

Classificazione microprocessori

Slide 14→16

Un microprocessore è un'implementazione fisica di un processore (CPU) realizzata su un singolo circuito integrato.

Esso è utilizzato, oltre che in quasi tutti i computer, in molti dispositivi digitali quali ad esempio telefoni cellulari e stampanti.

Esistono 3 grandi gruppi di sistemi di microprocessori:

- General Purpose: sono processori ad alte prestazioni e versatili, non limitati ad un solo possibile utilizzo. Fanno parte di questo gruppo i processori dei personal computer.
- Special Purpose: sono processori realizzati per compiti specifici. Ne fanno parte i microprocessori delle schede grafiche
- Embedded: sono sistemi a microprocessore progettati per una determinata applicazione, spesso con una piattaforma hardware ad hoc. Sono quindi integrati completamente nel sistema, di cui ne gestiscono tutte o quasi le funzionalità.

Ad oggi il mercato è costituito per il 98% da microprocessori per applicazioni embedded come elettrodomestici, periferiche computer e altro. I microprocessori per personal computer, pur rappresentando solo lo 0.2% del numero totale di pezzi prodotti, assorbono però la metà del fatturato totale.

Sistemi e prestazioni

Slide 3→7

I vari microprocessori sono classificati in base a differenti parametri che possono essere più o meno incidenti a seconda dell'applicazione per cui vengono usati.

	Server	Desktop	Embedded
Prezzo	500\$-5000\$	5000\$-5M\$	10\$-100K\$
Prezzo/μP	50\$-500\$	200\$-10K\$	0,01\$-100\$
Aspetti Critici	<ul style="list-style-type: none">• Rapporto prezzo/prestazioni• Prezzo• Prestazioni Grafiche	<ul style="list-style-type: none">• Throughput• Dependability (MTBF)• Scalabilità	<ul style="list-style-type: none">• Consumo Energetico• Prestazioni in Task Specifici

Nella tabella precedente si può vedere una classificazione in base al prezzo medio sul mercato, che può influire molto nelle applicazioni embedded le quali richiedono un elevato numero di microprocessori dalle basse prestazioni ma altamente specializzati nelle loro funzioni.

Per i server, invece, un parametro che ne guida il progetto architetturale è il MTBF (Mean Time Between Failures), da cui si può determinare un Costo orario di Disservizio, cioè la perdita di denaro che si subisce fintanto che il server non è nuovamente operativo.

Più in particolare $MTBF = MTTF$ (Mean Time To Failure, tempo di individuazione del guasto) + $MTTR$ (Meant Time To Repair, tempo di riparazione del guasto).

Prestazioni ed evoluzione

Slide 8→13

I principali parametri per la valutazione dello sviluppo e delle prestazioni sono:

- Throughput (ampiezza di banda) = (Operazioni, Azioni, Compiti) / Tempo, corrisponde alla quantità di operazioni svolte in un'unità di tempo.
- Latenza, ovvero il tempo impiegato per la completa esecuzione di un'operazione.

Nella valutazione dello sviluppo nel campo della microelettronica è molto importante la Potenza dissipata.

Essa è definita come:

$$P_{Dissipata} = \frac{1}{2} \cdot C_{Carico} \cdot Tensione^2 \cdot frequenza = \frac{Energia}{Tempo}$$

Per esempio, quindi, basterebbe diminuire la tensione di alimentazione per avere una notevole riduzione del consumo.

Inoltre, l'energia dissipata (potenza per tempo) è un parametro essenziale per tutti i sistemi portatili alimentati a batteria (ne determina la durata).

A parità di microelettronica sono stati apportati miglioramenti anche grazie all'utilizzo del parallelismo (numero di pipeline per gestire le istruzioni in parallelo) e lo sfruttamento del principio di localizzazione spaziale (inserimento nella cache delle istruzioni adiacenti in memoria a quella in esecuzione) e temporale (inserimento nella cache delle istruzioni chiamate più frequentemente).

Un errore comune è però considerare la CPU, in quanto a prestazioni, l'elemento più influente dell'intera architettura di un elaboratore. In realtà le prestazioni di un sistema sono determinate anche e soprattutto da:

- Architettura del sistema (BUS, gerarchia delle memorie, I/O, interrupt)
- Prestazioni e tipologia delle periferiche (rete, memorie di massa)
- Software di base

Per esempio in un sistema per applicazioni web il 60% del tempo è occupato dalle operazioni di I/O, mentre solo il restante 40% dall'esecuzione delle istruzioni.

Quindi, sostituendo il processore di questo sistema con uno 10 volte più veloce, lo speed-up, ovvero l'aumento di prestazioni è pari a:

$$SpeedUp = \frac{Prestazioni_{MIGLIORAMENTO}}{Prestazioni_{ORIGINALI}} = \frac{T \cdot (\%Tempo_{I/O} + \%Tempo_{Elaborazione})}{T \cdot (\%Tempo_{I/O} + \frac{\%Tempo_{Elaborazione}}{Fattore_{MIGLIORAMENTO}})}$$

Ecco perchè in qualunque progetto di server il costo del processore non è mai dominante, mentre è molto rilevante quello dei dischi.

Architettura dei microprocessori

Slide 17→26

Un microprocessore è composto da più unità:

- Unità di Controllo + ALU (Algebrical Logic Unit)
- Unità di Decodifica Istruzioni (determina la tempistica dovuta ai tempi di Fetch, Decode ed Execute. Nelle architetture CISC questi tre tempi sono variabili per la diversa lunghezza delle istruzioni, dal numero di microistruzioni necessarie per la loro implementazione e dal tipo di istruzione (1 o più argomenti, ecc...)
- Unità di Gestione Indirizzi
- Registri
- Unità di Gestione BUS

Il chipset, un'appendice della CPU posta di solito sulla scheda madre, si occupa della gestione periferiche, del DMA, ecc...

Per migliorare le prestazioni si sono introdotti dei miglioramenti architetturali: si è scelto di disaccoppiare il bus di sistema da quello della CPU demandando la fase di prelevamento istruzioni alla BIU (Bus Interface Unit) ed inserendo una coda di prefetch, il che consiste nel precaricare l'istruzione successiva del programma in esecuzione; è anche possibile utilizzare una memoria cache L1 apposta per il codice.

Si è inoltre scelto di aumentare il parallelismo di esecuzione introducendo una o più pipeline o utilizzando un'architettura superscalare, cioè un'architettura che permette di eseguire più istruzioni contemporaneamente.

Per quanto riguarda i sistemi multi-core il miglioramento delle prestazioni è conseguito grazie ad un più ampio utilizzo del parallelismo tra thread (strategia Thread Level Parallelism) piuttosto che ad uno tra istruzioni (Instruction Level Parallelism): quest'ultima strategia infatti ha bisogno di più complessità hardware oltre che di nuovi criteri di programmazione.

Differenze architetturali

Slide 27→28

	IA-16 (x86-16)	IA-32 (x86-32)	X64 (x86-64)
Indirizzamenti	16	16,32	16,32,64
Registri	8,16	8,16,32	8,16,32,64
Natività	8086 e 286	386,486,Pentium	Pentium

L'IA-64 non è presente nella precedente tabella in quanto è una architettura Intel completamente differente, chiamata Itanium, basata su tecnologia VLIW (Very Long Instruction Word). Le VLIW sono macroistruzioni già ottimizzate per essere eseguite con il massimo parallelismo.