (1964-1971)

Tecnología básica

Durante la generación anterior los equipos electrónicos estaban compuestos en su mayoría por componentes discretos -transistores, resistencias, condensadores, etc.- cada uno de los cuales se fabricaba separadamente y se soldaban o cableaban juntos en tarjetas de circuitos. Todo el proceso de fabricación resultaba caro y difícil, especialmente para la industria de computadores, que necesitaba colocar juntos cientos de miles de transistores que había que soldar, lo cual dificultaba enormemente la fabricación de máquinas nuevas y potentes.

Por eso, la invención del circuito integrado a finales de los 50 (J. Kilby de Texas Instruments construye el primero en 1958 y R. Noyce de Fairchild Semiconductor construye otro en 1959) fue la gran clave para el crecimiento de la industria de computadores, y suele tomarse como punto de inicio de la Tercera Generación de computadores.

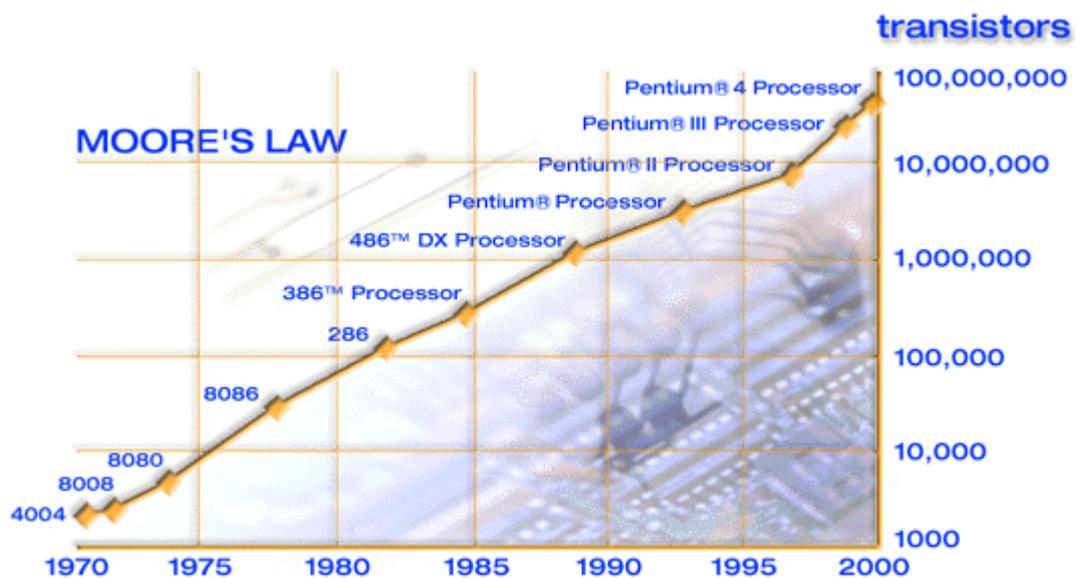


Figura 13: Ley de Moore según Intel

La introducción de circuitos integrados comerciales empezó en 1961 con componentes RTL (resistor-transistor logic), que fueron pronto sustituidos por componentes TTL (transistor-transistor logic). Posteriormente pequeños grupos de dispositivos de tecnologías TTL SSI (Small Scale Integration) fueron reemplazados por dispositivos de tecnologías TTL MSI (Medium Scale Integration) y LSI (Large Scale Integration). Entre 1961 y 1971 los chips se fueron haciendo mayores y los transistores cada vez más pequeños, de modo que el número de transistores en un chip casi se duplicaba anualmente –esta predicción se ha denominado posteriormente “ley de Moore” (ver Figura 13: Ley de Moore según Intel). Así las funciones lógicas que podían realizar los circuitos también se habían complicado considerablemente. De esta forma era posible realizar módulos y unidades de control aún más complejas, sin que el precio de los circuitos se incrementase, y el tamaño de los computadores se redujo considerablemente, a la vez que aumentó su velocidad y disminuyó el consumo.

Los computadores B2500 y B3500 de Burroughs usaron circuitos integrados y fueron construidos en 1968. Los CIs también disminuyeron el coste de los controladores de discos y de la electrónica para controlar los brazos, de forma que se podían incluir dentro de la caja del disco y ésta se podía sellar. Así surgieron los primeros discos que no eran extraíbles: en 1965 aparece el disco Winchester. Después, en 1970 aparecen los discos flexibles (floppy).

Avances en arquitectura

En este periodo también tienen lugar importantes innovaciones arquitectónicas. La principal de ellas es la microprogramación, es decir, Describir las operaciones involucradas en la ejecución de una instrucción máquina mediante un conjunto de bits, que representan a las distintas señales de control que es necesario activar. Dicho conjunto de bits o microinstrucción se almacena en un memoria, denominada memoria de control. Este enfoque

había sido propuesto por Wilkes a principios de los años 50, pero la tecnología de memorias disponible no permitió su materialización.

Es realmente en 1964 cuando IBM introduce la microprogramación en la industria de los computadores al lanzar la familia IBM System/360 [AmBB64], en la que todos los modelos, excepto los más avanzados, eran microprogramados. El concepto de familia de computadores, todos con la misma arquitectura pero con distintos precios y prestaciones, contribuyó a que el IBM/360 fuera el mainframe más popular en los 70. Su éxito fue tan grande que los mainframes actuales de IBM todavía son compatibles con él. Entre otras prestaciones, merecen citarse la posibilidad de programar la prioridad de las interrupciones, los mecanismos de protección de memoria y la inclusión de controladores de DMA.



Figura 14: Imagen del IBM 360

Por otra parte, Wilkes propone en 1965 la memoria cache: se trata de añadir un nivel de memoria intermedio entre el procesador y la memoria principal, con una capacidad inferior a la memoria principal pero un tiempo de aproximación mucho menor. La primera implementación comercial la llevó a cabo IBM en su modelo 360/85 en el año 1968, y pronto se hizo común en las grandes máquinas y minicomputadores. Actualmente es uno de los métodos más utilizados para mejorar el rendimiento debido a la creciente diferencia

entre la velocidad del procesador y la de la memoria. El problema de diseño de caches es uno de los compromisos dirigidos por la tecnología, porque los valores óptimos de los tres principales parámetros de diseño de las caches (tamaño, asociatividad y tamaño de bloque) están muy interrelacionados entre sí y dependen de los detalles de cada diseño en particular. Como el tiempo de aproximación a cache está casi siempre en el camino crítico del diseño de un procesador, el tiempo necesario para saber si un dato está en cache suele influir en el tiempo de ciclo del computador y este tiempo suele ser dependiente del tamaño de cache y del grado de asociatividad.

Al nivel de los minicomputadores también se produjo un paso importante, con la presentación en 1965 del PDP-8 de DEC. Cuando la mayoría de los computadores requerían una habitación con aire acondicionado, el PDP-8 podía colocarse encima de una mesa de laboratorio. Los últimos modelos del PDP-8 usan por primera vez estructura de bus.



Figura 15: El NEC-PDP 8 se podía situar en una mesa de laboratorio.

También hubo avances importantes en el campo de los sistemas operativos. IBM crea el OS/360, primer sistema operativo multiprogramado. Además, aparecen el sistema operativo Multics (1965) y después D. Ritchie y K. Thomson crean el Unix (1970) en los laboratorios Bell. Con esta generación de computadores se consiguieron velocidades de procesamiento de 1 millón de instrucciones por segundo (1 MIPS).

CUARTA GENERACIÓN: LOS MICROPROCESADORES (1971-1980)

Tecnología básica

En 1970 tanto la industria de computadores como la de semiconductores habían madurado y prosperado y su unión permitió el desarrollo de la denominada Cuarta Generación de computadores: basados en microprocesador. Esta etapa viene caracterizada nuevamente por un avance tecnológico, como es el desarrollo de la técnica de integración LSI, que permite incluir hasta 100.000 transistores en un único chip. En 1973 se consiguen integrar 10.000 componentes en un chip de 1cm².

El primer microprocesador, el 4004 de Intel [Fagg96b], surge en 1971 ideado por T. Hoff y construido por F. Faggin. Era un procesador de 4 bits con 2300 transistores en tecnología de 8 micras. Fue fabricado en obleas de 2 pulgadas y empaquetado con 16 pines. Podía direccionar 8 Kbytes de ROM y 640 bytes de RAM. Un año después apareció el 8008, un procesador de 8 bits con 3500 transistores, que podía direccionar 16 Kbytes de memoria y trabajar a 0.5 MHz [Tred96].

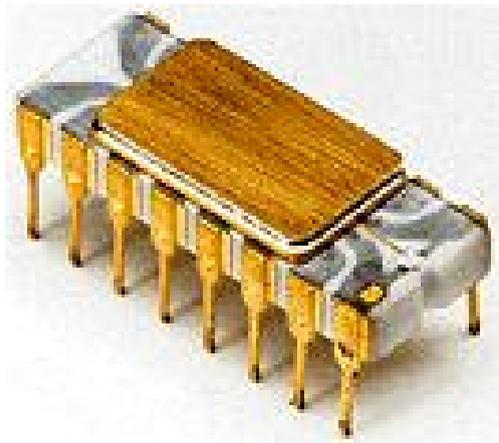


Figura 16: El primer microprocesador, el 4004 de Intel.

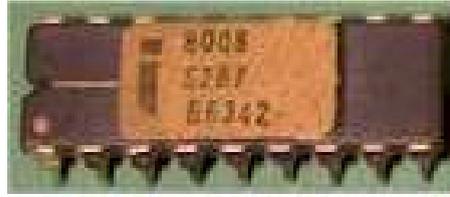


Figura 17: El 8008 contenía ya 3500 transistores en 1972

La primera aplicación del 4004 fue una calculadora de escritorio [Fagg96a]. Sin embargo, dos años después el 4004 se utilizaba en una gran variedad de sistemas empotrados como ascensores, etc. A partir de ese momento cada dos o tres años aparecía una nueva generación de microprocesadores, y los diseñadores los usaban para cualquier producto que pudiera beneficiarse de alguna cantidad de inteligencia, desde juguetes a calculadoras de bolsillo y a computadores personales. Durante los últimos 25 años, a una velocidad impresionante, el microprocesador ha cambiado la estructura de muchas de las industrias existentes e incluso ha empezado a cambiar también el aspecto de la sociedad.

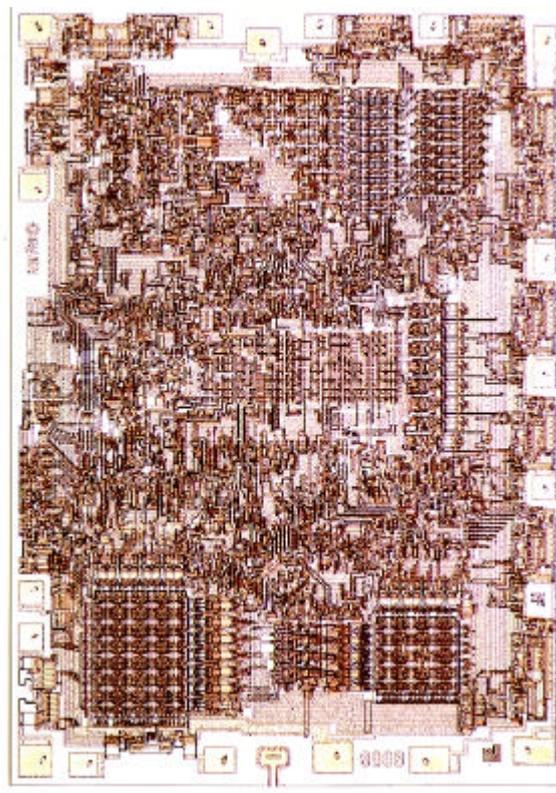


Figura 18: Layout del 8008

Otras tecnologías

Otro de los factores tecnológicos que permiten este abaratamiento de los computadores es la introducción de las memorias de semiconductores. Las memorias de ferritas se caracterizaban principalmente por ser voluminosas, caras y de lectura destructiva. Por eso, un gran avance fue la aplicación de la tecnología de CIs a la construcción de memorias. En el año 1970 Fairchild produjo la primera memoria de semiconductores de una capacidad apreciable. Este primer chip era del mismo tamaño que un único núcleo de ferrita y podía contener 256 bits de memoria. Además presentaba un tiempo de aproximación mucho menor que el de la memoria de ferritas. Sin embargo, su coste por bit era mayor que el del núcleo de ferrita.

En el año 1974, el coste por bit de la memoria de semiconductores cayó por debajo del coste de la memoria de ferritas. Ese año se construyó un chip de DRAM de 4 Kbits. Desde entonces, la capacidad de almacenamiento de las memorias no ha dejado de incrementarse año tras año. Este crecimiento ha ido acompañado por una disminución de su coste y un incremento en la velocidad de aproximación.

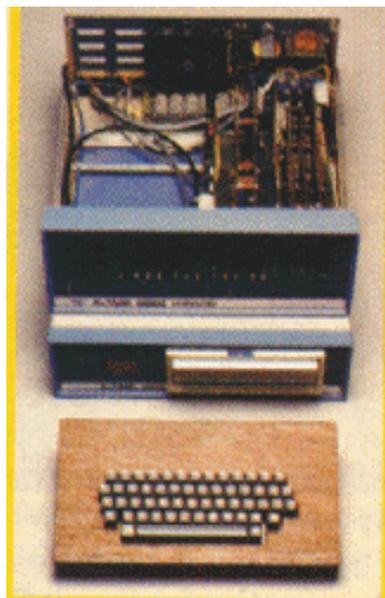


Figura 19: MITS Altair 8800

Avances en arquitectura

La arquitectura de los primeros microprocesadores [Fagg96a] fue una adaptación de las ideas usadas con anterioridad en los minicomputadores y los mainframes. Las compañías incorporaban estas ideas en los microprocesadores tan pronto como la rápida evolución de las capacidades dadas por la tecnología y el coste lo permitían. Por eso esta generación se caracteriza principalmente por las mejoras en la tecnología de circuitos integrados, que los microprocesadores aprovechan más que otros computadores debido a su mayor integración, y no tanto por las mejoras arquitectónicas.

Ya en 1974 el 6800 contenía alrededor de 5000 transistores en tecnología NMOS de 6 micras. Operaba a 2 MHz y podía direccionar 64 Kbytes de memoria. También aparecieron el MC6502 y el Intel 8080 entre otros.

La disminución del coste de los CIs conduce a un gran abaratamiento de los computadores, lo cual permite la fabricación de los primeros computadores personales. En 1975 apareció el primer sistema de computador popular basado en microprocesador: el MITS Altair 8800. Estaba basado en el Intel 8080, un microprocesador de 8 bits que trabaja a 2 MHz introducido en 1974. El Apple II se introdujo en 1977, basado en el MC6502, junto con un terminal CRT, un teclado y una disquetera. Fue el primer computador personal con gráficos en color.

En 1978 Intel introdujo el microprocesador de 16 bits 8086, con 29000 transistores, tecnología HMOS de 3 micras, un rango de direcciones de 1 Mbyte y una velocidad de 8MHz. Este diseño fue utilizado por IBM para el computador personal (IBM PC) que se presentó en 1981, para el que elige el PC-DOS de Microsoft como sistema operativo.



Figura 20: El Apple II

En 1979, pensando que la memoria seguiría reduciendo su coste y que los futuros programas se escribirían en lenguajes de alto nivel, Motorola incrementó el ancho de banda con un bus de datos de 16 bits y un bus de direcciones de 32 bits para el MC68000. Este microprocesador podía dar un rendimiento pico de 2 MIPS. Debido a limitaciones de empaquetamiento (tenía 64 pines) los 32 bits se redujeron a 24 en los primeros productos. También se añadieron registros de propósito general de 32 bits, hasta un total de 16. El número total de transistores era de alrededor de 68000, podía trabajar a 5V y a una velocidad de 8 MHz. Apple Computer seleccionó el 68000 para la línea de computadores personales Macintosh.

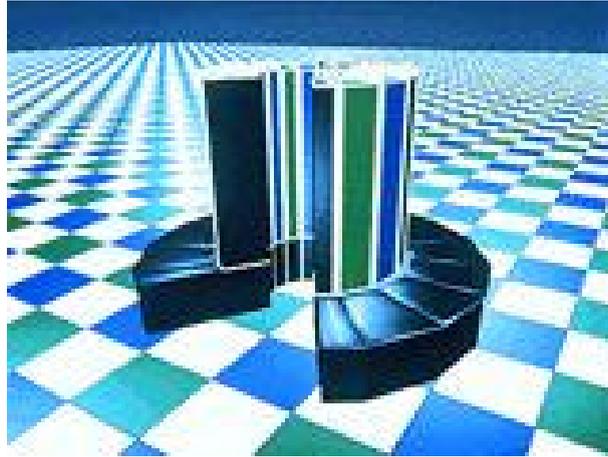


Figura 21: Imagen publicitaria del Cray-1

En 1976 aparece el Cray-1 de Cray Research, el primer supercomputador vectorial. En 1978 DEC presenta el VAX 11/780, un computador de 32 bits que se hace popular para aplicaciones científicas y técnicas. Los diseñadores del VAX buscan simplificar la compilación de lenguajes de alto nivel –en 1972 habían aparecido C, SmallTalk y Prolog, que se sumaron a los ya existentes Fortran, Cobol y Lisp- y para ello crean una arquitectura ortogonal de instrucciones complejas. Además optimizan el tamaño del código para que ocupe menos memoria.

En el sentido opuesto al de los diseñadores del VAX, J. Cocke trabaja en el proyecto 801 de IBM para desarrollar un minicomputador que será origen de las futuras arquitecturas RISC de la siguiente generación.



Figura 22: Imagen del VAX 11/780

QUINTA GENERACIÓN: DISEÑO VLSI 1981-?

Año tras año el precio de los computadores disminuye forma drástica, mientras las prestaciones y la capacidad de estos sistemas siguen creciendo. El incremento de la densidad de integración ha permitido pasar de circuitos con unos pocos miles de transistores a principios de los años 70 a varios millones en la actualidad. Por ello podemos afirmar que la aparición de la tecnología VLSI a principios de los 80 puede considerarse como el origen de la Quinta Generación, que se caracteriza fundamentalmente por la proliferación de sistemas basados en microprocesadores[BuGo97] [Gei90].

Tecnología básica

Los tamaños mínimos de fabricación (minimum feature size, λ) han disminuido desde las 50 micras de los años 60 a las 0.13 micras actuales, mejorándose de este modo tanto la densidad de integración (con un crecimiento anual aproximado del 35%) como la velocidad de los circuitos integrados. En paralelo con esta reducción, las dimensiones máximas del chip también han aumentado, aunque en este caso la evolución es menos predecible, rondando entre el 10% y el 20% anual. El efecto combinado del aumento de la densidad de transistores y del tamaño de los chips ha dado lugar a un aumento en el número de transistores por chip en torno a un 55% anual [HePa02], consiguiéndose integrar en la actualidad del orden de 200 millones de transistores en un único chip. Por citar un ejemplo destacable, el Power4 de IBM integra unos 174 millones de transistores¹[Dief99].

No obstante, a consecuencia de dicha disminución han surgido nuevos desafíos. Al reducirse el tiempo de conmutación de los transistores cobra especial relevancia los retardos de propagación de las señales dentro del propio circuito integrado. Un ejemplo destacable lo tenemos en el diseño del Pentium 4, en el que dos de las 20 etapas de que consta su pipeline se utilizan exclusivamente para propagar las señales dentro del chip [HePa02].

¹ El chip incluye 2 procesadores

Otro aspecto importante a tener en cuenta es el consumo de potencia. En los primeros microprocesadores, el consumo era de tan solo unas decenas de Watios. Actualmente, la potencia máxima disipada por los microprocesadores de gama más alta puede alcanzar entre los 100 y los 150 Watios, siendo probable que en un futuro próximo sean factores relacionados con el consumo los que limiten o bien la cantidad de hardware que pueda ser incluido en el microprocesador o bien la frecuencia de reloj a la puedan trabajar[HePa02].

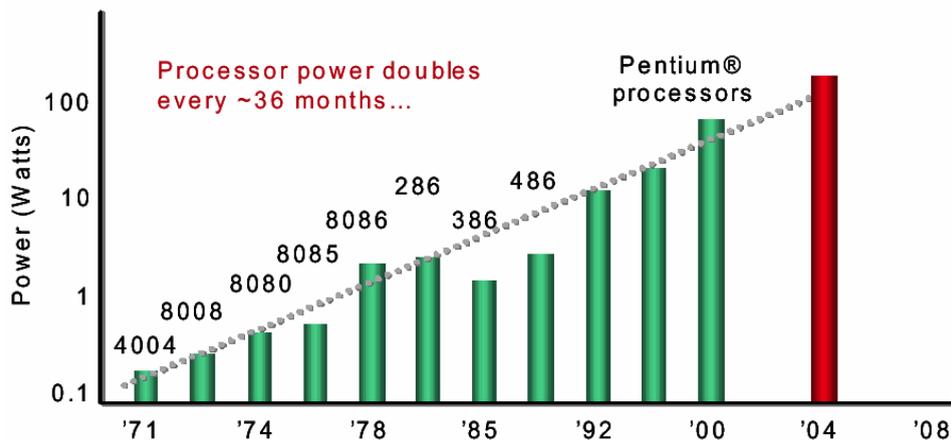


Figura 23: Evolución histórica del consumo de los microprocesadores desarrollados por Intel [MNWS02].

Del incremento en la densidad de integración también se ha beneficiado la tecnología de memorias dinámicas (DRAMs). La densidad (en número de bits por unidad de área) de los módulos DRAM ha aumentado entre un 40% y un 60% anual. Lamentablemente, el tiempo de ciclo se ha ido reduciendo muy lentamente, aproximadamente en un tercio cada diez años [HePa02].

Modelo	Tipo	Ciclo de Reloj Efectivo	Bus de Datos	Ancho de Banda Pico
PC66	SDRAM	66 MHz	64 Bit	0,5 GB/s
PC100	SDRAM	100 MHz	64 Bit	0,8 GB/s
PC133	SDRAM	133 MHz	64 Bit	1,06 GB/s
PC1600	DDR200	100 MHz	64 Bit	1,6 GB/s
	DDR200 Dual		2 x 64 Bit	3,2 GB/s
PC2100	DDR266	133 MHz	64 Bit	2,1 GB/s

	DDR266 Dual	133 MHz	2 x 64 Bit	4,2 GB/s
PC2700	DDR333	166 MHz	64 Bit	2,7 GB/s
	DDR333 Dual	166 MHz	2 x 64 Bit	5,4 GB/s
PC3200	DDR400	200 MHz	64 Bit	3,2 GB/s
	DDR400 Dual	200 MHz	2 x 64 Bit	6,4 GB/s
PC4200	DDR533	266 MHz	64 Bit	4,2 GB/s
	DDR533 Dual	266 MHz	2 x 64 Bit	8,4 GB/s
PC800	RDRAM Dual	400 MHz	2 x 16 Bit	3,2 GB/s
			2 x 32 Bit	6,4 GB/s
PC1066	RDRAM Dual	533 MHz	2 x 16 Bit	4,2 GB/s
			2 x 32 Bit	8,4 GB/s
PC1200	RDRAM Dual	600 MHz	2 x 16 Bit	4,8 GB/s
			2 x 32 Bit	9,6 GB/s

Figura 24 Frecuencias de trabajo y los anchos de banda pico de los módulos de memoria comercializados en mayo de 2002[Inst02].

Respecto al ancho de banda proporcionado por cada chip de memoria, durante el mismo periodo de tiempo (diez años) se ha observado un incremento del orden de dos tercios, si bien se han logrado mejoras adicionales mediante el diseño de nuevas interfaces. Los interfaces basados en protocolos asíncronos como page mode, fast page mode o EDO RAM han sido desplazados, compitiendo actualmente por el mercado dos tecnologías con protocolo síncrono (SDRAM): Rambus DRAM (RDRAM) y double-data-rate (DDR) DRAM [Paul02]. En la tabla 1.2.1 se muestran las frecuencias de trabajo y los anchos de banda pico de los módulos de memoria que se comercializaban en mayo de 2002 [Inst02].

Otros avances tecnológicos

Paralelamente al incremento de la densidad de integración, ha aumentado notablemente la capacidad de los sistemas de almacenamiento, y ha disminuido su coste. Hasta 1990 la capacidad de almacenamiento de los discos se incrementaba en un 30% anual. Recientemente, la densidad de integración se ha aumentado en más del 100% anual, aunque como en el caso de las memorias, el tiempo de acceso sólo se reduce por un factor de un tercio cada diez años [HePa02].

Este desequilibrio supone un verdadero problema, sobre todo en los sistemas de memoria virtual. Una de las principales soluciones aportadas a este problema es la técnica RAID (*Redundant Array of Independent/Inexpensive Disks*) [CLGK94], que surgió por primera vez en 1987. El objetivo de esta técnica es aprovechar la reducción de tamaño y coste de los discos para aumentar la fiabilidad y el rendimiento de los sistemas de almacenamiento masivo. Consiste en utilizar varias unidades de discos que operen independientemente y en paralelo. De esta forma se puede acceder a un bloque de datos en paralelo siempre que los datos de dicho bloque estén adecuadamente distribuidos a lo largo de varios discos, consiguiendo una velocidad de transferencia mucho mayor. Desde el punto de vista del usuario o del sistema operativo este conjunto de discos físicos opera como una única unidad lógica. Además existen discos que almacenan información redundante que permite garantizar la correcta recuperación de los datos en caso de fallo.

También han surgido otras tecnologías de almacenamiento: la óptica y la magneto-óptica [Stal97]. El CD-ROM, introducido por Sony y Philips en 1984, es una memoria de sólo lectura basada en tecnología óptica, que se escribe mediante un rayo láser que realiza hoyos microscópicos sobre una superficie reflectante. Los discos ópticos borrables, de tecnología magneto-óptica, utilizan la energía de un rayo láser junto con un campo magnético para grabarse. Otro

disco óptico, el DVD (*Digital Versatil Disk*), tiene las mismas dimensiones que un CD-ROM pero puede almacenar hasta 17 Gbytes [Dutt99].

En general estas nuevas tecnologías permiten sistemas de almacenamiento de alta capacidad, seguras e intercambiables, pero su tiempo de acceso es mayor que el de los discos magnéticos, por lo que no suponen una alternativa a aquellos como almacenamiento secundario durante la ejecución de los programas.

Por último, debemos destacar también la mejora que ha experimentado durante estos últimos años la tecnología de red. Tanto la latencia como el ancho de banda han mejorando con el tiempo, si bien durante los últimos años la evolución se ha acelerado notablemente, habiéndose puesto especial énfasis en el ancho de banda. Por citar un ejemplo, hicieron falta unos 10 años para pasar de Ethernet de 10 Mb a 100 Mb, mientras que tan solo cinco años después estuvo disponible la tecnología Ethernet de 1Gb. Esta mejora es debida tanto a la generalización de los dispositivos ópticos como a la mayor densidad de integración de los conmutadores [HePa02]. Aunque hemos dicho que estamos en la quinta generación, dentro de ella podemos distinguir tres etapas en función de los avances arquitectónicos que han tenido lugar. Veamos con un poco más de detalle cada una de estas etapas y los cambios que las delimitan.

Avances arquitectónicos: primera etapa

La primera etapa de esta generación dura hasta mediados de los 80, cuando el número de transistores en un CI se aproxima a 250.000. En este periodo no hay grandes novedades en la arquitectura de los computadores personales. Para mejorar el rendimiento la secuencia de instrucciones se segmenta a 5 ó más etapas.

Un ejemplo de esta generación es el MC68020, introducido en 1984, el primer microprocesador de 32 bits tanto en los buses de datos y direcciones

como en registros y ALU. Tenía 200.000 transistores, incorporaba por primera vez una cache interna de 256 bytes y una segmentación de hasta 5 etapas [Tred96]. Funcionaba a una frecuencia de 16 MHz.

Por su parte Intel incrementó el ancho de bits de su serie x86 a 32 bits con el Intel 80386 [CrGe87] [Cat90] [Brey95] (1985). Tenía 275.000 transistores y reloj de 33 MHz. Incluía la lógica de gestión de memoria en el chip. No utilizó segmentación.

En esa época los fabricantes de minicomputadores como Hewlett-Packard (HP), Digital Equipment Corp. (DEC), Tektronix y Apollo comienzan a usar microprocesadores para sus CPUs, creando el mercado de las estaciones de trabajo [Tred96], que pronto consiguieron ser mucho más potentes que los mainframes de las generaciones anteriores.

A principios de los años 80, John Hennessy, de la Universidad de Stanford, y David Patterson, de la Universidad de Berkeley, definen la base de los procesadores actuales. Estudios dinámicos sobre ejecución de las cargas de trabajo habituales revelaron que las instrucciones y los modos de direccionamiento complejos se usaban muy poco. Además, las instrucciones eran muy largas y eso aumentaba el tiempo necesario para leerlas de memoria, cada vez más crítico. De estos estudios surgió la idea de diseñar computadores de repertorio de instrucciones reducido o RISC (*Reduced Instruction Set Computer*), nombre acuñado por D. Patterson. Los repertorios simples facilitaron la labor de construir microprocesadores segmentados ya en 1980-81 [Kogg81][bhcl91]. Los primeros prototipos RISC eran segmentados y el primero en llegar al mercado fue el MIPS R2000 en 1986, seguido del Sparc [Cata91] de Sun, 29000 de AMD, etc... [Feel94] [Gimi87] [Henn96] [Henn99][ibm94] [Pase82] [Kate85].

En el campo de los supercomputadores se presenta en 1986 el Cray XP de cuatro procesadores, que alcanza 713 MFLOPs.



Figura 25: Cray XP

Avances arquitectónicos: segunda etapa

La segunda etapa comienza cuando se pueden integrar un millón de transistores en un CI, a finales de los 80. Los mayores fabricantes de microprocesadores para computadores personales, Intel y Motorola, tenían absoluta necesidad de compatibilidad, por lo que no modificaban significativamente sus arquitecturas. Así para obtener mayor rendimiento con una arquitectura fija se aumentaba la frecuencia de reloj y se usaban los transistores disponibles para añadir memoria cache interna, coprocesadores matemáticos, segmentaciones más profundas y algoritmos de predicción de saltos [Dani96].

En 1989 se lanzan al mercado el i486 y el MC68040 con 1.2 millones de transistores cada uno y rendimiento similar [Yu96]. El MC68040 estaba segmentado en 6 etapas, con memoria cache de instrucciones y datos de 4 Kbytes cada una y coprocesador matemático. El i486 funcionaba a 25 MHz con tecnología de 1 micra y a 50 MHz con la de 0.8 micras, incluía un coprocesador matemático, una cache de 8 Kbytes y soporte para memoria virtual, además de segmentación. El éxito de Intel en el mercado de computadores personales atrae a competidores (AMD, Cyrix, etc.) a crear soluciones compatibles con la familia x86.

En los microprocesadores para estaciones de trabajo aparecen evoluciones de la segmentación como la ejecución supersegmentada y la ejecución

superscalar. En los procesadores supersegmentados se disminuye el tiempo de ciclo a costa de aumentar el número de etapas del pipeline. Los procesadores superescalares son capaces de lanzar más de una instrucción por ciclo con el objeto de explotar el paralelismo a nivel de instrucción (*Instruction Level Parallelism*, ILP). No obstante, los primeros procesadores capaces de ejecutar múltiples instrucciones por ciclo fueron dos computadores de los años 60, el CDC 6600 [Thor64], del que hemos hablado al describir los avances arquitectónicos de la segunda generación, y el IBM 360/91 [AnST67], donde ya se incluye etiquetado de instrucciones y renombramiento de registros. Ejemplos de esta generación son Sun SuperSPARC, el HP-PA 7100, MIPS R4000, Intel 80960CA y Motorola 88110.

Otros hitos importantes de esta generación son la estación de trabajo Sun 3/260 que en 1986 incorpora por primera vez dos niveles de cache, uno interno y otro externo. En 1991 el Cray Y-MP C90, que tiene 16 procesadores, consigue alcanzar 16 GFLOPs.

Destacar por último, por su impacto en las metodologías de diseño actuales, que a finales de los 80 se funda la organización SPEC (*System Performance and Evaluation Company*) (<http://www.specbench.org/spec/>) con el objetivo inicial de proporcionar un método de comparación más realistas para el mercado de servidores y estaciones de trabajo que los hasta la fecha populares MIPS, MFLOPS o los kernels sintéticos tipo Whetstone. La primera propuesta de SPEC hoy se conoce con el nombre de SPEC89 y estaba orientada casi exclusivamente al rendimiento de procesador. A esta versión inicial han seguido tres nuevas entregas, SPEC92, SPEC95 y SPEC CPU2000, así como nuevos benchmarks destinados a medir otros elementos del computador como el sistema gráfico (SPECviewperf y SPECcapc) o el sistema de ficheros (SPECsfc). Desde su aparición ha sido un referente importante utilizado tanto por la industria como dentro del mundo académico donde es uno de los bancos de pruebas más utilizados para explorar nuevas ideas o analizar diferentes alternativas de diseño.

Avances arquitectónicos: tercera etapa

A finales de los 90 hemos asistido a la vertiginosa expansión de Internet y del *Word Wide Web* (WWW), al éxito comercial de los PDAs (*personal digital assistants*) y a la proliferación de otros productos electrónicos de gran consumo como las consolas de video juegos, las cámaras digitales o los reproductores de MP3. Estos cambios, han dado lugar a tres segmentos claramente diferenciados dentro del mercado de los computadores, cada uno de ellos con diferentes aplicaciones objetivo, diferentes requisitos y diferentes tecnologías: computadores personales, servidores y procesadores empotrados. Vemos un poco más en detalle cada uno de ellos

Computadores Personales y Estaciones de Trabajo (*Desktop Computing*)

Es el segmento con mayor volumen de negocio. Comprende desde equipos de gama baja, de menos de 1000 dólares, hasta potentes estaciones de trabajo de más de 10.000 dólares. Para la mayoría de los usuarios, las decisiones de compra vienen determinadas por la relación entre el rendimiento (tanto del procesador como del sistema gráfico) y el precio del sistema. Por lo tanto, el factor precio-rendimiento es el principal objetivo de diseño.

Dentro de este segmento, hemos asistido a una intensa batalla por el dominio del mercado de los PCs compatibles, lo que ha incentivado una importante mejora de prestaciones. Actualmente sólo sobreviven la propia Intel y Advanced Micro Devices (AMD), si bien en el ámbito de los equipos portátiles, lucha por hacerse un hueco, aunque con bastantes dificultades económicas, Transmeta Corporation, con su procesador Crusoe del que hablaremos en la sección 0. Es una familia revolucionaria x86-compatible especialmente diseñada para el mercado de ordenadores móviles de mano y de peso ligero. El procesador Crusoe de alto rendimiento consume 60 al 70 por ciento menos de potencia (según el fabricante) y trabaja mucho más refrigerado que los chips que compiten con él, transfiriendo la parte más compleja del trabajo del

procesador -la determinación de instrucciones a ejecutar y cuando - a software en un proceso llamado *Code Morphing*.

Como consecuencia de esta competencia, el último procesador de la familia x86, el Pentium 4, se ha puesto a la altura de los mejores procesadores RISC incluso en el proceso en punto flotante (Tan solo el IBM Power4, con *multiprocessor-on-chip* supera al Pentium 4 en SPEC CPU2000fp). No obstante, el éxito de este nuevo componente de la familia x86 no está en contradicción con las ventajas atribuidas a la filosofía RISC, ya que aunque la arquitectura Pentium mantiene por cuestiones de compatibilidad la ilusión de una arquitectura x86, internamente se utiliza un núcleo RISC. Para ello, en la fase de decodificación se traducen dinámicamente las complejas instrucciones x86 (IA-32) en microoperaciones más sencillas, que se pueden ejecutar fuera de orden por el núcleo RISC de este procesador. Entre las propuestas del ámbito académico recogidas por el Pentium 4 destacan la Trace Cache o el *Multithreading* Simultáneo (SMT).

El concepto de Trace Cache fue presentado por primera vez en 1996 [RoBS96]. La idea básica es la de capturar el comportamiento dinámico de las secuencias de instrucciones, almacenando trazas de instrucciones en lugar de bloques contiguos (contiguos tras una ordenación estática).

El SMT es otra novedad que lleva bastante tiempo dentro del mundo académico [YaNe95][EELS97]. La idea es permitir que haya instrucciones de diferentes flujos de ejecución conviviendo dentro del procesador. En cada ciclo se realiza la búsqueda de instrucciones para diferentes threads, manteniendo separados recursos como el banco de registros mediante un exhaustivo control en la asignación. En las primeras implementaciones del Pentium 4, el SMT (limitado a 2 *threads* en este procesador) estaba desactivado. Actualmente comienzan a salir unidades (Pentium Xeon) [Inte9] que hacen uso de esta potente posibilidad.

La importancia de las aplicaciones multimedia dentro de este segmento ha motivado la inclusión en todos los microprocesadores de mejoras y extensiones para acelerar específicamente este tipo de aplicaciones. La clave está en que mientras los microprocesadores de propósito general están optimizados para manejar datos de 32 ó 64 bits, en las aplicaciones multimedia es habitual tratar con flujos continuos de datos más cortos (pixels de 8 bits, señal de audio de 16 bits, etc). Para explotar esta característica se han incluido operaciones tipo SIMD, aprovechándose así el ancho de los datapath y las unidades funcionales. Además, es habitual incluir mecanismos automáticos o semi-automáticos para realizar prebúsqueda de datos y operaciones de carga y almacenamiento que evitan (hacen un *bypass*) los diferentes niveles de cache a fin de paliar los problemas de localidad. Recientemente se han añadido nuevas extensiones para tratamiento de gráficos en 3D, extendiéndose las operaciones SIMD a datos en punto flotante (en simple e incluso en doble precisión). El primero en incorporarlas fue AMD, con la extensión 3DNow! para los procesadores K6-II y posteriores, aunque no tardaron en aparecer las extensiones SSE (Streaming SIMD Extensions) y SSE2 al repertorio x86 de Intel y AltiVec de Motorola en el PowerPC G4.

Servidores

El mercado de servidores esta dominado por multiprocesadores simétricos de memoria compartida (SMPs) y por clusters. En este segmento la relación coste-rendimiento no es tan decisiva, siendo en este caso mucho más relevantes factores como la alta disponibilidad, la escalabilidad o la productividad (*throughput*). La alta disponibilidad hace referencia a la necesidad de que los sistemas estén operativos en todo momento, lo cual lleva inherente la necesidad de algún tipo de redundancia, ya que en servidores de gran escala son inevitables los fallos. La escalabilidad es también un aspecto importante, ya que las necesidades de cómputo, memoria, disco o entrada-salida de los servidores suele crecer durante el tiempo de vida de estos sistemas. Por último, aunque es importante el tiempo de respuesta que

ofrecen, las métricas de rendimiento más importante a la que deben hacer frente los diseñadores de estos sistemas, como los SPECrate, los TPC o los recientes SPECintFS y SPECweb, son medidas de productividad.

Los SPECrate se obtienen a partir de los SPEC CPU2000 ejecutando múltiples instancias de dichos benchmarks (habitualmente una por procesador). No obstante, como la productividad de los servidores no sólo depende de la capacidad de cómputo, SPEC desarrolló los SPECintFS, diseñados para evaluar no sólo el procesador sino también el sistema de entrada-salida, tanto el almacenamiento secundario (discos) como el interfaz de red. La importancia del WWW dentro de este segmento también ha sido recogida por SPEC mediante el SPECweb, otro benchmark orientado a medir la productividad en el que se simula un entorno de múltiples clientes que solicitan páginas (estáticas y dinámicas) y envían información a un servidor Web.

Los benchmarks con más historia dentro de este segmento son los TPC. De hecho, el *Transaction Processing Council (TPC)* fue creado con anterioridad a la organización SPEC (a mediados de los 80), con el objetivo de crear benchmarks realistas para el procesamiento de transacciones. Como en el caso de SPECintFS y SPECweb, los TPC evalúan el comportamiento global del sistema, es decir, no solamente el procesador sino también el subsistema de entrada-salida, el sistema operativo y el gestor de base de datos utilizado. Las métricas utilizadas por esta organización son transacciones por minuto (TPM) y TPMs por dólar, aunque también incluyen requisitos para el tiempo de respuesta (sólo se consideran las transacciones que satisfacen dichos requisitos). Para modelar configuraciones realistas, tanto el número de usuarios como el tamaño de la base de datos se escala con la potencia del servidor.

Entre los microprocesadores utilizados en este segmento se encuentra la gama Pentium Xeon de Intel, y otros microprocesadores tipo RISC (también disponibles en el mercado de estaciones de trabajo) como el UltraSPARC III de SUN, el Power 4 de IBM, el HP PA-8700, el Alpha 21264 y el MIPS R14000

utilizado en las máquinas de SGI. No obstante, algunos de estos procesadores tienen un futuro incierto. La tecnología Alpha fue adquirida por Intel, que ha apostado por una nueva arquitectura conocida como **IA-64**, en cuyo desarrollo también participa HP y en la que han demostrado gran interés otros fabricantes como SGI. Esta arquitectura ha suscitado un gran debate dentro del área por hacer uso de **VLIW** (*Very Long Instruction Word*) del que hablaremos también en la sección 0.

Finalmente destacar que durante los últimos años hemos asistido a la irrupción de los **clusters** como una alternativa económica, especialmente a los multiprocesadores de gran escala.

Procesadores Empotrados

Los procesadores empotrados representan el segmento del mercado con un mayor crecimiento. Están presentes en multitud de dispositivos, desde tarjetas inteligentes y controladores industriales a sofisticados conmutadores de red o consolas de videojuegos. Es por ello el segmento en el que se aprecia una mayor variedad tanto en prestaciones como en coste. El factor de diseño más importante es en este caso el precio. Existen obviamente algunos requisitos relativos con el rendimiento, a menudo relacionados con alcanzar tiempos de respuesta en tiempo real, pero el primer objetivo suele ser alcanzar dichas prestaciones con el mínimo coste posible. De hecho, a diferencia de los computadores personales o los servidores, los benchmarks para este segmento puede considerarse que están aún en su infancia. El intento de estandarización que ha tenido más éxito hasta la fecha son los denominados EEMBC (*EDN Embedded Microprocessor Benchmark Consortium*). Sin embargo, muchos fabricantes siguen facilitando el resultado de kernels sintéticos ya obsoletos en los otros segmentos como Dhrystone o medidas basadas en MIPS.

Para garantizar tiempo real, los diseñadores de este tipo de sistemas tratan de optimizar el peor caso posible, a diferencia de los microprocesadores de propósito general donde siempre se intenta favorecer las situaciones más

probables a costa incluso de penalizar al resto. Otros factores de diseño importantes en algunas aplicaciones el tamaño de la memoria y el consumo de potencia. El primer factor esta relacionado directamente con el coste del sistema, aunque también se relaciona con el consumo de potencia al ser la memoria uno de los componentes de mayor consumo. Este factor se traslada a menudo en un énfasis por reducir el tamaño de los códigos, existiendo en algunos casos soporte hardware para este propósito. La preocupación por el consumo de potencia esta relacionada en la mayoría de los casos por el uso de baterías. No obstante también guarda relación con el coste, ya que un menor consumo permite por ejemplo utilizar empaquetados plásticos, más económicos que los cerámicos, y evita la necesidad de incorporar ventiladores.

Dentro de este segmento, el diseñador puede optar por tres aproximaciones diferentes:

- Combinación Hardware/Software que incluye algún circuito de propósito específico integrado junto a algún core.
- Procesador empotrado genérico (*off-the-self*) con el software específico para resolver el problema.
- DSP (Digital signal processor) o un procesador multimedia (media processor) con el software específico para resolver el problema.

En En los 90 la diferencia entre microprocesadores y microcontroladores se ha ensanchado cada vez más. Los primeros buscan aumentar su rendimiento y los segundos se concentran en disminuir el coste y aumentar la integración del sistema, incluyendo en el mismo chip funciones analógicas, todo tipo de memorias y sensores.

Conclusiones

En los primeros años de su existencia, los computadores tenían un coste muy elevado, por lo que su uso y comercialización estaban restringidos para cálculos complejos y por lo tanto los usuarios eran especialistas.

En la segunda generación se produce un gran avance tecnológico con la invención del transistor. Los computadores se usan, como anteriormente, para cálculos científicos complejos. Por eso se aprovecha la mejora tecnológica para aumentar las capacidades de cálculo de los computadores con segmentación y paralelismo y se crean sistemas operativos por lotes y lenguajes de alto nivel para mejorar el aprovechamiento de los mismos. Aparecen los supercomputadores.

En la tercera generación se amplía la gama de los computadores con la creación de los minicomputadores y mainframes. Las mayores capacidades que ofrecen los circuitos integrados permiten o bien aumentar las funciones del computador para resolver problemas aún más complejos, o disminuir su coste, de forma que se abre el mercado a nuevos usuarios, todavía del entorno de la ciencia o empresarial.

Hasta esta época, aproximadamente 1970, los cursos de arquitectura de computadores impartían fundamentalmente aritmética de computadores [Patt98].

A partir de entonces se produce un enorme cambio en el uso de los computadores con la invención de los microprocesadores y la memoria de semiconductores en la cuarta generación. Esto reduce los costes de los computadores, de forma que están disponibles para muchas nuevas aplicaciones. Aparecen los microcontroladores y los computadores personales. Simultáneamente se aprovecha la mayor densidad de circuitos integrados para construir mainframes cada vez más potentes para uso científico y empresarial.

También se crean los supercomputadores vectoriales, para cálculos científicos complejos.

Esto se refleja en que durante la década de los 70 los cursos de arquitectura estudiaban básicamente diseño de la arquitectura del repertorio de instrucciones, especialmente repertorios apropiados para facilitar la tarea de los compiladores (CISC).

Finalmente, en la quinta generación se ha producido la invasión de la sociedad por parte de los sistemas basados en microprocesador. La tecnología proporciona circuitos cada vez más complejos y rápidos, pero la desigual evolución de la velocidad del procesador, la de memoria y la de E/S sugiere cambiar el enfoque seguido para el diseño de la arquitectura (sólo ocuparse del procesador) y considerar el diseño del sistema completo. Se amplía la jerarquía de memoria y se usan buses jerárquicos (local, del sistema, de expansión). Hay mejoras en la gestión de E/S con DMA y procesadores de E/S.

En resumen, los sistemas basados en microprocesador, tanto los computadores personales como las estaciones de trabajo, han aumentado de rendimiento y complejidad de forma vertiginosa. Esto permitió en un primer momento que las estaciones de trabajo sustituyeran a los minicomputadores y las mainframes y posteriormente que los mismos PCs se incorporen a este mercado. De hecho, la mayor parte de los supercomputadores y sistemas multiprocesador actuales se construyen con componentes del mercado de servidores y de PCs de gama alta. Por ello, los cursos de arquitectura actuales imparten diseño de CPU, del sistema de memoria, del de E/S y multiprocesadores. También es necesario ampliar la oferta de materias optativas para complementar la formación y los conocimientos.