

Arquitecturas multiprocesadores

Pedro O. Pérez M., Phd.

Multiprocesadores
Tecnológico de Monterrey

pperezm@tec.mx

08-2022

Arquitecturas de computadoras

RISC

CISC

ARM

GPU

Cache

Coherencia del cache

Arquitecturas de procesadores

Core i5

Ryzen

ARM

GPU

Paralelismo

Definición

Metas

Tipos de paralelismo

Paralelismo a nivel de instrucción

Nivel de bits

Pipelining

Super-escalar

Paralelismo vs. Concurrencia

Cálculo de desempeño

Arquitecturas de computadoras

- ▶ Se define como el diseño conceptual y la estructura operacional de un sistema de computadoras, especialmente todo lo relacionado con la forma en que trabaja el CPU y cómo accede a la memoria.
- ▶ También suele definirse como la forma en que se interconectan los componentes de hardware, para crear computadoras según los requerimientos de funcionalidad, rendimiento y costo.

Arquitecturas de computadoras

Existen muchas arquitecturas de computadoras, sin embargo mencionaremos solo las tres más importantes:

- ▶ RISC (Reduced Instruction Set Computing).
- ▶ CISC (Complete Instruction Set Computing).
- ▶ ARM (Advanced RISC Machine).
- ▶ GPU.

Un Computador con Conjunto de Instrucciones Reducidas (RISC) es un CPU generalmente utilizado en microprocesadores o microcontroladores con las siguientes características fundamentales:

- ▶ Instrucciones de tamaño fijo y presentadas en un reducido número de formatos.
- ▶ Sólo las instrucciones de carga y almacenamiento acceden a la memoria de datos.

Su objetivo es posibilitar la segmentación y el paralelismo en la ejecución de instrucciones y reducir los accesos a memoria. Ejemplos de esta arquitectura son los microprocesadores. PowerPC, DEC Alpha, MIPS, ARM y SPARC.

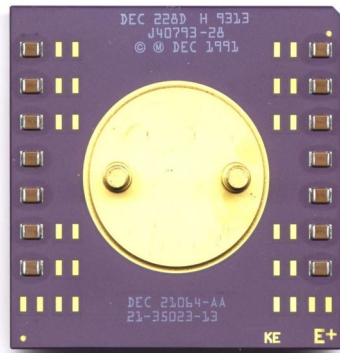


Figura: <https://bit.ly/2zb9Gj4>

Un Computador de Conjunto de Instrucciones Complejas (CISC) tienen un conjunto de instrucciones que se caracteriza por ser muy amplio y permitir operaciones complejas entre operandos situados en la memoria o en los registros internos, en contraposición a la arquitectura RISC.

- ▶ Este tipo de arquitectura dificulta el paralelismo entre instrucciones.

Ejemplos de esta arquitectura son los microprocesadores Motorola 68000, Zilog Z80, Intel x86 y AMDx86-64.

ARM

ARM es una arquitectura RISC de 32 bits y, con la llegada de su versión V8-A, también de 64 Bits, desarrollada por ARM Holdings.

- ▶ Se llamó Advanced RISC Machine, y anteriormente Acorn RISC Machine.
- ▶ La arquitectura ARM es el conjunto de instrucciones de 32 y 64 bits más ampliamente utilizado en unidades producidas.

Ejemplos de esta arquitectura son los microprocesadores Applied Micro Circuits Corporation X-Gene, DEC StrongARM, Freescale i.MX, Marvell Technology Group XScale, NVIDIA Tegra, Qualcomm Snapdragon, Texas Instruments OMAP, Samsung Exynos, Apple Ax, ST-Ericsson NovaThor, Huawei K3V2 e Intel Medfield.



Figura: <https://bit.ly/2TPTmhy>

- ▶ La arquitectura NVIDIA GPU se basa en una matriz escalable de multiprocesadores. Un multiprocesador está diseñado para ejecutar cientos de hilos al mismo tiempo. Para administrar una cantidad de hilos tan grande se emplea una arquitectura única llamada SIMT (Single-Instruction, MultipleThread).
- ▶ Las instrucciones se canalizan para aprovechar el paralelismo a nivel instrucción dentro de un único subproceso, así como el paralelismo a nivel de subprocesos a través de multiprocesamiento de hardware simultáneo.

<https://bit.ly/1mKiP2s>

Cache

- ▶ La memoria cache es un componente usado por el CPU para reducir el costo promedio (en tiempo o energía) de acceso a los datos que se encuentran en memoria principal.
- ▶ La mayoría de los CPUs tienen varias memorias cache independiente, de datos o instrucciones, que se encuentran organizados en una jerarquía de varios niveles (L1, L2, etc.).
- ▶ La información que se cargan en la memoria cache depende de algoritmos sofisticados y ciertas suposiciones sobre el código del programa. El objetivo del sistema cache es garantizar que el CPU tenga el siguiente bloque de datos que necesitará ya cargado en memoria.

Coherencia del cache

Para mayor rendimiento en un sistema multiprocesador, cada procesador generalmente tiene su propio cache. La coherencia de cache se refiere al problema de mantener la coherencia de los datos de estas caches. El principal problema es lidiar con las escrituras de un procesador.

<https://bit.ly/2TPTmhy>

Coherencia del cache

Hay dos estrategias generales para tratar las escrituras en un cache:

- ▶ Escritura: Todos los datos escritos en el cache también se escriben en la memoria al mismo tiempo.
- ▶ Reescritura: Cuando los datos se escriben en un cache, se marca como “sucio” el bloque afectado. El bloque modificado se escribe en memoria solo cuando se reemplaza el bloque.

<https://bit.ly/2Hkqt86>

Core i5

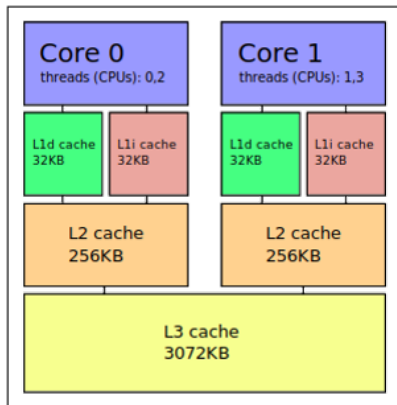


Figura: <https://bit.ly/2Zib3ai>

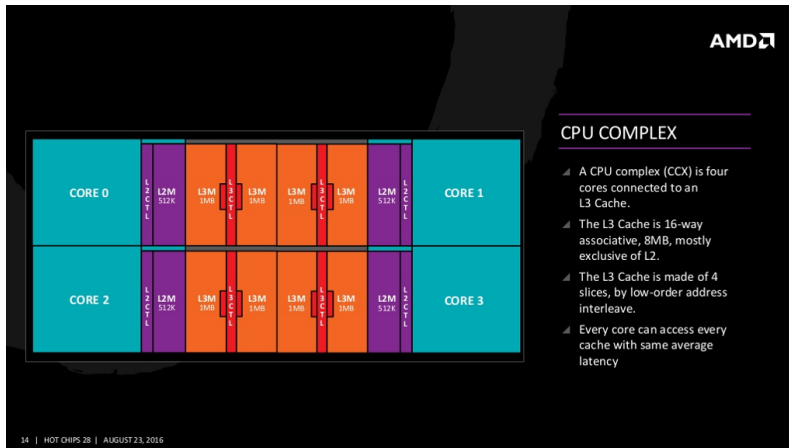


Figura: <https://bit.ly/2Zib3ai>

ARM

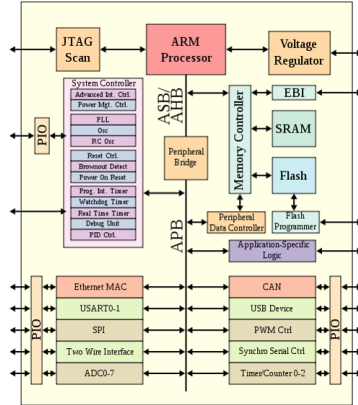


Figura: <https://bit.ly/33LLvWH>

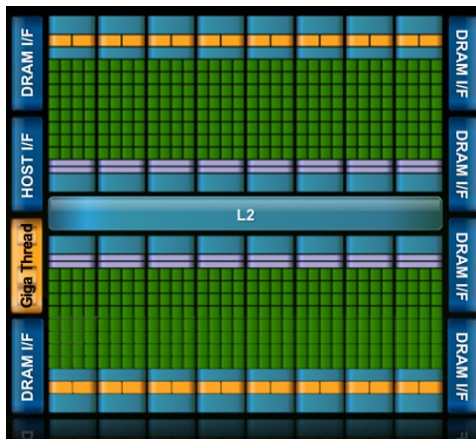


Figura: <https://bit.ly/31Qmntnr>