

Clase 7

El chipset y la CPU

7.1 Introducción

Muchas veces llamado “el cerebro del ordenador”, la CPU o microprocesador es el encargado de controlar las operaciones que realiza la máquina, en gran medida de él dependen las demás funciones que realice la computadora.

Sin embargo, la CPU por sí sola no podría procesar toda la información que requerimos, o al menos no con la misma eficiencia. Requiere de un conjunto de elementos denominados buses y chipset, junto con la estructura que los sostenga y vincule, concentrados en la placa madre.

En esta clase ampliaremos algunas características de la CPU y veremos los conceptos de buses y chipset de la placa madre.

7.2 Los buses

En un sistema de computadora, los diversos subsistemas deben tener interfaces entre sí; por ejemplo, la memoria y la CPU necesitan comunicarse, así como la CPU y los dispositivos de E/S. Esto, normalmente, se realiza con un bus. Digamos que, básicamente, los buses son líneas que transmiten bits (información) entre los dispositivos.

Las dos principales ventajas de la organización con buses son:

- El bajo costo: El costo es bajo, ya que un simple conjunto de cables es un camino múltiple compartido.
- La versatilidad: Al definir un sencillo esquema de interconexión, se pueden añadir fácilmente nuevos dispositivos y los periféricos pueden incluso compartirse entre sistemas de computadoras que utilicen mismos tipos de buses.

La principal desventaja de un bus es que crea un cuello de botella de comunicación, limitando posiblemente la máxima productividad.

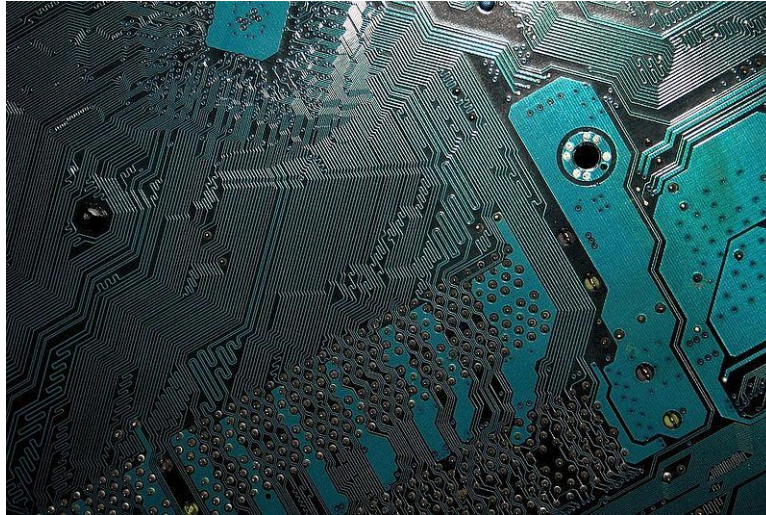
Como la entrada y salida de información debe pasar a través de un bus central, esta limitación de ancho de banda es importante.

Una razón, por la que el diseño del bus es difícil, es que la máxima velocidad del bus está limitada por factores físicos: la longitud del bus y el número de dispositivos (y, por consiguiente, la carga del bus). Estos límites físicos prohíben velocidades de bus arbitrarias. El deseo de altas velocidades de E/S (baja latencia) y alta productividad de las E/S también puede conducir a requerimientos de diseño conflictivos.

7.2.1 Clasificación de los buses

Los buses tradicionalmente se clasifican en:

- **Buses de CPU-memoria**
- **Buses de E/S.**



Buses en una placa madre

Los buses de E/S: pueden ser más largos, pueden tener muchos tipos de dispositivos conectados a ellos (USB, PCI, PCIe, SATA, etc.), tienen un amplio rango en el ancho de banda de datos de los dispositivos conectados a ellos y, normalmente, siguen un estándar de bus.

Los buses CPU-memoria: son cortos, generalmente de alta velocidad, y adaptados al sistema de memoria para maximizar el ancho de banda memoria-CPU.

Durante la fase de diseño, el diseñador de un bus CPU-memoria conoce todos los tipos de dispositivos que deben conectarse, mientras que el diseñador del bus de E/S debe aceptar dispositivos que varían en posibilidades de latencia y ancho de banda.

Por otro lado, los buses también se clasifican entre buses de control, de direcciones y de datos.

- **El bus de control** gobierna el uso y acceso a las líneas de datos y de direcciones.
- **El bus de direcciones:** como vimos, la memoria RAM es direccionable, es decir, cada celda de memoria tiene su propia dirección. El bus de direcciones transmite la dirección de memoria del dato en tránsito.

- **El bus de datos** permite el intercambio de datos entre la CPU y el resto de unidades.

7.2.2 Principales Buses de datos

Actualmente un dispositivo informático utiliza una variedad de buses de datos, cada uno de ellos con unas características y una finalidad concreta, incluso varios han ido desapareciendo, dejando lugar a otros nuevos y más veloces. Veremos a continuación los buses de datos actuales más utilizados:

El Bus de memoria

Como su propio nombre indica, este bus es el encargado de comunicar el controlador de memoria de la CPU con la memoria RAM (actualmente el controlador del bus de memoria se encuentra insertado en el procesador, en el IMC son las siglas de «Integrated Memory Controller», o controlador de memoria integrado).

Este bus ha tomado diferentes nombres según la marca fabricante del procesador, ya fueran Hyper Transport (HT), Quick Path Interconnect (QPI) o Direct Media Interface (DMI).

Peripheral Computer Interconnect (PCI)

El famoso bus PCI es uno de los que más soportó el paso del tiempo, aunque finalmente fue reemplazado por las diversas variantes del bus PCIe. Este bus de datos, que tuvo su origen en el año 1992, se empleaba originalmente para dar servicio tanto a las tarjetas gráficas como a las tarjetas auxiliares que se podían conectar a un ordenador. A diferencia del bus PCIe, se trata de un bus paralelo (que envía una serie de bits al mismo tiempo) y no en serie (que envía los bits uno tras otro) como este.

Bus Peripheral Computer Interconnect Express (PCIe)

Precedido por la corta existencia del bus AGP, el bus PCIe se diseñó para ser el sustituto de dicho bus y del PCI, con una interfaz que fuera modular en su diseño.

Esto significa que no existe un único tamaño de ranura para las tarjetas de expansión, sino que éstas pueden ser más largas o cortas, en función de la cantidad de tráfico que se vaya a enviar por ellas. Así, tenemos ranuras PCIe x1, x4, x8 y x16. Generalmente, es el procesador el encargado de gestionar el bus PCIe del sistema.

Actualmente, la mayoría de componentes se conectan al bus PCIe del sistema, ya sean la tarjeta gráfica, la tarjeta de sonido, la tarjeta de

red (alámbrica o inalámbrica) o los USB adicionales. Por lo tanto, es junto con el bus de memoria, el más importante dentro de todos los buses del ordenador.

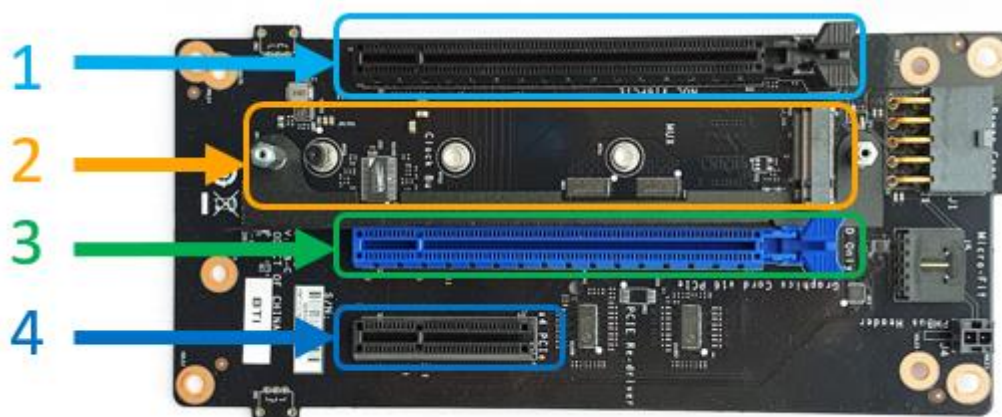
A continuación, algunos ejemplos de conectores PCIe:

1) Ranura PCIe 3.0 x16

2) Ranura PCIe 3.0 x4 M.2, configurada para conectar discos de estado sólido en formato M.2

3) Ranura PCIe 3.0 x16

4) Ranura PCIe 3.0 x4



(Los conectores 1 y 3 tienen igual especificación, solo cambia el color)

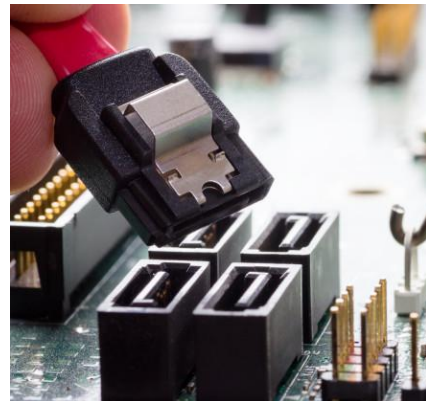
Y por último, una placa de video Nvidia RTX 3090, en donde podemos observar su conector PCIe:



Bus Serial AT Attachement (SATA)

El bus SATA es al cual se conectan las unidades de almacenamiento de nuestra computadora (discos rígidos y SSD). Este bus llegó al mercado en el año 2003 y, a diferencia del bus PATA (Parallel AT Attachement), es un bus en serie.

Como tal, ya en su primera revisión presentaba grandes mejoras sobre el antiguo bus PATA, como la mayor velocidad alcanzable por cada uno de los canales de comunicación, o la posibilidad de efectuar hot plugging (conexión en caliente, es decir, estando conectada la PC) con las unidades de almacenamiento.



Bus USB

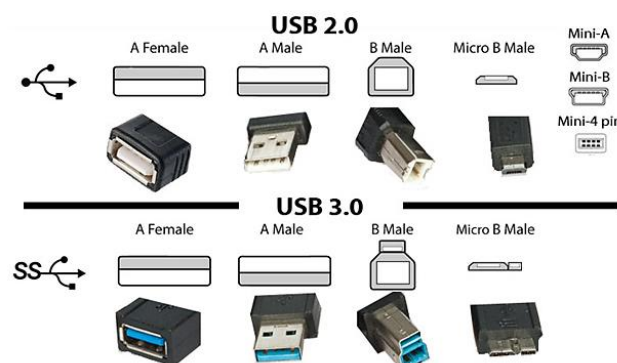
Este bus de datos seguramente es el más conocido. Se encarga de transmitir o recibir los datos de los dispositivos conectados a él. Generalmente, este bus se utiliza para conectar periféricos a la computadora, como el teclado, el mouse, unidades de almacenamiento externas, celulares, cámaras de video, y un largo etcétera.

La especificación más moderna de los USB (USB 3.1 2ª Generación) permite la transmisión de datos hasta a 10 GB/s.

Algo bueno de los puertos USB es que sus distintas versiones son compatibles entre sí. Si bien no se aprovechará la velocidad al conectar un dispositivo preparado para USB 3.0 en un puerto USB 2.0.

Sin embargo, en algunos casos pueden surgir problemas de velocidad o de conexión, ya que el estándar USB 3.0 soporta la entrega de 900mA de corriente, mientras que el USB 2.0 solo puede entregar hasta 500mA de corriente continua.

Una forma rápida de diferenciar entre USB 3.0 y 2.0 es observar el color de su conector. El USB 3.0 se caracteriza por su color azul.



Últimamente se está popularizando el USB tipo C, en reemplazo del micro USB. Sobre todo, por su resistencia a la rotura del conector.

7.3 El Chipset

Hasta principios de la década del 90 el término chipset aún no se utilizaba en informática. Las plataformas de computadoras personales Commodore y Atari, diseñaron en el año 1985 un conjunto de chips auxiliares que se encargaban del manejo de la memoria, el sonido, los gráficos o el control de unidades de almacenamiento masivo dejando a la CPU bastante más liberada, para realizar otras tareas.

A este conjunto de chips se los denominó Chipset.

El chipset, también conocido como circuito integrado, es la médula espinal de la computadora: está integrado en la placa base, lo que permite que ésta funcione como eje del sistema, permitiendo el tráfico de información entre el microprocesador (CPU) y el resto de los componentes de la placa base, interconectándolos a través de diversos buses.

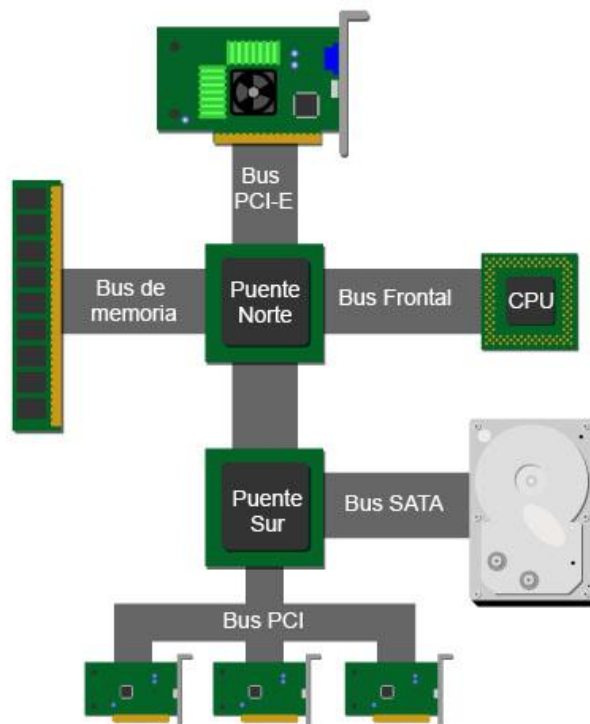
Hasta hace poco tiempo los dos componentes que formaban el chipset eran el Northbridge (puente norte) y Southbridge (puente sur). De hecho, es muy probable que en nuestras computadoras tengamos esos componentes.

El Northbridge era un circuito integrado que funcionaba como puente de enlace entre el microprocesador y la memoria, al igual que entre el microprocesador y las tarjetas gráficas y las tarjetas de video PCI-Express, así como las comunicaciones con el Southbridge.

El Southbridge es un circuito integrado que coordina los dispositivos de entrada y salida, además de realizar otras funciones de baja velocidad dentro de la placa base; y se comunicaba con la CPU a través del Northbridge.

El puente norte tenía un bus bastante más grande que el puente sur, ya que los datos más importantes se gestionaban desde él. Al puente sur llegan las comunicaciones de los controladores de USB, discos duros, SSD, audio, Ethernet, etc.

A continuación, un esquema en donde se observan los buses y su función de comunicar distintos dispositivos, a través de los puentes norte y sur.



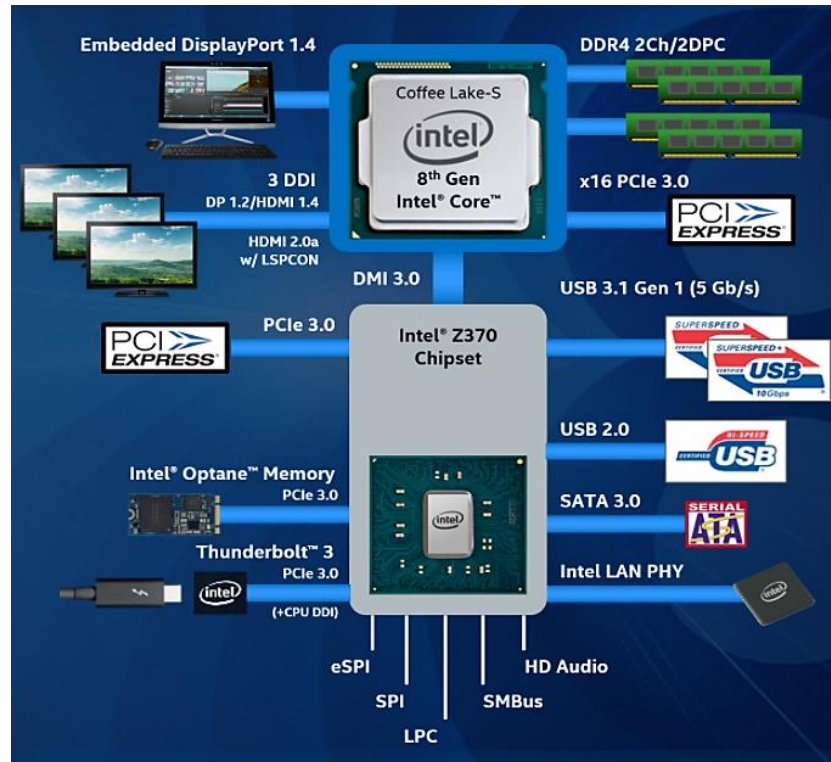
Sin embargo, esta arquitectura de chipset ya no se utiliza más, específicamente no se utiliza el puente norte.

Básicamente lo que se ha buscado, y conseguido, es que el puente norte desaparezca y las conexiones que ante se hacían desde él ahora se hagan directamente desde el procesador.

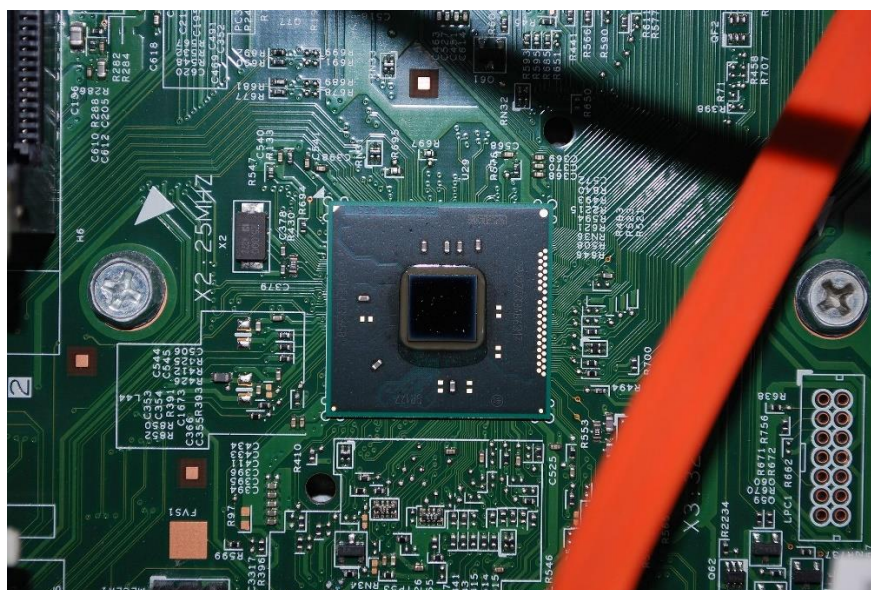
Solo queda el puente sur, al cual frecuentemente se lo nombra de otra manera. En el caso de los procesadores Intel se lo llama PCH (Platform Controller Hub, o concentrador del controlador de plataforma). En el caso de AMD simplemente **chipset**.

Este cambio consigue que los datos se transporten mucho más rápido, pues no hay intermediario entre la CPU y la RAM, y de esa manera el procesador maximiza aún más el rendimiento.

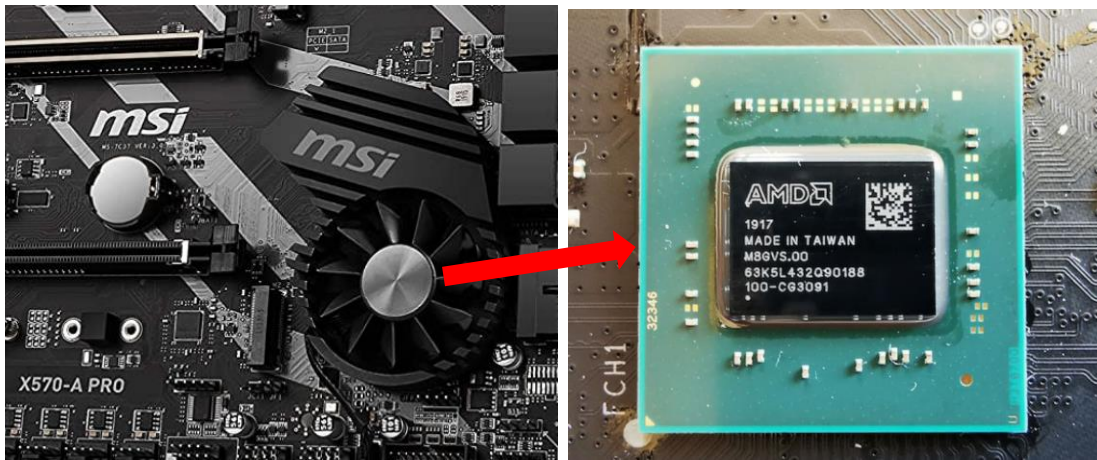
En la siguiente imagen observamos un esquema de conexión en el cual la memoria RAM y el video está conectado directamente a la CPU, mientras que los demás dispositivos son organizados y controlados por el chipset; que también suele incluir dispositivos de Audio HD y la placa de red.



A continuación, observamos un chipset o PCH Intel DH82H81 con su núcleo expuesto, pues en la placa madre viene cubierto por un disipador.

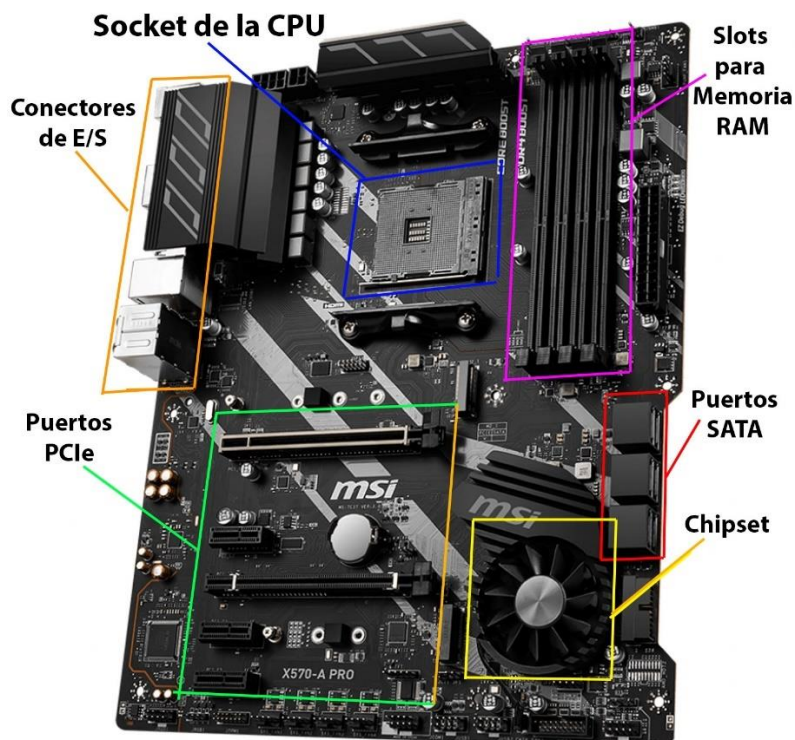


A continuación, observamos un chipset AMD X570, tan “potente” y con una generación tal de calor que requiere disipación activa mediante un ventilador:



Cómo podrán deducir, los chipsets actuales son, en términos de potencia, superiores a microprocesadores de décadas pasadas (pero no cumplen la misma función, claramente).

Para concluir con el tema del chipset, podríamos decir que la placa madre se compone de los buses, el chipset y los demás dispositivos electrónicos y conectores que permiten interconectarlos entre ellos y con el exterior. En la siguiente imagen vemos el motherboard MSI X570-A PRO:



7.4 La CPU

En clases anteriores hemos comentado varios aspectos y definiciones de la CPU, en esta clase vamos a ver algunas características propias de las CPU actuales.

7.4.1 Palabra (dword) del procesador

Un término que ha sido importante en la historia de la informática, en cuanto a la caracterización de la arquitectura de la CPU, es el de palabra del ordenador.

Seguramente alguien recordará que en su época las consolas de videojuegos se identificaban por ser de 8 bits, luego de 16 bits, 32 bits, 64 bits, etc.

Una palabra es una cadena de bits que son manejados como un conjunto por la computadora. El tamaño o longitud de una palabra hace referencia al número de bits contenidos en ella, y es un aspecto muy importante al momento de diseñar una arquitectura de ordenadores.

El tamaño de una palabra se refleja en muchos aspectos de la estructura y las operaciones de las computadoras:

- La mayoría de los registros en un ordenador normalmente tienen el tamaño de la palabra.
- El valor numérico típico manipulado por un ordenador es probablemente el tamaño de palabra.
- La cantidad de datos transferidos entre la CPU del ordenador y el sistema de memoria a menudo es más de una palabra.
- Una dirección utilizada para designar una ubicación de memoria a menudo ocupa una palabra. Es por esto que los procesadores de 64 bits pueden direccionar 2^{64} posiciones de memoria, mucho más que las 2^{32} posiciones disponibles para una CPU de 32 bits.

Los ordenadores modernos normalmente tienen un tamaño de palabra de 16, 32 o 64 bits.

Este término, hoy en día, ya no se usa tan “cotidianamente” como en el pasado, sin embargo, sigue siendo parte característica de las arquitecturas de los procesadores.

7.4.1 Arquitectura de los procesadores

La arquitectura de la CPU es un modelo y una descripción funcional de los requerimientos y las implementaciones sobre la forma en que la CPU trabaja internamente y accede a las direcciones de memoria.

Básicamente, la arquitectura está basada en tres grandes principios: instrucciones, velocidad, capacidad y tipo de conexión.

Las distintas arquitecturas de CPU se caracterizan por el conjunto de instrucciones que implementan, y por lo tanto manejan e interpretan.

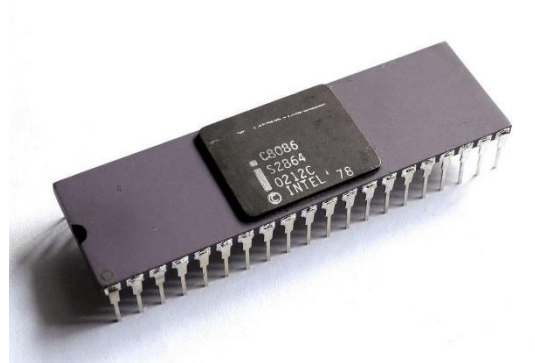
Dicho conjunto de instrucciones es operable por un programador en el lenguaje assembler e incluye, entre otros, comandos para:

- Sumar, restar, multiplicar o dividir dos registros y colocar el resultado en alguno de ellos.
- Realizar operaciones bit a bit, almacenando el AND y el OR de cada bit en un par de registros, o el NOT de cada bit en un registro.
- Comparar dos valores que se encuentren en registros (por ejemplo, si son iguales o si uno es mayor que otro).
- Establecer un registro de la CPU a un valor constante.
- Mover datos desde una posición de memoria a un registro de la CPU y viceversa.
- Leer y escribir datos desde dispositivos de hardware.
- Saltar a otra posición del programa y ejecutar instrucciones allí.
- Saltar a otra posición si se cumple cierta condición.
- Saltar a otra posición, pero salvando la posición actual para poder volver.

7.4.2 La arquitectura x86

La arquitectura de procesadores x86 es la responsable detrás de la mayoría de los sistemas de PC del mundo. En su momento fue una de las grandes revoluciones del mundo informático.

La x86 es una familia de instrucciones de procesamiento agrupadas en un conjunto que, como ya mencionamos, se le conoce como arquitectura. Esta fue desarrollada por Intel basada en su microprocesador 8086 a finales de la década del 70.



El mítico CPU Intel 8086

En 1978, Intel comenzó a comercializar el procesador 8086, un ambicioso chip de 16 bits. El 8086 se comercializó en versiones desde 4,77 y hasta 10 MHz.

El nombre de la arquitectura proviene del diseño y disposición del conjunto de instrucciones del microprocesador, ya que todos terminan en números de serie 86.

El éxito de esta serie de CPU fue tal que, a partir de ese momento, todos los CPU de Intel mantuvieron una estricta política de retrocompatibilidad. Todo CPU fabricado por Intel desde ese momento y hasta la actualidad, año 2021, es capaz de ejecutar código compilado para cualquiera de sus predecesores.

Al 8086 lo sucedió el 80286, comúnmente conocido como 286, en 1982. Este chip, de 24/16 bits, implementó el modo protegido de ejecución, sentando las bases para la aparición de los sistemas multitarea de escritorio. El 80286 apareció a 6 MHz, y a lo largo de los años llegó hasta los 12MHz.

Finalmente, en el año el 1985, llegaría el Intel 80386, o simplemente el Intel 386, el cual fue el primer procesador de Intel de 32 bits, y manejaba velocidades de 16 a 33 MHz.

A partir del procesador 486 un coprocesador numérico, y la memoria cache L1, fueron incorporadas en el mismo chip que la CPU.

Seguramente, los de más edad, no recordarán ninguna CPU de nombre 586; esto es porque a partir de esta línea de CPU Intel decidió cambiar la nomenclatura, surgiendo el nombre de Pentium.

Este hecho anecdótico pero histórico, tiene su fundamento en que un número no se puede hacer marca registrada, por lo que otras empresas, como AMD o Cyrix, aprovecharon la popularidad y sacaron procesadores también denominados 486.



A partir de aquí ya (casi) es historia moderna.

- En 1993 surge el Pentium, que llegó desde los 60 hasta los 233Mhz.
- Con la salida de los Pentium II (1997), Pentium III (1999) y Pentium 4 (2000) se consiguió la integración de más componentes, la adición de instrucciones específicas de multimedia y a elevar la velocidad tanto como sea posible. El Pentium II llegó desde 233 hasta 450 MHz. El Pentium III desde 450 a 1400 (1,4 GHz). El Pentium 4 debutó a 1,3 GHz y en noviembre del 2005 llegaba ya a los 3,80 GHz.

En este punto vamos a detenernos, como vemos el Intel Pentium 4 ya llegaba a los 3,8Ghz en el año 2005. Si bien los procesadores actuales mejoraron en muchos otros aspectos, esta velocidad no está muy alejada a la de las CPU actuales. El procesador AMD Ryzen 9 5950X salió en el año 2020 con una frecuencia base de 3.4Ghz, alcanzando una frecuencia máxima de 4.9Ghz (en un núcleo).

Es así que, en un punto, la carrera por aumentar exponencialmente los GHz se vio frenada debido al creciente consumo de energía y la generación de calor producida por los microprocesadores a altas frecuencias de reloj, además de limitaciones físicas por la miniaturización de los componentes (en teoría, el límite físico del silicio, como material actual de integración ronda los 10 GHz).

Con la octava generación de procesadores compatibles x86, los **x86-64**, presentados en el año 2005. La especificación x86-64 fue creada por AMD, y adoptado por Intel y otras compañías, se comienza a que utilizar a nivel doméstico la arquitectura y los buses de 64 bits, añadiendo la posibilidad de usar múltiples núcleos.

Además, soporta una cantidad mucho mayor de memoria virtual y memoria física, permitiendo a los programas almacenar grandes cantidades de datos en la memoria. También provee registros de uso general de 64 bits y muchas otras mejoras.

Con esta tecnología de 64bit y múltiples núcleos se da paso a los procesadores que actualmente tenemos entre nosotros.

7.4.3 Los múltiples núcleos

Un procesador con múltiples núcleos, o multinúcleo, es un sistema con al menos dos procesadores. En contraste, el monoprocesador tiene un único procesador.

Los primeros procesadores de doble núcleo para PC salieron, como mencionamos, en el año 2005, siendo los Intel Dual Core y los Athlon 64 x2.

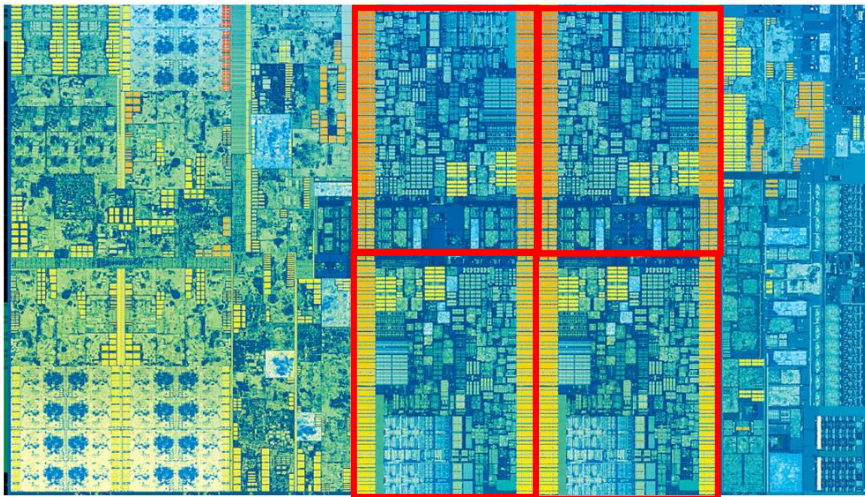
El concepto se basa en crear sistemas potentes simplemente conectando muchas CPU. Idealmente, se desearía que el procesador sea cada vez más rápido, sin embargo, otra posibilidad para mejorar la performance es que un procesador pueda realizar varias operaciones al mismo tiempo, en paralelo. Estos es lo que se logra con los múltiples núcleos.

Para esto, el software de los procesadores debe diseñarse para funcionar correctamente con un número variable de núcleos. En los inicios de esta tecnología esto no era tan común, sin embargo, hoy en día los entornos de desarrollo ya contemplan que los programas aprovechen esas ventajas.

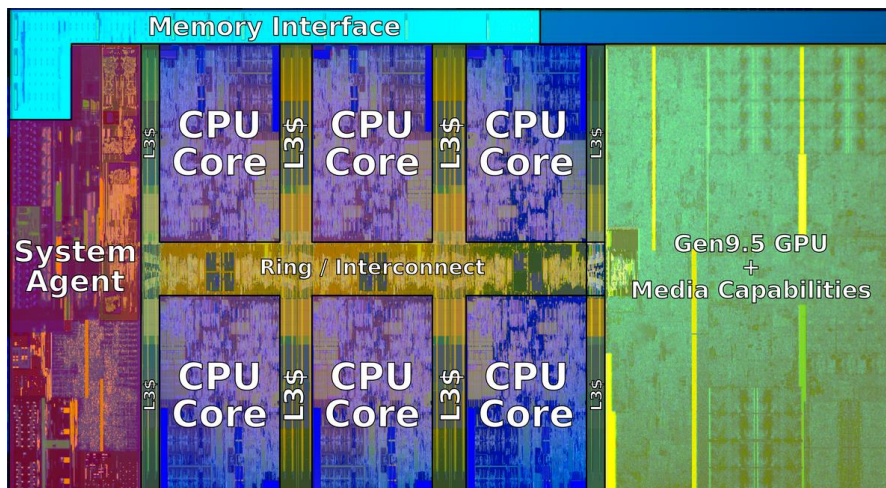
La tendencia actual es aumentar constantemente el número de núcleos de la CPU.

Vamos a dar algunos ejemplos en donde se señalan los núcleos en una vista interna (die shot) de los CPU:

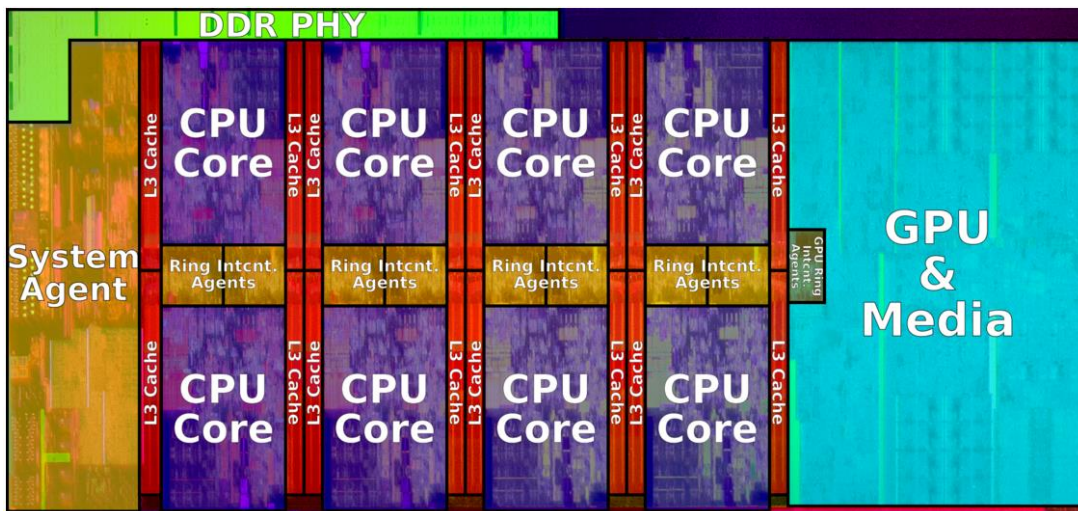
El Intel i7-7700k (2017) – 4 núcleos / 8 hilos



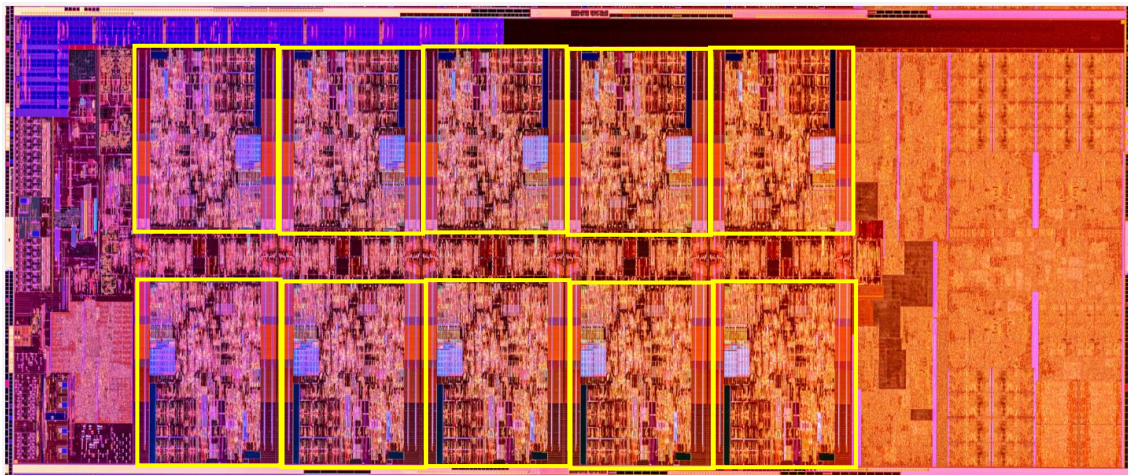
El Intel i7-8700k (2018) – 6 núcleos / 12 hilos



El Intel i7-9900k (2019) – 8 núcleos / 16 hilos



El Intel i9-10900k (2020) – 10 núcleos / 20 hilos



El AMD Ryzen 9 5950x (2020) – 16 núcleos / 32 hilos **(no encontré imagen interna)**

7.4.4 El Multithreading (múltiples hilos)

El Multithreading es una forma de dividir el trabajo de una CPU en procesos simultáneos. En la actualidad la mayoría de los programas dividen el trabajo en múltiples subprocesos, cada uno ejecutado por un núcleo de la CPU. Distintos núcleos del CPU procesan estos subprocesos en paralelo para ahorrar tiempo.

Algunos motores de videojuegos antiguos dependían de un único subproceso, lo que significa que la mayoría usan un solo núcleo de la CPU y obtienen una gran mejora con velocidades más altas del reloj.

Sin embargo, en la actualidad, los videojuegos y en general los programas complejos tienen varios subprocesos ejecutándose al mismo tiempo. Algunos para las físicas, al mismo tiempo que requiere otros para la inteligencia artificial, para enviar y recibir instrucciones a la placa de video y una diversidad de cálculos.

En un principio las CPU multinúcleo eran de dos núcleos, con los cual podían ejecutar dos procesos en paralelo. Esto, si bien fue un avance, muchas veces no era suficiente. Por lo que se implementó el concepto de hilos. Resumidamente, cada núcleo ejecuta tiene dos hilos, es decir que una CPU de 2 núcleos tendría 4 hilos, con lo cual podría ejecutar 4 subprocesos en simultáneo.

Los términos HyperThreading y SMT son muy conocidos en la industria del hardware, y más concretamente en la de los procesadores. Ambas son tecnologías de Intel y AMD respectivamente que se basan en la patente de 1994 de Sun Microsystems llamada MultiThreading, que permite duplicar las unidades lógicas de cada núcleo, es decir, permite a un núcleo realizar dos tareas simultáneas.

Si el procesador fuera quad-core, estaríamos hablando de ocho hilos de proceso. Así, si un procesador tiene Hyper Threading o SMT, siempre tendrá el doble de hilos de proceso que núcleos tenga.

Muchas veces se hace referencia a los núcleos reales como núcleos físicos y a los hilos como núcleos lógicos. Con HT y SMT, el sistema operativo verá dos núcleos de CPU «lógicos» por cada núcleo físico, y será el sistema operativo quien reparta el trabajo entre ellos. En otras palabras, si uno de los núcleos ya está ocupado haciendo sus operaciones, se le asignará el trabajo al siguiente. Y claro, teniendo dos hilos de proceso por núcleo, realmente el sistema operativo ve, de manera literal, el doble de núcleos a los que asignar trabajo.

A continuación, vemos dos imágenes de hilos llegando a un procesador. La primera, sin hyperthreading ejecuta los hilos uno tras otro, en simultáneo. La ejecución del hilo 2 debe esperar a que termine el hilo 1.



Por otro lado, al tener hyperthreading el procesador puede dividir su carga de trabajo para ejecutar los dos hilos al mismo tiempo, en paralelo.



Con esta clase concluimos la introducción al hardware de PC y su arquitectura. Por cuestiones de no extender demasiado la clase no hablamos de la arquitectura ARM, muy utilizada en nuestros días pues es la propia de los procesadores móviles con sistema operativo Android.

Como podrán observar, el tema del hardware ha estado en constante evolución a lo largo de la historia y lo sigue haciendo. Sin ir más lejos, el 1 de junio de 2021 AMD confirmó una nueva tecnología, la 3D V-Cache Stack, que llegará a las CPU actuales con arquitectura Zen 3 logrando una cantidad de memoria cache L3 de 192 MB y con la posibilidad de ampliarse hasta 512MB, algo impensable hasta hace poco tiempo. Además, las especificaciones indican que posiblemente por primera vez la cache L3 supere en velocidad a la cache L1, lo que aumentaría el rendimiento de las CPU en gran medida.