

Buses, procesadores y memorias

Índice

3.1. Buses	103
3.1.1. Capacidad de transmisión de datos	103
3.1.2. Sistema de Buses de un ordenador	106
3.1.3. Jerarquía de Buses	109
Bus único	111
Buses local, del sistema y de expansión	112
Otras topologías: Buses para dispositivos de distinta velocidad	114
3.1.4. Bus del sistema	116
Intel: Bus frontal (FSB)	116
AMD: Hypertransport	117
3.1.5. Buses de expansión	119
ISA (Industry Standar Architecture)	121
EISA (Extended ISA)	121
MCA (Micro Channel Architecture)	122
PCI (Peripheral Component Interconnect)	122
PCIe (PCI Express)	123
3.1.6. Buses Dedicados	128
AGP (Accelerated Graphics Port)	128
IDE (Integrated Drive Electronics)	129
SCSI (Small Computer System Interface)	130

3. Buses, procesadores y memorias

SATA (Serial ATA)	130
USB (Universal Serial Bus)	132
IEEE 1394 o FIREWIRE	133
DVI (Digital Visual Interface)	134
HDMI (High-Definition Multi-media Interface)	134
3.2. Procesador	136
3.2.1. Magnitudes principales de un microprocesador	137
Frecuencia de reloj	137
Tecnología de integración	138
Paralelismo a nivel de instrucción	139
Memoria interna.	141
Conjunto de instrucciones: CISC y RISC	143
3.2.2. Estructura del microprocesador	144
Unidad de control	144
Unidad de proceso	146
Ejecución de una instrucción	146
3.3. Memoria	148
3.3.1. Memoria Principal - RAM Dinámica	151
Parámetros funcionales	151
Paridad	152
Zócalos de conexión	152
Tipos de organización	153
3.3.2. Memoria Caché - RAM Estática	158
3.3.3. Memoria ROM (Read Only Memory)	160
3.4. Anexo. Las generaciones de Microprocesadores	161
3.5. Ejercicios	163

Buses

Podríamos definir a grandes rasgos el sistema de bus de un ordenador como una especie de red de comunicaciones similar a una red de autopista con muchos carriles a través de los cuales se intercambian datos los diferentes elementos del computador.

La estructura del sistema de bus consta de 3 redes: Bus de datos, de direcciones o de control.

Físicamente los buses están constituidos por finísimas líneas conductoras impresas sobre la placa base, por donde se transmiten los datos e instrucciones por diferentes de tensión en forma de señales digitales.

El ordenador cuenta con un sistema de regulación de tráfico, es decir, con un controlador de buses: **El chipset**.

El chipset es el encargado de regular el enorme tráfico de datos que circula a través de los buses.

3.1.1 Capacidad de transmisión de datos

La capacidad de transmisión de datos de los buses depende de dos parámetros:

1. **Ancho del bus.** Lo constituye el número de líneas que lo forman, por cada línea o carril solamente puede circular un bit, la unidad mínima de información digital. Si el bus tiene 8 líneas se dice que el bus tiene 8 bits, si tiene 16 se dice que tiene 16 bits, así sucesivamente en 32, 64, 128. . . . El ancho del bus lo determina el microprocesador. Fíjese en la imagen 3.1
2. **Frecuencia del bus.** Se determina como el número de ciclos por segundo, es decir:

$$Frecuencia = \frac{Numerodeciclos}{Numerodesegundos} \quad (3.1)$$

Siguiendo la ecuación 3.1 se obtienen los siguientes resultados:

- 1 Hz → 1 ciclo/seg

3. Buses, procesadores y memorias

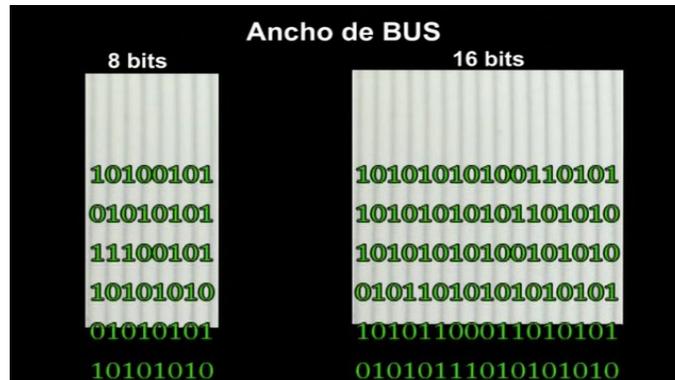


Figura 3.1: Ancho de bus de 8 y 16 bits

- 1 KHz → 1000 ciclos/seg
- 1 Mhz → 1.000.000 ciclos/seg
- 1 Ghz → 1000.000.000 ciclos/seg

Ejemplo. Si el bus permite transmitir 33 millones de señales o datos por segundo diríamos que dicho bus tiene una frecuencia de 33 Mhz. Evidentemente cuanto mayor sea la frecuencia del bus más rápido será el bus transmitiendo datos.

La fórmula para calcular la capacidad de transmisión de los buses es:

$$Capacidad\ en\ MB/seg = \frac{Frecuencia \cdot ancho\ del\ bus}{8} \quad (3.2)$$

Donde frecuencia se expresa en Mhz, ancho de bus en número de bits, y 8 es el número de bits que forman un byte.

Ejemplo 1. El bus PCI que trabaja a una frecuencia de 33 Mhz y un ancho de bus de 32 bits/ciclo permite transmitir:

$$Capacidad = \frac{(33 \cdot 32)}{8} = 133 Mb/seg$$

Ejemplo 2. El bus AGPx2. Su frecuencia es (66 Mhz x 2 datos por ciclo) y su ancho de bus 32 bits/ciclo.

$$Capacidad = \frac{(133 \cdot 32)}{8} = 528 Mb/seg$$

3. Buses, procesadores y memorias

Ejemplo 3. Memoria SDRAM de frecuencia 133 Mhz y ancho de bus de 32 bits/ciclo.

$$Capacidad = \frac{(133 \cdot 32)}{8} = 532 Mb/seg$$

Llegados a este punto es obligado hablar del concepto de velocidad interna y externa de un PC.

- **Velocidad interna.** Es la velocidad a la que trabaja el microprocesador.
- **Velocidad externa.** Es la velocidad a la que trabaja el bus del sistema, es decir, la velocidad a la que se comunica el microprocesador con la placa base. Se pueden encontrar velocidades de 400, 800, 1066, etc. . .

A partir de estos conceptos se deriva otro concepto conocido como **multiplificador**. El multiplicador es la relación que existe entre la velocidad interna y externa, véase la figura 3.2:

$$Multiplicador = \frac{Velocidad\ Interna}{Velocidad\ Externa}$$

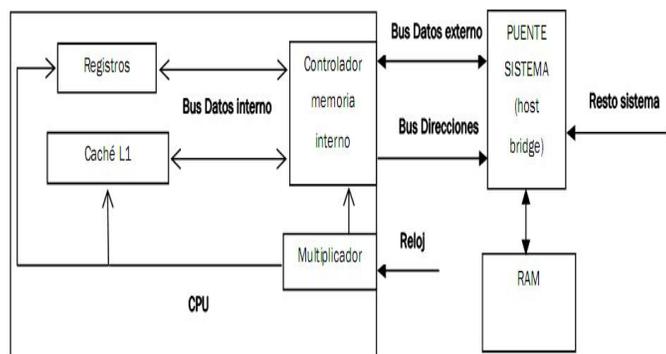


Figura 3.2: Estructura interna de la CPU

Así, por ejemplo, un Pentium 4, de 3200 Mhz en una placa base con un bus del sistema de 800 Mhz tiene un multiplicador de 4X:

$$4X = \frac{3200}{800}$$

Aparte del bus del sistema han ido apareciendo los diferentes estándares: **ISA, VESA, EISA, PCI, AGP, IDE, PCI Express.**

3.1.2 Sistema de Buses de un ordenador

El sistema de buses varía en función de la familia de procesadores, pero en todas las familias encontraremos el siguiente sistema de buses:

- **Bus del sistema.** El cual está formado por Bus de datos, direcciones y control.
- **Buses subsidiarios.** Donde encontramos buses de alta frecuencia y de baja frecuencia.
 - **De alta frecuencia.** AGP, PCI (cada vez más cercano a bus de baja frecuencia), PCIe
 - **De baja frecuencia.** ISA, IDE, USB

El bus del sistema es aquel a través del cual la placa base mediante el chipset northbridge(puente norte) se comunica con el procesador y la memoria principal. Figura 3.3

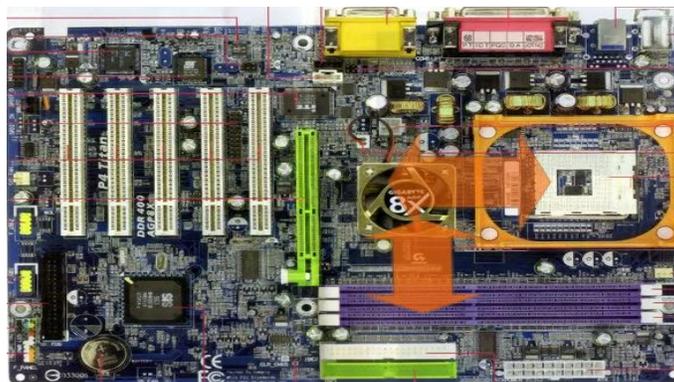


Figura 3.3: Bus del sistema que comunica la placa base con el procesador y la memoria

Los buses secundarios o subsidiarios pueden ser de alta frecuencia que son aquellos que cuelgan del chipset northbridge y de baja frecuencia que son los gobernados por el chipset southbridge. Figura 3.4.

3. Buses, procesadores y memorias

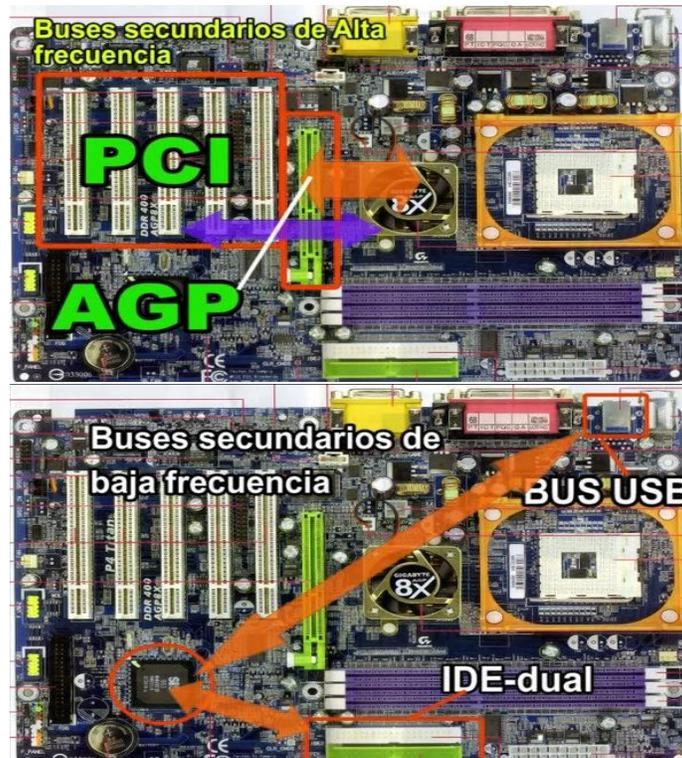


Figura 3.4: Buses subsidiarios

Ejemplos

Para comprender bien lo anteriormente expuesto vamos a ilustrarlo con varios ejemplos.

Ejemplo 1. Pentium 4 con chipset SiS 648 FX

En la gama del Pentium 4 tenemos un bus del sistema de 400/533/800 Mhz que resultan de multiplicar por 4 las velocidades reales del bus de la CPU que pueden ser de 100/133/200 Mhz. Esto se explica porque en el Pentium 4 puede sincronizar tanto el procesador como el chipset 4 operaciones por ciclo.

Esta frecuencia del bus da soporte a los módulos de memoria que trabajan a 400 Mhz puesto que en realidad físicamente el bus del sistema es de 100/133/200 Mhz también puede utilizarse placas bases que soporten memoria de 100 y 133 Mhz.

3. Buses, procesadores y memorias

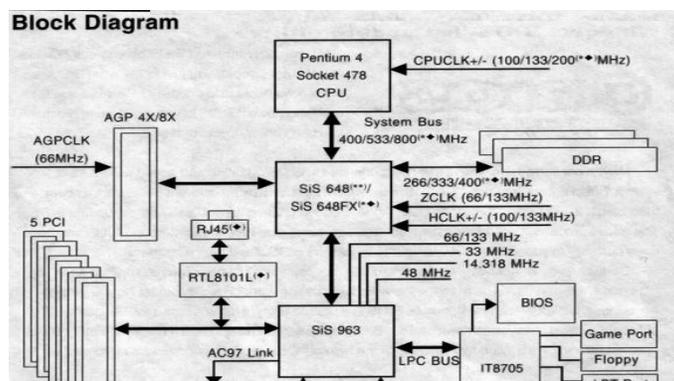


Figura 3.5: Diagrama de bloques del Pentium 4. Imagen 1

Según el diagrama que estamos observando, en la figura 3.5, el único bus secundario de alta frecuencia, es decir el que cuelga del chipset northbride, es el bus AGP que aunque tiene una frecuencia básica de 66 Mhz lo normal es encontrarlo como 4X y 8X lo que en realidad implican unas frecuencias de 264 y 528 Mhz.

En cuanto a los buses secundarios de baja frecuencias, es decir, los que cuelgan del chipset southbride, tenemos 5 ranuras PCI a 33 Mhz, 6 puertos USB a 48 Mhz, 2 puertos IDE ATA 33/66/100/133 Mhz, una tarjeta de sonido integrada AC97, un chip controlador IT8705 que trabajan a 24 y 33 Mhz que a su vez cuelgan la BIOS, una disquetera, un puerto LPT, dos puertos PS/2 y dos puertos COM.

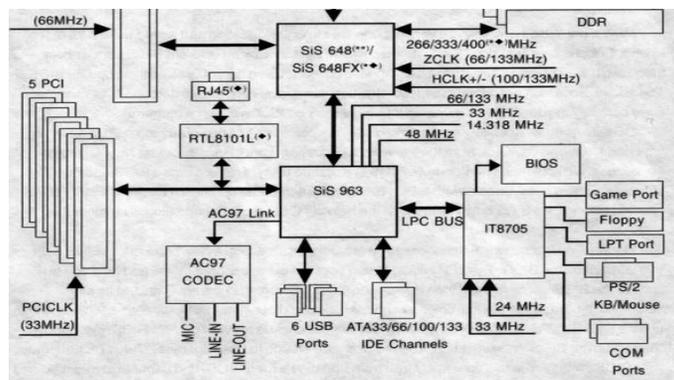


Figura 3.6: Diagrama de bloques del Pentium 4. Imagen 2

3. Buses, procesadores y memorias

Ejemplo 2. AMD K7 en placa base 7VT600P-RZ(-C)

Vemos en la figura 3.7 que el bus del bus del sistema puede trabajar a 100/133/166/200 Mhz.

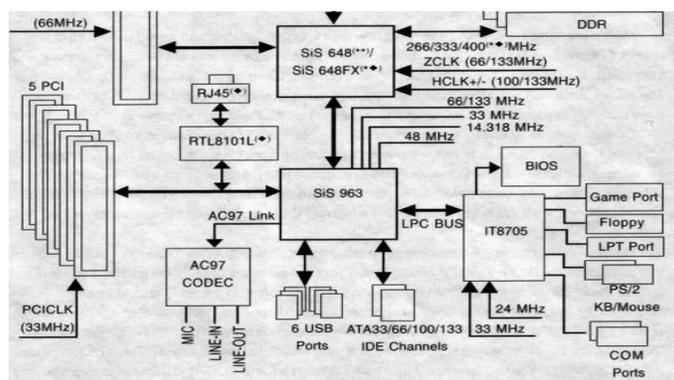


Figura 3.7: Diagrama de bloques del AMD K7 con placa base 7VT600P-RZ

También podemos observar que del diagrama Northbridge cuelga el bus AGP que trabaja a frecuencias 66 Mhz a 4x y 8x.

Las memorias DDR RAM que pueden trabajar a la misma velocidad del sistema y el chipset Southbridge que trabaja a frecuencias con 14,318Mhz, 33Mhz, 48Mhz y 66 Mhz.

Del chipset Southbridge cuelgan 5 ranuras PCI con frecuencia de 33Mhz, 2 puertos Serial ATA, 8 puertos USB a 48Mhz, dos puertos ATA 66/100/133 Mhz, 2 puertos PS/2, una tarjeta de sonido integrada, y un chip controlador de puertos locales IT8705 de este cuelgan a su vez el BIOS, disquetera, 1 puerto LPT y 2 puertos COM

3.1.3 Jerarquía de Buses

La jerarquía de buses del PC es bastante grande y su variedad continúa aumentando, para que un equipo pueda usar un tipo de bus concreto su placa base debe llevar implementado el controlador para ese bus y suministrar el zócalo correspondiente.

Es posible, pues, definir un bus como una línea de comunicación entre dispositivos. Particularizando, lo que se denomina ‘dispositivos’ de forma genérica

3. Buses, procesadores y memorias

pueden ser registros de la propia CPU, dispositivos propios del computador (teclado, memoria, etc.) u otros computadores propiamente dichos.

Los buses se pueden clasificar según diversos parámetros:

- **Según el método de transmisión.**
 - **Serie.** Los bits de información se envían uno tras otro.
 - **Paralelo.** Se pueden enviar varios bits simultáneamente.

- **Según el tipo de información que transportan.**
 - **Bus de datos.** Transportan datos e instrucciones.
 - **Bus de direcciones.** Dirección de memoria o periféricos (n bits direccionan 2^n).
 - **Bus de control.** Señales de estado y control.

- Según el nivel de jerarquía.
 - **Buses internos.**
 - **Buses de placa.**
 - **Buses dependientes del equipo: bus local, FSB.**
 - **Buses independientes del equipo:**
 - ◇ **Buses de expansión.** PCI y PCIe.
 - ◇ **Buses dedicados.** IDE, SCSI, SATA, AGP.
 - **Buses externos.** Interfaz paralelo, el serie, USB, Firewire, eSATA.

Dado que en el sistema coexisten multitud de buses, muchas veces es necesario realizar una conversión de información de uno a otro, para este menester se ubican unos chips especiales en los puntos de interconexión entre ambos que reciben el nombre de **puentes de conexión**.

En las primeras arquitecturas para PC, todos los componentes del equipo funcionaban a una misma velocidad: el procesador, su memoria, los distintos periféricos... conforme la frecuencia del microprocesador va emprendiendo su meteórica progresión, los elementos que tienen necesidad de seguirle tratan de hacerlo, mientras que los más lentos mantienen su velocidad y se van agrupando en estratos inferiores que conforman un diálogo común, de esta forma, se va produciendo una

3. Buses, procesadores y memorias

fragmentación de las vías de comunicación del PC, pasándose pronto de una arquitectura de bus único, figura 3.8, a otra con una jerarquía de buses, unos muy anchos y rápidos y otros más estrechos y/o lentos, siempre en función del número y la velocidad de los dispositivos que los utilizan, figura 3.9.

Bus único

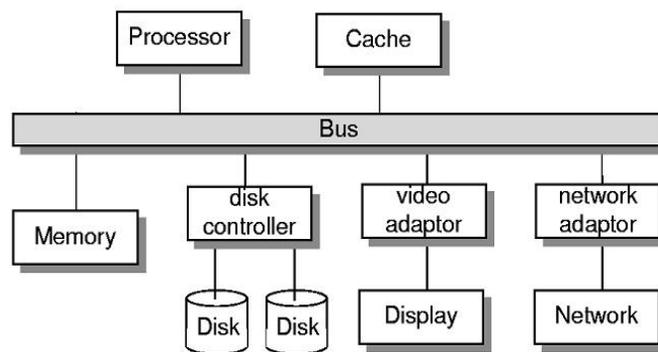


Figura 3.8: Jerarquía de buses. Bús único

Cuando queremos conectar un gran número de dispositivos a un mismo bus, nos encontramos con dos problemas fundamentales.

- **Disminución del rendimiento global del sistema.**

- Aumenta el **retardo de propagación** de las señales porque
 1. El bus debe tener **mayor longitud** para soportar mayor número de dispositivos.
 2. Las **señales de arbitraje** (GRANT), si son encadenadas, debe propagarse a través de un mayor número de posible maestros (masters).
- El bus puede actuar como un **cuello de botella** si la demanda de la transferencia es mayor que la capacidad del bus los dispositivos deberán esperar mucho tiempo para poder transmitir.
- La **diferencia de velocidad de los dispositivos** afecta negativamente al rendimiento global en el mismo tiempo que un dispositivo lento realiza una transferencia, uno rápido podría haber realizado miles de transferencias.

3. Buses, procesadores y memorias

- **Incompatibilidad del bus con los dispositivos** porque existen dispositivos de E/S diseñados para un determinado bus, que son incompatibles con otros computadores que utilizan un bus distinto.
 - **Solución ideal.** Que todos los computadores utilizaran un estándar de bus uniforme.
 - **Problema.** Cada fabricante diseña sus propios buses optimizados para sus arquitecturas, por lo que es muy difícil que todos se pongan de acuerdo.

Veamos a continuación un ejemplo de reducción del rendimiento del sistema de un sistema con bus único. Imaginemos un procesador de 200Mhz (tiempo ciclo = 5 ns), ciclo medio por instrucción: CPI = 2 ciclo. Por tanto una instrucción tarda en promedio $2 \times 5 \text{ ns} = 10 \text{ ns}$. El computador puede ejecutar aproximadamente 100 MIPS (Millones de Instrucciones Por Segundo). Si el procesador se conecta a la cache y al resto de dispositivos a través de un único bus del sistema cuando se realiza una operación de E/S se detiene la actividad del procesador, ya que no puede leer instrucciones de la caché mientras el bus está ocupado. Imaginemos ahora que disponemos de un disco duro que tiene un tiempo de acceso de 10 ms y una velocidad de transferencia de 10 MB/seg. Si quisiéramos realizar una transferencia de 512 KB de disco a memoria el bus del sistema se ocuparía durante el siguiente periodo:

$$\text{Tiempo} = 10\text{ms} + \frac{512\text{KB}}{10,000\text{KB/s}} = 61,2\text{ms}$$

En ese tiempo, la CPU podría haber ejecutado:

$$(0,0612\text{segundos}) \times (100 \times 10^6 \text{instrucciones/segundo}) = 6,12 \text{ millones de instrucciones}$$

Buses local, del sistema y de expansión

Observe la figura 3.9, como existe una jerarquía de buses donde podemos encontrar:

- **Bus local y bus del sistema.** Son buses rápidos y cortos, aunque son buses propietarios de empresas privadas (no estándares), el hecho de que sean buses propietarios proporcionan las características de que estén optimizados para la arquitectura en cuestión y que tenga un número fijo de dispositivos de prestaciones conocidas.

3. Buses, procesadores y memorias

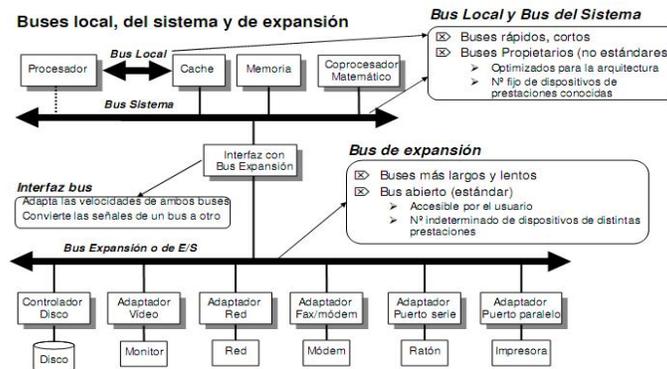


Figura 3.9: Jerarquía de buses. Buses local, del sistema y de expansión

- **Bus de expansión.** Son buses más largos y lentos aunque son buses abiertos, es decir, que existen estándares y este hecho conlleva las características de que puedan ser accesibles por el usuario y que el número indeterminado de dispositivos de distintas prestaciones. Se pueden conectar los controladores de disco, adaptadores de vídeo, de red, fax, módem, puerto serie, paralelo, ...
- **Función del interfaz o adaptador de bus.** Es el encargado de adaptar las velocidades de ambos buses. El bus del sistema es, en general, más rápido que el bus de expansión por tanto el adaptador debe actuar como buffer de almacenamiento intermedio para evitar la pérdida de datos.

Las ventajas que conlleva tener una jerarquía de buses:

- El bus local entre el procesador y la cache aísla el tráfico de E/S de procesador. Se puede transferir información entre la memoria y la E/S sin interrumpir la actividad del procesador.
- El bus de expansión reduce el tráfico en el bus del sistema. La transferencia entre caché y memoria principal se pueden realizar de forma más eficiente puesto que:
 - Se pueden realizar una transferencia de memoria cache a memoria principal al mismo tiempo que el interfaz recibe datos desde un dispositivo de E/S.

3. Buses, procesadores y memorias

- El procesador+caché o el coprocesador tienen la misma ‘prioridad’ en el acceso al bus que todos los dispositivos conectados al bus de expansión de forma conjunta.
- Se elimina el problema de la incompatibilidad.
 - El bus local y del sistema suelen ser propietarios (no estándar) y están optimizados para cada arquitectura en particular.
 - Los buses de expansión son buses estándares o abiertos (ISA, EISA, PCI, PCIe, ...). Esto conlleva ciertas características:
 - Los buses estándares son independientes del computador.
 - Estos buses tienen unas características y especificaciones perfectamente definidas.
 - Existe una amplia gama de controladores o adaptadores para periféricos compatibles con estos buses.
 - La conexión de un controlador a un bus estándar es sencilla y rápida (mediante conectores estándares).
 - Podemos utilizar los mismos controladores y periféricos en otro computador que disponga del mismo bus estándar.

Otras topologías: Buses para dispositivos de distinta velocidad

Observe la figura 3.10 donde se puede observar la arquitectura de esta topología donde encontramos:

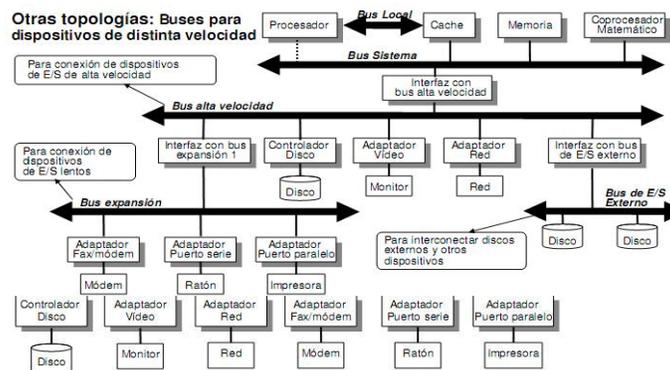


Figura 3.10: Jerarquía de buses. Otras topologías

3. Buses, procesadores y memorias

- **Bus local.**
- **Bus del sistema.**
- **Bus de alta velocidad.** Para conexión de dispositivos de E/S de alta velocidad como son el controlador del disco, adaptador de vídeo, red, etc.
- **Bus de E/S externo.** Para interconectar discos externos y otros dispositivos como pueden ser discos.
- **Bus de expansión.** Para conexión de dispositivos de E/S lentos como pueden ser el módem, ratón, impresoras. . .

La principal ventaja de usar buses de expansión de distintas velocidades es la de **aumentar la eficiencia del sistema** puesto que:

1. Con un único bus de expansión los dispositivos lentos pueden degradar el rendimiento global.
2. Con varios buses de expansión, para dispositivos de distintas velocidades, un dispositivo rápido tiene la misma ‘probabilidad’ de acceder al bus que todos los dispositivos lentos conjuntamente.

Para optimizar el rendimiento de los buses del sistema debemos tener siempre presente que cuanto más rápido sea un dispositivo, más cerca del procesador debemos conectarlo, por ello, las memorias se enganchan al bus local, mientras que los periféricos van enganchados al bus de expansión, normalmente al PCI, bien mediante zócalos dedicados (como ocurre con el disco flexible, el disco duro y el CD-ROM), o a través de los zócalos de expansión PCI (caso de las tarjetas gráficas que no son AGP). Los zócalos de expansión tipo PCI son cada vez menos usados porque el bus PCI es demasiado lento para los requerimientos de muchos de los dispositivos actuales.

Los buses del PC constituyen hoy por hoy uno de los cuellos de botella de la arquitectura del equipo en general, debido a que las continuas mejoras introducidas en los microprocesadores han provocado una brecha considerable en términos de velocidad (frecuencia) con respecto a la de los buses.

3.1.4 Bus del sistema

Intel: Bus frontal (FSB)

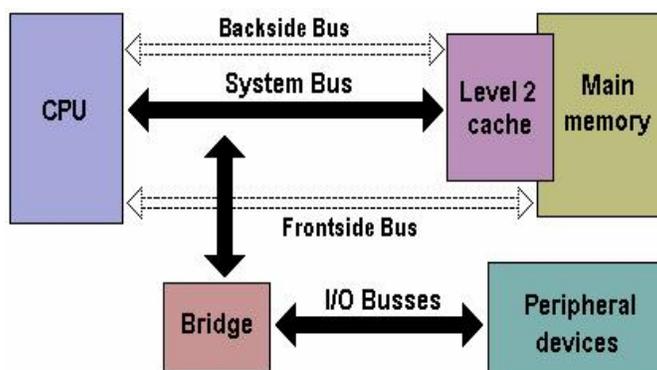


Figura 3.11: Bus del sistema. Bus frontal (FSB)

Bus de la parte frontal o más conocido como FSB (Front Side Bus), es el termino usado para referirse a uno de los buses más importantes que existen en un computador. Este bus es bidireccional y comunica a la CPU con el puente norte (northbridge). Este bus se descompone en tres líneas que transmiten datos, direcciones y control, así como señales de reloj que sincronizan el funcionamiento del mismo.

La frecuencia de trabajo del microprocesador se obtiene como resultado de multiplicar la frecuencia de reloj del FSB por un factor multiplicador. Este factor multiplicador, así como la frecuencia de reloj del FSB pueden alterarse a través de la configuración de la placa base, generalmente a través de la BIOS, permitiendo así el overclocking. Por ejemplo una CPU de 1400 MHz, podría funcionar con una frecuencia de reloj de 133 MHz y un factor multiplicador de 10,5.

La máxima velocidad teórica del FSB está determinada por su ancho de banda (que puede ser distinto de un sistema a otro). Por ejemplo, un FSB de 32 bits de ancho, funcionando a 200MHz ofrece un máximo de 800 MB/s

$$4\text{Bytes} \cdot 200\text{MHz} = 400\text{MB/s}$$

Por otra parte si se usa la tecnología **Quad Pumping**, si el bus funciona a 100

3. Buses, procesadores y memorias

MHz de señal de reloj, en cada ciclo de reloj hay cuatro transferencias de datos. Se dice entonces que el bus funciona a **400 MT/s**, y su ancho de banda es:

$$4\text{Bytes} \cdot 100\text{MHz} \cdot 4 = 1600\text{MB/s}.$$

Veamos algunos ejemplos en la siguiente tabla, donde se muestran distintas versiones de bus de sistemas usados por Intel:

Bus del sistema	Pentium	PII-III	PIV
Frecuencia (MHz)	66	100/133	100/133/200/267/333/400
Transacciones por ciclo	1	1	4
Frecuencia efectiva (MT/s)	66	100/133	400/533/800/1067/¿?/¿?
Bus de datos	64 bits	64 bits	64 bits
MB/s máximo	503,5	763/1015	¿?

Ejercicio 3.1 En la tabla anterior, hallar las velocidades marcadas con ¿?

Algunos ordenadores tienen una Memoria Caché L2 o L3 externa a la propia CPU conectados mediante un *Back Side Bus* (Bus trasero o bus de la parte de atrás). El acceso a la memoria caché, conectada a este bus, es más rápido que el acceso a la Memoria RAM por el FSB. En la actualidad, la caché L2, ha sido incluida en el chip del microprocesador, junto con la caché L1.

AMD: Hypertransport

En la arquitectura usada por AMD el controlador de memoria se encuentra encapsulado en el micro, y el bus del sistema es el **Hypertransport.**, fíjese en la figura 3.12.

Hypertransport (formalmente LTD: Lightning Data Transport) es un bus de alta velocidad de transferencia registrado por HyperTransport Technology Consortium para la interconexión de circuitos integrados. Está pensado para la conexión entre chips de alta velocidad como procesador y chipset o conexión entre procesadores en sistemas multiprocesador. Es utilizado, por ejemplo, por toda la familia de procesadores AMD.

Existen tres versiones de Hypertransport 1.0, 2.0 y 3.0, que puede llegar a funcionar hasta 2,6 GHz.

Su velocidad de transferencia máxima utilizando líneas de 32 bits, es de 20,8 GB/s (2.6GHz x (32 bits/8)). Al ser el bus bidireccional se puede conseguir una

3. Buses, procesadores y memorias

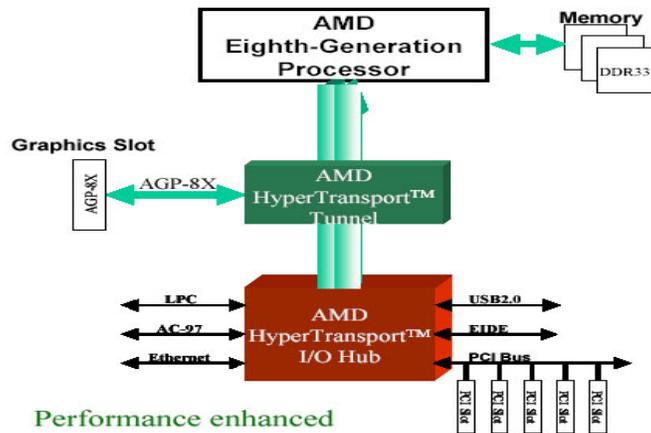


Figura 3.12: Bus del sistema. Hypertransport

transferencia total de 41,6 GB/s sumando las dos direcciones. Lo cual consigue desbancar totalmente al FSB.

Las principales ventajas son:

- Baja latencia, alta velocidad.
- Diseño simple que permite flexibilidad en el número de conexiones.

Sus prestaciones son:

- Frecuencia de reloj ajustable (200, 300, 400, 500, 600, 800, ... MHz) con 2 datos por ciclo.
- Puede ampliarse la anchura del bus añadiendo más enlaces punto a punto (2, 4, 8, 16 y 32 bits).

Veamos la siguiente tabla donde se puede ver las velocidades y anchos de bandas máximos para las diferentes versiones de hypertransport.

Versión de Hypertransport	Año	Max.Frecuencia	Enlace	Max. BW
1.0	2001	800MHz	32 Bit	12.8GB/s
1.1	2002	800MHz	32 Bit	12.8 GB/s
2.0	2004	1.4 GHz	32 Bit	22.5 GB/s
3.0	2006	2.6 GHz	32 Bit	41.6 GB/s

3. Buses, procesadores y memorias

Ejercicio 3.2 *Completa la tabla anterior con la especificación de Hypertransport 3.1, aparecido en 2008.*

Un ejemplo de HyperTransport 3.0. Supongamos una velocidad de reloj de 2,6Ghz, enviando 2 datos por ciclo y usando un ancho del bus de 16 bits tendríamos un ancho de banda de

$$2,6 \cdot 10^9 \text{ ciclos/s} \times 2 \text{ datos/ciclo} \times 16 \text{ bits/dato} = 83,2 \text{ Gb/s} = 10,4 \text{ GB/s} \text{ cada subenlace} \rightarrow 20,8 \text{ GB/s} \text{ cada enlace (upstream + downstream).}$$

Ejercicio 3.3 *Construye una tabla sobre el ancho de banda de Hypertransport 3.0 (32 bits) si la frecuencia del reloj fuese de 1,6GHz, 1,8GHz, 2,0GHz, 2,4GHz.*

Ejercicio 3.4 *Construye una tabla comparativa sobre las velocidades y el ancho de banda de los buses locales (FSB e Hypertransport) actuales en la gama Pentium y AMD.*

3.1.5 Buses de expansión

Los buses de expansión permiten la comunicación de la CPU y de la memoria con los distintos periféricos del sistema: Vídeo, impresora, modems, discos duros, etc.

Existen tres técnicas básicas de controlar la transferencia de datos hacia/desde un periférico:

- E/S programada.
- Interrupciones.
- DMA (Acceso directo a memoria).

E/S Programada.

Mediante esta técnica la CPU envía una orden al dispositivo (lectura, escritura, configuración, chequeo,...), y se mantiene ocupada hasta que concluye la operación. El dispositivo no interrumpe a la CPU para comunicar que ya ha terminado la operación, sino que es la CPU quien debe ocuparse de comprobarlo periódicamente. Como puede deducirse por lo expuesto, esta técnica consume mucho tiempo de CPU.

Interrupciones hardware.

Cada dispositivo pide servicio a través de una línea de interrupción. Desde el 8086 se dispone de una entrada de petición de interrupción (INTR - INTRerrupt Request) y una salida de concesión de interrupción (INTA - INTRerrupt Acknowledge). Ante la posibilidad de que varios dispositivos activen su interrupción simultáneamente, las interrupciones no las trata directamente la CPU, sino un controlador (Desde los primeros PC's, el circuito integrado 8259 o compatible que ahora se encuentra en el puente sur).

Además, existen las llamadas 'interrupciones software'. Se 'llaman' desde un programa para hacer que se ejecute una función del BIOS o del DOS que accede al hardware. Estas funciones son contempladas por la CPU como subrutinas, que una vez finalizadas, devuelven el control al programa que las llamó.

DMA (Acceso directo a memoria) La técnica DMA (Direct Memory Access) permite transferir datos entre un dispositivo y memoria o de una zona de memoria a otra sin la intervención del micro (excepto para arbitrar la transferencia). Esto es más rápido que el método tradicional donde cada dato llega a la CPU que lo graba en memoria.

- Requiere de un dispositivo que controle la operación en sustitución del micro.
- El bus PCI ya contempla entre sus características la posibilidad de acceder a memoria, por lo que no necesita de hardware adicional.

Las ventajas e inconvenientes son:

- Es útil cuando se deben transferir grandes bloques de datos (P. ej. lectura/escritura en disco).
- No es útil cuando la CPU tiene que procesar uno por uno los datos que se leen/escriben.

Posterior a DMA surgió **Ultra DMA** en 1996 que entre sus características primordiales se encontraba la de duplicar la tasa de transferencia a 33 MB/s.

3. Buses, procesadores y memorias

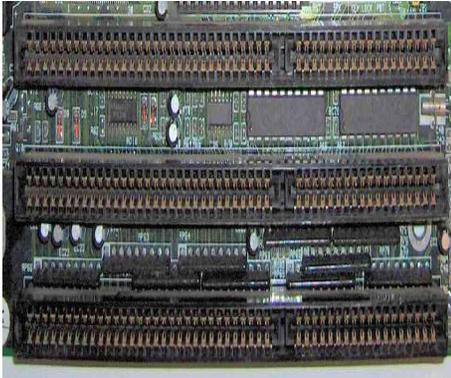


Figura 3.13: Ranura ISA



Figura 3.14: Ranura Vesa

ISA (Industry Standar Architecture)

Desarrollado por IBM en 1980 en versión de 8 bits y 4.77 MHz para el PC XT original, y más tarde en versión de 16 bits y 8.33 MHz para el PC AT en 1984, el bus ISA ha sido siempre el medio básico en la interconexión de periféricos.

Cuando el microprocesador trabajaba a frecuencias bajas, todo en el interior del PC funcionaba a su misma velocidad, incluido el tramo ISA del bus de expansión. Para cuando llegó el bus EISA, la frecuencia del procesador ya se había disparado, y así este bus quedó pronto relegado en los niveles inferiores de una jerarquía de buses progresivamente más rápidos según iban acercándose al microprocesador.

EISA (Extended ISA)

Este bus es, tal y como nos indica su nombre (Enhanced Industrial Standard Architecture), una extensión del primitivo bus ISA. Tal y como hacía el MCA, su bus de direcciones era de 32 bits. Mantuvo la compatibilidad con las tarjetas de expansión de su antecesor ISA, motivo por el cual tuvo que adoptar la velocidad de éste (8.33 Mhz).

Una gran cantidad de compañías contribuyeron a su desarrollo: AST, Compaq, Epson, Hewlett Packard, Olivetti, Tandy, Wyse, y Zenith.

3. Buses, procesadores y memorias

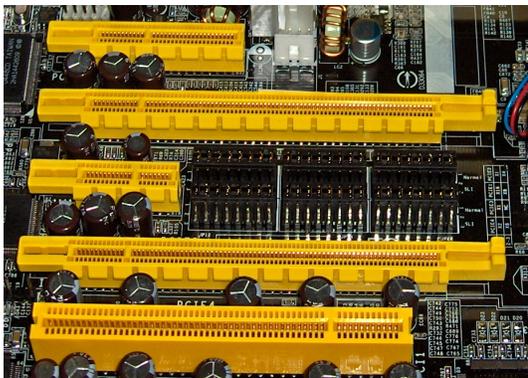


Figura 3.15: Ranuras PCIe 4x, 16x, 1x, 16x, PCI tradicional

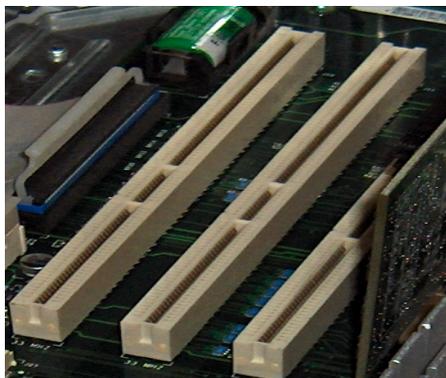


Figura 3.16: Ranura PCI

MCA (Micro Channel Architecture)

IBM, creador de ISA en 1980, se convirtió unos años más tarde en su principal enemigo. Con la llegada de los procesadores con bus de datos de 32 bits (80386 en adelante), en los que la versión ISA se quedaba algo pequeña, IBM aprovechó su peso específico en el mercado para diseñar un nuevo bus con derechos de propiedad para que todos aquellos fabricantes que se decidieran a adoptarlo en sus productos tuvieran que pagarle el correspondiente canon.

PCI (Peripheral Component Interconnect)

Especificado por un amplio grupo de fabricantes liderado por Intel, PCI estuvo en todo momento pensado para dar una respuesta eficiente de interconexión a los sistemas equipados con microprocesador Pentium.

La primera versión del bus PCI se especificó con una anchura de 32 líneas de datos y una frecuencia de funcionamiento de 33 MHz. Muy pronto apareció otra versión que ampliaba la anchura hasta las 64 líneas, y una frecuencia de funcionamiento en los 66 Mhz. Por lo que la versión de 32 bits podía suministrar un ancho de banda aproximado de $33 \text{ MHz} \times 32 \text{ bits} = 1056 \text{ Mbits/s} = 132 \text{ MBytes/s}$ y la versión de 64 bits suministraría $33 \text{ Mhz} \times 64 \text{ bits} = 2112 \text{ Mbits/s} = 264 \text{ MBytes/s}$.

Los zócalos PCI suelen venir en número de tres, cuatro o cinco en las placas base actuales. Pueden reconocerse por su color crema y longitud bastante inferior a la de los zócalos EISA (casi la mitad), se sitúan en paralelo con éstos y un poco

3. Buses, procesadores y memorias

desplazados con respecto a su punto de comienzo.

Una de las características más importantes del bus PCI es su compatibilidad con la especificación PnP (Plug & Play) de Intel, que permite autoconfigurar los dispositivos a él conectados sin necesidad de interactuar con jumpers y/o menús de la BIOS.

Ante la creciente necesidad de ancho de banda surgió una mejora de PCI denominada PCI-X que a menudo se confunde con PCIe. Las características que destacan de PCI-X son:

- PCI-X 66: 64 bits, 66 MHz.
- PCI-X 133: 64 bits, 133 MHz.
- PCI-X 266: Utiliza doble transferencia de datos (Double Data Rate), por lo que la frecuencia efectiva es de 266 MHz.
- PCI-X 533: Cuádruple transferencia de datos (Quad Data Rate), por lo que la frecuencia efectiva es de 533 MHz.

PCIe (PCI Express)

PCI-Express fue conocido al principio por las siglas 3GIO provenientes de 3rd Generation IO. PCIe se basa en un sistema de comunicación serie mucho más rápido que el antiguo PCI. PCIe ha sido apoyado por un conjunto de compañías de las que destaca Intel.

El bus PCIe está estructurado como enlaces punto a punto, full-duplex, trabajando en serie. En PCIe 1.1, 2007, cada enlace transporta 250 MB/s en cada dirección. PCIe 2.0 dobla esta tasa y PCIe 3.0 la dobla de nuevo. De este modo cada enlace de PCIe 2.0 transporta 500 MB/s y cada enlace de PCIe 3.0 transporta 1 GB/s.

Cada slot de expansión puede llevar uno, dos, cuatro, ocho, dieciséis o treinta y dos enlaces de datos entre la placa base y las tarjetas conectadas.

De este modo, un bus PCIe 1.1 con 16 enlaces puede transporta un ancho de banda de 4 GB/s (250 MB/s x 16) en cada dirección.

La velocidad superior del PCI-Express permitirá reemplazar casi todos los demás buses, AGP y PCI incluidos. La idea de Intel es tener un solo controlador PCI-Express comunicándose con todos los dispositivos, en vez de con el actual sistema de puente norte y puente sur.

3. Buses, procesadores y memorias

Este conector es usado mayormente para conectar tarjetas gráficas.

PCI-Express no es todavía suficientemente rápido para ser usado como bus de memoria.

PCI-Express en 2006 es percibido como un estándar de las placas base para PC, especialmente en tarjetas gráficas. Marcas como Ati Technologies y nVIDIA entre otras tienen tarjetas gráficas en PCI-Express.

Las principales características de transmisión de PCIe son:

- **Transmisión diferencial.** Al igual que otros buses del ordenador que han evolucionado a la transmisión serie (USB, Serial ATA) utiliza la técnica LVDS (Low Voltage Differential Signaling)
- **Transmisión síncrona.** La señal de reloj está mezclada con la propia información. Para ello, usa una codificación 8b/10b que transmite 10 bits por cada 8 de información, por lo que se genera una sobrecarga del 20 %.
- **Transmisión bidireccional.** Una conexión está compuesta de 2 canales, uno de ida y otro de vuelta que transmiten simultáneamente (dos canales simplex).
- **Conexión múltiple.** PCIe puede utilizar varias conexiones para la misma comunicación dando lugar a configuraciones llamadas x1, x2, x4, x8, x12, x16 y x32. Las conexiones x16 y x32 están pensadas para conectar dispositivos como tarjetas gráficas.
- **Transmisión isócrona.** Es posible reservar y garantizar un ancho de banda bajo demanda consiguiendo una transmisión en tiempo real. A esto se le conoce como transferencia isócrona, pues se puede garantizar el tiempo que durará una transmisión de datos (i.e. tiempo real).
- **Distancia.** La inclusión de la señal de reloj permite mayores distancias respecto a los buses paralelo cuya señal discurre por una línea separada. En el diseño de este bus se ha marcado como objetivo permitir hasta 50 cm de distancia entre dispositivos con tecnología de placa de circuito impreso de 4 capas y conectores estándar. Se podría aumentar la distancia usando componentes de mayor calidad.

SLI.

SLI significa **Scalable Link Interface** y es un sistema que permite conectar

3. Buses, procesadores y memorias

dos o más tarjetas gráficas simultáneamente para que produzcan una sola señal sumando la potencia de ambas.

Esta idea no es nueva, sino que proviene de 1998 por 3DFX y sus tarjetas gráficas Voodoo2, que al sumar la capacidad de procesamiento de ambas tarjetas genera un incremento en la capacidad de procesamiento igual a la suma de ambas tarjetas (teóricas).

Hay que recordar que estas tarjetas gráficas iban conectadas a slots PCI, ya que en esa época el puerto AGP estaba en sus comienzos y además solo permitía un puerto en placa base.

En el año 2004 la compañía NVidia relanzó esta idea, pero esta vez aprovechando las prestaciones que ofrecen las tarjetas gráficas montadas sobre puertos PCIe de 16x y 32x.

El sistema SLI solamente funciona en placas bases con dos o más puertos PCIe 16x para gráficas, y por supuesto que el chipset permita la tecnología SLI.

Las tarjetas se conectan mediante un pequeño conector de circuito impreso a un conector que tiene estas tarjetas. Este conector sirve de enlace para transmitir datos de sincronización, visualización y píxeles entre las dos GPUs. Proporcionando comunicación entre los dos procesadores gráficos.

Para que SLI tenga un buen funcionamiento las dos tarjetas gráficas deben ser exactamente iguales, marca, modelo y capacidad.

El incremento que se obtendrá será el de aumentar las capacidades entre un 75 % y un 100 % (rara vez se alcanzará el 100 % teórico).

Otra aplicación de la tecnología SLI es la de visualización en varios monitores. Es decir, si se configura en modo multi GPU, sólo se puede utilizar un monitor, pero en modo de una sola GPU es posible emplear un total de 4 monitores (dos por tarjeta) de forma simultánea.

A continuación se mostrará el montaje de dos tarjetas SLI mediante fotografías, estas fotografías han sido extraídas de la Web oficial de NVidia. 'http://www.slizone.com/object/slizone_howto_install.html'.

CrossFire.

CrossFire es el nombre con el que han bautizado ATI/AMD al sistema de doble GPU, diseñado para dar respuesta al sistema SLI.

Aunque en principio el objetivo de CrossFire es similar al de SLI el método para lograrlo ha sido diferente al de SLI. Sobre todo tienen en común que solo fun-

3. Buses, procesadores y memorias



Figura 3.17: Instalación de dos tarjetas PCIexpress para un sistema SLI



Figura 3.18: Dos tarjetas gráficas conectadas (Master y Slave) montadas en un sistema Crossfire

cionan en gráficas PCIe, en el caso de CrossFire en placas base y tarjetas gráficas certificadas CrossFire. En principio esta tecnología lo único que posibilita es que ambas tarjetas compartan la carga de trabajo de la renderización de las imágenes.

Para la renderización ATI/AMD se utiliza tres sistemas diferentes:

- **AFR o Alternate Frame Rendering.** Es el método que proporciona un mayor incremento en el rendimiento, consiste en que cada tarjeta gráfica renderiza fotogramas alternos.
- **Scicorring.** Divide cada frame en dos partes. Estas partes no tienen por qué ser iguales ya que la extensión de imagen que renderiza cada tarjeta se asigna dinámicamente.
- **STB o Super Tile Board** Divide la imagen en pequeñas porciones de 32x32 píxeles creando una especie de malla o tablero. En este caso cada tarjeta renderiza pequeños cuadrados alternos de la imagen dividida, superponiendo después las imágenes generadas por cada tarjeta, creando así la imagen que se mostrará por pantalla.

En cuanto al sistema de conexión de las tarjetas también hay diferencia entre ambos sistemas.

En el sistema CrossFire se utilizan hasta tres sistemas diferentes para realizar esta conexión:

- Las tarjetas de gama baja utilizan el propio bus PCIe para transmitir los datos visuales entre las dos GPUs. Este sistema fue desechado para tarjetas

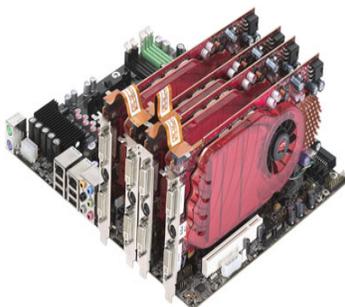


Figura 3.19: Cuatro tarjetas gráficas montadas por puentes en un sistema Crossfire

de gama superior debido al excesivo consumo de ancho de banda de PCIe para resoluciones muy grandes.

- La forma más utilizada para montar CrossFire es utilizar una tarjeta CrossFire Master y otra CrossFire Slave. La primera sustituye una de sus conexiones DVI por una conexión especial que mediante un cable externo nos permite enlazar ambas tarjetas gráficas entre sí y a la vez con el monitor. Véase la figura 3.18.
- El más novedoso consiste en usar un puente similar al SLI, pero en este caso con dos conectores de circuito impreso en vez de uno solo. Véase la figura 3.19.

3.1.6 Buses Dedicados

AGP (Accelerated Graphics Port)

Es la especificación de un bus desarrollado por Intel como solución a los enormes requerimientos de ancho de banda que necesitan los gráficos tridimensionales, el objetivo prioritario de esta especificación es el de proporcionar altas velocidades de transmisión de datos entre el procesador y la memoria de vídeo. Este bus permite que las texturas de los gráficos tridimensionales se puedan almacenar, no sólo en la memoria de vídeo sino también en la memoria principal.

3. Buses, procesadores y memorias

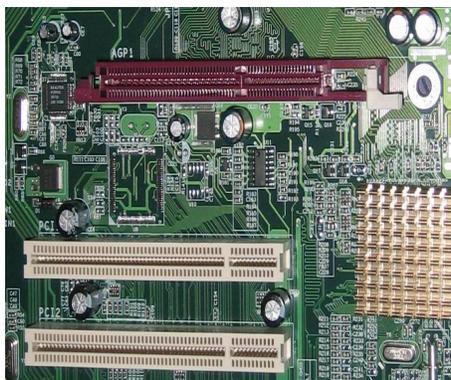


Figura 3.20: Ranura AGP

AGP utiliza un conector similar al PCI con 32 líneas multiplexadas para direcciones y datos y otras 32 líneas exclusivas para datos con objeto de completar los 64 bits del procesador.

Como es un bus especialmente dedicado a los gráficos, no tiene que compartir con otros dispositivos el ancho de banda. Donde realmente está la fuerza del bus AGP es en la especificación de su protocolo, que permite, entre otras cosas, transferir datos tanto en el flanco de subida como en el flanco de bajada del ciclo de reloj del bus. Dota de una mayor estabilidad al sistema al eliminar del bus PCI el intenso tráfico generado por las aplicaciones gráficas.

La primera versión de AGP fue la 1X, con un ancho de banda de 254 Mb/s. Sin embargo, rápidamente se comercializó el AGP 2X, con 508 Mb/s. Le siguieron las versiones 4X y 8X. El aumento en la velocidad se consigue haciendo varias transferencias por cada ciclo de reloj. Así se obtiene un ancho de banda de hasta 2032 Mb/s.

IDE (Integrated Drive Electronics)

IDE es la especificación de un bus para la transferencia de información con los dispositivos de almacenamiento masivo (discos flexibles, discos duros, CD-ROM...). IDE significa literalmente que el controlador del disco no está instalado en el PC, sino en el propio dispositivo, lo que se implementa en la placa base es un pequeño circuito integrado que contiene la funcionalidad para pasar la información del bus IDE a los buses del sistema.

Gracias a que este controlador se integra directamente en el disco, su elec-

3. Buses, procesadores y memorias

trónica puede optimizarse para las características específicas de cada dispositivo.

IDE es en realidad un cambio de nombre del interfaz ATA (AT Attachment: Conexión AT) utilizado en los primeros PC.

Se produjo un importante incremento en la velocidad de transferencia cuando el bus IDE incorporó la tecnología DMA (Direct Memory Access o Acceso Directo a Memoria) que permitía las transferencias directas a memoria sin intervención del procesador. Esto dió lugar al EIDE (Enhanced IDE o IDE Mejorado).

La aparición de la especificación Ultra-DMA duplicó las velocidades de transferencia con los discos. El Ultra-DMA introduce un límite de dos dispositivos por canal. El primer canal se denomina IDE primario, mientras que el segundo recibe el nombre de IDE secundario. El disco duro suele conectarse como maestro en el IDE primario, aunque existen múltiples variantes de conexión dependiendo de la configuración y las necesidades de cada usuario.

SCSI (Small Computer System Interface)

El interfaz SCSI, pronunciado ‘escasi’, surgió con el objetivo de desarrollar un interfaz común para los controladores de dispositivos de almacenamiento, de forma que éstos pudieran soportar un direccionamiento lógico para los bloques de datos, o sea, sin tener en cuenta cómo esté construido físicamente (los cilindros, cabezas, sectores...).

El interfaz SCSI-2 unificó el protocolo para todo tipo de discos, cintas y CDROM. Se pueden conectar un máximo de 7, 15 y 31 dispositivos al bus para las versiones de 8, 16 y 32 bits de anchura de datos, respectivamente.

La versión SCSI-3 permiten anchos de banda de 80 Mbytes/seg. para las transferencias, pero lo que marca las diferencias con el IDE es la inteligencia de que está dotado el protocolo SCSI: en lugar de enviar los datos al microprocesador, el controlador del bus SCSI y la unidad SCSI de cada dispositivo dialogan internamente a través de un conjunto de comandos propios

SATA (Serial ATA)

Serial ATA o S-ATA (acrónimo de Serial Advanced Technology Attachment) es una interfaz de transferencia de datos entre la placa base y algunos dispositivos de almacenamiento, como puede ser el disco duro, u otros dispositivos de altas prestaciones que están siendo todavía desarrollados. Serial ATA sustituye a la tradicional Parallel ATA o P-ATA (estándar que también se conoce como IDE o

3. Buses, procesadores y memorias

ATA). El S-ATA proporciona mayores velocidades, mejor aprovechamiento cuando hay varios discos, mayor longitud del cable de transmisión de datos y capacidad para conectar discos en caliente (con la computadora encendida).

Actualmente es una interfaz extensamente aceptada y estandarizada en las placas base de PC.

La primera generación especifica en velocidades de 1.5 Gbit por segundo, también conocida por SATA 1.5 Gb/s o Serial ATA-150. Actualmente se comercializan dispositivos SATA II, a 3 Gb/s, también conocida como Serial ATA-300. Se está desarrollando SATA 6 Gbit/s que incluye una velocidad de 6.0 Gbit/s estándar, pero que no entrará en el mercado hasta 2009.

Los discos que soportan la velocidad de 3Gb/s son compatibles con un bus de 1,5 Gb/s.

Los conectores y los cables son la diferencia más visible entre las unidades SATA y las PATA. Al contrario que los PATA se usa el mismo conector en los discos duros de equipos de escritorio o servidores (3,5 pulgadas) y los de los portátiles (2,5 pulgadas). Esto permite usar los discos duros de 2,5 pulgadas en los sistemas de escritorio sin necesidad de usar adaptadores a la vez que disminuyen los costes.

Las principales ventajas de SATA son:

- Velocidades de transferencias de datos más rápidas.
- Más ancho de banda.
- Más potencial para los aumentos de velocidad en generaciones futuras.
- Mejor integridad de los datos gracias al nuevo set de comandos avanzado.
- Cables más compactos que facilitan la ventilación interna de los ordenadores.
- Longitud máxima del cable de hasta 2 metros.
- Diseño de conector que permite HotPlug.
- Reducción de pineado que permite la escalabilidad RAID.
- Compatibilidad software y drivers existentes de Parallel ATA.

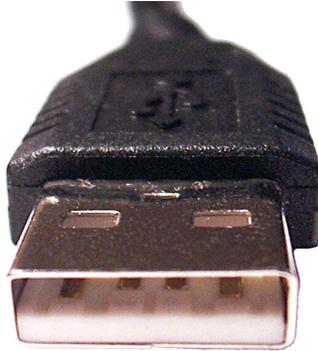


Figura 3.21: Conector USB

USB (Universal Serial Bus)

El bus USB ha surgido como una iniciativa para la interconexión estándar de periféricos. Dada la heterogeneidad de los dispositivos actuales, la versatilidad de una conexión es un aspecto muy valorado por el usuario. Y debido a que cada vez existen más usuarios neófitos, la comodidad de uso es otro de los valores en alza. USB combina muy acertadamente estas dos características en una especificación cuyos rasgos más sobresalientes se resumen en:

- **Unificación de conectores.** A la hora de ampliar nuestro equipo, con USB no es necesario liarse con el tipo de puerto a utilizar. Todos los periféricos pueden usar el mismo conector y el interfaz permite la conexión de hasta 127 periféricos utilizando un adaptador uno a muchos similar a los ladrones que se emplean para multiplicar la conexión a los enchufes eléctricos de las casas.
- **Facilidad de montaje.** A la hora de adquirir nuevos periféricos, como un monitor, impresora, escáner, módem o cámara digital, no se necesitan tarjetas adicionales, ni abrir la caja. La mayoría de los dispositivos más novedosos se venden ya con el conector USB incorporado.
- **Configuración automática.** El software MIS/IT se encarga de la gestión remota y el mantenimiento local de dispositivos de forma automática (PnP, Plug & Play). Esto evita que el usuario tenga que pelearse con jumpers, conmutadores DIP, configuradores IRQ, canales de DMA y direcciones de entrada/salida.

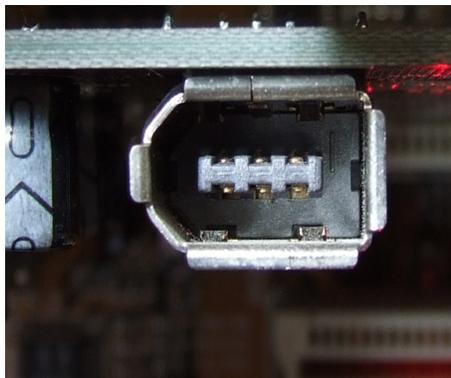


Figura 3.22: Conector Firewire

- **Conexión en caliente.** Sin duda, una de sus principales ventajas. Antes, cada vez que conectábamos un nuevo dispositivo era necesario reiniciar el equipo, configurar el hardware e instalar el driver correspondiente. Ahora, el software es capaz de detectar al instante que el usuario acaba de enchufar un nuevo dispositivo al bus e, inmediatamente, procede a la instalación del driver correspondiente.
- **Altas prestaciones.** El USB 1.1 puede transferir datos a una velocidad de 1'2 MB/s y el USB 2.0 puede alcanzar los 60 MB/s y USB 3.0 alcanza hasta 600 MB/s.

IEEE 1394 o FIREWIRE

Fire Wire (cable de fuego) es un nuevo bus que surge para aumentar la velocidad de las transferencias de datos por el puerto serie. FireWire soporta velocidades de más de 40 MB/s con ciertos periféricos y se ha establecido como el bus estándar para las transferencias de datos multimedia con todo tipo de dispositivos digitales, especialmente con cámaras de vídeo.

El conector FireWire es alargado y plano, similar en aspecto al USB. Dos de ellos se utilizan para la transmisión de datos y el tercero porta el suministro de corriente para el periférico. Las transferencias de datos son semi-dúplex, esto es, unidireccionales, y cada periférico puede elegir la velocidad a la que enviar sus datos. El bus permite la conexión de hasta 16 dispositivos, que pueden también conectarse/desconectarse sin necesidad de apagar el equipo.



Figura 3.23: Conector DVI

DVI (Digital Visual Interface)

DVI proviene de Digital Visual Interface o Interfaz Visual Digital, es una interfaz de vídeo diseñada para obtener la máxima calidad de visualización posible en pantallas digitales, tales como monitores de cristal líquido de pantalla plana o proyectores digitales.

El estándar anterior, VGA, fue diseñado para dispositivos CRT (tubos de rayos catódicos). Las características físicas de VGA se amoldaban a la física de los dispositivos CRT mientras que DVI se amolda a la física de las pantallas planas.

HDMI (High-Definition Multi-media Interface)

High-Definition Multi-media Interface (HDMI) -Interfaz multimedia de alta definición- es una norma de audio y vídeo digital cifrado sin compresión apoyada por la industria para que sea el sustituto DRM del euroconector. HDMI provee un interfaz entre cualquier fuente DRM de audio y vídeo digital como podría ser un sintonizador TDT, un reproductor de Blu-ray, un ordenador (con Windows, Linux, etc.).

El conector estándar de HDMI tipo A tiene 19 pines. Se ha definido también una versión de mayor resolución -tipo B-, pero su uso aún no se ha generalizado. El tipo B tiene 29 pines, permitiendo llevar un canal de vídeo expandido para pantallas de alta resolución. El tipo B fue diseñado para resoluciones más altas que las del formato 1080p.

El HDMI tipo A es compatible hacia atrás con un enlace simple DVI, usado por los monitores de ordenador y tarjetas gráficas modernas. Esto quiere decir que



Figura 3.24: Conector HDMI

una fuente DVI puede conectarse a un monitor HDMI, o viceversa, por medio de un adaptador o cable adecuado, pero el audio y las características de control remoto HDMI no estarán disponibles.

Ejercicio 3.5 *Calcula el ancho de banda (en megabytes por segundo) de un bus que dispone de 32 líneas y funciona a 1066Mhz.*

Ejercicio 3.6 *¿Qué componentes se interconectan mediante el bus local (FSB)?*

Ejercicio 3.7 *Contesta V (verdadero) o F (falso). Si la respuesta es F, explica por qué.*

- 1. La placa base puede llevar varias ranuras de expansión de diferentes tipos pero sólo contiene un zócalo para el procesador.*
- 2. La señal del reloj que marca el ritmo de funcionamiento de todos los elementos que se encuentran en la placa base tiene forma senoidal.*
- 3. El puerto AGP está controlado por el puente norte.*
- 4. Un jumper, como su nombre indica, sirve para que la información salte de un dispositivo a otro.*
- 5. En la misma placa base puede haber tanto ranuras PCI como ranuras ISA.*
- 6. El ancho de banda de un bus mide el grosor en micras de la línea por donde circula la información.*

3. Buses, procesadores y memorias

7. *En los primeros ordenadores, todos los componentes funcionaban a la misma velocidad que la CPU.*
8. *El bus PCI es más rápido que el bus ISA.*
9. *El bus ISA incorporó la tecnología Plug and Play para configurar de forma automática los dispositivos.*
10. *La ranura del bus PCI puede identificarse por su color negro mate.*
11. *El puerto AGP (Advanced Graphics Port) está especialmente diseñado para conectar impresoras.*
12. *El DMA permite la transferencia de información de un dispositivo de almacenamiento a la memoria principal directamente, sin la intervención del procesador.*
13. *IEEE 1394 es sinónimo de USB.*
14. *La conexión en caliente se refiere a que los dispositivos que se conectan de esta forma necesitan refrigeración porque trabajan a altas frecuencias.*

Ejercicio 3.8 *Calcula el ancho de banda de las dos versiones del bus ISA, la de 8 bits y la de 16 bits; teniendo en cuenta que el de 8 bits va a 4'77 MHz y el de 16 bits funciona a 8'33 Mhz.*

SECCION 3.2

Procesador

El microprocesador es el cerebro del ordenador, el que interpreta y ejecuta todas las instrucciones que forman parte de las aplicaciones en funcionamiento. Está constituido por un conjunto de unidades funcionales que forman la Unidad Central de Proceso (CPU) cuyos módulos más importantes son:

- **Unidad de control.** Encargada de gestionar las instrucciones que se ejecutan en el microprocesador.
- **Unidad de proceso.** Encargada de ejecutar dichas instrucciones.

3.2.1 Magnitudes principales de un microprocesador

A continuación se estudiarán los parámetros fundamentales que influyen en el rendimiento de un microprocesador. Son los siguientes:

- Frecuencia de reloj.
- Tecnología de integración.
- Paralelismo a nivel de instrucción.
- Memorias caché.
- Conjunto de instrucciones.

Frecuencia de reloj

Indica la velocidad a la que trabaja el microprocesador. Aunque se ha extendido la idea que especifica el número de instrucciones que puede ejecutar por segundo, no es del todo cierto. Veamos realmente lo que quiere decir este parámetro.

El reloj es un pequeño cristal de cuarzo, ubicado en la placa base, que proporciona una señal de onda digital, binaria, cuadrada, síncrona y periódica, tal y como se representa en la figura 3.25.

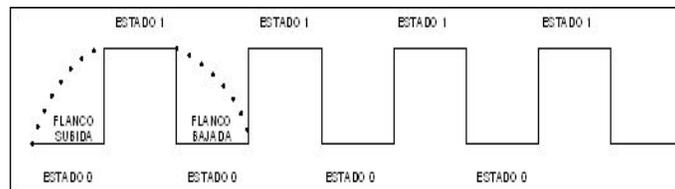


Figura 3.25: Señal periódica de un reloj

El período o ciclo de reloj es la duración de un pulso u oscilación del reloj. Por el contrario, la frecuencia expresa el número de ciclos producidos por el reloj en una unidad de tiempo, por lo que puede expresarse como la inversa del período. Las unidades utilizadas para medir la frecuencia son los hertzios (Hz) y determinan el número de ciclos por segundo. Esta unidad puede resultar pequeña, por lo que disponemos de los siguientes múltiplos:

3. Buses, procesadores y memorias

Megahertzios (MHz): $1 \text{ MHz} = 1.000.000 \text{ Hz}$.

Gigahertzios (GHz): $1 \text{ Ghz} = 1.000 \text{ MHz} = 1.000.000.000 \text{ Hz}$.

Debido a que nuestro sistema informático incorpora dispositivos muy dispares, que operan a diferentes velocidades, existen multiplicadores y divisores de frecuencia, generando, de esta manera, señales con frecuencias distintas, destinadas a sincronizar las operaciones tanto en los dispositivos más lentos como en los más rápidos.

El período es la unidad básica de tiempo. Cualquier operación debe consumir un número entero de períodos para que puedan sincronizarse. Normalmente, cada instrucción que se ejecuta en el microprocesador utiliza varios ciclos de reloj, ya que se descomponen en operaciones atómicas que consumen un período de reloj cada una de ellas. Por consiguiente, cuanto mayor sea la frecuencia del microprocesador, más rápido ejecutará las instrucciones y menor tiempo invertirá en finalizarlas.

Tecnología de integración

El microprocesador es un conjunto de componentes que procesan una señal eléctrica que puede tomar dos estados: ausencia o presencia de corriente (0 y 1 = lenguaje binario). De esta forma, todo programa es traducido a este lenguaje para que pueda ser procesado.

El microprocesador está compuesto por millones de transistores que son minúsculos interruptores que permiten o niegan el paso de corriente. Combinando estos elementos se pueden construir estructuras más complejas para construir puertas lógicas (AND, OR, NOT, XOR...) y circuitos integrados.

Conforme avanza la tecnología se tiende a disminuir la distancia de integración, proporcionando los siguientes beneficios:

- Los transistores construidos son más pequeños, por lo que podemos integrar mayor número en el mismo espacio. Esto conlleva un aumento de las prestaciones del microprocesador, como la incorporación de nuevas unidades funcionales o el aumento de las memorias caché.
- Aumenta la velocidad de operación del transistor, ya que al ser menor la distancia que tiene que recorrer el flujo eléctrico también es menor el tiempo que tarda en recorrerla. Esto conlleva un aumento de la frecuencia debido a que las operaciones se efectúan más rápidamente.

3. Buses, procesadores y memorias

- Disminuye el voltaje que necesita el microprocesador para ser alimentado. De esta forma se consigue una disminución del calor disipado pudiendo aumentar la frecuencia.
- Disminuye el coste de fabricación, ya que la cantidad de silicio empleada es menor.

La distancia de integración ha evolucionado desde 1 micra de los 486, hasta las 0,45 nanómetros de los actuales procesadores.

Paralelismo a nivel de instrucción

Mediante esta característica se intenta eliminar la ejecución secuencial de instrucciones, posibilitando la ejecución simultánea. Para ello se aprovecha la existencia de las distintas unidades funcionales del microprocesador, ya que una misma instrucción no puede trabajar simultáneamente sobre varias unidades funcionales. Por ello, cuando una instrucción está haciendo uso de una unidad funcional, otra instrucción puede trabajar simultáneamente en otra unidad funcional. Veamos las distintas formas en que puede ser tratado el paralelismo: Segmentación, superescalaridad y supersegmentación.

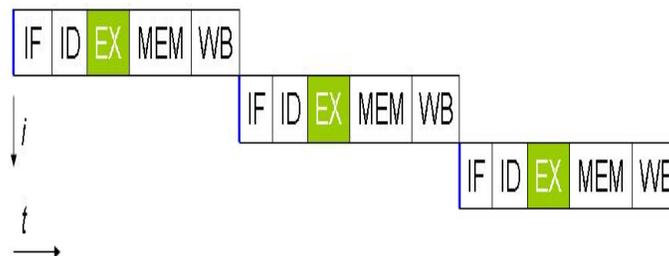


Figura 3.26: Procesador sin segmentar

- **Segmentación o pipelining.** Con esta técnica cada instrucción que entra a ejecutarse en el microprocesador se divide en una serie de etapas. Utilicemos, por ejemplo, el siguiente modelo de cinco etapas:
 - Etapa 1: Búsqueda de la instrucción (E1).
 - Etapa 2: Decodificación (E2).
 - Etapa 3: Lectura de operandos (E3).

3. Buses, procesadores y memorias

- Etapa 4: Ejecución (E4).
- Etapa 5: Escritura del resultado (E5).

De esta manera, mientras una instrucción está ejecutando su etapa 1, otra puede ejecutar su etapa2, una tercera puede ejecutar su etapa 3 y así sucesivamente. Destaquemos que cada instrucción se ejecutará secuencialmente de principio a fin, pero solapándose con otras instrucciones que también se ejecutan secuencialmente en el resto de unidades funcionales. Gráficamente quedaría de la siguiente forma:

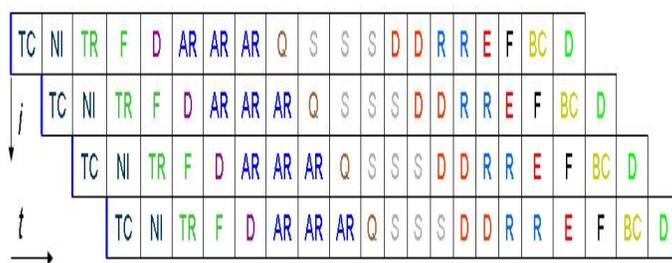


Figura 3.27: Procesador segmentado

El número de etapas en la segmentación varía según el modelo de microprocesador. En la actualidad pueden llegar a las 20 etapas.

- **Superescalaridad.** Las actuales técnicas de integración permiten incluir más componentes en el área del microprocesador. Es por ello que también se han efectuado réplicas de unidades funcionales para que varias instrucciones puedan ejecutar simultáneamente en una misma etapa.

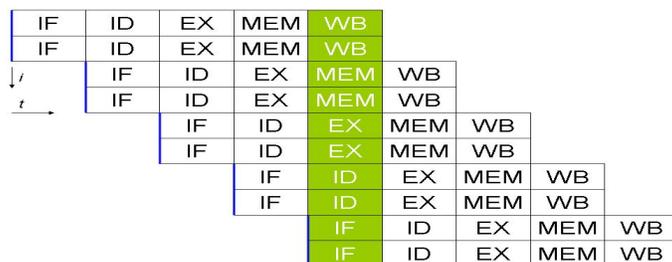


Figura 3.28: Procesador Superescalar

El inconveniente del paralelismo: **las dependencias.**

3. Buses, procesadores y memorias

La posibilidad de ejecutar instrucciones de un mismo programa de forma simultánea, choca con la propia estructura secuencial que poseen los comandos de las aplicaciones. Los riesgos asociados a las técnicas de paralelismo:

- **Dependencias de datos.** Si tenemos dos instrucciones en que la segunda necesita el resultado de la primera, la segunda no podrá comenzar su ejecución hasta que no haya finalizado la primera.
- **Dependencias de control.** Una instrucción de salto condicional interrumpe la ejecución concurrente de instrucciones, ya que hasta que no sea evaluada la condición no conoceremos la siguiente instrucción a ejecutar.
- **Dependencia funcional.** Se están empleando unidades funcionales (ALU, Memoria, . . .) y no pueden ser empleadas de nuevo.

Las soluciones que se plantean para esta situación son:

- **Ejecución fuera de orden.** Cuando una instrucción no puede ejecutarse por dependencia de datos el microprocesador continúa con la siguiente.
 - **Predicción de salto.** Una instrucción de salto condicional produce dos caminos por los que el programa puede continuar su ejecución. Para eliminar las dependencias de control el microprocesador elige uno de estos caminos para continuar la ejecución. Si acierta habremos avanzado en la ejecución de la aplicación, si falla simplemente tendremos que comenzar la ejecución por el otro camino.
- **Supersegmentación.** La combinación de las dos técnicas anteriores producen microprocesadores supersegmentados.

Memoria interna.

El microprocesador necesita tener un rápido acceso a los datos para poder ejecutar las instrucciones, ya que en el interior del microprocesador se producen mayor número de peticiones de datos que operaciones en sí. Por otro lado, el acceso a la memoria principal es muy lento, podemos decir que dicho acceso es unas 3.000 veces más lento que el intercambio de datos en el interior del microprocesador. Es por ello que es necesario que exista una jerarquía o niveles de memoria que acerquen los datos al microprocesador conforme vayan siendo necesarios. Los

3. Buses, procesadores y memorias

diferentes niveles comenzando por la capa más interna, la más próxima al núcleo de ejecución:

- **Banco de registros.** Dependiendo del microprocesador se incluyen entre 32 y 512 registros en los que, en cada uno de ellos, es posible almacenar un dato que son resultados intermedios de las operaciones que se están ejecutando.
- **Buffer de prebúsqueda de instrucciones.** Almacena las instrucciones candidatas a ser ejecutadas aproximándolas al máximo al núcleo de ejecución.
- **Caché de primer nivel L1.** Posee dos módulos independientes, uno para datos y otro para instrucciones. Esto es debido a la segmentación de los microprocesadores y a la división de las instrucciones en etapas. Si en un instante un proceso se encuentra en la etapa 1 (búsqueda de instrucción) y otro se encuentra en la etapa 3 (lectura de operandos) es necesario que ambas instrucciones puedan acceder de forma simultánea a la caché, por lo que debe implementarse en módulos independientes.
- **Caché de segundo nivel L2.** Conforme evolucionan los microprocesadores la caché L1 es insuficiente para suministrar instrucciones y operandos al núcleo de ejecución, por lo que se incorpora una segunda caché, más grande pero más lenta, que suministra datos e instrucciones a la caché L1 conforme van siendo necesarios.
- **Caché de tercer nivel L3.** La aparición de un nuevo nivel de caché está ligado al distanciamiento que se puede producir entre la memoria principal y las memorias caché. Si las caché de nivel L1 y L2 se encuentran muy próximas al microprocesador en cuanto a velocidad se produce un salto muy grande entre la memoria principal y las memorias caché, hueco que debe ser ocupado por un nuevo nivel, la caché de tercer nivel L3. En la actualidad prácticamente han desaparecido, ya que el aumento de la velocidad de la memoria principal ha supuesto un acercamiento de los datos al microprocesador siendo innecesaria la implementación de este tipo de caché.

La Ubicación de la caché respecto al microprocesador puede tener las siguientes ubicaciones:

- **Externa.** Se encuentra en la placa base, por lo que no forma parte del microprocesador.

3. Buses, procesadores y memorias

- **Interna.** Forma parte del microprocesador pero se encuentra en un chip independiente.
- **Integrada.** Incluida en el microprocesador e integrada en el propio chip de la CPU, trabaja a la misma velocidad que el microprocesador. En la actualidad todas las memorias caché son integradas.

Conjunto de instrucciones: CISC y RISC

Son dos vertientes totalmente opuestas para configurar y definir el conjunto de instrucciones, aunque veremos que en la actualidad existe una tendencia de incluir ambas técnicas en los microprocesadores, para aprovechar lo mejor de cada una de ellas.

- **CISC (Complex Instruction Set Computer).** El conjunto de instrucciones es complejo, incorporando multitud de instrucciones potentes y especializadas en tareas concretas (gráficos, sonido, Internet...). Son microprocesadores con circuitería compleja.
- **RISC (Reduced Instruction Set Computer).** El conjunto de instrucciones es reducido. Las instrucciones, al ser más simples se ejecutan más rápido, aunque una instrucción compleja deberá desglosarse en varias instrucciones simples. Son microprocesadores con circuitería sencilla.

Los primeros ordenadores nacieron con un conjunto de instrucciones simples que fue aumentando poco a poco, acorde con los avances tecnológicos de la época, acercándonos hacia una tecnología CISC. Pero a finales de los 80 el proceso se invierte, acercándonos a microprocesadores con tecnología RISC. En la actualidad vuelve la tendencia de microprocesadores CISC, debido al auge de entornos gráficos, aplicaciones multimedia e Internet, pero en realidad se fusionan ambas tecnologías para aprovechar al máximo los beneficios de cada una de ellas.

Los microprocesadores están evolucionando hacia una tecnología CISC en la que el conjunto de instrucciones se ve incrementado por instrucciones específicamente multimedia. Esto es debido a la demanda del mercado, ya que el usuario ejecuta en su máquina aplicaciones que incluyen vídeo, gráficos tridimensionales, animación, sonido... Los fabricantes de microprocesadores se adaptan a la evolución del software y a la propia tecnología que brinda al usuario máquinas con más posibilidades.

3.2.2 Estructura del microprocesador

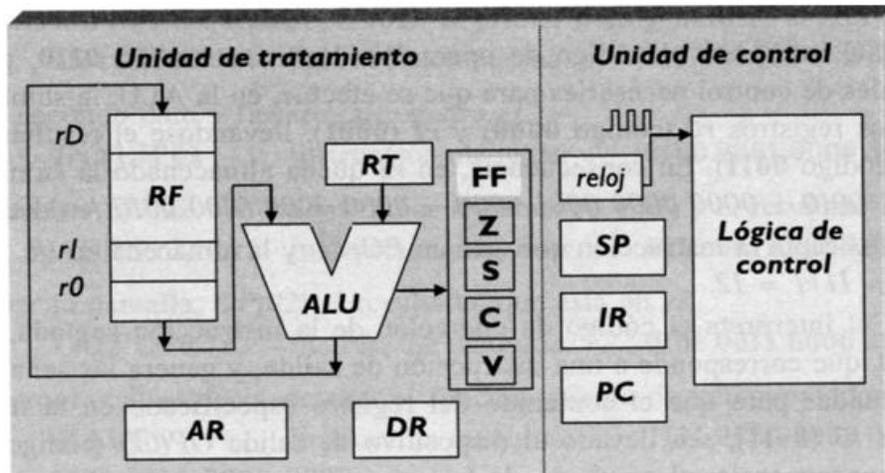


Figura 3.29: Esquema de un microprocesador

A continuación se estudiará cada una de las partes en las que se divide el microprocesador, estudiando cada uno de sus elementos: **Unidad de Control**, **Unidad de Proceso y Buses**.

Unidad de control

Su función principal es sincronizar el funcionamiento combinado de todos los elementos que intervienen en la ejecución de una instrucción. Consta de las siguientes partes:

- **Contador de programa.** Contiene la dirección de memoria de la siguiente instrucción a ejecutar. Normalmente estará ubicada en la siguiente posición de memoria, ya que es un proceso secuencial, pero no olvidemos que los programas pueden contener saltos, bifurcaciones y bucles.
- **Registro de instrucciones.** Contiene la instrucción que se está ejecutando en cada momento. Su formato es el siguiente:

- **Código de operación.** Indica qué tipo de operación se va a realizar, entre el conjunto de instrucciones que posee el microprocesador en su repertorio.
- **Operandos.** Los datos sobre los que actúa la instrucción, que pueden estar ubicados en los registros del microprocesador o en la memoria. Debido a que cada instrucción necesita de un conjunto distinto de operandos, el formato para cada una de ellas es diferente. Este tipo de codificación se denomina formato variable y la unidad de control interpretará los operandos en función del código de operación. Otra alternativa es el formato fijo, que consiste en que todas las instrucciones poseen el mismo formato, aunque cada instrucción sólo utilizará un subconjunto de los operandos disponibles. Este formato facilita la decodificación pero incrementa el tamaño de las instrucciones en memoria.
- **Decodificador.** Se encarga de extraer el código de operación de la instrucción en curso. Cada tipo de microprocesador posee un conjunto de instrucciones diferente, por lo que el decodificador debe averiguar la correspondencia entre la instrucción del programa y la que se incluye en el repertorio de instrucciones del microprocesador. En el caso de un microprocesador CISC (microprogramado) esta correspondencia suele existir. Por el contrario, en un microprocesador RISC (cableado), el decodificador tiene más trabajo, ya que debe descomponer la instrucción del programa en un conjunto de instrucciones (incluidas en el repertorio) que al ser ejecutadas constituyan la instrucción del programa.
- **Reloj.** La señal de reloj, producida en la placa base y que normalmente es multiplicada para el microprocesador, proporciona una sucesión de impulsos eléctricos o ciclos a intervalos constantes, que marcan los instantes en que han de comenzar los distintos pasos en los que se divide cada instrucción.
- **Secuenciador.** Este elemento pertenece a microprocesadores microprogramados (CISC). Recibe del decodificador el código de operación que es utilizado para leer de la memoria microprogramada el conjunto de microinstrucciones que corresponden con la instrucción a ejecutar. El secuenciador las va ejecutando sincronizadas por los impulsos de reloj.

Unidad de proceso

Formada por la Unidad Aritmético Lógica (ALU), que se encarga de las operaciones elementales de tipo aritmético (sumas, restas, productos, divisiones) y de tipo lógico (AND, OR, NOT...), y por la Unidad de Punto Flotante (FPU) especializada en realizar operaciones matemáticas con números reales.

La Unidad Aritmético Lógica está formada por los siguientes elementos:

- **Circuito operacional.** Contiene los circuitos necesarios para la realización de las operaciones con los datos procedentes de los registros de entrada. Este circuito tiene unas entradas de órdenes para seleccionar la clase de operación que debe realizar en cada momento (suma, resta, multiplicación...).
- **Registros de entrada.** En ellos se almacenan los datos u operandos que intervienen en una instrucción antes de la realización de la operación por parte del circuito operacional. También se emplean para el almacenamiento de resultados intermedios o finales de las operaciones respectivas.
- **Registros acumuladores.** Almacena los resultados de las operaciones llevadas a cabo por el circuito operacional. Está conectado con los registros de entrada para realimentación en el caso de operaciones encadenadas. Asimismo tiene una conexión directa al bus de datos para el envío de los resultados a la memoria central o a la unidad de control.
- **Registro de estado (flags).** Se trata de unos registros de memoria en los que se deja constancia de algunas condiciones que se dieron en la última operación realizada y que habrán de ser tenidas en cuenta en operaciones posteriores.
- **Unidad de Punto Flotante.** A partir del 486, la unidad de proceso también incorpora el coprocesador matemático o float point unit (unidad de punto flotante), especializada en realizar operaciones matemáticas con números reales. Anteriormente al 486 era un chip que se podía adquirir separadamente e insertarlo en el correspondiente zócalo de la placa base.

Ejecución de una instrucción

Una vez conocidas todas las partes del microprocesador podemos pasar a estudiar cómo se produce la ejecución de una instrucción del programa. Simplificando,

3. Buses, procesadores y memorias

el proceso consiste en leer la instrucción de memoria, ejecutarla y almacenar el resultado en memoria. Para realizar este intercambio entre memoria y microprocesador se utilizan los siguientes registros:

- **Registro de Direcciones de Memoria (RDM).** Contiene la dirección de memoria donde se va a leer o escribir.
- **Registro de Intercambio de Memoria (RIM).** Almacena el dato que se va a leer o escribir.

Los pasos para la ejecución de una instrucción en el microprocesador son los siguientes:

1. La Unidad de Control debe acceder a la siguiente instrucción del programa determinada por el Contador de Programa. Por ello, el contenido del Contador de Programa, que es la dirección de la siguiente instrucción a ejecutar pasa al RDM y el contenido de dicha dirección se lee en el RIM.
2. El contenido del RIM pasa al Registro de Instrucción, analizando el tipo de operación a realizar (obteniendo el código de operación y los operandos). En este momento el Contador de Programa localiza la siguiente instrucción a ejecutar.
3. La dirección de cada operando se introduce en el RDM para poder acceder a su valor a través del RIM. Ya disponemos del código de operación y de los operandos.
4. La Unidad de Control envía a la Unidad de Proceso el código de operación y los operandos, almacenándolos en los Registros de Entrada. La Unidad de Proceso ejecuta la instrucción y almacena el resultado en los Registros Acumuladores.
5. La Unidad de Control busca en el Registro de Instrucciones la dirección de memoria donde debe ser almacenado el resultado colocándola en el RDM. El resultado se almacena en el RIM y se pasa a memoria.

En realidad el proceso es más complejo, ya que, en los microprocesadores RISC, cada instrucción se descompone por el Decodificador en un conjunto de instrucciones, mientras que en los microprocesadores CISC es el Secuenciador quien realiza esta tarea.

Memoria

La memoria es un dispositivo que almacena información dentro del ordenador. Esta sección se centra en la memoria principal y la memoria caché, siendo los niveles fundamentales dentro de la jerarquía de memoria del PC. No obstante también se pueden encontrar celdas de memoria dentro del microprocesador, en chips de la placa base e incluso en periféricos. Por otro lado, los dispositivos que almacenan información con carácter permanente (discos duros, unidades de cd) no se estudian en este apartado.

La memoria principal tiene como función almacenar la información que es utilizada por el microprocesador para ejecutar un programa. Aloja tanto las instrucciones como los datos que éstas necesitan para ejecutarse.

A nivel histórico, el microprocesador evolucionó rápidamente incrementando su velocidad de forma destacada con respecto a otros dispositivos, incluida la memoria. Pero una elevada potencia de proceso no es efectiva si no se dispone de un sistema de memoria que facilite los datos e instrucciones al microprocesador conforme éstos son necesitados. Los fabricantes se encuentran con dos problemas:

1. Estructuralmente la memoria está fabricada con condensadores, mientras los microprocesadores se construyen con transistores. El tiempo de respuesta de los condensadores es más elevado que el de los transistores.
2. El software demanda más capacidad de memoria. La fabricación de chips de memoria de mayor capacidad implica un aumento en el tiempo de respuesta, al existir mayor número de celdas que direccionar.

La solución es construir memorias fabricadas con transistores, obteniendo tiempos de respuestas acordes a las necesidades del microprocesador. Este tipo de memorias son caras por lo que se adopta una combinación de ambas: La memoria con condensadores (memoria principal) que dota al sistema de una elevada capacidad de almacenamiento y la memoria con transistores (memoria caché) que aporta la velocidad de respuesta que necesita el microprocesador.

Pronto se necesitan añadir cachés adicionales que aproximen los datos al microprocesador, implementándose cachés internas y externas y de niveles L1, L2 y L3. Nace así una jerarquía en la memoria de PC perfectamente estructurada y que incluye los siguientes niveles:

3. Buses, procesadores y memorias

1. Banco de registros del microprocesador.
2. Memorias caché de nivel L1, L2 y L3.
3. Memoria principal.
4. Disco duro (memoria virtual).
5. Soportes para copias de seguridad (unidades cd, cintas).

La memoria se comunica con el microprocesador a través del bus local, que está compuesto por el bus de datos (transporta los datos), el bus de direcciones (contiene la posición de memoria a la que se va a acceder) y el bus de control (establece el tipo de operación: lectura o escritura).

Por lo tanto, en este tema se profundizará en el estudio de la memoria RAM, distinguiendo la memoria RAM dinámica (memoria principal) y memoria RAM estática (memoria caché), incluyendo sus características y una clasificación detallada de los modelos que nos podemos encontrar en el mercado actual.

Las características fundamentales de la memoria son 4:

1. **Tamaño o capacidad.** La memoria está dividida en celdas pudiendo tomar cada una de ellas el valor 0 ó 1. Éstas ocupan cada una un bit y se reúnen en las llamadas palabras de memoria que son el menor conjunto de celdas de memoria que se pueden leer o escribir simultáneamente, es decir una palabra es el conjunto de bits que se leen o escriben en memoria de una vez. Normalmente, la longitud de palabra coincide con un número exacto de bytes.

Su longitud (número de bits que la componen) depende de la arquitectura del ordenador, es decir, de su estructura interna, por lo que a lo largo del tiempo se han utilizado distintos números de bits, siendo los más habituales: 16 bits, 32 bits o 64 bits. Para expresar la capacidad de la memoria, utilizaremos las siguientes unidades: KB, MB, GB ...

2. **Velocidad de la memoria.** Son tres los factores a tener en cuenta:
 - **Tiempo de acceso.** Es el tiempo máximo que se tarda en leer el contenido de una posición de memoria o escribir un dato en una posición de memoria. Se expresa en nanosegundos. Según la velocidad de acceso a cada tipo de memoria podemos establecer la siguiente jerarquía de

3. Buses, procesadores y memorias

memorias de mayor a menor velocidad: Registros internos → Memoria Caché → Memoria RAM → Memoria ROM.

- **Tiempo de ciclo.** Teniendo en cuenta que el proceso de lectura requiere una serie de pasos y que, además, la memoria tiene que ‘refrescarse’, se puede definir como el tiempo mínimo que transcurre entre una lectura y la siguiente.
- **Ancho de banda.** Es la cantidad de información, expresada en bytes, que puede ser enviada de una unidad funcional a otra (en este caso entre memoria y la CPU), por unidad de tiempo.

3. **Duración de la información.** Es el tiempo que la información se mantiene en el soporte y es un factor fundamental para su posterior utilización. Podremos distinguir los siguientes tipos de memorias:

- **Duradera.** Son soportes no volátiles, es decir, la información se mantiene hasta que es sustituida por el usuario. Por ejemplo: soportes magnéticos, ópticos. . .
- **Volátil.** La información se pierde cuando se le deja de suministrar corriente eléctrica. Por ejemplo: memorias de semiconductores (RAM).
- **Con refresco.** Su información se va perdiendo poco a poco, por lo que deben refrescarse periódicamente como se indicó antes. Por ejemplo: memorias DRAM.
- **Permanente.** Contienen siempre la misma información y no pueden modificarse o una vez grabadas. Por ejemplo: memoria ROM.

4. **Cantidad de información a la que se accede.** En un octeto se almacena un carácter, sin embargo, si la CPU tuviese que leer o escribir la información octeto a octeto, el proceso sería muy lento, por lo que el acceso a la información se realiza según el tamaño de la ”palabra” del computador, que suele ser similar al número de bits que se pueden almacenar, por ejemplo, en un registro. La longitud de la palabra puede ser fija o variable, según el tipo de microprocesador.

3.3.1 Memoria Principal - RAM Dinámica

Está organizada en celdas. Cada celda almacena un bit mediante un condensador que se carga positiva o negativamente y que determina el valor lógico 0 ó 1 de la celda. Los bits se agrupan en bytes, cada uno de los cuales tiene asignada una dirección de memoria.

RAM son las iniciales de Random Access Memory (memoria de acceso aleatorio), es decir, un tipo de memoria en la que se puede leer y escribir, especificando la dirección de la celda donde se desea efectuar la operación. Este término también incluye la volatilidad de este tipo de memorias, perdiendo el contenido de las mismas cuando no se les suministra fluido eléctrico. No sólo la memoria principal es RAM, lo es también la memoria caché, el banco de registros del microprocesador o la que se implanta en las tarjetas gráficas.

El concepto RAM dinámica (DRAM) proviene de las características propias de los condensadores que forman las celdas de memoria. Cada condensador (que almacena un 0 ó 1) va perdiendo la carga a lo largo del tiempo, por lo que debe refrescarse periódicamente, consumiendo ciclos de reloj para realimentar la información almacenada. Por el contrario, las memorias caché (RAM estática SRAM) almacenan el estado del bit sin necesidad de consumir ciclos de refresco, siendo, por lo tanto más rápidas, pero mucho más caras.

Parámetros funcionales

Los parámetros que determinan las características de la memoria principal son los siguientes:

- **Tamaño.** Suele indicarse en Megabytes (Mb).
- **Latencia o tiempo de acceso a memoria** Se expresa en nanosegundos (mil millonésima de segundo). Según el fabricante puede expresar una de las siguientes mediciones:
 - **Tiempo de respuesta.** El tiempo que transcurre desde que la memoria recibe la dirección de la celda hasta que devuelve el dato que hay en ella.
 - **Tiempo de ciclo de memoria.** Desde que la memoria acepta una petición y queda lista para aceptar la siguiente.

3. Buses, procesadores y memorias

- **Ancho de banda.** Determina la velocidad de transporte de los datos entre la memoria principal y el microprocesador (memorias caché). Lógicamente este parámetro no depende únicamente de la memoria, sino también de las características del bus que transporta la información entre ambos elementos. Recordemos que el ancho de banda se puede obtener multiplicando la frecuencia por la anchura del bus.

Paridad

Este concepto se utiliza para la detección de errores en los datos transmitidos a memoria. La paridad es un bit que se añade a cada byte de información existiendo dos tipos de paridad:

- **Paridad par.** El bit de paridad se coloca a 0 ó 1 para que el byte resultante tenga un número par de unos.
- **Paridad impar.** El bit de paridad se coloca a 0 ó 1 para que el byte resultante tenga un número impar de unos.

De esta forma, cuando se detecta algún byte que, junto con el bit de paridad, no cumple la citada regla, se detecta el error y se solicita de nuevo la transmisión de ese mismo dato. La memoria con paridad posee chips adicionales para poder realizar estas comprobaciones, por lo que se puede distinguir una memoria con paridad contando el número de chips que posee. Si dicho número no es potencia de dos, entonces es una memoria con paridad.

El zócalo donde se inserta la memoria también debe disponer de esta característica. Una memoria con paridad funciona tanto en un zócalo que tenga o no tenga paridad, pero una memoria sin paridad sólo funciona en zócalos sin paridad.

Pero la paridad puede ocultar ciertos errores, por lo que se ha implementado un nuevo tipo de protección denominado ECC (Error Code Protection). Este sistema genera unos bits de información adicionales para cada byte que permite no sólo detectar los errores sino también corregirlos.

Zócalos de conexión

Dependiendo del zócalo existente en la placa base para insertar los chips de memoria. Estudiado en el tema de placa base.

Tipos de organización

Es necesario estructurar de forma óptima el acceso a una dirección de memoria para invertir el mínimo tiempo posible en la lectura y escritura de datos. Por ello, desde que el microprocesador comenzó a aumentar considerablemente su velocidad de proceso, empezaron a aparecer diferentes tipos de organización de memoria para suministrar al microprocesador los datos de la forma más rápida posible. La evolución es la siguiente:

FPM.

Fast Page Mode (modo de página rápido). Esta implementación se caracteriza por estructurar las celdas de memoria en formato de tabla. De esta forma, cada dirección de memoria se descompone en dos:

- **RAS (Row Access Strobe).** Determina la fila donde se encuentra el dato. El término página hace referencia al concepto fila, pero seguiremos utilizando fila ya que resulta más intuitivo.
- **CAS (Column Access Strobe).** Determina la columna donde se encuentra el dato.

Lógicamente, la intersección de ambos parámetros proporciona la ubicación de la dirección de memoria a la que se quiere acceder. El posicionamiento en fila es mucho más lento que el posicionamiento en columna, pero como el acceso a memoria para leer o escribir suele efectuarse sobre un conjunto de datos consecutivos, FPM permite dejar fija la fila e ir modificando el valor de la columna para acceder a las siguientes posiciones de memoria consecutivas, hasta que tengamos que efectuar una variación en el valor de la fila.

EDO.

Extended Data Output (salida de datos extendida). Aporta la reducción de tiempo de acceso a columnas consecutivas. Para ello se solapa el posicionamiento en la siguiente columna con la operación de lectura/escritura del dato anterior.

BEDO.

Burst Extended Data Output (EDO en ráfaga). Incluye una circuitería adicional que genera automáticamente las 3 direcciones siguientes a la solicitada, por lo

que el acceso a las direcciones consecutivas (siguientes columnas) es realizado directamente por el propio chip de memoria sin que intervenga el controlador de memoria.

SDRAM.

Synchronous Dynamic RAM (RAM dinámica síncrona). Todas las organizaciones estudiadas hasta ahora tienen un comportamiento asíncrono, es decir, la memoria trabaja con un ritmo independiente del reloj del sistema y la CPU debe esperar hasta que la memoria le entregue el dato, mientras tanto, las señales del bus de control y de direcciones deben permanecer activas a la entrada de la memoria hasta que finalice la instrucción. Existen dos tipos:

- **SDR SDRAM (Single Data Rate Synchronous Dynamic RAM.** El mercado las rebautizó como memoria SDRAM omitiendo el SDR y creando confusión con las memorias DDR SDRAM que también son memorias SDRAM pero se quedó el nombre mal para el mercado.

Fue utilizada en los Pentium 2, así como en los AMD K7. Se implementa en módulos DIMM de 168 contactos. Dependiendo de la frecuencia de trabajo se dividen en:

- **PC66.** La velocidad de bus de memoria es de 66 Mhz, temporización de 15 ns y ofrece tasas de transferencia de hasta 533 MB/s.
- **PC100.** La velocidad de bus de memoria es de 100 Mhz, temporización de 8 ns y ofrece tasas de transferencia de hasta 800 MB/s.
- **PC133.** La velocidad de bus de memoria es de 133 Mhz, temporización de 7,5 ns y ofrece tasas de transferencia de hasta 1066 MB/s.

RDRAM.

Rambus Dynamic RAM. La compañía Rambus decide dar un salto en el lento desarrollo de las arquitecturas de memoria. Asociada con Intel lanza la memoria RDRAM basada en el Canal Rambus Directo, un nuevo bus que admite una frecuencia de 400 MHz, pudiendo proporcionar datos tanto en el flanco de subida como en el de bajada de la señal de reloj, proporcionando una frecuencia efectiva de 800 MHz.

Utiliza el zócalo de conexión RIMM, cuya estructura es semejante a un zócalo DIMM de 184 contactos (de hecho podemos conectar un módulo

3. Buses, procesadores y memorias

RDRAM en un zócalo DIMM, o un módulo DDR en un zócalo RIMM, pero ninguno de los dos funcionará). La incompatibilidad es debida tanto a la frecuencia de reloj como a la forma en que queda insertado el módulo, ya que el zócalo RIMM consigue una conexión como si el módulo se encontrara soldado a la placa base. Estas frecuencias tan elevadas se consiguen gracias a que el bus de datos interno es de 16 bits. En la actualidad se empiezan a construir módulos de 32 bits, realizados por la empresa Kingston Technology que los ha bautizado con el nombre ValueRAM.

Las placas base con memoria RDRAM suministran zócalos RIMM que tienen que ser instalados por parejas si utilizamos RDRAM de 16 bits. En los módulos de 32 bits no ocurre esto, pudiendo efectuar la instalación de un único módulo en el sistema. Los zócalos forman una cadena en serie conectada al controlador de memoria, por ello, si no se llenan todos los zócalos RIMM, se tienen que insertar módulos de continuidad, que no incluyen memoria, pero que permiten mantener el canal de comunicación.

	Frec. bus	Frec. efect.	Anch. bus	T ciclo	Anch. banda	Volt.
PC600	266 MHz	532 MHz	16 bits	1,8 ns	1,06 Gb/s	2,5 v
PC700	356 MHz	712 MHz	16 bits	1,4 ns	1,42 Gb/s	2,5 v
PC800	400 MHz	800 MHz	16 bits	1,2 ns	1,6 Gb/s	2,5 v
RIMM3200	400 MHz	800 MHz	32 bits	1,2 ns	3,2 Gb/s	2,5 v
RIMM4200	533 MHz	1066 MHz	32 bits	0,9 ns	4,2 Gb/s	2,5 v

DDR SDRAM (Double Data Rate Synchronous Dynamic RAM).

Es una memoria SDRAM que responde a una sincronización tanto con los flancos de subida de la señal de reloj como con los flancos de bajada, es decir, es una SDR SDRAM 2x. De esta forma, una memoria DDR instalada en un sistema con bus de datos a 100 MHz sería equivalente a un sistema con memoria SDRAM instalada sobre un bus de datos de 200 MHz (si se pudiera implementar). Podemos comprobar, por lo tanto, que las frecuencias de las memorias DDR las podemos llamar efectivas pero no reales.

Se implementan en módulos DIMM de 184 contactos existiendo tres modelos:

3. Buses, procesadores y memorias

	Frec. bus	Frec. efectiva	Anch. bus	T. ciclo	Anch. banda	Vol.
PC1600(PC200)	100 MHz	200 MHz	64 bits	5 ns	1,6 Gb/s	2,5 v
PC2100(PC266)	133 MHz	266 MHz	64 bits	3,7 ns	2,1 Gb/s	2,5 v
PC2700(PC333)	166 MHz	333 MHz	64 bits	3 ns	2,7 Gb/s	2,5 v
PC3200(PC400)	200 MHz	400 Mhz	64 bits	2,5 ns	3,2 Gb/s	2,5 v

La memoria DDR fue implantada por AMD en los microprocesadores K7, ya que alcanzan velocidades de 200 MHz en su bus. La memoria SDRAM sólo alcanza los 133 MHz, por lo que hay que encontrar una arquitectura que soporte este aumento de la velocidad. Pero pronto Intel atacó con la memoria RDRAM, llegando a frecuencias de 800 MHz, lo que no quiere decir que el ancho de banda se multiplique bajo el mismo factor que la frecuencia. Es por ello que la memoria DDR cambió su nomenclatura (PC200 pasa a denominarse PC1600) expresando el ancho de banda y no la frecuencia.

DDR2 SDRAM.

Son capaces de trabajar con 4 bits por ciclo, es decir 2 de ida y 2 de vuelta en un mismo ciclo mejorando sustancialmente el ancho de banda potencial bajo la misma frecuencia de una DDR SDRAM tradicional (si una DDR a 200 MHz reales entregaba 400 MHz nominales, la DDR2 por esos mismos 200 MHz reales entrega 800 MHz nominales). Este sistema funciona debido a que dentro de las memorias hay un pequeño buffer que es el que guarda la información para luego transmitirla fuera del modulo de memoria, este buffer en el caso de la DDR convencional trabajaba tomando los 2 bits para transmitirlos en 1 sólo ciclo, lo que aumenta la frecuencia final. En las DDR2, el buffer almacena 4 bits para luego enviarlos, lo que a su vez redobla la frecuencia nominal sin necesidad de aumentar la frecuencia real de los módulos de memoria.

La tasa de transferencia varía desde 400 hasta 1024MB/s y sus capacidades llegan hasta 2 x 2 GB. Algunas marcas de estas memorias son: STD, Transcend, Kingston, Buffalo, NEC, Elixir, Vdata, TRCND, OCZ, Corsair, G. Skill, Mushkin.

Para usar en PCs, las DDR2 SDRAM son suministradas en tarjetas de memoria DIMMs con 240 pines y una localización con una sola ranura. Las tarjetas DIMM son identificadas por su máxima capacidad de transferencia (usualmente llamado ancho de banda).

3. Buses, procesadores y memorias

Nombre	Frec. reloj	Frec. efectiva	Max. capacidad de transferencia
DDR2-400/PC2-3200	100 MHz	400 MHz	3.200 MB/s
DDR2-533/PC2-4200	133 MHz	533 MHz	4.264 MB/s
DDR2-667/PC2-5300	166 MHz	667 MHz	5.336 MB/s
DDR2-800/PC2-6400	200 MHz	800 MHz	6.400 MB/s
DDR2-1.066/PC2-8500	266 MHz	1066 MHz	8.500 MB/s

DDR3 SDRAM.

DDR 3 es el nombre del nuevo estándar DDR, que viene siendo el sucesor del DDR 2. En febrero, Samsung Electronics anunció un chip prototipo de 512 MB a 1066 MHz (La misma velocidad de bus frontal del Pentium 4 Extreme Edition más rápido) con una reducción de consumo de energía de un 40 % comparado con los actuales módulos comerciales DDR 2.

Teóricamente, estos módulos pueden transferir datos a una tasa de reloj efectiva de 800-1600 MHz, comparado con el rango actual del DDR 2 de 533-800 MHz ó 200-400 MHz del DDR. .

Los DIMMS DDR3 tienen 240 pines, el mismo número que DDR 2; sin embargo, los DIMMs son físicamente incompatibles, debido a una ubicación diferente de la muesca.

Nombre	Frec. reloj	Frec. efectiva	Max. transferencia
DDR3-800/PC3-6400	100 MHz	800 MHz	6400 MB/s
DDR3-1.066/PC3-8500	133 MHz	1066 MHz	8533 MB/s
DDR3-1.333/PC3-10600	166 MHz	1333 MHz	10667 MB/s
DDR3-1.600/PC3-12800	200 MHz	1600 MHz	12800 MB/s

GDDR (Graphics Double Data Rate.

Esta memoria se diseñó específicamente para el uso en tarjetas gráficas. Actualmente existen los siguientes:

- **GDDR2.**
- **GDDR3.**
- **GDDR4.**
- **GDDR5.**

3. Buses, procesadores y memorias

El origen de la memoria DDR2 fue empleado por las tarjetas gráficas nVidia GeForce FX 5800. Sin embargo, es importante aclarar que la memoria DDR2 usada en las tarjetas gráficas (llamada oficialmente GDDR2) no es DDR2, sino un punto intermedio entre las DDR y DDR2. De hecho, no incluye el doble ratio del reloj de entrada/salida, tiene problemas de calentamiento etc.

Ejercicio 3.9 *Construye tablas similares a las anteriores pero con las características propias de las memorias GDDR2, GDDR3, GDDR4 y GDDR5.*

3.3.2 Memoria Caché - RAM Estática

La memoria RAM estática (SRAM) es la memoria más rápida, implementada por medio de transistores que permiten almacenar los bits de información sin necesidad de consumir ciclos de refresco; por el contrario es mucho más cara que la memoria RAM dinámica.

La memoria caché se estructura en líneas o grupos de bytes consecutivos que representan la unidad mínima de intercambio con la memoria principal, por lo que se optimizan las siguientes operaciones:

- **Acceso.** Al solicitar a memoria principal una línea de caché se efectúan lecturas de direcciones consecutivas, optimizando, como hemos visto, la gestión de señales RAS y CAS.
- **Transporte.** El ancho de banda del bus de datos se aprovecha más cuanto más grande sea el bloque de datos a transmitir. Al transportar líneas en vez de bytes individuales optimizamos el uso del bus.
- **Direccionamiento.** Si el elemento más pequeño es la línea de caché es menor el número de elementos a gestionar, optimizando el direccionamiento o localización de los mismos.

Existe una jerarquía en niveles implícita en las memorias caché, que viene determinada por su ubicación:

- **Externa.** Se encuentra en la placa base, por lo que no forma parte del microprocesador.

3. Buses, procesadores y memorias

- **Interna.** Forma parte del microprocesador pero se encuentra en un chip independiente.
- **Integrada.** Incluida en el microprocesador e integrada en el propio chip de la CPU. En la actualidad todas las memorias caché son integradas.

Con respecto a la tipología de las memorias caché podemos observar la siguiente evolución histórica:

- **Caché asíncrona.** Las primeras memorias caché no podían trabajar síncronamente con el microprocesador, sino que éste debía esperar a que la operación se completase, aunque el tiempo de espera era inferior que trabajar directamente con la memoria principal.

Este tipo de memorias caché eran externas y estaban asociadas a microprocesadores 386, con un tamaño que oscila entre los 256 y los 512 Kb y una latencia entre los 20 y 15 ns. Solían existir zócalos de expansión para ampliar la cantidad de memoria caché.

- **Caché síncrona.** Al incorporar la memorias caché dentro del microprocesador se consigue trabajar a la misma frecuencia de reloj, naciendo así las memorias caché síncronas montadas sobre microprocesadores Pentium. Podemos encontrar dos tipos:

- **Burst SRAM (caché ráfaga).** Incorpora, al igual que la memoria BEDO, circuitería adicional que retiene la dirección de acceso a memoria caché y genera automáticamente las siguientes posiciones consecutivas, con el consiguiente ahorro de tiempo.

Trabaja sobre placas base con bus de datos a 66 MHz, alcanzando tiempos de acceso entre 8 y 12 ns, aunque debido a su coste de fabricación pronto desaparecieron del mercado.

- **Pipeline SRAM (caché segmentada).** Además de incorporar circuitería a la entrada de la caché relativa al bus de direcciones, como hemos visto en la burst SRAM, incluye circuitería a la salida, acoplada al bus de datos, incorporando la técnica vista en la memoria EDO: Mientras la caché accede a la siguiente posición de memoria se efectúa la operación de lectura/escritura del dato anterior. Este el tipo de memoria caché que se incorpora en los microprocesadores actuales pudiendo llegar hasta los 533 MHz de frecuencia de bus.

Una mejora que se ha efectuado a la memoria caché segmentada viene aportada por el diseño Claymore SRAM, que incorpora tecnología DDR, efectuando operaciones de transferencia de datos tanto en el flanco de subida como en el de bajada de la señal de reloj, proporcionando una frecuencia efectiva que duplica a la frecuencia real del bus.

Hemos comentado que la unidad mínima de transferencia en las memorias caché es la línea, constituida por varios bytes. En la actualidad nace un nuevo concepto, denominado conjunto, que agrupa a un conjunto de líneas que son transferidas como una unidad. La utilización de conjuntos mejora el rendimiento, pudiendo encontrar configuraciones de 1, 2 u 8 líneas por conjunto. Superar 8 líneas por conjunto no mejora sustancialmente el rendimiento y aumenta el coste de fabricación de las memorias caché.

3.3.3 Memoria ROM (Read Only Memory)

ROM (Read Only Memory, Memoria de Sólo Lectura) formada también por semiconductores, permiten el acceso directo, pero sólo puede accederse a ella para la operación de lectura de la información que contiene.

A diferencia de la memoria RAM, estas memorias no son volátiles. Son memorias programadas durante el proceso de su fabricación, almacenándose en ellas los programas necesarios para el funcionamiento del sistema. Como ejemplo más representativo está el programa de arranque (boot), encargado de hacer un chequeo del hardware para posteriormente iniciar la carga del sistema operativo. El tiempo de acceso a la información contenida en la memoria ROM es superior a la RAM.

Desde el punto de vista hardware, existen distintos tipos de ROM:

- **Memoria PROM (Memoria de Sólo Lectura Programable).** Es una variante de la memoria ROM y puede ser programada por el usuario mediante un dispositivo externo de escritura (programador de memorias). Una vez que esa información ha sido grabada, ya no puede cambiarse, por lo que se convierte en memoria ROM, permitiendo sólo su lectura.
- **Memoria EPROM (Memoria de Sólo Lectura Programable y Borrable).** También llamada PROM reprogramable, permite grabar y borrar su con-

tenido tantas veces como quiera el usuario. Esta memoria está sin programar y se graba fuera del ordenador, permitiendo, después, sólo su lectura. El inconveniente es que si queremos modificar el contenido de una celda solamente, tiene que borrarse toda la información de la memoria.

- **Memoria EEPROM (Memoria de sólo lectura programable y borrable electrónicamente).** Su mayor ventaja frente a las otras es que se puede escribir en cualquier momento sin borrar su contenido anterior; sólo se actualiza la parte concreta que se quiere modificar.
- **Memoria Flash.** Son memorias ROM programables por software. Actualmente las BIOS de la tarjeta gráfica, placa base y de otros dispositivos están contenidas en memorias flash, que son memorias de tipo ROM, pero ésta tiene la ventaja de poder ser actualizada conforme el software evoluciona, es decir, permiten modificar su contenido mediante la ejecución de un determinado programa. Estas se emplean en sustitución de las ROM y EPROM.
Por sus características podríamos decir que es una pseudo-memoria RAM, ya que permite tanto su lectura como escritura en cualquier momento sin necesidad de dispositivos externos, pero con la ventaja adicional de no ser volátil, por lo que no la englobamos en ninguno de los dos grupos. Estas memorias también son utilizadas como dispositivos de almacenamiento externo en cámaras digitales, reproductores de MP-3 portátiles o los PDAs.

SECCION 3.4

Anexo. Las generaciones de Microprocesadores

La evolución histórica de los microprocesadores se agrupa por generaciones, concepto utilizado para destacar grandes cambios producidos en la arquitectura de los mismos. Se ha incluido en cada generación el modelo correspondiente al fabricante Intel, dado que es la compañía más representativa.

- **Primera generación (1978-1982) 8086.** Viene representada por el primer microprocesador, el 8086. Se caracteriza por la potencia de cálculo, al compararlo con sus predecesores, el 4004 y el 8008. Fue la arquitectura de 16 bits la que permitió incorporar al microprocesador unidades funcionales más complejas que permitieron mejorar las prestaciones. Una arquitectura de 16 bits quiere decir que el banco de registros del microprocesador son

3. Buses, procesadores y memorias

de 16 bits, que su unidad aritmético lógica trabaja con enteros de 16 bits y que el direccionamiento a memoria se realiza a través de 16 bits (bus de direcciones).

- **Segunda generación (1982-1985) 286.** La memoria es la protagonista en esta generación. El software demanda cada vez más espacio, por lo que se aumenta el bus de direcciones, pudiendo direccionar hasta 1 Gb de memoria. También se utiliza memoria virtual.
- **Tercera generación (1985-1989) 386.** La utilización de memoria virtual conlleva efectuar traducciones de direcciones virtuales y físicas para cada acceso a memoria, siendo un proceso costoso, incrementado por la ausencia de memorias caché internas. Para su optimización se decide instalar en el propio microprocesador la circuitería que lleve a cabo dichas traducciones. Se implanta la arquitectura de 32 bits.
- **Cuarta generación (1989-1993) 486.** Vuelve a centrar el protagonismo la potencia de cálculo. Se decide acoplar la unidad de punto flotante o coprocesador matemático dentro de la estructura del microprocesador. Aparecen las primeras memorias caché internas.
- **Quinta generación (1993-1997) Pentium.** La frecuencia del microprocesador se dispara, diferenciándose del resto de los componentes del ordenador. La jerarquía de buses que empezó a implantarse en los 486 queda consolidada favoreciendo la modularidad del sistema: cada dispositivo funciona a su velocidad, dispone de su zócalo correspondiente de conexión y puede adquirirse y montarse independientemente. La incorporación y replicación de unidades funcionales en el microprocesador permiten el paralelismo a nivel de instrucción.
- **Sexta generación (1997-2000) Pentium II - Pentium III.** Nace el concepto SMP (Symetric MultiProcessing = multiprocesamiento simétrico) que consiste en incorporar al microprocesador la circuitería necesaria para realizar las operaciones conjuntamente con otros microprocesadores. No nos encontramos ante un multiprocesador puro, ya que se comparten elementos comunes, como la memoria, los periféricos y los buses. Esta técnica, incorporada en los Pentium Pro, es un fracaso a nivel comercial, ya que su coste es muy elevado y el software no saca el verdadero partido al SMP. Por otra parte se incorpora una nueva memoria caché interna de nivel 2.

- **Séptima generación (2000) Pentium 4.** Se supera la barrera del Gigahercio en la frecuencia de reloj. Pasamos a una arquitectura de 64 bits, aunque los primeros Pentium 4 no fueran una arquitectura de 64 bits pura. El paralelismo a nivel de instrucción provoca una serie de conflictos que generan la puesta en marcha de mecanismos que garantizan la ejecución de las instrucciones de un programa. Este papel lo desarrollaba hasta ahora el microprocesador, invirtiendo parte de su tiempo en dichas comprobaciones, pero ahora es trasladado a la capa software, siendo los compiladores los encargados de detectar los posibles conflictos que se pudieran producir en la ejecución paralela de las instrucciones de un programa.

SECCION 3.5

Ejercicios

1. ¿Cual es la velocidad de un bus paralelo, que tiene capacidad para enviar 8 bits, si trabaja a una frecuencia de 11KHZ?.
2. Si una CPU tiene un bus de direcciones de 8 bits, ¿cuál será la dirección de memoria más alta posible?
3. ¿Qué es overclocking?
4. ¿Qué es el hypertransport?
5. ¿Qué diferencias existen entre Hypertransport e Hyperthreading? ¿Y entre Hypertransport y FSB? ¿Y entre FSB y Hyperthreading?
6. ¿Qué es la memoria caché? ¿Existe memoria caché de datos y de instrucciones? ¿Cuántos niveles están integrados en el microprocesador?
7. ¿Cuales son las ventajas e inconvenientes de los repertorios de instrucciones CISC y RISC. ¿Cuál se usa actualmente? Da un ejemplo de procesadores que lo usen.
8. ¿Qué es Plug and Play? ¿Es lo mismo Hot Plug que Plug and Play? ¿Qué buses soportan Plug and Play? ¿Y cuáles soportan Hot Plug? ¿Existen buses que soporten Plug and Play y Hot Plug?
9. ¿Qué elementos se modifican al hacer overclocking?

3. Buses, procesadores y memorias

10. ¿Qué es la velocidad interna y velocidad externa de un procesador?. ¿Qué es el multiplicador? ¿Que relación existe entre ellos?
11. Utiliza la documentación proporcionada sobre microprocesadores o Internet para responder a las siguientes cuestiones:
 - a) ¿Cuál fue el primer microprocesador de Intel de 32 bits? ¿Y el de 64 bits?
 - b) ¿Cuál es la diferencia entre un disipador y un ventilador?
 - c) ¿Cómo se acopla el disipador al microprocesador?
 - d) ¿Qué son los benchmarks?
12. ¿Qué diferencia existe entre la DRAM y la SRAM?
13. ¿Qué es la memoria SDR SDRAM? ¿Qué es la memoria DDR SDRAM?
14. ¿Que significa PC100? en memria SDR SDRAM
15. Si tenemos una SDR SRAM con ancho de bus de datos igual a 64 bits y una velocidad de 100 Mhz. ¿cual es su velocidad de transferencia?
16. ¿Qué es el tiempo de ciclo de memoria?. ¿En qué unidad se mide?.
17. En la memoria principal (RAM) se cargan únicamente los programas y datos que se están utilizando en un momento determinado. Explica los dos motivos fundamentales por los que NO se utiliza la memoria principal para almacenar todos los programas y todos los datos que nos pueden hacer falta.
18. ¿Cómo se lleva a cabo el proceso de lectura de información de la memoria DRAM?
19. ¿Cómo se lleva a cabo el proceso de escritura de información de la memoria DRAM?
20. ¿Cuál es el objetivo de la ‘arquitectura jerárquica de memoria’?.Describela.
21. ¿Cuál es la mínima unidad de información a la que se puede tener acceso en una memoria?¿Qué otras unidades de acceso hay?
22. ¿Unidades en las que se mide la capacidad de memoria?

3. Buses, procesadores y memorias

23. Características de la memoria SDR SDRAM.
24. ¿Son compatibles los módulos SIMM y DIMM? Razona la respuesta
25. ¿Para que es necesario la memoria caché?
26. ¿Cuántos tipos de caché existen? ¿Cuántos niveles se integran en el microprocesador? ¿cómo varia la capacidad de la caché según el nivel que ocupa?
27. ¿Qué capacidad de memoria tenemos si poseemos un bus de direcciones de 32 bits y un ancho de palabra de 1 byte?
28. Si tenemos dos memorias con el mismo tiempo de acceso a la las posiciones y una es de 1 Mb y la otra de 4MB. ¿En cual de las dos se tarda mas en acceder a la posición 79?