# Compoñentes do sistema informático

## Introdución

Os dous compoñentes fundamentais dun computador, imprescindibles para que este funcione, son o hardware, parte física ou tanxible, e o software, parte lóxica ou intanxible. Esta unidade describe o hardware dun sistema informático, os seus compoñentes, configuración, as liñas polas que circula a información, os periféricos, etcétera.

## 1.1 Compoñentes hardware dun ordenador

É evidente que, debido á rapidísima evolución do hardware e do propio software, os compoñentes dos computadores tamén evolucionan rapidamente. Cada vez son máis rápidos, teñen maior capacidade de almacenamento, son máis pequenos, etc. Nesta Unidade exponse as características básicas e xenerais dos compoñentes hardware de calquera computador persoal e farase referencia ás características xenéricas, sen detallar demasiado as prestacións específicas de cada un dos compoñentes.

Os compoñentes básicos dun computador son os seguintes:

* Carcasa do computador. Dentro dela atópanse:

-Placa base.

-tarxeta gráfico.

-tarxeta de son.

-Procesador.

-Memoria RAM.

-Disco/s duro/s.

-DVD-ROM/CD-ROM.

-Unidade de disco flexible.

-Módem interno.

-tarxeta de rede.

-Etcétera.

* Monitor.
* Teclado e rato.
* Impresora, módem externo, escáner e periféricos en xeral.

## 1.2 Carcasa do ordenador

A carcasa do computador é unha parte importante do mesmo, dado que nela alóxanse a maioría dos dispositivos hardware do equipo.

A carcasa do computador ten que adaptarse ás características que cada usuario vaia a necesitar; non conterá o mesmo un computador para un usuario que non necesite conectar ningún tipo de periférico que a carcasa de grandes ordenadores, como os servidores de rede ou servidores de Internet.

O tamaño das carcasas vén dado polo factor de forma da placa base. Con todo o factor de forma só especifica o tamaño interno da caixa.

As características que se deben ter en conta o elixir unha carcasa do ordenador son:

* Espazo para posibles ampliacións de hardware (baias, ranuras de expansión, etc)
* Ventilación.
* Protección física (golpes, contactos accidentais, interferencias electromagnéticas).
* Estética.

Os tipos de carcasa, ou caixas, que podemos atopar para os computadores adoitan ser os seguintes:

* Barebone: Carcasas de pequeno tamaño cuxa función principal é a de ocupar menor espazo e crea un deseño máis agradable. Son útiles para persoas que queiran dar boa impresión como unha persoa que teña un despacho no que reciba a moita xente. Os barebone teñen o problema de que a expansión é complicada debido a que admite poucos (ou ningún) dispositivos. Outro punto en contra é o quecemento ao ser de tamaño reducido aínda que para unha persoa que non esixa moito traballo ao computador pode estar ben. Este tipo de caixas teñen moitos portos USB para compensar a falta de dispositivos, como unha disqueteira (xa obsoleta), para poder conectar dispositivos externos como un disco USB ou unha memoria.
* Minitorre: Dispón dunha ou dúas baías de 5 ¼ e dúas ou tres baías de 3 ½. Dependendo da placa base pódense colocar bastantes tarxetas. Non adoitan ter problema cos USB e véndense bastantes modelos deste tipo de torre xa que é pequena e á súa vez permiten a expansión do equipo. O seu quecemento é normal e non ten o problema dos barebone.
* Sobremesa: Non se diferencian moito das minitorres, fóra de que en lugar de estar en vertical colócanse en horizontal sobre o escritorio. Antes usábanse moito, pero agora están cada vez máis en desuso. Adoitábase colocar sobre ela o monitor da computadora.
* Mediatorre ou semitorre: A diferenza desta é que aumenta o seu tamaño para poder colocar máis dispositivos. Normalmente son de 4 baías de 5 ¼ e 4 de 3 ½ e un gran número de ocos para poder colocar tarxetas e demais aínda que isto depende sempre da placa base.
* Torre: É o máis grande. Podes colocar unha gran cantidade de dispositivos e é usado cando se precisa unha gran cantidade de dispositivos.
* Servidor: Adoitan ser carcasas máis anchas que as outras e dunha estética inexistente debido a que van destinadas a lugares nos que non hai moito tránsito de clientes como é un centro de procesamento de datos. O seu deseño está baseado na eficiencia onde os perifé
* a (SAI) que protexe aos equipos dos picos de tensión e consegue que en caso de caída da rede eléctrica o servidor siga funcionando por un tempo limitado.
* Rack: Son outro tipo de servidores. Normalmente están dedicados e teñen unha potencia superior que calquera outro ordenador. Os servidores ''rack'' se aparafusan a un moble que ten unha medida especial: a "U". Unha "U" é o ancho dunha ranura do moble. Este tipo de servidores adoita colocarse en salas climatizadas debido á temperatura que alcanza.
* Modding: O ''modding'' é un tipo de carcasa que é totalmente estético mesmo se podería dicir nalgúns casos que son pouco funcionais. Normalmente este tipo de carcasas leva incorporado unha chea de luces de neón, ventiladores, debuxos e cores estrañas pero tamén os hai con formas extravagantes que fan que moitas veces sexa difícil a expansión (como unha torre (informática) torre en forma de Pirámide (xeometría)pirámide na que colocar compoñentes complícase.
* Portátiles: Son equipos xa definidos. Pouco se pode facer para expandilos e adoitan quentarse moito se son moi esixidos. O tamaño adoita depender do monitor que trae incorporado e cos tempos son cada vez máis finos. A súa utilidade baséase en que temos todo o equipo integrado no carcasa: Teclado (informática)Teclado, monitor, e rato, e por tanto fano portátil.

## 3. O MODELO DE VON NEUMANN.

A idea principal do modelo para unha computadora ideado por John Von Neumann é almacenar as instrucións do programa dunha computadora na súa propia memoria, logrando así, que a máquina siga os pasos indicados polo seu programa almacenado.

Unha computadora que funciona baixo o modelo de Von Neumann tamén recibe o nome de computadora de programa almacenado e ten a configuración presentada na figura 1, así como as relacións estruturais que existen entre as diversas partes que compoñen a máquina.

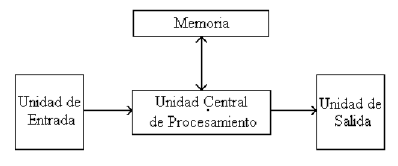


Figura 1. Representación básica dunha computadora.

A unidade central de procesamento (CPU) contén a unidade aritmética e lóxica (ALU) e a unidade de control.

Os pasos necesarios para poder executar estas dúas operacións son os seguintes:

Lectura.

1. Proporcionar a dirección da cela que será lida.

2. Esperar un tempo fixo para obter o dato contido na cela.

3. Tomar ese dato e indicar que terminou a operación.

Escritura.

1. Proporcionar o dato que se desexa depositar na cela.

2. Especificar a dirección da cela sobre a que se escribirá o dato.

3. Esperar un tempo para que poida ser escrito o dato e indicar que terminou a operación.

Agora débese resolver o problema de como almacenar as instrucións en memoria.

Se nas celas de memoria só se poden gardar números, entón haberá que traducir as instrucións a números para poder almacenalas.

A codificación de instrucións contempla cantas e cales son as instrucións dispoñibles e o esquema de codificación a empregar.

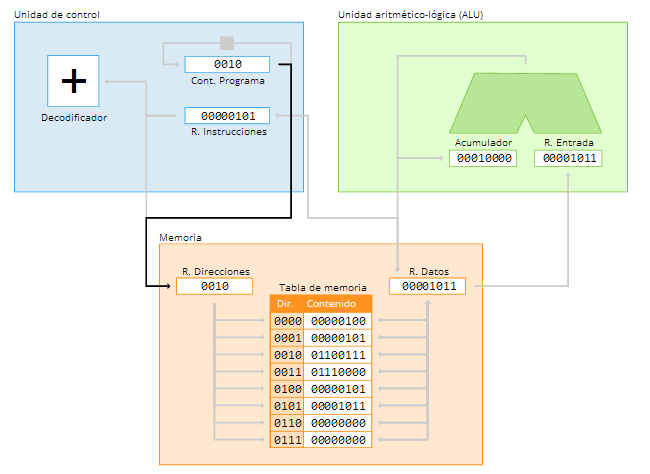
Pola súa banda, para a CPU, a función principal da unidade de control é dirixir a secuencia de pasos para que a computadora leve a cabo un ciclo completo de execución dunha instrución, e facer isto con todas as operacións que conteña o programa. Os pasos para executar unha instrución calquera son os seguintes:

1. Traer da memoria o código da seguinte instrución.

2. Decodificar a instrución lida, é dicir, interpretar o seu código.

3. Executar a instrución de acordo á interpretación dada.

4. Preparar a lectura da seguinte cela de memoria, para a si volver ao paso 1.



A unidade de control executará en múltiples ocasións este ciclo composto de catro operacións a unha gran velocidade. Estas operacións non se atopan en memoria, senón que a máquina execútaas mediante circuítos electrónicos.

Salvo raras excepcións, todas as computadoras modernas teñen o mesmo funcionamento e estrutura xeral, polo cal se pode referir a elas como máquinas de Von Neumann.

## 4. COMPOÑENTES FÍSICOS DE UN COMPUTADOR.

O computador non está formado por un só bloque, senón que o forman diferentes partes encargadas cada unha delas dun labor moi específico dentro de todo o conxunto.

Hai unha parte do computador que lle serve para comunicarse co exterior, é dicir, para recibir e emitir información. Son as unidades de entrada saída, unidades periféricas ou simplemente periféricos.

Outra parte fundamental é a memoria, que se encarga de memorizar as instrucións, datos e resultados.

A terceira parte fundamental é a unidade central de proceso (U.C.P. ou C.P.U.), que se divide en dous partes: unha parte que controla todo o proceso (unidade de control, U.C.) e outra parte na que se realizan as operacións aritméticas e lóxicas que ordene a U.C. (é a unidade aritmético lóxica, O.A.L. ou A.L.U. ).

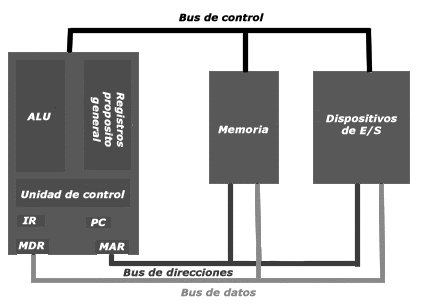


Fig. Esquema xeral da estrutura dun computador.

## 4.1. UNIDADE CENTRAL DE PROCESO.

A unidade central de proceso ou CPU é o verdadeiro motor dun computador. Encárgase de realizar as tarefas fundamentais:

* Operacións aritméticas e lóxicas
* Direcionamento de Memoria
* Xestión de instrucións
* Control do transporte dos datos a través dos buses.

O microprocesador ou simplemente procesador, é o circuíto integrado central e máis complexo dun sistema informático. O procesador pode definirse, como un circuíto integrado constituído por millóns de compoñentes electrónicos agrupados nun paquete. Constitúe a unidade central de procesamento (CPU) dun PC catalogado como microcomputador.

Desde o punto de vista funcional é, basicamente, o encargado de realizar toda operación aritmético-lóxica, de control e de comunicación co resto dos compoñentes integrados que conforman un PC, seguindo o modelo base de Von Neumann. Tamén é o principal encargado de executar os programas, sexan de usuario ou de sistema; só executa instrucións programadas a moi baixo nivel, realizando operacións elementais, basicamente, as aritméticas e lóxicas, tales como sumar, restar, multiplicar, dividir, as lóxicas binarias e accesos a memoria.

Esta unidade central de procesamento está constituída, esencialmente, por rexistros, unha unidade de control e unha unidade aritmético lóxica (ALU), aínda que actualmente todo microprocesador tamén inclúe unha unidade de cálculo en coma flotante, (tamén coñecida como co-procesador matemático), que permite operacións por hardware con números decimais, elevando polo tanto notablemente a eficiencia que proporciona só a ALU co cálculo indirecto a través dos clásicos números enteiros. Aparece en computadoras de cuarta xeración.

A capacidade de proceso do microprocesador (mal chamada velocidade) adoita medirse pola cantidade de operacións por ciclo de reloxo que pode realizar e nos ciclos por segundo que este último desenvolve, ou tamén en MIPS. O elemento que marca o ritmo ao que se executan estas instrucións é un reloxo oscilador de cuarzo, de tal forma que canta máis alta sexa a súa frecuencia de oscilación, maior velocidade de traballo obterase en termos xerais. Cabe destacar que a frecuencia de reloxo non é o único factor determinante no rendemento, pois só se podería facer comparativa entre dous microprocesadores dunha mesma microarquitectura. É importante notar que a frecuencia de reloxo efectiva non é a suma da frecuencia de cada núcleo físico do procesador, é dicir, un de 6 núcleos físicos con 3 GHz cada un nunca terá 18 GHz, senón 3 GHz, independentemente do seu número de núcleos. Outra unidade habitual de medida dá capacidade de proceso son os MegaFLOPS (millóns de operacións en coma flotante que se realizan por segundo).

As instrucións por ciclo ou **IPC** (eninglés *instructions per cycle*) indica a cantidade deinstrucións que un procesador executa nunciclo de reloxo. Noutras palabras, é un indicador máis de rendemento dun microprocesador.

As instrucións dun procesador sóense executar dividas en variasetapas ou fases. Cando se teñen varias instrucións en execución, unha en cada etapa, en condicións ideais se da o resultado aparente de obter unha instrución por ciclo de reloxo (IPC = 1), posto que cando unha instrución acaba a súa última fase, a instrución seguinte entra na súa última fase de maneira que o seguinte ciclo acaba tamén, como si se tratara da cadea de montaxe dun coche.

Hai outros factores moi influentes no rendemento, como pode ser a súa memoria caché, a súa cantidade de núcleos, sexan físicos ou lóxicos, o conxunto de instrucións que soporta, a súa arquitectura, etc; polo que sería dificilmente comparable o rendemento de dous procesadores distintos baseándose só na súa frecuencia de reloxo.

Un sistema informático de alto rendemento pode estar equipado con varios microprocesadores traballando en paralelo, e un microprocesador pode, á súa vez, estar constituído por varios núcleos físicos ou lóxicos. Un núcleo físico refírese a unha porción interna do microprocesador case-independente que realiza todas as actividades dunha CPU solitaria, un núcleo lóxico é a simulación dun núcleo físico a fin de repartir de maneira máis eficiente o procesamento.

A tendencia dos últimos anos foi a de integrar o maior número de elementos do PC dentro do propio procesador, aumentando así a súa eficiencia enerxética e o seu rendemento. Unha das primeiras integracións, foi introducir a unidade de coma flotante dentro do encapsulado, que anteriormente era un compoñente aparte e opcional situado tamén na placa base, logo introduciuse tamén o controlador de memoria, e máis tarde un procesador gráfico dentro da mesma cámara, aínda que non dentro do mesmo encapsulado. Posteriormente chegáronse a integrar completamente no mesmo encapsulado (''die''). Respecto disto último, compañías tales como Intel xa planean integrar a ponte sur dentro do microprocesador, eliminando completamente ambos, os circuítos auxiliares da placa.

### 4.1.1. Rexistros da CPU. Memoria interna do procesador

Todo o funcionamento do microprocesador levase a cabo sobre os seus rexistros internos, que constitúen a memoria máis rápida que existe no computador. O tamaño destes rexistros é o que determina a arquitectura do computador. Así, os antigos procesado res 8086 e 80286 eran microprocesadores de 16 bits, xa partir de aí fabricáronse procesadores de 32 bits, como os Pentium IV. Actualmente apareceron os procesadores de 64 bits, como, por exemplo, o Itanium de Intel e o Athlon 64 de AMD. Esta recente arquitectura de 64 bits estivo ligada ás contornas profesionais e aos microprocesadores RISC.

Os procesadores non adoitan operar case nunca directamente sobre a memoria principal. Collelos datos e operar directamente sobre a memoria principal sería o ideal pero tecnicamente sería moi caro de construír e seguramente moi lento. Resulta moito máis conveniente que os operandos sobre os cales vai actuar o procesador sexan traspasados previamente a unhas pequenas memorias auxiliares fixas e ultrarápidas. Estas memorias de altísima velocidade destinadas a memorizar os datos esenciais de cada instante do proceso chámanse rexistros, e están situados dentro do propio procesador. O tamaño dos

rexistros adoita ser o mesmo que o bus de datos.

En cada procesador concreto haberá un número determinado deles, segundo os usos específicos que se teñan previstos. Estes usos poden ser, por exemplo, os seguintes:

* Instrución que se estea executando (RI)
* Primeiro operando da instrución que se estea executando (OP1)
* Segundo operando da instrución que se estea executando (OP2)
* Resultado da operación (Acumulador)
* Información que será utilizada para calcular a dirección de memoria á que se quere acceder.
* Información sobre a situación na que quedou a CPU despois da execución dunha instrución (por exemplo: resultado da última comparación efectuada ) Información que permita controlar o funcionamento da CPU. (Rexistro de estado)
* Dirección da seguinte instrución a executar (é unha dirección de memoria) (CP)

Os rexistros constitúen o nexo entre a CPU e a memoria. Os operandos sobre os cales se aplica unha instrución deberán ser transportados previamente aos correspondentes rexistros desde a memoria principal e o resultado obtido no rexistro de resultado deberá transportarse á localización desexada da memoria principal.

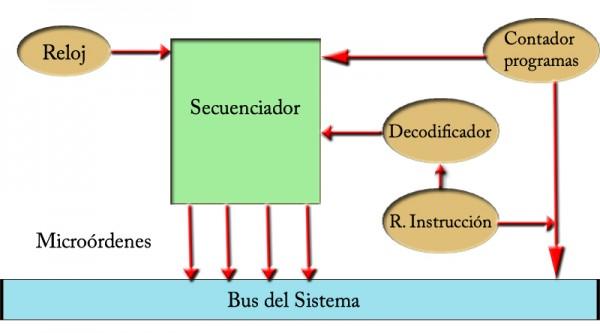
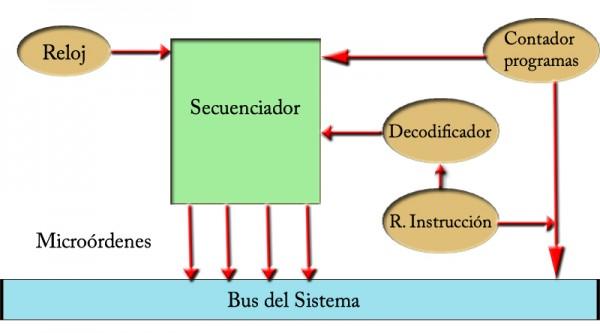
## 4.1.2. Unidade de control

A unidade de control dirixe todas as actividades do computador. Actúa como o corazón do sistema, enviando impulsos eléctricos (sinais de control) para secuenciar (pór en orde) e sincronizar (marcar o tempo) o funcionamento dos restantes compoñentes.

A súa función é buscar as instrucións na memoria principal, decodificalas (interpretación) e executalas, empregando para iso a unidade de proceso.

As funciones básicas da UC son tres:

* Analiza e interpreta instrucións do programa que se está executando.
* Controla os demais compoñentes físicos do ordenador mediante ordes dirixidas a estes compoñentes.
* Atende e decide sobre posibles interrupcións que se poden producir no proceso á unidade de control cada vez que se pulsa por exemplo unha tecla do teclado, a impresora tamén envía sinais á unidade de control, por exemplo cando queda sen papel.



Os compoñentes que a forman son os que se nomean a continuación:

* **Rexistro de instrución (RI):** É o encargado de almacenar a instrución que se está executando nese momento. Dita instrución soe estar composta por varias partes entre as que destacan o código de operación, e dicir que operación debe realizar a unidade de control e as direccións de memoria onde se encontran os operandos que pode precisar esa instrución.
* **Rexistro contador de programas (CP ou CI):** Tamén chamado contador de instrución, contén a dirección de memoria da seguinte instrución a executar.
* **Decodificador (D):** Encárgase de interpretar a instrución en curso ou a dirección de memoria, a que a Unidade de Control, debe acceder para ler ou escribir en nela. Analiza e emite os sinais necesarios o resto de elementos para a súa execución a través do secuenciador.
* **Secuenciador(S):** Xera unha serie de microórdenes o resto de elementos que, sincronizadas cos pulsos de reloxo, fai que se execute paso a paso a instrución que está aloxada no rexistro de instrución.
* **Reloxo (R):** indica en que momento debe comezar unha determinada operación e en que momento debe finalizar, para isto emite uns impulsos a intervalos fixos de tempo que serven como referencia o resto dos compoñentes.

Para marcar os intervalos precisos de tempo das actividades de todos os dispositivos que controla, a unidade de control dispón tamén cadenciómetro para lograr a sincronización. Nun microprocesador este labor lévaa a cabo o reloxo.

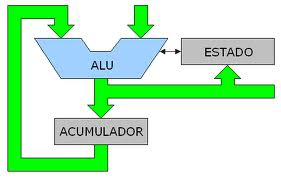
## 4.1.3. Unidade aritmético-lóxica

A unidade aritmético-lóxica é a parte do microprocesador que realiza os cálculos e as operacións cos datos indicados nas instrucións. Poderiamos chamala "calculadora" se unicamente realizase operacións aritméticas, pero como é capaz de realizar operacións de varios tipos é mellor chamala unidade aritmético-lóxica para destacar que é capaz de realizar operacións lóxicas.

Estas operacións lóxicas que a ALU pode realizar son, en realidade, a simple aplicación dunhas regras simples de comparación de dous datos. Poden establecerse comparacións de igual, maior que, menor que, menor ou igual que e maior ou igual que. Usando estes tipos de instrucións o computador poderá simular o comportamento lóxico humano e adoptar actuacións diferentes fronte a situacións diferentes. Pero isto non é intelixencia automática ou intelixencia artificial: debe de ser o programa realizado polo programador o que contemple convenientemente o uso destas instrucións para simular a capacidade de decisión que a intelixencia comporta.

* **Circuíto operacional (COP):** encargase da realización de operacións cos datos subministrados polos rexistros de entrada. Este circuíto dispón dunhas entradas (microórdenes) para seleccionar o tipo de operación a realizar en cada momento.
* **Rexistro de Entrada 1 e 2 (R.EN.A R.EN.B):** son utilizados tanto para almacenar os datos u operandos que interveñen nunha instrución antes da realización da operación por parte del COP como para o almacenamento dos resultados intermedios.
* **Rexistro de Estado:** no que queda constancia de algunhas condicións que se deron na última operación realizada e que se terá que ter en conta en outras operacións posteriores.
* **Rexistro Acumulador (RA):** nel se depositan os resultados das operacións levadas a cabo polo circuíto operacional.

O circuito combinacional (COP) proporciona a ALU a capacidade de realizar operacións aritméticas e lóxicas cos datos. O conxunto de operacións que pode levar a cabo a ALU están concibidas e fixadas durante o seu deseño. No caso dunha ALU elemental, esta poderá: sumar, restar, realizar operacións lóxicas (E, Ou , NON) e relacionais (=, >, <, >=, <=, <>). O resto de operacións realizaranse a partir das operacións elementais, por exemplo multiplicar dous números faríase mediante unha sucesión de sumas.



### 4.1.4 Comunicación interna e externa do procesador

Os microprocesadores dispoñen tamén de determinados buses de datos ou camiños polos que flúe a información. Estes adoitan ser dous: o bus interno, polo que se comunica a información dentro do propio microprocesador, e o bus externo, que se comunica co exterior.

O microprocesador controla todo o computador e utiliza buses (camiños) que permiten levar a cabo este control. Estes buses permiten comunicar o microprocesador co subsistema de memoria RAM e memoria caché, así como co resto de dispositivos.

O bus-procesador-memoria componse de tres partes: un bus de datos, un bus de direccións e un bus de control. Este bus é, hoxe en día, un dos principais pescozos de botella na arquitectura das placas base, debido a que os microprocesadores son cada vez máis rápidos, pero a memoria evoluciona máis lentamente.

### 4.1.5. Repertorio ou xogo de instrucións

A potencia dun microprocesador dependerá da súa velocidade na execución das instrucións pero tamén do tipo, variedade e riqueza de operacións que sexa capaz de realizar. Cada microprocesador dispón dun repertorio de instrucións propio, que se coñece como xogo de instrucións do microprocesador. Este xogo depende da circuitería interna coa cal foi deseñado. A programación íntima e verdadeira do procesador debe facerse usando ese xogo de instrucións da máquina concreta coa que esteamos traballando. Afortunadamente, existen linguaxes de programación que nos evitan ter que aprender a linguaxe de cada máquina particular.

Dentro do repertorio completo de instrucións do microprocesador existirá un conxunto asociado as distintas operacións que pode realizala ALU. Cando a UC identifica un código de instrución que indica unha determinada operación da ALU, xerará as sinais oportunas para activala operación correspondente dentro da ALU.

O formato xeral das instrucións que son executadas polo microprocesador axustase o formato:

*<Código de operador» <Operando 1> <Operando 2>*

Poden existir operacións que só conteñan un operador o incluso ningún.

#### 4.1.5.1 Procesadores RISC

RISC é unha filosofía de deseño de CPU para computadora que está a favor de conxuntos de instrucións pequenos e simples que toman menor tempo para executarse. O obxectivo de deseñar máquinas con esta arquitectura é posibilitar a [segmentación](https://es.wikipedia.org/wiki/Segmentaci%C3%B3n_de_cauce) e o [paralelismo](https://es.wikipedia.org/wiki/Procesador_superescalar) na execución de instrucións e reducir os accesos a memoria. As máquinas RISC protagonizan a tendencia actual de construción de microprocesadores.

Por isto a filosofía RISC foi crear instrucións pequenas, implicando que hai poucas, de aí o nome *conxunto reducido de instrucións*. O código é imprementado como series de esas instrucións simples, en vez dunha soa instrución complexa que dera o mesmo resultado. Isto fixo posible ter máis espacio dentro da instrución para transportar datos, resultando isto na necesidade de menos rexistros en la memoria. O mesmo tempo a interface coa memoria era considerablemente simple, permitindo ser optimizada.

Mais RISC tamén tiña súas desvantaxes. Debido a que unha serie de instrucións son necesarias para completar incluso as tarefas máis sinxelas, o número total de instrucións para a lectura da memoria é máis grande, e polo tanto leva máis tempo. O mesmo tempo non estaba claro onde habería o non unha ganancia neta no desempeño debido a esta limitación, houbo unha batalla case continua no mundo da prensa e do deseño sobre os conceptos de RISC.

#### 4.1.5.2 Procesadores CISC

Os microprocesadores CISC teñen un [conxunto de instrucións](https://es.wikipedia.org/wiki/Conjunto_de_instrucciones) que se caracteriza por ser moi amplo e permite operacións complexas entre operandos situados na memoria ou nos [rexistros](https://es.wikipedia.org/wiki/Registro_(hardware)) internos, en contraposición á arquitectura [RISC](https://es.wikipedia.org/wiki/Reduced_instruction_computing).

Este tipo de arquitectura dificulta o paralelismo entre instrucións, polo que, na actualidade, a maioría dos sistemas CISC de alto rendemento imcorporan un sistema que converte ditas instrucións complexas en varias instrucións simples do tipo RISC, chamadas xeralmente [microinstrucciones](https://es.wikipedia.org/w/index.php?title=Microinstrucci%C3%B3n&action=edit&redlink=1).

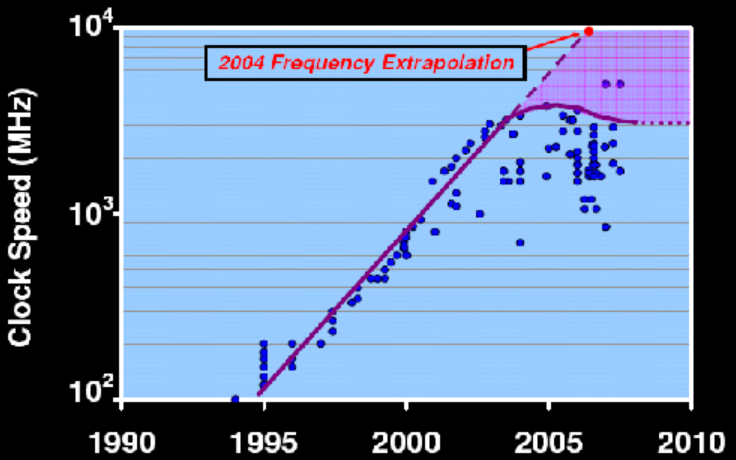
Os CISC pertencen á primeira corrente de construción de procesadores, antes do desenvolvemento dos RISC.

### 4.1.6 Melloras para a optimización do rendemento

A continuación, se describen algunhas características en canto o modo de funcionamento dos microprocesadores, que se introduciron en épocas máis modernas.

#### 4.1.6.1 Overclocking

Denomínase overclocking o feito de provocar que un microprocesador funcione a unha frecuencia maior que aquela para a que foi deseñado. Por exemplo, o configurar a placa base para que funcione a 2 Ghz, en vez dos 1,8 Ghz de fábrica. O primeiro efecto que se pode apreciar é un aumento considerable da temperatura, que poderíamos corrixir cun axeitado sistema de refrixeración. O overclocking tamén pode realizarse para que afecte o chipset, aínda é probable que aparezan erros de funcionamento sobre dispositivos conectados á placa, como tarxetas de expansión.



A idea básica é conseguir que o procesador realice máis operacións por segundo, forzando a que cada instrución se realice a un ritmo máis rápido. Ese ritmo se marca coa frecuencia de reloxo. O problema é que isto pode producir problemas de estabilidade, polo que non todos os procesadores e resto dos compoñentes o soportan.

#### 4.1.6.2 Multiproceso

O multiprocesamento ou multiproceso consiste na execución de programas ou procesos nun sistema con máis dun microprocesador. En teoría isto debería dobrar as capacidades en xeral: dobre velocidade e metade de tempo, pero na práctica non é así, aínda que é certo que se consegue un aumento considerable do rendemento. O multiproceso obtén todo seu rendemento cando dispoñemos dun software (un sistema operativo) especialmente deseñado para tal fin. O multiproceso pode ser de dous tipos, simétrico o asimétrico, en función de como o sistema operativo reparta a carga de traballo. O máis eficaz é o simétrico SMP (Symetric MultiProcessing, Multiprocesamento simétrico), onde a carga se reparte de maneira equitativa entre o número de microprocesadores utilizados.

#### 4.1.6.3 Hyperthreading

A nomenclatura Hyperthreading, tamén coñecido como Hyperthread, utilízase para designar a unha tecnoloxía desenvolvida por Intel. Un procesador que a icorpore será capaz de ofrecer o sistema operativo o dobre de [núcleos](http://computadoras.about.com/od/preguntas-frecuentes/a/Que-Es-El-Nucleo-De-Un-Procesador.htm) dos que ten fisicamente. Para iso Intel duplica certos bloques, en concreto algúns [rexistros](http://computadoras.about.com/od/conoce-procesadores/a/Como-Funcionan-Los-Registros-Del-Procesador.htm), pero sen chegar a crear un duplicado perfecto.

Os sistemas operativos e os programas verán dous núcleos onde só hai un. De esta forma, certas aplicacións, aquelas deseñadas para traballar con varios de eles o mesmo tempo, conseguirán aumentar seu rendemento.

Apareceu por primeira vez nun Pentium IV pero deixouse de usar o non conseguir as melloras de rendemento que Intel esperaba. Nalgúns sistemas, Windows 2000 e anteriores, podería incluso chegar a enlentecer teu sistema. Mais esta tecnoloxía tivo un renacemento coa aparición dos Intel Core. Sendo case a diferencia cualitativa máis importante entre un Core i7 e un i5.

Como ocorre cando aumentamos el número de núcleos, as melloras dependen moito de como foran programadas as aplicacións. En todo caso, é mellor ter un i7 con 4 núcleos reais que o mesmo con 2 e Hyperthread.

Hyperthread se centra sobre todo en intentar paliar o máximo os problemas de rendemento os que levan os bloqueos. Estes se producen cando unha aplicación, ou parte de ela, queda parada á espera de datos que teñen que ser procesados. Para outro tipo de programas máis específicos, totalmente deseñados para sacar partido a un maior número de núcleos, esta tecnoloxía non é tan efectiva.

AMD ten algo parecido e incorpora nos seus procesadores baseados en [Bulldozer](http://computadoras.about.com/od/Procesador/a/Que-Es-Amd-Bulldozer.htm) e [Piledriver](http://computadoras.about.com/od/Procesador/a/Que-Es-Amd-Trinity-Y-Piledriver.htm) a tecnoloxía [CMT](http://computadoras.about.com/od/preguntas-frecuentes/a/En-Que-Consiste-La-Tecnologia-Cmt-De-Amd.htm). Neste caso, o fabricante, crea bloques con dous núcleos no seu interior. Ambos comparten a unidade de punto flotante pero non os outros elementos.

A diferencia é que Intel só duplica algúns rexistros para ser capaz de eliminar bloqueos e AMD trata de acelerar tódalas aplicacións multinúcleo.

#### 4.1.6.4. A segmentación de canle ou pipeline.

A **segmentación** (en inglés *pipelining*, literalmente 'cano' ou 'cañería') é un método polo cal se consegue aumentar o rendemento de algúns sistemas electrónicos dixitais. Usase principalmente nos microprocesadores.

Máis coñecida como pipeline, é unha técnica empregada no deseño de procesadores, baseada na división da execución das instrucións en etapas, conseguindo así que unha instrución empece a executarse antes de que terminasen as anteriores e, por tanto, que haxa varias instrucións procesándose simultaneamente.



Figura. Como se conforma a segmentación dunha instrución

A segmentación consiste en descompoñer a execución de cada instrución en varias etapas para poder empezar a procesar unha instrución diferente en cada unha delas e traballar con varias á vez.

No caso do procesador [DLX](https://es.wikipedia.org/wiki/DLX) podemos encontrar as seguintes etapas nunha instrución:

* IF (*instruction fetch*): procura.
* ID (*instruction decode*): decodificación.
* EX (*execution*): execución na unidade aritmético lóxica.
* MEM (*memory*): memoria.
* WB (*writeback*): escritura.

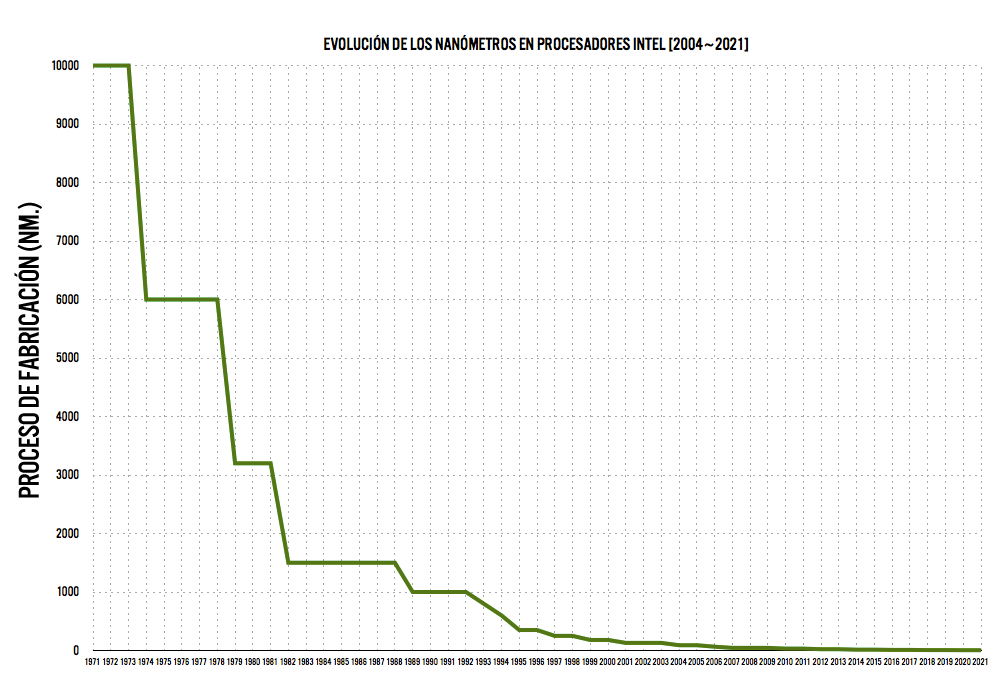
Cada unha destas etapas da instrución usa en exclusiva un hardware determinado do procesador, de tal forma que a execución de cada unha das etapas en principio non interfire na execución do resto.

No caso de que o procesador non puidese executar as instrucións en etapas segmentadas, a execución da seguinte instrución só se podería levar a cabo tras a finalización da primeira. En cambio nun procesador segmentado, salvo excepcións de dependencias de datos ou uso de unidades funcionais, a seguinte instrución podería iniciar súa execución tras acabar a primeira etapa da instrución actual.

### 4.1.7 Fabricación dos microprocesadores

Para a fabricación dos chips utilízanse materiais moi específicos: os denominados semicondutores, materiais que só conducen a electricidade baixo determinadas condicións. De entre todos eles, o máis empregado é o silicio, así como o arseniuro de galio (GaAs).

O silicio cristalízase en grandes obleas ou cristais que, posteriormente, divídense en seccións rectangulares máis pequenas, das que finalmente se obterán os chips. Despois, durante unha fase posterior, implantar a lóxica que o dotará das funcións de cálculo. Os microprocesadores fabrícanse mediante un proceso denominado fotolitografía. A seguinte fase é a do encapsulado, que consiste en ensamblar o diminuto chip nunha placa de circuíto impreso cos diferentes contactos cos que logo se inserirá na placa base. Finalmente, pásase á fase de proba e de calibración de velocidade.

A técnica actual permitiu crear transistores cada vez máis pequenos, ata os 14 nanómetros. Estes foron precedidos polos 22, e antes por 32. e 45, 65, 90... así ata los **10.000 nanómetros** (10 micrómetros) cos que empezaron os primeiros chips Intel do mercado, os 4004 e 8008 dos años 71 e 72, os que comentaron a carreira da arquitectura x86.

Isto ten como consecuencia unha **mellora global no rendemento**. Noutras palabras: os novos chips son máis rápidos. En que grado? É difícil de cuantificar, aínda que Intel asegura que entre los 45 nanómetros de 2007 e los 14 de 2014 conseguiron duplicar o rendemento. Entre eles, catro xeracións e unha mellora cércana a ser liñal.

Esta mellora no rendemento chega acompañada de outro factor determinante no mundo dos ordenadores, seu consumo enerxético é menor. A súa vez, o menor consumo enerxético implica unha maior potencia por vatio, que é o *ratio* que se utiliza en moitas ocasións á hora de comparar varias xeracións.

### 4.1.8 Potencia e refrixeración dos procesadores

A tensión de alimentación que utiliza un microprocesador é importante, xa que a maior voltaxe, maior cantidade de calor xera. Para que os microprocesadores estean perfectamente refrixerados e disipen a calor de maneira eficaz, e necesario dotalos de refrixeradores, xa que un sobrequencemento pode provocar un funcionamento inestable, bloqueos, erros de cálculo e, en casos extremos, que acabe queimado el propio chip. O método de refrixeración consiste nunha peza metálica, xeralmente de aluminio, colocada sobre o chip, que pode disipar a calor xerada polo mesmo. Esta peza ten unha serie de láminas que axudan neste proceso. Estes disipadores soen ir pegados directamente o chip, ben cunha cola especial, ben usando unha pasta termocondutora.

Os disipadores activos dispoñen ademais dun ventilador que introduce aire, que posteriormente, é expulsado pola parte lateral. O encapsulado é o envoltorio dun microprocesador, que o fai máis resistente o uso e, o mesmo tempo, permite unha conexión sinxela dos seus contactos a placa base. Estes últimos envoltorios se preparan en material plástico ou metálico, dado que ofrece menos peso e unha maior disipación da calor.

Os procesadores utilizan un índice para indicar a potencia que consume este, que se denomina TDP (acrónimo de Thermal Design Power). O TPD se mide normalmente en Watios. Serve de indicación para os fabricantes e montadores de computadoras, e desta forma saben como deseñar a refrixeración que deben usar.

## 4.2. BUSES DO SISTEMA

O **bus** (ou canal) é un sistema dixital que transfire datos entre os compoñentes dunha computadora ou entre varias computadoras. Está formado por cables ou pistas nun circuíto impreso, dispositivos como resistores e condensadores ademais de circuítos integrados.

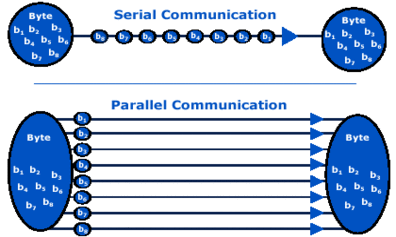
Nos primeiros computadores electrónicos, todos os buses eran de tipo paralelo, de maneira que a comunicación entre as partes do computador facíase por cintas ou moitas pistas no circuíto impreso, nos cales cada condutor ten unha función fixa e a conexión é sinxela requirindo unicamente portos de entrada e de saída para cada dispositivo.

A tendencia nos últimos años facía o uso de buses serie como o [USB](https://es.wikipedia.org/wiki/Universal_Serial_Bus), [Firewire](https://es.wikipedia.org/wiki/IEEE_1394) para comunicación con periféricos, reemprazando os buses paralelos, incluíndo o caso do microprocesador co [chipset](https://es.wikipedia.org/wiki/Chipset) na placa base, a pesar de que o bus serie posúe unha lóxica complexa (requirindo maior poder de cómputo que o bus paralelo) prodúcese a cambio de velocidades e eficacias maiores.

Existen diversas especificacións de buses, nas que se definen un conxunto de características mecánicas como conectores, cables e tarxetas, ademais de protocolos eléctricos e de sinais.

### 4.2.1. Tipos de buses

Existen dous grandes tipos clasificados polo el método de envío da información: **bus paralelo** e **bus serie**. Hai diferencias nas prestacións e ata fai uns anos se consideraba que o uso apropiado dependía da lonxitude física da conexión: para curta distancia o bus paralelo, para longa o serie.



### Bus paralelo

É un bus no cal os datos son enviados por bytes o mesmo tempo, coa axuda de varias liñas que teñen funcións fixas. A cantidade de datos enviada é bastante grande cunha frecuencia moderada é igual o ancho dos datos pola frecuencia de funcionamento. Nos computadores foi usado de maneira intensiva, dende o bus do procesador, os buses de discos duros, tarxetas de expansión e de vídeo, ata as impresoras.

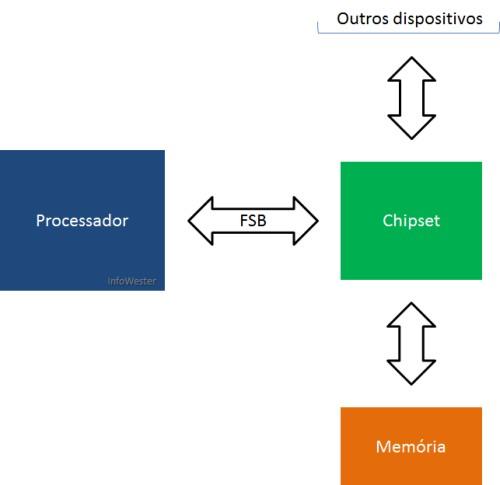
### Bus serie

Neste tipo de bus, os datos son enviados, bit a bit e se reconstrúen por medio de rexistros ou rutinas. Está formado por poucos condutores e seu ancho de banda depende da frecuencia. Neste momento é usado en buses para discos duros, unidades de estado sólido, tarxetas de expansión e para o bus do procesador.

### 4.2.2. Bus interno de la placa base (FSB, Front Side Bus)

O front-side bus, tamén coñecido polo seu acrónimo FSB (do inglés literalmente "bus da parte frontal"), é o tipo de bus usado como bus principal nalgúns dos antigos microprocesadores da marca Intel para comunicarse co circuíto integrado auxiliar ou chipset. Ese bus inclúe sinais de datos, direccións e control, así como sinais de reloxo que sincronizan o seu funcionamento. Nos novos procesadores de Intel, desde Nehalem, e hai tempo nos de AMD úsanse outros tipos de buses como o Intel QuickPath Interconnect e o HyperTransport respectivamente.

O FSB empezou a formar parte da arquitectura de computadoras estándar dende que as aplicacións requiren máis memoria da que o procesador podería reter.



Os máis modernos FSB utilízanse a modo de conexión exclusiva principal entre a unidade central de procesamento e o circuíto integrado auxiliar (chipset). Este (xeralmente composto polo traballo en conxunto da ponte norte ou northbridge e a ponte sur ou southbridge) é o encargado de interconectar o resto de buses do sistema. Os buses como PCI, PCI Express, e buses de memoria comunícanse co chipset para permitir o correcto fluxo de datos entre os diferentes dispositivos. Xeralmente estes buses secundarios funcionan a unha velocidade derivada da velocidade do FSB.

A pesar da solución que se lle deu ao problema, sempre se pensou en que o FSB debería ser unha tecnoloxía con tendencia a desaparecer. Empresas como AMD sempre criticaron o FSB, xa que limita moito as capacidades reais dun sistema xerando moita latencia e un tempo de resposta maior, creando un auténtico pescozo de botella para o resto de dispositivos. Non foi até 2001 e a aparición da tecnoloxía HyperTransport cando se puido deseñar unha tecnoloxía capaz de substituír o uso do FSB. Actualmente empresas fabricantes de chipsets como NVIDIA, Silicon Integrated Systems ou VIA Technologies, xa comezaron a eliminar o uso do FSB substituíndoo coa versión 3.0 de HyperTransport.

O [*front-side bus*](https://es.wikipedia.org/wiki/Front-side_bus) dos procesadores Intel é un bus deste tipo e como calquera bus presenta unhas funcións en liñas dedicadas:

* As **liñas de dirección** so as encargadas de indicar a posición de memoria ou o dispositivo co que se desexa establecer comunicación.
* As **liñas de control** son as encargadas de enviar sinais de arbitraxe entre os dispositivos. Entre as máis importantes están as liñas de interrupción, DMA y os indicadores de estado.
* As **liñas de datos** transmiten os bits de forma aleatoria de maneira que polo xeral un bus ten un ancho que é potencia de 2.

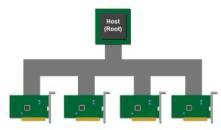
### 4.2.3 Outros tipos de buses

A continuación, descríbense os tipos de bus mais habituais, enumerados por orden de aparición en no mercado:

#### 4.2.3.6 Bus PCI.

Consiste nun bus paralelo de ordenador estándar utilizad para conectar dispositivos periféricos directamente a súa placa base. Estes dispositivos poden ser circuítos integrados insertados nesta ou tarxetas de expansión que se axustan en conectores. O bus PCI permite a configuración dinámica dun dispositivo periférico. No tempo de arranque del sistema, as tarxetas PCI e a BIOS interactúan e negocian los recursos solicitados pola tarxeta PCI. Isto permite asignación de las [IRQ](https://es.wikipedia.org/wiki/IRQ) e direccións do porto por medio dun proceso dinámico.

O bus PCI ofrece un ancho de funcionamento desde 32 até 64 bits, e unha taxa de transferencia de 133 MB/s (32 bits a 33 MHz) até 533 MB/s (64 bits a 66 Mhz). Existe algunha variante deste bus que permite maior velocidade de transferencia; PCI-X, con 4,3 GB/s como máximo.



Unha característica fundamental da placa base e que incorpore este tipo de bus é a posibilidade de configurar automaticamente os diferentes dispositivos que se conectan a ela, a condición de que dispoñamos dun sistema operativo adecuado. É o que se coñece co nome de Plug and Play (PnP), que permite ao procesador conectar calquera compoñente, recoñecelo e incluílo na súa configuración básica.

Os conectores destes buses son máis pequenos que os buses ISA e EISA, xeralmente son de cor branca e neles insírense tarxetas de 32 bits.

**4.2.3.8 Bus USB**.

O Universal Serial Bus (bus universal en serie) pode conectar periféricos como ratos, teclados, escáneres, cámaras dixitais, teléfonos móbiles, reprodutores multimedia, impresoras, discos duros externos, tarxetas de son, sistemas de adquisición de datos e compoñentes de rede.

O USB, como estándar inclúe a transmisión de enerxía eléctrica ao dispositivo conectado para dispositivos de pequeno consumo.

Os dispositivos USB clasifícanse en catro tipos segundo a súa velocidade de transferencia de datos:

* Baixa velocidade (1.0): Taxa de transferencia de até 1,5 Mbps (192 KB/s). Utilizado no seu maior parte por dispositivos de interface humana (Human Interface Device, en inglés) como os teclados, os ratos (mouse), as cámaras web, etc.
* Velocidade completa (1.1): Taxa de transferencia de até 12 Mbps (1,5 MB/s) segundo este estándar. Esta foi a máis rápida antes da especificación USB 2.0, e moitos dispositivos fabricados na actualidade traballan a esta velocidade. Estes dispositivos dividen o ancho de banda da conexión USB entre eles, baseados nun algoritmo de impedancias LIFO.
* Alta velocidade (2.0): Taxa de transferencia de até 480 Mbps (60 MB/s) pero polo xeral de até 125Mbps (16MB/s). Está presente case no 99% dos PC actuais. O cable USB 2.0 dispón de catro liñas, un par para datos, unha de corrente e un cuarto que é o negativo ou retorno.
* Super alta velocidade (3.0): Ten unha taxa de transferencia de até 4.8 Gbps (600 MB/s). A velocidade do bus é dez veces máis rápida que a do USB 2.0, debido a que incluíron 5 conectores adicionais, e é compatible cos estándares anteriores.
* Usb 3.1: Substitúe o estándar 3.0. 10 Gb/s. Se definen dous modos de transferencia, un baseado no estándar existente 3.0 denominado USB 3.1 Gen 1 (SuperSpeed) de 5 Gbit/s (450 MB/s). Inclúe a nova especificación USB 3.1 Gen 2 (SuperSpeed+) de 10 Gbit/s (1250MB/s). É compatible con USB 3.0 e USB 2.0.
* USB 3.2. é unha "actualización incremental" que trae como principal novidade a posibilidade de aproveitar dúas pistas de 5 o 10 Gbps para acadar **velocidades de transferencia máximas de ata 20 Gbps** nos futuros dispositivos con conetores USB-C (USB Tipo C).

### 

|  |  |  |  |
| --- | --- | --- | --- |
| **Modo** | **Abrev** | **Taxa de tranferencia** | **Versión** |
| Low Speed | LS | 1.5 Mbit/s (187.5 KB/s) | USB 1.0 |
| Full Speed | FS | 12 Mbit/s (1.5 MB/s) | USB 1.0 |
| High Speed  Also, Hi-Speed | HS | 480 Mbit/s (60 MB/s) | USB 2.0 |
| SuperSpeed | SS | 5 Gbit/s (625 MB/s) | USB 3.0 |
| SuperSpeed+ | SS+ | 10 Gbit/s (1.25 GB/s) | USB 3.1 |
| SuperSpeed+ | SS+ | 20 Gbit/s (2.5 GB/s) | USB 3.2 |

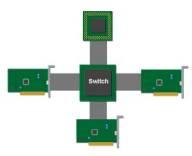
A conexión entre dispositivos realizase mediante un par trenzado en modo half-duplex nas versións de USB 2.0 ou menores. A partir da 3.0 existen dous pares que permiten a comunicación full-duplex entre dispositivos.

USB ten unha lonxitude máxima de conexión de 5 m e USB 3.0 permite unha distancia de 3m.

El estándar USB 1.x e 2.0 proporciona 5 V de alimentación e unha conexión para los dispositivos USB conectados. A unidade de carga se definiu con 100 mA en USB 1.x e 2.0, e 150 mA e USB 3.0. Un dispositivo pode consumir un máximo de cinco unidades de carga (500 mA) dende un porto USB 1.x e 2.0, ou 6 unidades de carga (900 mA) en USB 3.0.

**4.2.3.11 PCI-Express**

O bus PCI Express é un bus serie de alta velocidade que substitúe o antigo bus PCI/PCI-X. Unha das diferencias clave entre o PCI Express e o vello PCI é a topoloxía. PCI usa unha arquitectura de de bus paralelo compartido, no que o controlador PCI e todos os dispositivos comparten un conxunto de liñas de direccións, datos e control. Polo contrario, PCI Express está baseada nunha topoloxía punto-a-punto, con enlaces serie full-duplex, con cada dispositivo ata o concentrador do chipset. A comunicación do PCI Express é encapsulada en paquetes.

O Bus PCI Express soporta comunicación full-duplex entre dous puntos finais, sen limitar outros enlaces simultáneos entre outros dispositivos conectados. No mundo de PCI Express, o bus deixa de ser o recurso compartido para pasar a selo o switch. Cada dispositivo no sistema ten un acceso directo e exclusivo ao switch, a conexión de cada dispositivo co switch denomínase ligazón (link).

Cada slot de expansión leva un, dous, catro, oito, dezaseis ou trinta e dúas ligazóns de datos entre a placa base e as tarxetas conectados. O número de ligazóns escríbese cunha x de prefixo (x1 para unha ligazón simple e x16 para unha tarxeta con dezaseis enlaces. Trinta e dúas ligazóns de 250MB/s dan o máximo ancho de banda, 8 GB/s (250 MB/s x 32) en cada dirección para PCIE 1.1. No uso máis común (x16) proporcionan un ancho de banda de 4 GB/s (250 MB/s x 16) en cada dirección. En comparación con outros buses, unha ligazón simple é aproximadamente o dobre de rápido que o PCI normal, un slot de catro enlaces, ten un ancho de banda comparable á versión máis rápida de PCI-X 1.0, e oito enlaces teñen un ancho de banda comparable á versión máis rápida de AGP.

Slots PCI Express (de arriba a abaixo: x4, x16, x1 e x16), comparado cun tradicional PCI de 32 bits, tal como ven na placa DFI LanParty nF4 Ultra-D (imaxe da dereita). 

PCI-Express está pensado para ser usado só como bus local, aínda que existen extensores capaces de conectar múltiples placas base mediante cables de cobre ou mesmo fibra óptica. Debido a que se basea no bus PCI, as tarxetas actuais poden ser reconvertidas a PCI-Express cambiando soamente a capa física. A velocidade superior do PCI-Express permitirá substituír case todos os demais buses, AGP e PCI incluídos. A idea de Intel é ter un só controlador PCI-Express comunicándose con todos os dispositivos, no canto de con o actual sistema de ponte norte e ponte sur.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Versión de PCI Express** | **Código en liña** | **Velocidade de transferencia** | **Ancho de banda** | |
| **Por carril** | **En 16x** |
| 1.0 | 8b/10b | 2,5 GT/s | 2 Gbit/s (250 MB/s) | 32 Gbit/s (4 GB/s) |
| 2.0 | 8b/10b | 5 GT/s | 4 Gbit/s (500 MB/s) | 64 Gbit/s (8 GB/s) |
| 3.0 | 128b/130b | 8 GT/s | 7,877 Gbit/s (984,6 MB/s) | 126,032 Gbit/s (15,754 GB/s) |
| 4.0 | 128b/130b | 16 GT/s | 15,754 Gbit/s (1969,2 MB/s) | 252,064 Gbit/s (31,508 GB/s) |
| 5.0 | 128b/130b | 32 GT/s | 31,504 Gbit/s (3938 MB/s) | 504 Gbit/s (63.0 GB/s) |
| 6.0 (2021) | 128/130b | 64.0 GT/s | 63,016 Gbit/s (7877 MB/s) | 1008,24 Gbit/s(126.03 GB/s) |
| 7.0 (2025 |  | 128 GT/s | 121 Gbits/s | 1936 Gbits/s (242 GB/s) |

O Slot PCIE é usado como conector de tarxetas gráficas.

PCI-Express é abreviado como PCIE ou PCIX. Con todo, non ten nada que ver con PCI-X. PCI-X é unha evolución de PCI, na que se consegue aumentar o ancho de banda mediante o incremento da frecuencia, chegando a ser 32 veces máis rápido que o PCI 2.1. A súa velocidade é maior que PCI-Express, pero presenta o inconveniente de que ao instalar máis dun dispositivo, a frecuencia base redúcese e perde velocidade de transmisión.

**4.2.3.12 HyperTransport (HT).**

A tecnoloxía HyperTransport é unha conexión punto a punto de alta velocidade e baixa latencia, deseñada para aumentar a velocidade das comunicacións entre os circuítos integrados en computadoras, servidores, sistemas integrados e equipos de redes e telecomunicacións até en 48 veces máis que os sistemas existentes.

A tecnoloxía HyperTransport foi deseñado para reducir o número de buses nun sistema, o que pode diminuír os funiles e posibilitar que os microprocesadores máis rápidos da actualidade utilicen a memoria de maneira máis eficiente en sistemas máis sofisticados.

O desenvolvemento de HyperTransport fíxose sobre a base de querer eliminar o Front Side Bus (FSB). Non foi até a versión 3.0 cando varios fabricantes de chipsets decidiron utilizar HyperTransport para substituír o FSB con excelentes resultados. Esta foi o seu imprementación máis famosa.

Tamén deu grandes resultados noutras implantacións, tales como interconexiones entre microprocesadores MIPS, servidores, sistemas informáticos de alto rendemento, e en routers e switches.

HyperTransport é unha tecnoloxía escalable de interconexión punto a punto que achega unha ligazón de alta velocidade e elevado rendemento. Esta conexión universal reduce o número de buses a instalar nun sistema e permite aos chips presentes nos equipos informáticos, de networking ou de comunicacións comunicarse entre si até 48 veces máis rápido que as tecnoloxías do momento. Ademais, era compatible co moi difundido PCI, polo que non é necesario perder tempo e diñeiro en reescribir o código do software baseado neste estándar.

HyperTransport baséase nunha arquitectura escalable que incrementa enormemente a capacidade de proceso das transaccións de bus sobre as tecnoloxías E/S existentes, como PCI, PCI-X e AGP. Aínda que as arquitecturas de bus E/S herdadas utilízanse amplamente debido ao seu baixo prezo e facilidade de imprementación, alcanzan o seu punto máximo a ao redor dos 66 MHz. Con todo, os procesadores de hoxe corren a frecuencias de reloxo de 500 MHz a máis de 1 GHz, e xa están en camiño outros máis rápidos. Por sorte, a arquitectura de bus E/S de HyperTransport escala desde configuracións modestas con relativamente baixas velocidades de reloxo (200 MHz) a velocidades de reloxo de 800 MHz a máis de 32 bits.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Versión HyperTransport** | **Ano** | **Max. Frecuencia HT** | **Max. Ancho enlace** | **Max. Ancho de banda agregado (bidireccional)** | **Max. Ancho de banda a 16-Bit (unidireccional)** | **Max. Ancho de banda a 32-Bit (unidireccional)\*** |
| **1.0** | 2001 | 800 MHz | 32 Bit | 12.8 GB/s | 3.2 GB/s | 6.4 GB/s |
| **1.1** | 2002 | 800 MHz | 32 Bit | 12.8 GB/s | 3.2 GB/s | 6.4 GB/s |
| **2.0** | 2004 | 1.4 GHz | 32 Bit | 22.4 GB/s | 5.6 GB/s | 11.2 GB/s |
| **3.0** | 2006 | 2.6 GHz | 32 Bit | 41.6 GB/s | 10.4 GB/s | 20.8 GB/s |
| **3.1** | 2008 | 3.2 GHz | 32 Bit | 51.2 GB/s | 12.8 GB/s | 25.6 GB/s |

Con HyperTransport, pódese utilizar un bus estándar de 2, 4, 8, 16 e 32 bits para render funcións de bus E/S para aplicacións específicas. A súa flexibilidade inherente inclúe anchos de bus asimétricos a fin de soportar anchos de banda ascendentes e descendentes diferentes.

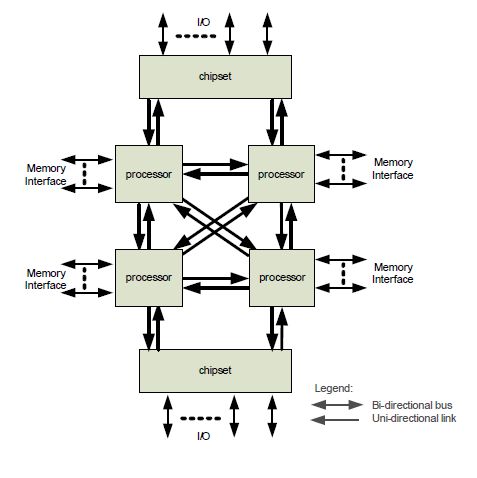
Unha opción para conseguir máis ancho de banda é adoptar PCI-X, que opera a 133 MHz. Pero PCI-X conta coa limitación de que só pode soportar un dispositivo periférico, mentres que o tradicional PCI admite catro e HyperTransport até 31! Ademais, cun conmutador é capaz de soportar mesmo máis dispositivos, creando unha enorme “fábrica” HyperTransport (túneles que actúan como bloques de E/S), preservando relativamente os custos do hardware e os sistemas.

**4.2.3.13 QuickPath.**

Intel QuickPath Interconnect (QPI) é unha interconexión punto a punto para procesadores desenvolvida por Intel para substituír ao front-side bus (FSB). QPI será substituído polo Intel UltraPath Interconect (UPI) nos futuros procesadores Skylake EX/EP Xeon baseados no socket [LGA 3647](https://en.wikipedia.org/wiki/LGA_3647).

O QPI é un elemento dun sistema de arquitectura que Intel chama ''QuickPath architecture'' que incorpora a chamada ''QuickPath technology''. Tal como o HyperTransport de AMD, a arquitectura QuickPath Architecture asume que o procesador ten un controlador de memoria integrado, obrigando así aos multiprocesadores a usar unha arquitectura NUMA.

Cada QPI consta de 2 conexións punto a punto de 20-bit, unha para cada dirección, para un total de 42 sinais. Cada sinal é un par diferencial, formando así un número de 84.



O ''QuickPath'' reporta velocidades de 4,8 a 6,4 MT/s|GT/s por segundo por dirección. O ancho de banda vai de 12,0 a 16,0 GB/s por dirección, ou 24,0 a 32,0 GB/s por conexión.

A taxa de transferencia se calcula da seguinte forma:

3.2 GHz

× 2 bits/Hz (double data rate)

× 16(20) (bits de datos/QPI ancho del enlace)

× 2 (envío e recepción unidireccional operando simultaneamente(full duplex))

÷ 8 (bits/byte)

= 25.6 GB/s

A implementación inicial no Nehalem usa unha conexión de 25,6 GB/s a 20-bit. Esta conexión prové exactamente o dobre do ancho de banda teórico dun FSB de Intel a 1600 MHz (usados no chipset Intel X48).

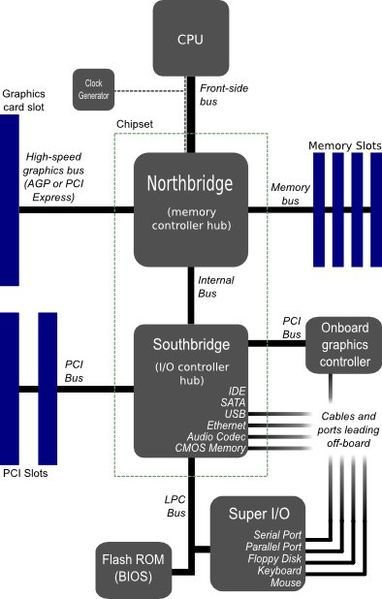
### 4.2.4. Circuitos integrados auxiliares (Chipset).

Se analizamos a arquitectura dun computador, podemos ver que existen varias canles polos que flúe a información e o centro neurálxico que é o microprocesador. As canles principais polos que adoita circular información nunha placa base son:

* Entre o microprocesador e a memoria RAM.
* Entre o microprocesador e a tarxeta gráfica
* Entre o microprocesador e as ranuras de expansión ás que adoitan estar conectadas as tarxetas (PCI, PCI Express, etcétera).
* Entre o microprocesador e os dispositivos de almacenamento.

O elemento que integra todas estas funcións de control, ademais de moitas outras, é o denominado chipset. A velocidade con que se desprazan os datos no interior dun computador está directamente relacionada con este compoñente. Tamén se pode definir o chipset como un conxunto de chips, é dicir, un grupo de circuítos integrados (xeralmente dous) que cumpren unha serie de funcións para o funcionamento da placa base:

* Soporte para o microprocesador.
* Unidade de control de memoria (MMU, Memory Management Unit), memoria RAM e memoria caché (arquitecturas de ponte norte-sur).
* Controlador IDE/ATA, SATA para discos duros e outros dispositivos de almacenamento.
* Control de periféricos e do bus de E/S.
* Controlador de interrupcións programable (PIC, Programmable Interrupt Controller).
* Reloxo de tempo real (RTC, Real Time Clock).
* Soporte para a xestión de enerxía.
* Controlador de acceso directo a memoria (DMA, Direct Memory Access).
* Controlador de infravermellos, de teclado e do porto PS/2, etcétera.

Circuíto integrado auxiliar ou chipset é o conxunto de circuítos integrados deseñados con base á arquitectura dun procesador (nalgúns casos deseñados como parte integral desa arquitectura), permitindo que ese tipo de procesadores funcionen nunha placa base. Serven de ponte de comunicación co resto de compoñentes da placa, as tarxetas de expansión, os portos USB, rato, teclado, etc.

As placas base antigas adoitan incluír dous integrados, denominados Norte e Sur, e adoitan ser os circuítos integrados máis grandes despois da GPU e o microprocesador. As últimas placas base carecen de Ponte Norte xa que os procesadores de última xeración lévano integrado.

O chipset determina moitas das características dunha placa base e polo xeral a referencia da mesma está relacionada coa do chipset.

O conxunto ou xogo de chips ou Chipset é un conxunto de circuítos integrados que se encarga de realizar as funcións que o microprocesador delega neles. Chipset traducido literalmente do inglés significa conxunto de circuítos integrados. Desígnase circuíto integrado auxiliar ao circuíto integrado que é periférico a un sistema pero necesario para o funcionamento do mesmo. A maioría dos sistemas necesitan máis dun circuíto integrado auxiliar.

A terminoloxía dos integrados cambiou desde que se creou o concepto do chipset a principio dos anos 90, pero aínda existe equivalencia facendo algunhas aclaracións:

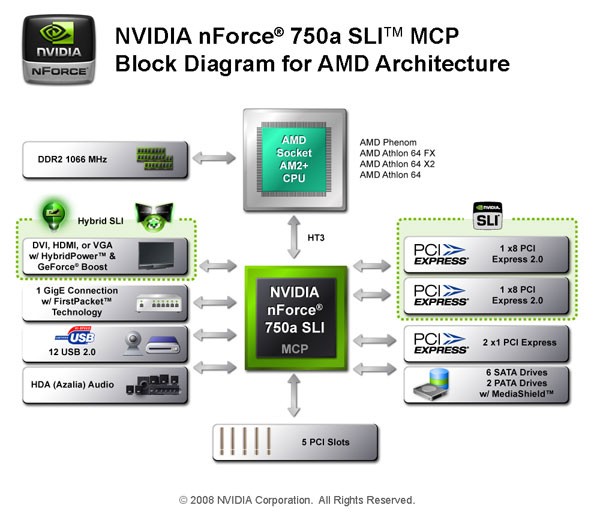
* O NorthBridge, ponte norte, MCH (Memory Controller Hub), GMCH (Graphic MCH), úsase como ponte de ligazón entre o microprocesador e a memoria. Controla as funcións de acceso cara a e entre o microprocesador, a memoria RAM, o porto gráfico AGP ou o PCI-Express de gráficos, e as comunicacións coa ponte sur. Ao principio tiña tamén o control de PCI, pero esa funcionalidade pasou á ponte sur.
* O SouthBridge ou ponte sur, ICH (Input Controller Hub), controla os dispositivos asociados como son a controladora de discos IDE, portos USB, FireWire, SATA, RAID, ranuras PCI, ranura AMR, ranura CNR, portos infravermellos, disqueteira, LAN, PCI-Express 1x e unha longa lista de todos os elementos que podamos imaxinar integrados na placa nai. É o encargado de comunicar o procesador co resto dos periféricos.

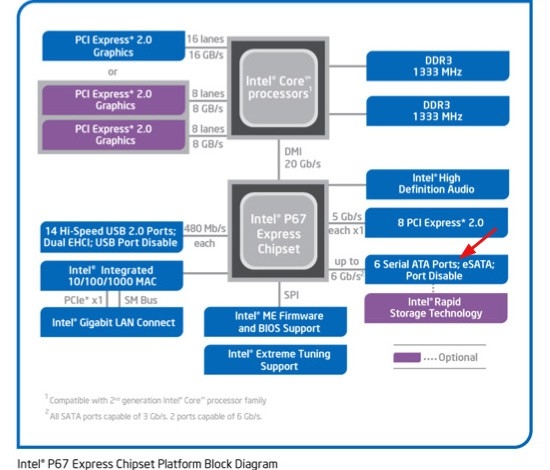
### 4.2.5 Evolución da arquitectura interna do computador

A evolución da arquitectura interna dun computador se pode apreciar estudando a evolución das funcións que foi perdendo o Chipset e que foi engadíndose o procesador.

Para comezar este estudo desta evolución comezaremos coa figura do apartado anterior, onde podemos apreciar unha arquitectura con ponte norte e sur. Nesta arquitectura a memoria e a tarxeta gráfica se conectaban co procesador a través da ponte norte. Esta a súa vez tamén permitía o acceso do resto dos dispositivos concentrados a través da ponte sur.

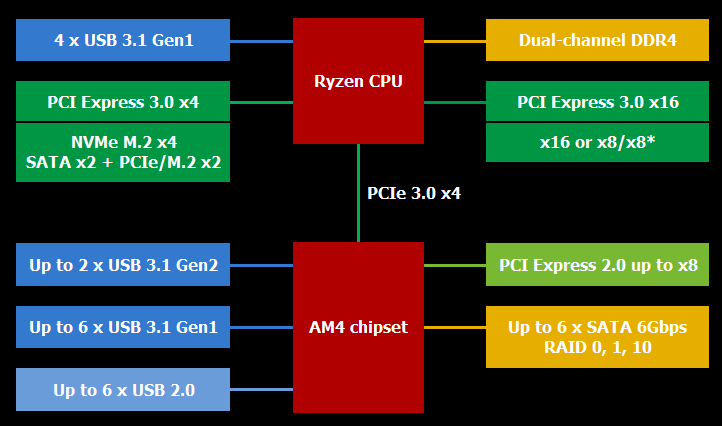
Isto evolucionou nos procesadores AMD que incorporaron a MMU no microprocesador polo que a ponte norte só conectaba a tarxeta gráfica ou, como no caso da figura seguinte, unha única ponte comunicaba a tarxeta gráfica e o resto dos periféricos.





Finalmente a comunicación coa tarxeta gráfica e memoria foi asumida plenamente polo procesador incorporando a MMU e controladores do bus PCIe. Quedando un só chip para conectar o resto dos dispositivos. Na figura anterior se pode apreciar o diagrama do chipset Intel P67 que ilustra isto.

Na seguinte figura pode apreciarse como o microprocesador Ryzen de AMD vai incorporando novos controladores e alixeirando as funcións do chipset.



## 4.3 Xestión de enerxía.

A xestión e o aforro de enerxía nas placas base e noutros compoñentes, como, por exemplo, o microprocesador ou o disco duro, foi unha das funcións que máis se desenvolveron nos últimos tempos. Estas funcións coñécense xenericamente coas siglas SMM (System Management Mode, Modo de administración do sistema).

Os circuítos de xestión de enerxía intégranse no chipset e, xeralmente, a través da BIOS, configúranse os tempos nos que devanditos compoñentes reducen a súa actividade en maior ou menor nivel. Os estándares principais son APM (Advanced Power Management, Administración avanzada de enerxía) e ACPI (Advanced Configuration and Power Interface, Interface avanzada de configuración e enerxía).

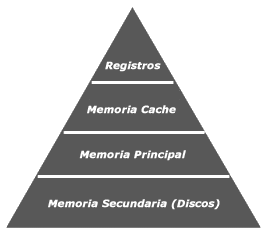
APM é un compoñente software integrado en determinados sistemas operativos que, xunto cos BIOS, permite certas funcións de aforro de enerxía. A diferenza fundamental entre APM e ACPI é que o primeiro realiza a xestión desde os BIOS, a un nivel máis baixo, e o segundo faio a nivel do sistema operativo.

A enerxía subminístrase á placa base a través duns conectores nos que se conecta a fonte de alimentación.

### 4.4 O sistema de Memoria

O sistema de memoria está composto de distintos dispositivos de almacenamento de datos, que forman un conxunto que permite o almacenamento de datos de forma temporal ou permanente, en tempo e forma necesario para que o sistema poida funcionar de forma esperada.

O problema xorde de que existen distintas características dos dispositivos de memoria que limitan a súa aplicación, como son o custo, velocidade, capacidade, etc, e imposibilitan a utilización de un só tipo. A solución é formar un sistema que permita aproveitar os aspectos máis vantaxosos dos dispositivos, suplindo as carencias entre eles. Isto define a pirámide da memoria que se mostra na figura.

Nesta pirámide se poden ver os distintos tipos de memorias clasificadas por velocidade, capacidade e custo. Sendo as de maior custo e mellor tempo de acceso as situadas na parte máis alta da pirámide como son os rexistros, cache, etc. A capacidade ven representada pola anchura da base do triangulo, tendo máis capacidade a memoria secundaria e a de menos os rexistros.

Todo o subsistema de memoria dun computador, tanto a memoria RAM como a memoria caché que se integra en diferentes niveis, é controlado por unha unidade que formaba parte do chipset da placa base e que nestes momentos está sendo incorporada nos directamente nos procesadores. A esta unidade denomínaselle unidade de administración de memoria (MMU, Memory Management Unit). O controlador de memoria, ou MMU, é un compoñente esencial en calquera parte do computador. É un chip que adoita integrarse como parte do chipset ou do propio microprocesador e cuxa función consiste en controlar o intercambio de datos entre a memoria e o microprocesador, así como a xestión do control de erros.

Este intercambio de datos entre memoria e microprocesador faise, de forma interna, mediante unha agrupación de bits denominada palabra. Así, un microprocesador con arquitectura de 32 bits pode ler ou escribir en cada momento 32 bits, polo que dicimos que ten unha lonxitude de palabra de 32 bits. A isto tamén se lle adoita chamar ciclo de bus e é un factor determinante á hora de medir as prestacións dun computador.

Cando se fala de memoria dun computador, faise referencia á súa memoria real ou física, é dicir, a memoria RAM; con todo, ás veces recórrese a un truco que permite facer crer que o computador ten máis memoria. Esta memoria adicional é a memoria virtual, que a administra un xestor de memoria virtual que crea un arquivo en disco duro ou utiliza unha partición a modo de memoria adicional para suplir a carencia da real. A este arquivo denomínaselle arquivo de intercambio ou partición swap.

Canto maior sexa a cantidade de memoria instalada nun computador, maiores serán as súas prestacións; con todo, existen unha serie de limitacións á cantidade de memoria que pode admitir un sistema informático:

* Límite de direccionamento do microprocesador. Vos microprocesadores teñen un límite non número de bits que son capaces de manexar para direccionar distintas posicións de memoria.
* O chipset. O deseño propio deste compoñente hardware, determina a cantidade máxima de memoria que se pode utilizar nun computador ou sistema informático. Isto esta mudando, sendo vos procesadores quen están levando a responsabilidade do direccionamento, polo que ou chipset deixaría de ser un límite.
* Límites no encapsulado físico da memoria. Os computadores actuais utilizan módulos de memoria de tipo DIMM teñen un tamaño físico determinado. Este tamaño limita a cantidade de memoria que son capaces de albergar. É evidente que a medida que se vaian miniaturizando os compoñentes, no mesmo tamaño poderanse almacenar moitos máis módulos de memoria.

### 4.4.1. Características da memoria

Algunhas das características fundamentais das memorias (de calquera tipo) son as seguintes:

a) Volatilidade. Dise que a información almacenada nunha memoria é volátil a condición de que corra o risco de verse alterada no caso que se produza algún fallo de subministración de enerxía eléctrica (memorias de biestables). Non son volátiles aquelas memorias nas que a información, independentemente ou non que exista algún fallo no fluído eléctrico, permanece inalterada.

b) Tempo de acceso (Retardo). É o tempo que transcorre desde o instante en que se lanza a operación de lectura na memoria e o instante en que se dispón da primeira información buscada. Na memoria principal este tempo é, en principio, independente da dirección na que se atope a información á que queremos acceder.

c) Capacidade. Número de posicións de memoria dun sistema (número de informacións que pode conter unha memoria). A capacidade total de memoria será un dato esencial para calibrar a potencia dun computador. A capacidade da memoria medirémola en múltiplos de byte (8 bits): Kilobytes (1024 bytes), Megabytes (1024 Kilobytes) e Gigabytes (1024 Megabytes).

d) Caudal (Ancho de banda). Número máximo de informacións lidas ou escritas por unidade de tempo.

### 4.4.2 Memoria principal

A Memoria Principal ou Memoria Central é o dispositivo que serve para almacenar os programas (instrucións) que se queiran executar (cando haxa que cargar o programa) e para almacenar os datos, os cálculos intermedios e os resultados (cando o programa xa se estea executando).É dicir, almacena todo aquilo que ha de ser procesado pola CPU. A CPU pode traer e levar datos directamente desde e cara á memoria.

A posibilidade de que en dous momentos diferentes estean dous programas diferentes na memoria é o que permite que unha mesma máquina poida servir para traballos distintos (computadores de propósito xeral).

Só os datos almacenados na memoria poden ser procesados pola CPU. Os datos que estean contidos nalgún dispositivo de almacenamento externo deben ser previamente introducidos á memoria, por medio dunha unidade periférica.

A cantidade de memoria usada para almacenar o programa dependerá da complexidade do mesmo (número de instrucións que o formen) e do tamaño dos datos que se queren procesar no programa.

### 4.4.3. Direccionabilidade da memoria

A memoria pódese comparar cos buzóns de correos que hai nos portais dos edificios. Están dispostos en forma matricial, ousexa en filas e columnas; cada un dos buzóns de correos é unha unidade de almacenamento e está identificada por un código de piso e porta ao que corresponde .

Na memoria cada un destes buzóns de correos correspóndese cunha cela de memoria ou posición de memoria. Estas celas teñen un tamaño (número de bits que poden almacenar) sempre igual para cada sistema. A cada cela de memoria asígnaselle un indicativo, que chamaremos dirección da cela (un número) que a identifica e permite referenciala pola súa posición dentro do conxunto total de celdillas que compoñen o total da memoria. Para ler ou escribir información nunha determinada cela será preciso identificala mediante a súa dirección de cela.

Como adoita ser o propio ordenador quen manexa continuamente os números que representan as direccións, estas direccións exprésanse en binario (que é o sistema con o que traballa o computador). Para maior comodidade humana, pódense representar en hexadecimal.

O número de bits (unidade binaria básica de información) que compoñen unha celdilla elemental de memoria, a unidade mínima direccionable, o que chamamos posición de memoria, depende da construción electrónica do deseño de cada computador. Hoxe día predomina o emprego da lonxitude de 8 bits (1 byte) para posición elemental de memoria.

Outra cousa distinta, que non hai que confundir, é a lonxitude de bits que son abarcados como operando dunha instrución. É un dato moi importante porque limita o valor máximo que pode chegar a ter un operando dentro do computador.

Ao conxunto de bits que forman un campo de memoria que contén un operando dunha instrución da máquina chámaselle **palabra**.

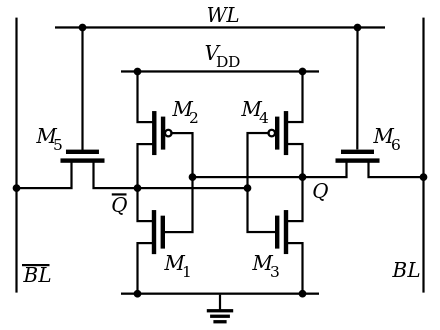
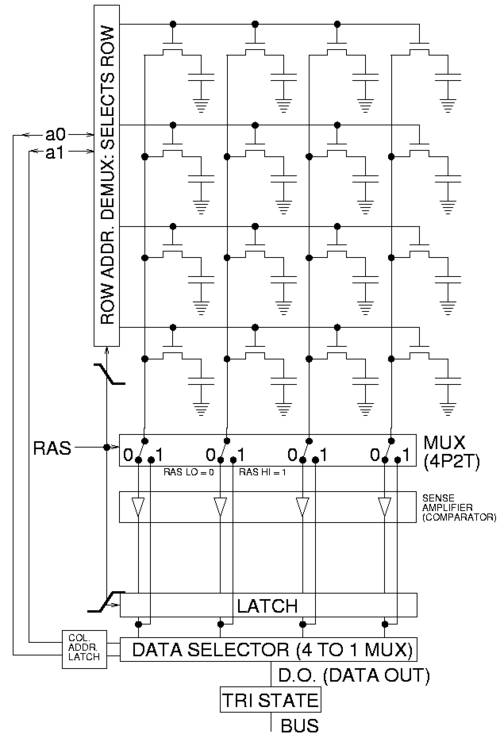
### 4.4.4. Espazo direccionable

En calquera computador existirá un espazo direccionable. É o intervalo de direccións que poden ser utilizadas polo microprocesador. O espazo direccionable está limitado polo tamaño do bus de direccións e este á súa vez depende do deseño do microprocesador.

De forma xeral, a cantidade de memoria máxima que pode direccionar o microprocesador será de 2n bytes, sendo n o número de liñas do bus de direccións; para un bus de direccións de 20 liñas, teriamos 220 bytes = 1.048.576 bytes = 1.024 Kb = 1 Mb.

### 4.4.5 Circuítos de memoria estáticos e dinámicos.

A memoria RAM soe estar formada por circuítos integrados, que poden ser de dous tipos:

* RAM estática o SRAM (Static RAM).
* 
* Celda de memoria estática
* RAM dinámica o DRAM (Dynamic RAM).
* 
* Grupo de celdas de memoria dinámica

A memoria SRAM non necesita actualización, frente a memoria DRAM, que sí require ser actualizada constantemente. Polo tanto, a memoria SRAM é ata 5 veces máis rápida que a DRAM, ainda que tamén é moito máis cara e, por iso, soe implementarse como memoria caché (memoria intermedia de acceso rápido). Polo contrario, a memoria DRAM soe implementase como memoria do sistema, é dicir, como memoria principal.

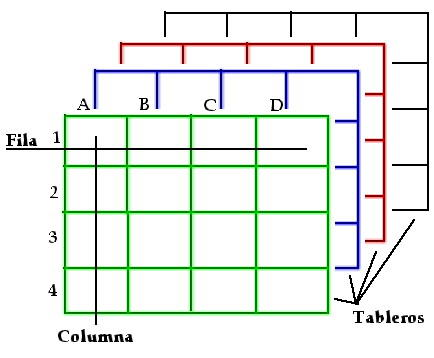
A diferenza fundamental entre estes dous tipos de memoria está no deseño. A memoria DRAM precisa de un só transistor e un condensador para poder fabricar unha célula e polo tanto se conseguen maiores densidades de celas. A memoria estática precisa de polo menos 6 transistores para a súa fabricación.

### 4.4.6 Funcionamento interno

Un módulo de memoria está fabricado a base de celas eléctricas (biestables). Cando escribimos información na memoria, esta permanece nas celas durante un tempo determinado. É dicir, gravado un dato en memoria, aínda que o computador siga aceso, se non se lle fornece periodicamente corrente eléctrica, a información desaparece das celas, xa que estas a o ser pequenos condensadores de corrente, descárganse pasado un período de tempo.

O proceso de actualización é unha técnica que consiste en cargar de novo a información na cela de memoria para que esta non perda a súa información. Os métodos de actualización máis habituais son a actualización distribuída, que carga unha fila cada vez en orde secuencial, e a actualización en refachos (ráfagas en castelán), que carga un grupo completo de filas. A velocidade de actualización coa que se poden recargar as celas, así como as veces que é necesario recargalas, é un parámetro esencial para determinar a velocidade da memoria.

A memoria está composta por un determinado número de celas, capaces de almacenar un dato ou unha instrución e colocadas en forma de taboleiro de xadrez. Non só existe un "taboleiro" senón que existen varios, desta forma a estrutura queda en forma de taboleiro de xadrez tridimensional .



Cando deséxase acceder a memoria, é imprescindible indicar o número de taboleiro, o número de fila dentro do taboleiro, e o número de columna ou cela dentro de esa fila, nesa orden.

#### 4.4.6.1 Tempos de espera (latency) dunha memoria

A latencia ou retardo, en poucas palabras, o tempo que a memoria leva para obter un dato solicitado.

Existen varios tipos de esperas (latencias) nas memorias, con todo, as máis importantes son:

* CAS: indica o tempo que tarda a memoria en colocarse sobre unha columna ou cela.
* RAS: indica o tempo que tarda a memoria en colocarse sobre unha fila.
* ACTIVE: indica o tempo que tarda a memoria en activar un taboleiro.
* PRECHARGE: indica o tempo que tarda a memoria en desactivar un taboleiro.

O proceso a seguir cando se desexa ler ou escribir na memoria será o seguinte:

* Mandar un sinal para activar o taboleiro e esperar a que termine o tempo de activación (ACTIVE).
* Mandar un sinal para saber cal é a fila na que se debe situar e esperar o seu tempo (RAS).
* Mandar un sinal para saber cal é a columna ou cela onde se debe situar e esperar (CAS).

O tempo que tarda a memoria en proporcionar o dato, é a suma de tres latencias: ACTIVE, RAS e CAS.

Como se comenta anteriormente, antes de enviar o dato/instrución onde deba ir, débense ler varias celas de memoria, por tanto hai que pasar dunha cela a outra, e ir esperando o seu correspondente latencia CAS.

Se cada taboleiro ten, por exemplo 64 celas, e vanse ler 20 posicións, as latencias totais a esperar son:

1 X ACTIVE (xa que se len menos de 64 celas, que son as que ten o taboleiro completo)

3 X RAS (xa que cada fila ten 8 posicións)

20 X CAS (xa que se van ler 20 celas)

A latencia máis importante, como queda patente, é a latencia CAS, e canto menor sexa esta, mellor rendemento terá o computador en xeral.

Canto menor ese número, mellor. Velaí as taxas máis comúns para cada tipo de memoria:

- DDR: 2, 2,5 e 3 ciclos de clock;

- DDR2: 3, 4 e 5 ciclos de clock;

- DDR3: 7, 6, 8 ou 9 ciclos de clock.

- DDR4: 15, 16 ou 17 ciclos de reloxo.

Entenda que, con iso, un módulo DDR2 pode gastar até 5 ciclos de reloxo para ter acceso a un determinado dato, mentres que no tipo DDR3 ese intervalo pode ser de até 9 ciclos.

#### 4.4.6.2 Memorias síncronas e asíncronas

A memoria que se utilizaba dende os primeiros tempos era do tipo asíncrona, é dicir, non estaba sincronizada co reloxo do sistema. Funcionaba razoablemente ben ata que apareceron buses máis rápidos, como os de 66 Mhz. Estas memorias son aquelas que se agrupan baixo a denominación de DRAM. Os tempos de acceso destas memorias se median en segundos.

Polo contrario, existe unha memoria denomina síncrona o SDRAM, que si está deseñada para sincronizarse coa frecuencia do microprocesador, o que significa que o controlador de memoria coñece o ciclo de reloxo exacto no que os datos estarán preparados, polo que o microprocesador non ten que entrar en estados de espera e, polo tanto, se mellora o rendemento.

### 4.4.7 Comunicación cos buses.

Como acabamos de ver as memorias síncronas están coordinadas co procesador a través dos buses. Antigamente se comunicaban con este por medio do FSB e modernamente por buses como HT ou QpI.

As memorias DDR destacan en relación o patrón anterior, memorias SDR SDRAM, porque son capaces de realizar dúas operacións de lectura ou escrita por ciclo de reloxo, a velocidade coa cal o procesador solicita operacións. As memorias DDR2, á súa vez, dobran esa capacidade, realizando catro operacións por ciclo de reloxo.

O tipo DDR3 segue o mesmo camiño: dobra a cantidade de operacións por vez en relación ao modelo anterior, ou sexa, realiza 8 procedementos de lectura ou gravación a cada ciclo de reloxo, catro no inicio deste e outros catro no final.

Para comprender mellor este aspecto, considera que, cando nos referimos ao ciclo de reloxo, estamos tratando da comunicación da memoria co exterior. Pero, a memoria traballa cunha frecuencia propia internamente.

Tendo esa característica en conta a cuestión das operacións por ciclo de reloxo, temos o seguinte escenario:

- Un módulo DDR-400, por exemplo, funciona internamente a 200 MHz, mais ofrece 400 Mhz, o traballar con dúas operacións por ciclo (2 x 200);

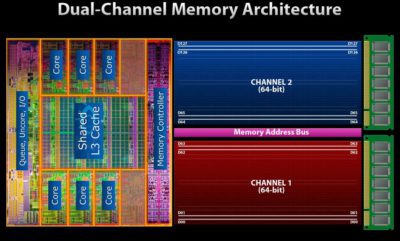
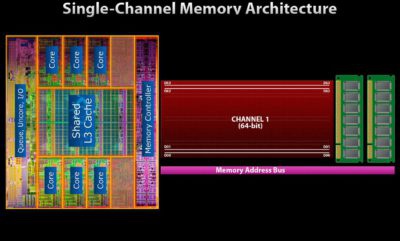
- Un modulo DDR2-800, que tamén funciona internamente a 200 MHz, pode ofrecer 800 MHz, xa que fai uso de catro operacións de cada vez (4 x 200).

- Un módulo DDR3-1600 que, funciona internamente a 200 MHz, por utilizar 8 operacións por ciclo de clock, pode ofrecer 1.600 MHz (8 x 200).

#### 4.4.7.1 Dual-Channel e Triple-Channel

É unha tecnoloxía para memorias aplicada nas computadoras ou ordenadores persoais, a cal permite o incremento do rendemento gracias o acceso simultáneo os módulos distintos de memoria. Isto conséguese mediante un segundo controlador de memoria na ponte norte do *chipset* ou bus adicado.

As melloras de rendemento son particularmente perceptibles cando se traballa con controladoras de vídeo integradas á placa base xa que estas, o non contar con memoria propia, usan a memoria RAM ou memoria principal do sistema e, gracias o dobre canal, poden acceder a un módulo mentres o sistema accede o outro.



Mais, a partir da liña de procesadores Intel Core i7, as memorias DDR3 pasaron a contar cunha nova modalidade: Triple-Channel. Como o nome indica, neste modo, as memorias pasan a traballar co triplo de datos por ciclo. Así, se cada canle transmite 64 bits, temos entón un total de 192 bits por vez. Alén diso, se no modo Dual-Channel é necesario utilizar dous peites de memoria coas mesmas especificacións, no Triple-Channel son necesarios tres. Iso indica que a placa-nai necesita contar non so cun chipset compatible (o mesmo vale para o procesador) como tamén posuír máis slots de memoria, tornando tales dispositivos máis caros ao usuario.

### 4.4.8 *Tecnoloxías de memoria RAM*

* DDRII SDRAM. SDRAM Double Data Rate, RAM síncrona dinámica de cuádruple velocidade de transferencia) Mellora a súa taxa de transferencia pudendo facer catro transferencias de datos por cada ciclo de reloxo. Reduce a tensión de traballo a 1,7 voltios
* DDR3 SDRAM. Mellora o consumo de enerxía fronte a DDRII o traballar a 1,5 voltios. Comparada coa DDR2 ten a capacidade de dobrar a taxa de datos de transferencia, e dicir, oito veces a frecuencia de traballo interna das matrices de memoria. Con dúas transferencias por ciclo de unha sinal cuadriplicada de reloxo, co ancho de 64 bits, un modulo de DDR3 pode acadar unha taxa de transferencia 64 veces por enriba da frecuencia de reloxo da memoria. Iso ven de que DDR3 SDRAM ten unha taxa de transferencia de: (frecuencia de reloxo interna da memoria) x 4 (multiplicador da frecuencia do bus) x 2 (data rate) x 64 (número de bits transferido)/ 8 (numero de bits/byte).
* DDR4 SDRAM. As principais vantaxes da DDR4 sobre o seu predecesor DDR3, é que os seus módulos permiten maior densidade, menores voltaxes de alimentación, unido a unha maior taxa de transferencia de datos. DDR4 traballa con unha voltaxe de 1.2 V cunha frecuencia entre 800 e 1600 MHz, comparados coas frecuencias entre 400 e 1067 Mhz e voltaxe de 1.5 ou 1.65 V de DDR3. O contrario que os anteriores desenvolvementos, DDR4 non incrementará o ancho das lecturas, que seguirá sendo de 8 bytes como na DDR3, senón que intercalará lecturas en diferentes bancos para acadar as velocidades de bus desexadas.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nome estándar | Velocidade de reloxo | Tempo entre sinais | Velocidade de reloxo de [E/S](https://es.wikipedia.org/wiki/E/S) | Operacións por segundo | Nome do módulo | [Taxa de bits](https://es.wikipedia.org/wiki/Tasa_de_bits) |
| DDR4-2133 | 266 MHz |  | 1066 MHz | 2133 millones | PC4-17000 | 17 066 [MB](https://es.wikipedia.org/wiki/Megabyte)/s |

* DDR5 desenvolveuse co obxectivo de **dobrar o ancho de banda da memoria DDR4, cun menor consumo**, mantendo a o tempo de acceso. Os chips están fabricados con *litografía de 10 nanómetros (nm)*, e están formados por 4 bancos de memoria que a súa vez pódense agrupar de 8 en 8. Duas das melloras que permiten aumentar o rendemento están na **ráfaga de datos** e **nos canais de memoria. DDR5 dobra o tamaño das ráfagas de datos de 8 a 16.** Además aporta **dous canais independentes de 40 bits por cada módulo DIMM de memoria**, mentres que DDR4 só tiña un canal, de 72 bits. A memoria DDR5 parte dunha velocidade de 4.800 MT/sg, o que significa que é **1,87 veces máis rápida** e pode chegar ata os **6.400 MT/sg**. A memoria DDR4 funciona a 1,2V, pero **DDR5 o fai a 1,1V**, é dicir, consume un pouco menos. Outra mellora é que os módulos de memoria poden integrar reguladores de voltaxe na propia placa de memoria, para aumentala velocidade e reducir o ruído eléctrico.

### 4.4.9 *Outros tipos de memoria*

#### 4.4.9.1 Memoria flash.

Tipo de memoria non volátil no que o contido se pode modificar. Almacena información binaria en forma de celdas, exactamente igual que a memoria SRAM e DRAM. Porén, súa característica máis importante é que actúa como un disco duro, é dicir, seu contido é permanente aínda despois de interromperse o subministro eléctrico. Tamén se caracteriza polo seu baixo consumo, súa durabilidade e a súa velocidade.

Parécese moito á memoria DRAM, xa que se presenta en forma de chips, módulos ou tarxetas de memoria, aínda que seu contido non necesita actualización constante nin alimentación eléctrica para manter seus datos. Esta memoria se soe utilizar moito en cámaras dixitais, impresoras e escáneres de altas prestacións, placas base, etcétera.

Os tipos de memoria flash son moi variados e dependen do seu construtor, pero, para ter unha idea xeral, se poden encontrar os seguintes tipos: Compact Flash de tipo I e II (compatibles PCMCIA e ATA), Smart Media (pouco utilizadas), Multimedia Card, Secure Digital (permite cifrado de datos e ata 256 Mb de capacidad), Memory Stick (as últimas acadan os 2 Gb no modelo Memory Stick Pro) e xD Picture (para cámaras dixitais; seu tamaño, dende 1 ata 8 Gb).

#### 4.4.9.2 Memoria Caché

Un tipo importante de memoria é a memoria caché. Funcionalmente, a memoria caché é igual á memoria principal. Con todo, fisicamente no computador é un compoñente distinto (non é imprescindible que estea nos computadores). Pódese definir como unha memoria rápida e pequena, situada entre a memoria principal e o procesador, especialmente deseñada para conter información que se utiliza con frecuencia nun proceso co fin de evitar accesos a outras memorias (principal), reducindo considerablemente o tempo de acceso ao ser máis rápida que o resto da memoria principal.

Cando o procesador le datos ou os almacena na memoria principal, os datos tamén se almacenan na memoria caché. Se o microprocesador necesítaos de novo, leos do caché e non da principal. A característica fundamental desta memoria é que, ao comezo dunha sesión, o computador atópase sen ningunha información. Unha vez que o procesador ten acceso á memoria RAM para cargar ou extraer información, esta pasa pola memoria caché e queda almacenada nela. Así, se é necesario acceder, á mesma información da memoria RAM, xa non será necesario ir até ela, senón que esa información atoparase xa almacenada na memoria caché. Ao ser esta moi rápida a velocidade increméntase drasticamente.

O principal problema da memoria caché é que non ten demasiada capacidade e é cara; o seu pequeno tamaño implica que a información que se almacenará nesta memoria será só aquela que utilicemos con maior frecuencia. Pódese pensar que nalgún momento a memoria caché pode chegar a saturarse. Isto é certo, pero os datos que contén iranse actualizando dependendo da antigüidade e uso dos mesmos. Eliminarase a información que se utilice pouco ou que teña pouca importancia.

Cando se atopa un dato no caché, dise que se produciu un acerto, sendo un caché xulgado pola súa taxa de acertos (hit rate). Os sistemas de memoria caché usan unha tecnoloxía coñecida por caché intelixente na cal o sistema pode recoñecer certo tipo de datos usados frecuentemente. As estratexias para determinar que información debe de ser posta no caché constitúen un dos problemas máis interesantes na ciencia das computadoras.

Actualmente, calquera sistema informático (computador) incorpora no propio microprocesador memoria caché moi rápida denominada L1 (nivel 1), xeralmente utilízase a metade para datos e a outra metade para instrucións. Tamén incorporan outro tipo de memoria caché denominada L2 (nivel 2) dentro do propio núcleo, polo que o bus funciona exactamente á mesma frecuencia do microprocesador, e obtense con iso o máximo rendemento.

#### 4.4.9.3 Memoria de vídeo

A memoria RAM de vídeo das diferentes tarxetas gráficas varía en calidade e cantidade: canto maior sexa a cantidade de memoria RAM de vídeo, maior poderá ser o número de colores simultáneos en pantalla e a resolución, así como a capacidade e velocidade gráfica en aplicacións 3D.

A continuación, enuméranse os diferentes tipos de memoria RAM de vídeo:

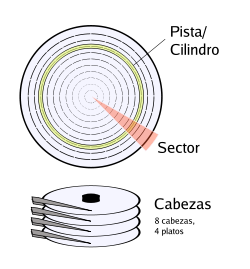
* GDDR. Funciona segundo o estándar [DDR](http://es.wikipedia.org/wiki/DDR_SDRAM), enviando dous bits por cada ciclo de reloxo, pero neste caso os módulos de memoria GDDR se optimizaron para lograr altas frecuencias de reloxo acortando os tempos de acceso das células de memoria en comparación coa memoria [DDR](http://es.wikipedia.org/wiki/DDR_SDRAM) convencional. É unha memoria DDR optimizada para computo gráfico, funciona a 2.5V. Permite un ancho de banda ata los 1.2 Gb/s a 300Mhz.
* GDDR2 funcionaba segundo o estándar [DDR](http://es.wikipedia.org/wiki/DDR_SDRAM),[1](http://es.wikipedia.org/wiki/GDDR2" \l "cite_note-GDDR2-00-1) enviando ou recibindo dous bits por ciclo de reloxo. As memorias GDDR2 foron realmente unha pequena optimización de [GDDR](http://es.wikipedia.org/wiki/GDDR) chegando a unha capacidade máxima de 1 GB, obtiveron máis frecuencia real, e a pesar de funcionar cun intervalo de voltaxe entre 1,8 a 2 voltios. Permite un ancho de banda ata os 2.0Gb/s a 500Mhz.
* GDDR3. Foi deseñada por ATI, e esta feita en base á DDR2, pero os requirimentos de enerxía e a dispersión da calor se melloraron, para así permitir módulos de maior rendemento. Para mellorar o rendemento da GDDR3, a transferencia de datos se fai en datos de 4 bits por pin en 2 ciclos. Unha vantaxe da GDDR3, é que ten a capacidade de baleirar tódolos datos da memoria (repouso [idle]) e comezar unha tarefa nova. Funciona a 1.5V. Permite un ancho de banda de 9.6Gb/s a 1Ghz.
* GDDR4. Deseñada por Samsung sendo a evolución de DDR2. Baseada na tecnoloxía [DDR2](http://es.wikipedia.org/wiki/DDR2), sucede á memoria [GDDR3](http://es.wikipedia.org/wiki/GDDR3), reducindo considerablemente o consumo eléctrico e a disipación térmica, aumentando a velocidade de transferencia de datos, permitindo aumentar a eficiencia de maneira notable. Os módulos traballan a 1.5V.
* GDDR5. Evolución de DDR2, con cambios nas especificacións de DDR3. GDDR5 permite unha transferencia de 32 bit dobres por ciclo que mais o bufer de 8 bits, permite un ancho de 256 bits, permitindo así un ancho de banda de 20Gb/s a 1.5V.

## 4.5. Unidades de disco

### 4.5.1 *Características xerais*

Un disco duro é un dispositivo de gran capacidade de almacenamento non volátil, que permite almacenar gran cantidade de datos de forma rápida e segura. Os datos permanecen gravados no seu interior permanentemente e poden ser lidos, borrados e gravados as veces que se desexe. A velocidade de lectura e escritura é elevada e o custo por Mb é cada día menor. O tamaño das primeiras unidades era de 10 ou 20 Mb, aínda que hoxe existen discos de centos e miles de Gigabytes, capacidades enormes debidas á alta densidade de gravación, que é posible grazas á miniaturización das cabezas de lectura e escritura e á calidade dos compoñentes.

Un disco duro componse internamente dun ou máis pratos de aluminio recubertos dunha capa de material con propiedades magnéticas en ambas as caras. Inicialmente utilizábase óxido férreo, pero actualmente adóitase usar cobalto polas súas mellores calidades. O tamaño físico dos discos duros oscila entre 3,5 e 5,25 polgadas. Existe un terceiro tamaño, de 2,5 polgadas, que se utiliza habitualmente nos computadores portátiles. A este formato adóitaselle denominar microdrive.

Os pratos do disco están montados ao redor dun eixo (spindle, en inglés) que xira sempre no mesmo sentido grazas a un motor de tipo plano. Normalmente, cada prato dispón de dúas cabezas de lectura e escritura, unha para cada cara. As cabezas atópanse agrupadas nunha armazón en forma de peite que as move conxuntamente sobre toda a superficie dos pratos. Ao xirar a gran velocidade, os pratos orixinan unha pequena corrente de aire que fai que as cabezas voen sobre os mesmos a unha distancia microscópica de 3 ou 4 micropulgadas.

Os datos grávanse en forma de fluxos magnéticos escritos en círculos ao redor do anel central do disco, cada un dos cales se denomina pista. Cada pista está dividida nun número igual de segmentos, chamados sectores, que son a unidade mínima de información direccionable e dun tamaño habitual de 512 bytes. Cada pista está dividida normalmente nunha cantidade de sectores que oscila entre 100 e 300.

Outro concepto importante é a denominada densidade de gravación; canto maior sexa a densidade de gravación, maior será a cantidade de datos que se poderán almacenar por unidade de superficie. Existen dúas formas de conseguir unha maior densidade de gravación: almacenar os bits máis xuntos uns doutros, ou ben diminuír a distancia entre pistas. A densidade de gravación nun disco duro mídese en bits por polgada (BPI, Bits Per Inch) e a densidade de pistas en cada prato mídese en pistas por polgada (TPI, Tracks Per Inch).

Cando un disco duro dispón de dous ou máis pratos xorde o concepto de cilindro. Para controlar o funcionamento do motor e xestionar como se grava e le a información, todos os discos actuais contan cun circuito controlador integrado na propia unidade.

As características dos discos duros dependen dunha serie de factores físicos, entre os que cabe destacar os seguintes:

* *Velocidade de rotación (Vr)*. Oscila entre 4500 E 7200 revolucións por minuto (rpm), aínda que algúns modelos SCSI alcanzan velocidades de até 15000 rpm. Canto maior sexa a velocidade de rotación maior será a velocidade de transferencia, aínda que isto implique unha maior nivel de ruído e de quecemento do disco. Os discos duros actuais teñen unha velocidade de rotación constante e len todos os sectores dunha pista nun único xiro. A isto denomínaselle factor de intercalación 1:1.
* *Número de sectores por pista (Spp).* Os discos duros actuais utilizan diferentes tamaños de pista. As zonas máis externas do disco dispoñen de máis espazo para os sectores que as zonas internas.
* *Tempo de procura.* O tempo de procura máis rápido nun disco duro ten lugar cando se le desde unha pista determinada á seguinte pista contigua; este é o chamado tempo de acceso pista a pista. Adoita oscilar entre 1 e 2 milisegundos. O máximo tempo de procura ten lugar cando se le desde a pista máis externa á máis interna, e oscila ao redor dos 12 milisegundos. O tempo medio de acceso é o tempo que se utiliza en situar as cabezas do disco en pistas aleatorias. Este adoita oscilar entre 5 e 8 milisegundos.
* *Latencia rotacional.* Unha vez que a cabeza se sitúa sobre a pista desexada, debe esperar ao sector correcto para levar a cabo a operación de lectura ou escritura. A este tempo de espera chámaselle latencia rotacional, e adoita oscilar entre 2 e 4 milisegundos, segundo a velocidade de rotación.
* *Taxa de transferencia interna (Tti).* A velocidade de rotación dun disco duro á que pode ler información do prato e transferila á controladora (xeralmente ao búfer ou memoria caché) denomínase taxa de transferencia interna, mídese en megabits por segundo (Mbps) e pódese calcular segundo a seguinte fórmula:

*Tti = (Vr /60 \* Spp \* 512 \* 8) / 1000000*

Onde Tti é a taxa de transferencia interna, Vr é a velocidade de rotación e Spp é o número de sectores por pista.

Tempo medio entre erros (MTBF, Mean Time Between Failures). Xeralmente mídese en horas e representa o tempo que ten que transcorrer para que se produzan dous erros aleatorios no disco duro. Este factor oscila entre 200 000 e 500 000 horas.

Para saber máis na dirección podedes ver esta animación, sobre almacenamento de bits en disco verticalmente. http://www.hitachigst.com/hdd/research/recording\_head/pr/perpendicularanimation.html

### 4.5.2 *Interface dun disco duro*

Sata y PCI-Express

Para ler e escribir correctamente a información nos soportes de almacenamento externo (discos duros, unidades de CD-ROM, disqueteiras convencionais, etc., é necesario un conector co que unir o periférico co bus do sistema. Este denomínase porto de comunicacións. Non só se encargan de administrar a transferencia de datos, senón que ademais controla as operacións de entrada e saída da información, dende e ata os periféricos ou unidades de disco. Por iso se lle soe chamar tamén controlador de disco. Súa función consiste en establecer a conexión entre os periféricos de información e o bus do sistema. Polo tanto, ten que ser compatible co bus do sistema e cos propios periféricos.

Tódolos discos duros incorporan unha placa de circuíto impreso integrada ou controladora. Esta placa encárgase de controlar todo o funcionamento do disco, así como unha serie de funcións adicionais, como control do motor de xiro, funciones de xestión de enerxía, control de tempo en operacións de lectura e escritura, etcétera.

A circuitería da controladora conten un microprocesador e unha memoria interna, ademais dunha memoria caché para axilizar as operacións de lectura e escritura. Tamén incorporan un firmware ou rutinas gravadas en memoria ROM de tipo flash que permiten xestionar o resto de compoñentes do disco duro.

### 4.5.3 *Modos de transferencia e protocolos utilizados nos discos duros*

A forma de realizar a transferencia de datos neste tipo de controladores depende en gran medida do BIOS e do chipset da placa base. Son os seguintes:

* PIO (programmable Input/Output, Entrada e saída programable)

Ten cinco modos de transferencia que varían entre 3,3 e 16,6 Mb/s. Porén, cando se conectan dous dispositivos transferindo datos mediante PIO nunha mesma banda de datos, a taxa de transferencia sempre se adapta a do dispositivo máis lento. Cando se utiliza PIO para a transferencia de datos, é o microprocesador o que se encarga de levar a cabo tódalas operacións, polo que para evitar unha diminución no rendemento do sistema se soe utilizar o modo de acceso directo a memoria (DMA).

* DMA (Direct Memory Access, Acceso directo a memoria)

Este método permite a transferencia de datos entre memoria e un dispositivo, sen necesidade de que interveña o microprocesador. Para que a transferencia sexa posible, se debe facer uso do controlador DMA que está integrado no chipset da placa base do ordenador.

* Ultra DMA

É unha función de transferencia de alta velocidade que permite os comandos DMA transferir datos a unha velocidade máxima de 33, 66 o 100 Mb/s, sempre empregando o conector EIDE de 40 hilos. Para dispoñer de todo o seu potencial, necesitamos un disco duro Ultra DMA, placa base compatible con Ultra DMA e drivers de sistema operativo apropiado para tal fin. Os últimos estándares aparecidos no mercado foron ATA-6 e ATA- 7, tamén coñecidas como UltraDMA100 e UltraDMA133, respectivamente que aumentan a taxa de transferencia orixinal ata 100 e 133 Mb/s. Porén, para iso necesitamos un cable de datos de 80 hilos co conector clásico de 40 patillas.

* Transferencia en bloques (Block Mode)

Tódolos BIOS actuais permiten unha configuración denominada Block Mode. Esta función permite agrupar varias ordes de lectura e escritura nun só bloque, de maneira que poidan ser controladas por unha única interrupción. Os discos duros actuais permiten a transferencia de ata 32 sectores de datos completos con unha soa interrupción, opción que se recomenda activar en tódolos BIOS.

#### *4.5.4 Interfaces*

* Adaptador Serial ATA (SATA)

Este recente estándar é o sucesor do clásico IDE e os modos de transferencia UltraDMA. A diferenza co seu predecesor estriba en a interface utilizada para a transmisión dende e ata o disco duro. Se pasa da tradicional tecnoloxía IDE paralela a unha en serie, da clásica banda dos 40 fíos a un conector que unicamente inclúe dous cables, un para a transmisión e outro para recepción de datos, ademais de outros cinco que se usan para o control da transmisión. Outra diferencia fundamental é que tratase dunha conexión punto a punto, é dicir, que cada disco duro debe ir conectado con un cable de datos Serial ATA a placa base, o que permite eliminar así a configuración de discos duros mediante jumpers en modo mestre e escravo. Tamén incrementa a lonxitude do cable, que pasa dos 40 cm da banda IDE ata 1 m. No referente a velocidades, este estándar parte de los 150 MB/s, chega a 300 MB/s en Serial ATA-II e acada os 600 MB/s para Serial ATA-III.

PCIe

Coa irrupción dos discos SSD comezouse a utilizar o bus PCIe como interface de conexións destes dispositivos. O problema é que os discos SSD están limitados polas conexións SATA III, este estándar admite unha velocidade máxima de transferencia de 600MB/sg. Coas velocidades que poden acadar os actuais discos SSD o conector SATA é un funil moi grande. Para evitar esta limitación optouse por utilizar conexións PCIe. Por exemplo, o SSD de Samsung NVMe 960 EVO pro, é un SSD M.2 PCIe 3.0 x4 que acada os 3.500 y 2.100 MB/sg en lectura/escritura secuencial, moito maior ancho de banda que os 600MB/sg de los SSD SATA III.

### 4.5.5 Dispositivos de estado sólido (SSD)

Unha memoria de estado sólido é un dispositivo de almacenamento secundario feito con compoñentes electrónicos de estado sólido para o seu uso en equipos informáticos en substitución dunha unidade de disco duro convencional, como memoria auxiliar ou para a fabricación de unidades híbridas compostas por SSD e disco duro.

Consta dunha memoria non volátil, en lugar dos pratos xiratorios e cabezal, que son atopados nas unidades de disco duro convencionais. Sen partes móbiles, unha unidade de estado sólido pretende reducir drasticamente o tempo de procura, latencia e son practicamente inaudibles, o que tradúcese nunha mellora do rendemento.

Ao ser inmune ás vibracións externas, faino especialmente apto para o seu uso en vehículos, computadoras portátiles, etc.

Pola contra, a súa vida útil pode ser inferior, xa que teñen un número limitado de ciclos de escritura, podendo producirse a perda absoluta dos datos de forma inesperada e irrecuperable. Porén, por medio do cálculo do [tempo medio entre fallos](https://es.wikipedia.org/wiki/Tiempo_medio_entre_fallos) e a administración de sectores defectuosos dito problema pode ser mitigado.

#### 4.5.5.1 Partes que o compoñen

Un SSD componse principalmente:

* Controladora: É un procesador electrónico que se encarga de administrar, xestionar e unir os módulos de memoria NAND cos conectores en entrada e saída. Executa software a nivel de Firmware e é con toda seguridade, o factor máis determinante para as velocidades do dispositivo.
* Caché: Un dispositivo SSD utiliza un pequeno dispositivo de memoria DRAM similar ao caché dos discos duros. O directorio da colocación de bloques e o desgaste de nivelación de datos tamén se mantén na memoria caché mentres a unidade está operativa.
* Condensador: É necesario para manter a integridade dos datos da memoria caché, se a alimentación eléctrica detívose inesperadamente, o tempo suficiente para que se poidan enviar os datos retidos cara á memoria non volátil.

#### 4.5.5.2 Vantaxes

Os dispositivos de estado sólido que usan flash teñen varias vantaxes únicas fronte aos discos duros mecánicos:

* Arranque máis rápido, ao non ter pratos que necesiten tomar unha velocidade constante.
* Gran velocidade de escritura.
* Maior rapidez de lectura, mesmo 10 veces máis que os discos duros tradicionais máis rápidos grazas a RAIDs internos nun mesmo SSD.
* Baixa latencia de lectura e escritura, centos de veces máis rápido que os discos mecánicos.
* Lanzamento e arranque de aplicacións en menor tempo. Resultado da maior velocidade de lectura e especialmente do tempo de procura. Pero só se a aplicación reside en flash e é máis dependente da velocidade de lectura que doutros aspectos.
* Menor consumo de enerxía e produción de calor. Resultado de non ter elementos mecánicos.
* Sen ruído. A mesma carencia de partes mecánicas fainos completamente inaudibles.
* Mellorado o tempo medio entre fallos, superando 2 millóns de horas, moi superior ao dos discos duros.
* Seguridade. Permitindo unha moi rápida "limpeza" dos datos almacenados.
* Rendemento determinístico. Diferenza dos discos duros mecánicos, o rendemento dos SSD é constante e determinístico a través do almacenamento enteiro. O tempo de "procura" constante.
* O rendemento non se deteriora mentres o medio se enche. (Véxase Desfragmentación).
* Menor peso e tamaño que un disco duro tradicional de similar capacidade.
* Resistente. Soporta caídas, golpes e vibracións sen estragarse e sen descalibrarse como pasaba cos antigos discos duros, grazas a carecer de elementos mecánicos.
* Borrado máis seguro e irrecuperable de datos; é dicir, non é necesario facer uso do Algoritmo Gutmann para asegurar totalmente o borrado dun arquivo.

#### 4.5.5.3 Limitacións

Os dispositivos de estado sólido que usan flash teñen tamén varias desvantaxes:

* Prezo. Os prezos das memorias flash son considerablemente máis altos en relación prezo/gigabyte, a principal razón da súa baixa demanda. Con todo, esta non é unha desvantaxe técnica. Segundo se establezan no mercado irá minguando o seu prezo e comparándose aos discos duros mecánicos, que en teoría son máis caros de producir ao levar pezas metálicas.
* Menor recuperación. Despois dun fallo físico pérdense completamente os datos pois a cela é destruída, mentres que nun disco duro normal que sofre dano mecánico os datos son frecuentemente recuperables usando axuda de expertos.
* Capacidade. A día de hoxe, teñen menor capacidade máxima que a dun disco duro convencional, que chega a superar os tres terabytes.
* O número de ciclos de lectura e escritura destas unidades, redúcese a medida que se reduce o tamaño dos transistores de memoria, polo que as máis recentes teñen un menor tempo de vida total. Espérase que este problema se solucione proximamente.

Antigas desvantaxes xa solucionadas:

* Degradación de rendemento ao cabo de moito uso nas memorias NAND (solucionado, en parte, co sistema TRIM).
* Menor velocidade en operacións E/S secuenciais. (Xa se conseguiu unha velocidade similar).
* Vulnerabilidade contra certo tipo de efectos. Incluíndo perda de enerxía abrupta (nos SSD baseado en DRAM), campos magnéticos e cargas estáticas, comparados cos discos duros normais (que almacenan os datos dentro dunha gaiola de Faraday).

### 4.5.5 Sistemas Raid

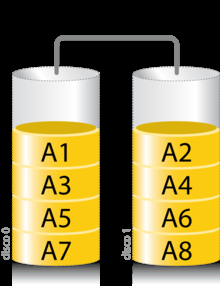
O acrónimo RAID (do inglés Redundant Array of Independent Disks), “conxunto redundante de discos independentes”, fai referencia a un sistema de almacenamento que usa múltiples discos duros ou SSD entre os que se distribúen ou replican os datos. Dependendo da súa configuración (á que adoita chamarse «nivel»), os beneficios dun RAID respecto dun único disco son un ou varios dos seguintes: maior integridade, maior tolerancia a fallos, maior rendemento e maior capacidade.

#### 4.5.5.1 Descrición xeral

Unha "matriz de discos" está formada por un grupo de dous (ou máis) unidades de disco que aparecen ante o sistema como un só disco. A vantaxe dunha matriz é proporcionar un mellor rendemento de fluxo e/ou unha tolerancia os fallos nos datos. Un mellor rendemento lograse compartindo a carga de traballo en paralelo entre varios discos físicos. A tolerancia os erros lógrase mediante unha operación redundante de datos onde si un (ou máis) discos fallan ou teñen un fallo de sector, se pode encontrar unha copia en espello dos datos en outro (ou outros) disco(s). Para obter resultados óptimos, se soen instalar nas matrices de discos Ultra ATA/100 idénticos. Os discos de rendemento equivalente permiten que a matriz funcione mellor como unha soa unidade de disco.

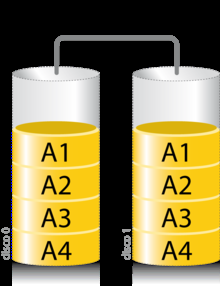
#### 4.5.5.2 Entrelazado (RAID 0)

Un RAID 0 (tamén chamado conxunto dividido ou volume dividido) distribúe os datos equitativamente entre dous ou máis discos sen información de paridade que proporcione redundancia. É importante sinalar que o RAID 0 non era un dos niveis RAID orixinais e que non é redundante. O RAID 0 úsase normalmente para incrementar o rendemento, aínda que tamén pode utilizarse, como forma de crear, un pequeno número de grandes discos virtuais, a partir dun gran número de pequenos discos físicos.



#### 4.5.5.3 Espello (mirroring) (RAID 1)

Un RAID 1 crea unha copia exacta (ou espello) dun conxunto de datos en dous ou máis discos. Isto resulta útil cando o rendemento en lectura é máis importante que a capacidade. Un conxunto RAID 1 só pode ser tan grande como o máis pequeno dos seus discos. Un RAID 1 clásico consiste en dous discos en espello, o que incrementa exponencialmente a fiabilidade respecto dun só disco; é dicir, a probabilidade de fallo do conxunto é igual ao produto das probabilidades de fallo de cada un dos discos (pois para que o conxunto falle é necesario que o fagan todos os seus discos).



Adicionalmente, dado que todos os datos están en dous ou máis discos, con hardware habitualmente independente, o rendemento de lectura increméntase aproximadamente como múltiplo lineal do número das copias; é dicir, un RAID 1 pode estar a ler simultaneamente dous datos diferentes en dous discos diferentes, polo que o seu rendemento duplícase. Para maximizar os beneficios sobre o rendemento do RAID 1 recoméndase o uso de controladoras de disco independentes, unha para cada disco (práctica que algúns denominan splitting ou duplexing).

Como no RAID 0, o tempo medio de lectura redúcese, xa que os sectores a buscar poden dividirse entre os discos, baixando o tempo de procura e subindo a taxa de transferencia, co único límite da velocidade soportada pola controladora RAID. Con todo, moitas tarxetas RAID 1 IDE antigas len só dun disco da parella, polo que o seu rendemento é igual ao dun único disco. Algunhas implementaciones RAID 1 antigas tamén len de ambos os discos simultaneamente e comparan os datos para detectar erros. A detección e corrección de erros nos discos duros modernos fan esta práctica pouco útil.

Ao escribir, o conxunto compórtase como un único disco, dado que os datos deben ser escritos en todos os discos do RAID 1. Por tanto, o rendemento non mellora.

#### 4.5.5.4 Raid 5

Un RAID 5 usa división de datos a nivel de bloques distribuíndo a información de paridade entre todos os discos membros do conxunto. O RAID 5 logrou popularidade grazas o seu baixo custo de redundancia. Xeralmente, o RAID 5 se implementa con soporte hardware para o cálculo da paridade. RAID 5 necesitará un mínimo de 3 discos para ser montado.

No gráfico de exemplo anterior, unha petición de lectura do bloque «A1» sería servida polo disco 0. Unha petición de lectura simultánea do bloque «B1» tería que esperar, pero unha petición de lectura de «B2» podería atenderse concorrentemente xa que seria servida polo disco 1.



Cada vez que un bloque de datos se escribe nun RAID 5, xérase un bloque de paridade dentro da mesma división (stripe). Un bloque componse a miúdo de moitos sectores consecutivos de disco. Unha serie de bloques (un bloque de cada un dos discos do conxunto) recibe o nome colectivo de división (stripe). Se outro bloque, ou algunha porción dun bloque, é escrita nesa mesma división, o bloque de paridade (ou unha parte do mesmo) é recalculada e volta a escribir. O disco utilizado polo bloque de paridade está graduado dunha división á seguinte, de aí o termo «bloques de paridade distribuídos». As escrituras nun RAID 5 son custosas en termos de operacións de disco e tráfico entre os discos e a controladora.

Os bloques de paridade non se len nas operacións de lectura de datos, xa que isto sería unha sobrecarga innecesaria e diminuiría o rendemento. Con todo, os bloques de paridade lense cando a lectura dun sector de datos provoca un erro de CRC. Neste caso, o sector na mesma posición relativa dentro de cada un dos bloques de datos restantes na división e dentro do bloque de paridade na división utilízanse para reconstruír o sector erróneo. O erro CRC ocúltase así ao resto do sistema. Da mesma forma, se falla un disco do conxunto, os bloques de paridade dos restantes discos son combinados matematicamente cos bloques de datos dos restantes discos para reconstruír os datos do disco que fallou «ao voo».

O anterior denomínase ás veces Modo Interino de Recuperación de Datos (Interim Data Recovery Mode). O sistema sabe que un disco fallou, pero só co fin de que o sistema operativo poida notificar ao administrador que unha unidade necesita ser substituída: as aplicacións en execución seguen funcionando alleas ao fallo. As lecturas e escrituras continúan normalmente no conxunto de discos, aínda que con algunha degradación de rendemento. A diferenza entre o RAID 4 e o RAID 5 é que, no Modo Interno de Recuperación de Datos, o RAID 5 pode ser lixeiramente máis rápido, debido a que, cando o CRC e a paridade está no disco que fallou, os cálculos non teñen que realizarse, mentres que no RAID 4, se un dos discos de datos falla, os cálculos teñen que ser realizados en cada acceso.

O RAID 5 require polo menos tres unidades de disco para ser implementado. O fallo dun segundo disco provoca a perda completa dos datos.

O número máximo de discos nun grupo de redundancia RAID 5 é teoricamente ilimitado, pero na práctica é común limitar o número de unidades. Os inconvenientes de usar grupos de redundancia maiores son unha maior probabilidade de fallo simultáneo de dous discos, un maior tempo de reconstrución e unha maior probabilidade de achar un sector irrecuperable durante unha reconstrución. A medida que o número de discos nun conxunto RAID 5 crece, o MTBF (tempo medio entre fallos) pode ser máis baixo que o dun único disco. Isto sucede cando a probabilidade de que falle un segundo disco nos N-1 discos restantes dun conxunto no que fallou un disco no tempo necesario para detectar, substituír e recrear devandito disco é maior que a probabilidade de fallo dun único disco. Unha alternativa que proporciona unha protección de paridade dual, permitindo así maior número de discos por grupo, é o RAID 6.

Algúns vendedores RAID evitan montar discos dos mesmos lotes nun grupo de redundancia para minimizar a probabilidade de fallos simultáneos ao principio e o final da súa vida útil.

As implementacións RAID 5 presentan un rendemento malo cando se someten a cargas de traballo que inclúen moitas escrituras máis pequenas que o tamaño dunha división (stripe). Isto débese a que a paridade debe ser actualizada para cada escritura, o que esixe realizar secuencias de lectura, modificación e escritura tanto para o bloque de datos como para o de paridade. Implementacións máis complexas inclúen a miúdo cachés de escritura non volátiles para reducir este problema de rendemento.

No caso dun fallo do sistema cando hai escrituras activas, a paridade dunha división (stripe) pode quedar nun estado inconsistente cos datos. Se isto non se detecta e repara antes de que un disco ou bloque falle, poden perderse datos debido a que se usará unha paridade incorrecta para reconstruír o bloque perdido na devandita división. Esta potencial vulnerabilidade coñécese ás veces como «buraco de escritura». Son comúns o uso de caché non volátiles e outras técnicas para reducir a probabilidade de ocorrencia desta vulnerabilidade.