

La memoria virtuale

Politecnico di Torino
Dip. di Automatica e Informatica



Introduzione

La *Memoria Virtuale* è un meccanismo attraverso il quale il processore

- **può fare accesso ad uno spazio di indirizzamento molto maggiore delle dimensioni della memoria principale realmente disponibile**
- **utilizza *indirizzi logici* che vengono poi trasformati in *indirizzi fisici* (quelli che giungono alla memoria).**

Il meccanismo mira a tenere nella memoria principale i blocchi di memoria che si prevede debbano venire utilizzati a breve dal processore, sfruttando il principio della *località dei riferimenti*.

Gli altri blocchi vengono tenuti nella memoria secondaria.

Indirizzi fisici ed indirizzi logici

Il processore fa riferimento alla memoria attraverso *indirizzi logici*.

L'accesso alla memoria fisica avviene invece tramite *indirizzi fisici*.

Il meccanismo della memoria virtuale crea una corrispondenza *dinamica* tra indirizzi fisici e logici.

La traduzione da indirizzi logici a fisici, attuata durante l'esecuzione dei programmi, viene realizzata in hardware dalla *Memory Management Unit* (MMU).

Funzionamento

Per ogni indirizzo generato dal processore, la MMU provvede a

- **Verificare se la parola richiesta si trova fisicamente nella memoria principale**
- **Se così è**
 - **la MMU produce l'indirizzo fisico corrispondente, con cui si accede alla memoria principale**
- **Se no (*page fault*)**
 - **la MMU richiede l'intervento del Sistema Operativo, che provvede a spostare nella memoria principale il relativo blocco di memoria.**

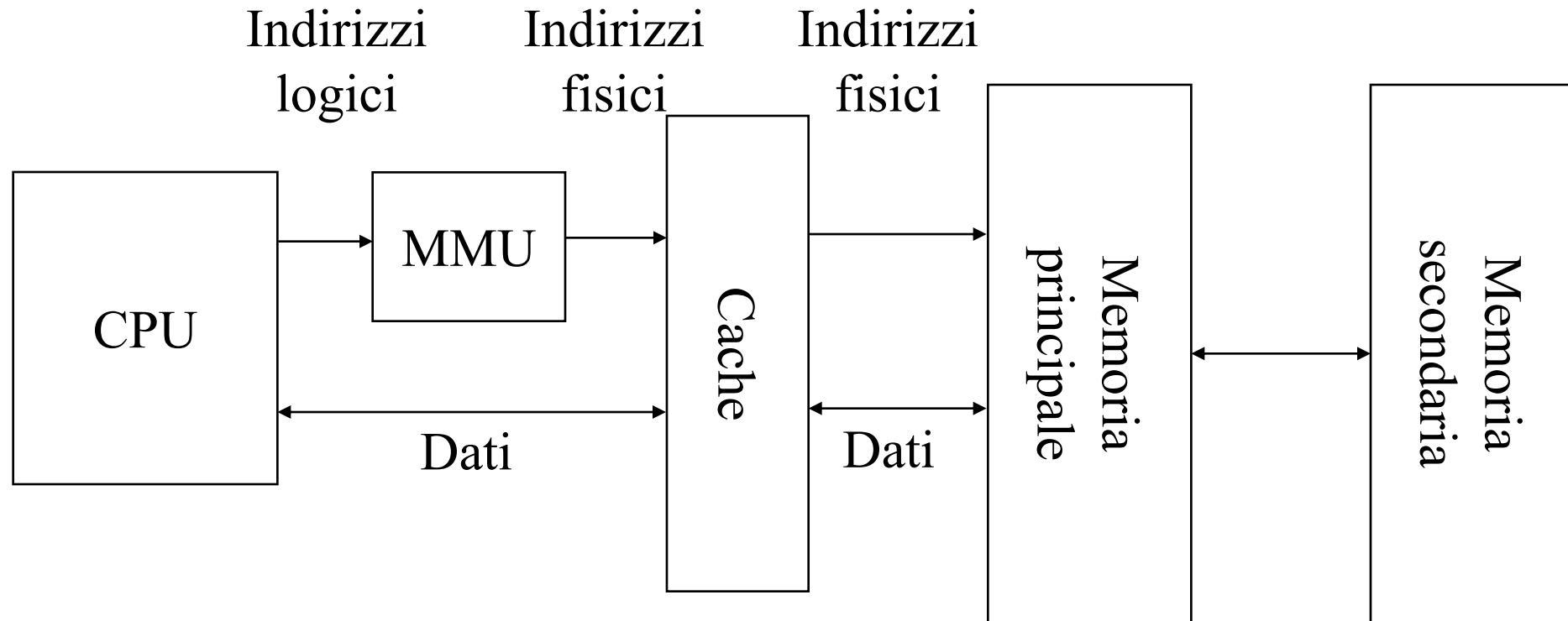
Suddivisione in pagine

Vista la località dei riferimenti, conviene fare in modo che:

- **la memoria sia suddivisa in blocchi, denominati *pagine***
- **quando la CPU fa accesso ad un dato in una pagina non contenuta nella memoria principale, l'intera pagina vi venga trasferita, di solito utilizzando il meccanismo del DMA.**

La dimensione tipica di una pagina varia tra 1 KB e 64 KB.

Architettura generale



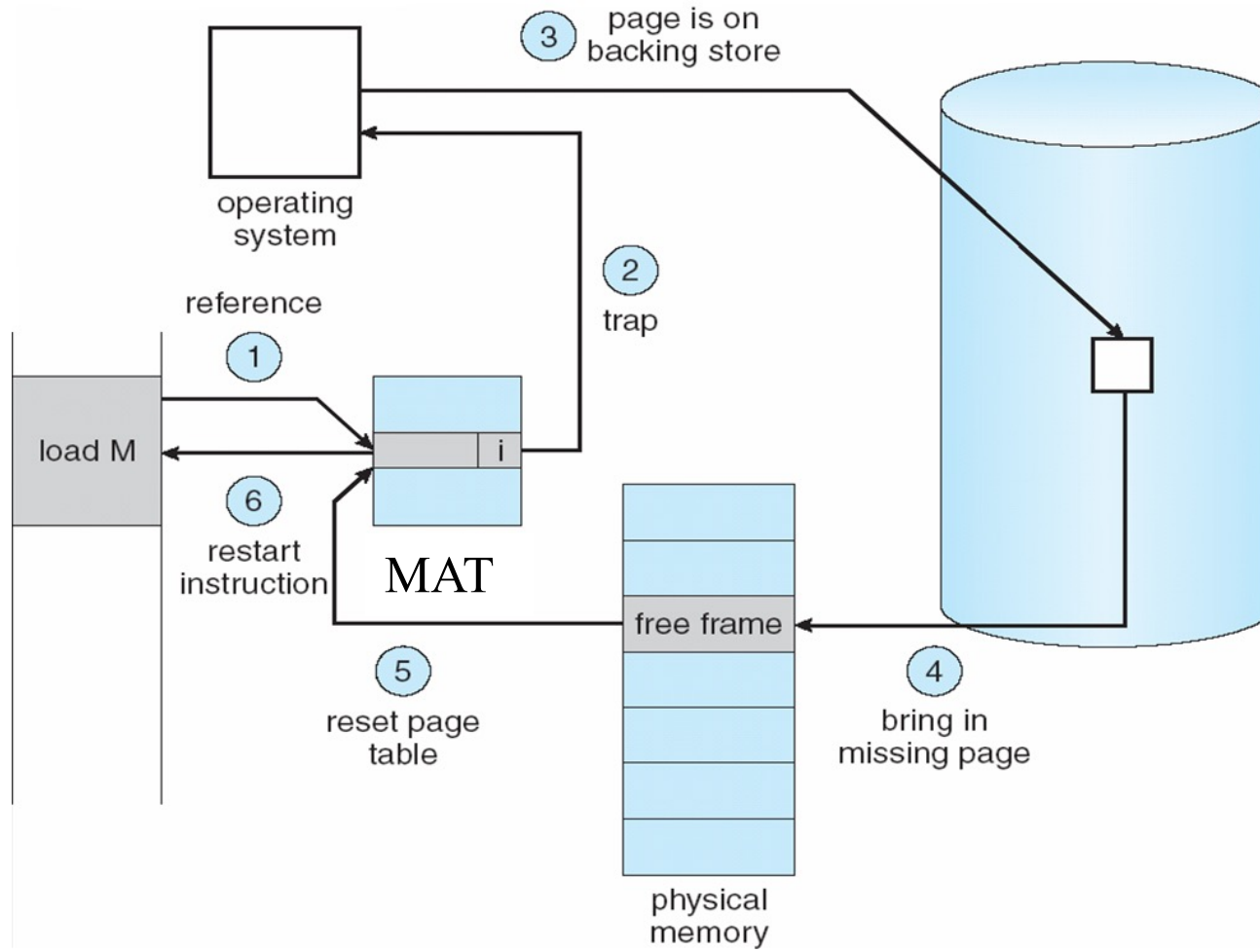
Memory Address Table

Per eseguire la trasformazione da indirizzi logici in fisici la MMU utilizza una tabella detta *Memory Address Table* (MAT).

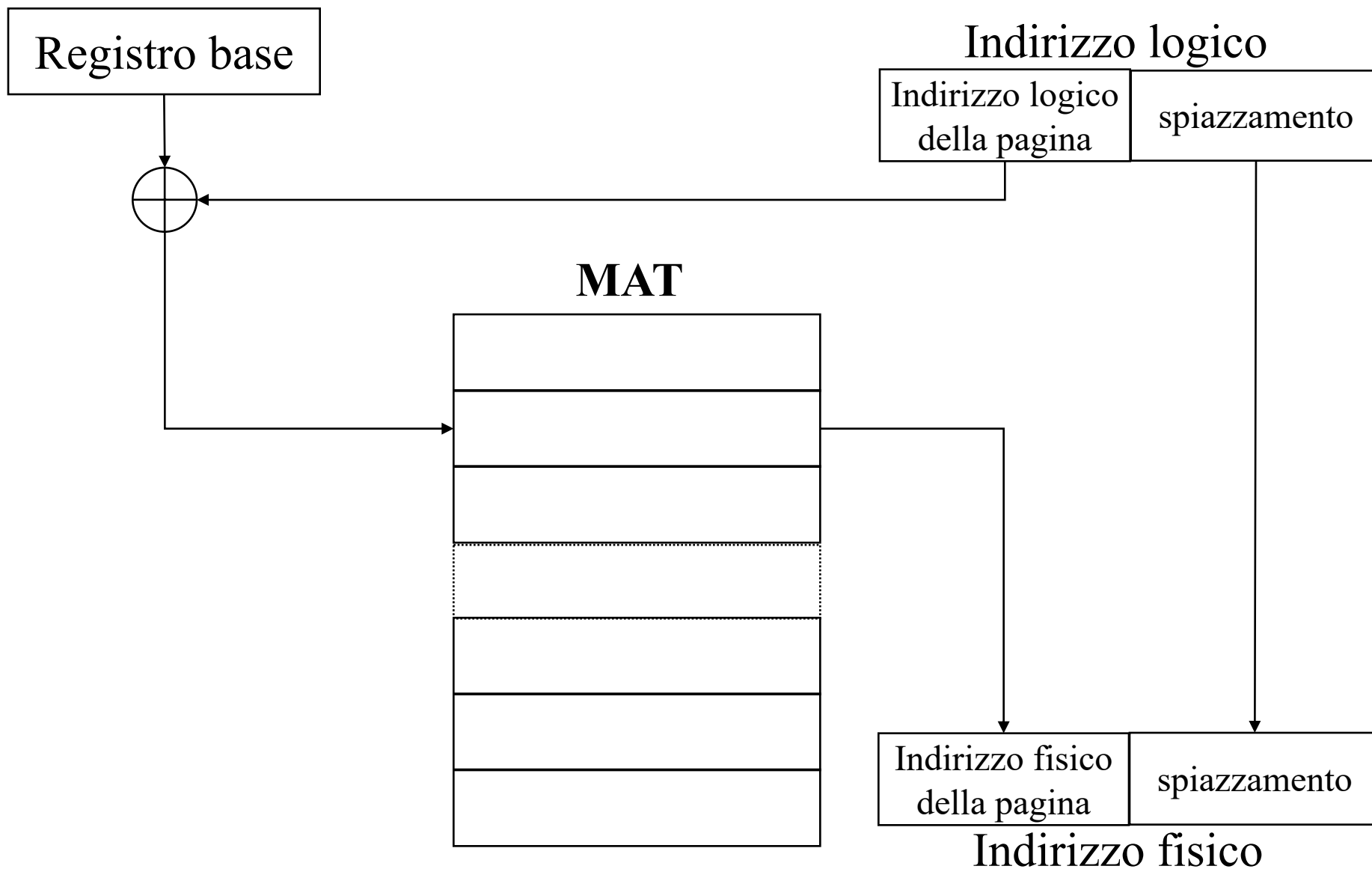
La MAT contiene una entry per ogni pagina virtuale. Per ogni entry la MAT memorizza:

- l'indirizzo fisico al quale la pagina si trova nella memoria principale, oppure**
- la posizione della pagina in memoria secondaria.**

Per accedere rapidamente all'elemento corretto della MAT la MMU somma al contenuto di un registro che contiene l'indirizzo di base della MAT la parte alta dell'indirizzo (corrispondente al numero della pagina virtuale).



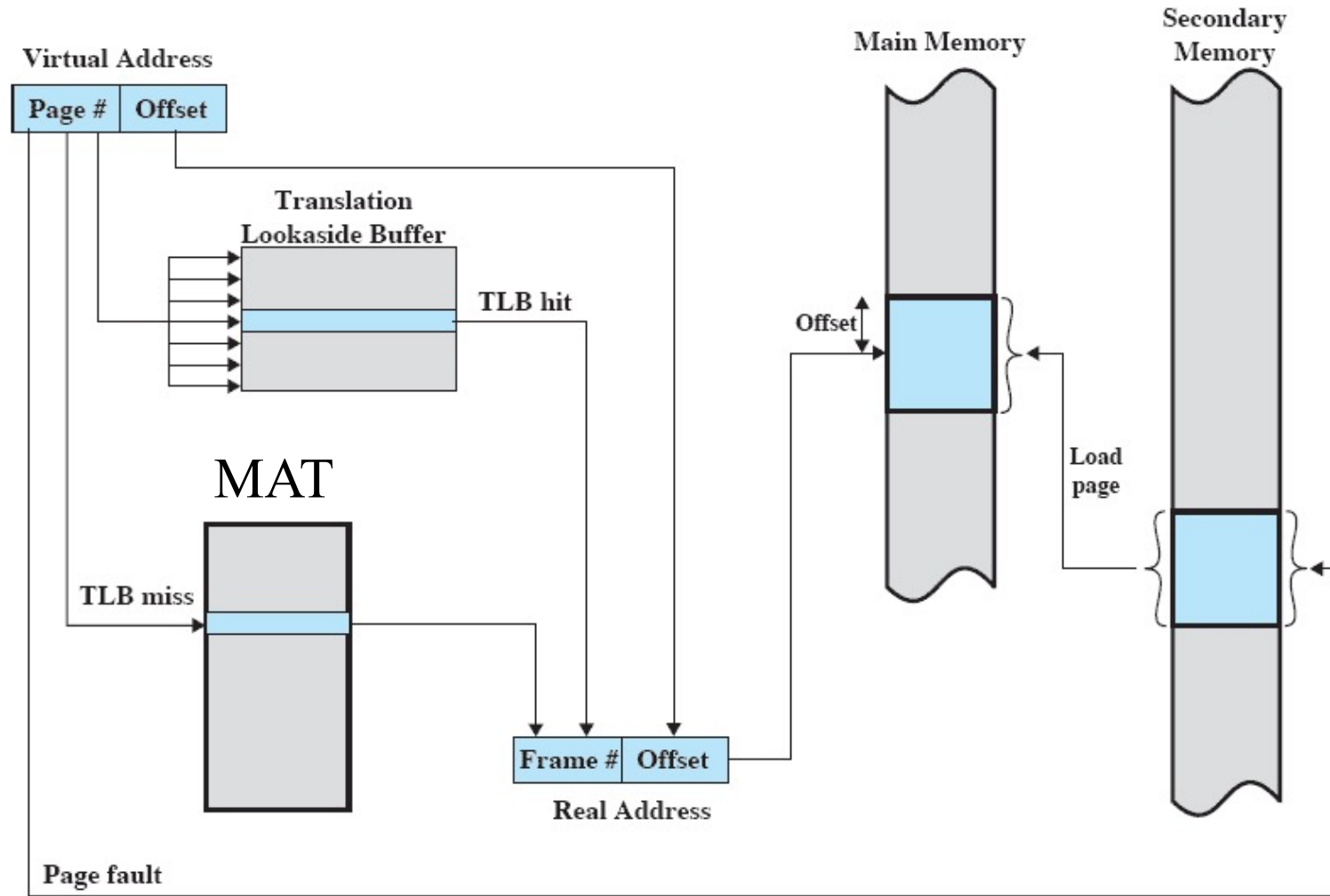
Traduzione degli indirizzi



TLB

- Nei sistemi con memoria virtuale ad ogni accesso in memoria si deve trasformare un indirizzo logico in fisico, facendo quindi accesso alla MAT
- Per velocizzare l'accesso alla MAT, le entry accedute più di recente sono memorizzate in un'apposita memoria denominata *Translation Lookaside Buffer (TLB)* posta direttamente a bordo della MMU
- Il TLB svolge il ruolo di cache della MAT.

TLB



Vantaggi

L'uso della memoria virtuale offre i seguenti vantaggi:

- **rende i programmi indipendenti (dal punto di vista funzionale) dalla configurazione reale di memoria del sistema e dalla posizione reale dei dati in memoria**
- **fa sì che i programmi vedano una memoria complessiva molto ampia, con tempi di accesso ridotti e con basso costo per bit**
- **permette un uso efficiente della memoria in ambienti multi-utente.**

Memoria virtuale rispetto a cache

Benché vi siano molti elementi di somiglianza, le seguenti caratteristiche differenziano il rapporto tra cache e memoria principale da quello tra memoria principale e memoria secondaria:

- rapporto nei tempi di accesso 5:1 per le cache, 1000:1 per la memoria virtuale**
- gestione realizzata interamente tramite hardware apposito per le cache, con il supporto del sistema operativo per la memoria virtuale**
- dimensioni delle pagine: da 4 a 32 byte per le cache, da 1 a 64KB per la memoria virtuale.**

Valutazioni quantitative

- I meccanismi di gestione dei page fault sono implementati via software dal sistema operativo
- il rapporto tra il tempo di accesso alla memoria secondaria e quello alla memoria principale è dell'ordine di $100 \div 1000$.

Quindi

- la durata delle operazioni conseguenti ad un miss è pari a centinaia di migliaia di colpi di clock
- la probabilità di miss deve essere molto bassa (dell'ordine dello 0,1% o minore)
- le pagine devono essere relativamente grandi (tra 1 e 64Kbyte)
- in caso di scrittura, si possono applicare solo meccanismi di tipo write-back.