



# Reti di Interconnessione per sistemi paralleli

# Interconnessione

- ◆ Modello di sistema parallelo cui ci si riferisce:
  - Un insieme di PE dotati di memoria locale, interagiscono per tramite di una rete di interconnessione mediante scambio di messaggi

# Principali indici utilizzati per caratterizzare una rete di interconnessione

- ◆  $L$  lunghezza media di un messaggio iniettato nella rete da un PE
- ◆  $\lambda$  tasso medio di spedizione dei messaggi da parte di un generico PE
- ◆  $T(\lambda, L)$  latenza
  - il tempo medio richiesto per consegnare un messaggio nella rete
  - $T(0, L)$  indica la latenza a rete scarica
- ◆  $\sigma$  throughput
  - il massimo tasso di messaggi,  $\lambda$ , che può essere sostenuto dalla rete (per  $\lambda > \sigma$  la rete si satura e la latenza tende ad infinito)
- ◆ Obiettivo da perseguire nella progettazione di una rete di comunicazione
  - minimizzare la latenza  $T(\lambda, L)$
  - massimizzare il throughput  $s$

# Una rete di comunicazione è caratterizzata da (1)

## ■ Topologia

- descritta dal grafo di interconnessione  $G(N,C)$
- $N$  l'insieme dei vertici che rappresentano i nodi di comunicazione e/o elaborazione
- $C$  l'insieme dei canali fisici che interconnettono i nodi
- $C \subset N \times N$

## ■ Flow Control

- il metodo utilizzato per regolare il traffico nella rete (se due messaggi, ad esempio, si contendono l'utilizzo dello stesso canale, il flow control arbitra l'assegnazione del canale e definisce le sorti del messaggio non selezionato)

# Una rete di comunicazione è caratterizzata da (2)

## ■ Routing

- il metodo usato per selezionare un percorso nella rete
- è una relazione  $C \times N \times N$  che mappa i canali occupati dalla testa di un messaggio ed il nodo di destinazione per il messaggio nell'insieme dei canali che possono essere usati successivamente dal messaggio.
- Il routing è una forma di allocazione a tempo di esecuzione di risorse. Data una topologia, una posizione iniziale ed una destinazione, la relazione di routing determina il percorso che il messaggio dovrà seguire

# Topologia

Parametri che caratterizzano la topologia di una rete di interconnessione:

- B ampiezza di bisezione
  - il minimo numero di canali che, se tagliati, separano la rete in due parti uguali
- $B_w$  wire bisection
  - il numero di linee che attraversano il taglio della rete.  $B_w = B \times W$  essendo  $W$  il parallelismo del canale in bit. Misura la densità delle linee di trasmissione nella rete
- $d$  grado di un nodo
  - il numero di canali afferenti in un nodo scomponibile in  $d_{in} + d_{out}$  canali. La complessità della logica di controllo di un router è legata a  $d$ , il numero di pin in un nodo è  $W \times d$
- $D$  diametro di una rete
  - il più lungo fra i cammini più corti fra tutte le coppie di nodi della rete.  $D$  è stata sempre usata come figura di merito della complessità della rete

# Topologia

- Lunghezza di una connessione
  - la lunghezza della connessione vincola, per i problemi di propagazione dei segnali, la velocità di trasferimento dei dati e la potenza dissipata dal driver per pilotarla
- Simmetria
  - una rete è detta simmetrica se essa è isomorfa a se stessa con ogni nodo marcato come origine. Una rete non simmetrica è detta asimmetrica

# Topologia

## Classificazione delle reti di comunicazione: Reti Dirette ed Indirette

- ◆ Con riferimento alle funzioni di processazione e comunicazione, una rete è detta:
  - Diretta (k-ary n-cube)
    - se tutti i suoi nodi sviluppano entrambe le funzioni di processazione e comunicazione
  - indiretta (k-ary n-fly)
    - se i nodi sono specializzati nelle funzioni di processazione e di comunicazione. In tal caso i nodi di processazione interagiscono per tramite di messaggi scambiati utilizzando una rete di nodi di comunicazione detti “switch”

# Topologia

## Classificazione delle reti di comunicazione: Reti Statiche e dinamiche

Con riferimento al tipo di connessioni una rete può essere

### ■ Statica

- se tutte le connessioni fra i nodi sono basate su link diretti, fissati all'atto della costruzione della rete ( e quindi a topologia assegnata e con pattern di comunicazione fra i nodi prefissati),
- Tale tipo di rete si adatta a sistemi particolarmente atti a risolvere preassegnate classi di problemi che ben si adattano alla topologia della rete stessa,
- Le reti statiche sono per la maggior parte dirette,
- Array lineare, Ring e Chordal Ring, Barrel Shifter, Tree e StarFat Tree, Mesh e Torus, Array Systolici, Hypercube, Cube-Connected Cycles

# Topologia

Classificazione delle reti di comunicazione  
Reti Statiche e dinamiche

## ■ Dinamica

- se le interconnessioni fra i nodi sono ottenute a tempo di esecuzione ricorrendo all'impiego di switch o arbitri in grado di stabilire dinamicamente il percorso dei messaggi,
- Tale tipo di rete si adatta a sistemi in grado di risolvere classi più generali di problemi,
- Le reti dinamiche sono per la maggior parte indirette,
- Bus, Reti di interconnessione multistadio (Omega, Baseline)

# Topologia

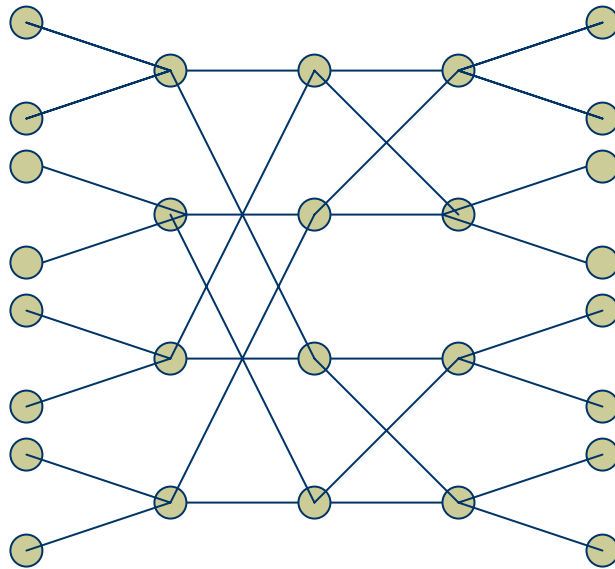
## Reti indirette k-ary n-fly

### Parametri delle reti k-ary n-fly

- $n$  è il numero di stadi di cui è composta la rete (la dimensione)
- $k$  è la radice, il numero di nodi di processazione afferenti al nodo di comunicazione
- $N = k^n$ ,  $k = \sqrt[n]{N}$ ,  $n = \log_k N$
- $B_i = N/2$
- $B_{wi} = NW/2$
- $d_{in} = d_{out} = k$
- $d = 2k$
- $D = n + 1$

# Topologia

## Esempio di rete 2-ary 3-fly



Rete indiretta

2-ary 3-fly

◆ Parametri della rete 2-ary n-fly:

- $n=3$
- $k=2$
- $N = k^n=8$
- $B_i=N/2=4$
- $B_{wi}=NW/2=2W$
- $d_{in}=d_{out}=k=2$
- $d=2k=4$
- $D=n+1=4$

# Topologia

## Caratteristiche di una rete indiretta (I)

- Elevata ampiezza di bisezione  $B$  ( $N/2$ )
- basso grado ( $2k$ )
- basso diametro ( $n+1$ )
- connessioni di lunghezza elevata
- simmetriche
- $B$  non riflette l'effettiva densità massima delle connessioni per tali tipi di rete

# Topologia

## Caratteristiche di una rete indiretta (II)

- il taglio verticale della rete seleziona  $N$  fili di connessione, valore elevato per quel che concerne l'implementazione,
- La superficie richiesta per la realizzazione di una tale rete è proporzionale ad  $N^2$  ciò le rende particolarmente costose in termini di fili di connessione Le reti indirette presentano la stessa complessità di realizzazione dei crossbar switch che hanno pochi switch ma richiedono lo stesso ammontare di fili
- $B$  è indipendente da  $k$  ed  $n$  per cui non è consentito al progettista di scegliere fra ampiezza di banda e diametro della rete

# Topologia

## Reti Dirette k-ary n-Cube

### ■ Reti k-ary n-Cube

- Ring, mesh, tori, binary n-cubes (hypercubes) e reti Omega sono fra loro topologicamente isomorfe ad una famiglia di reti dette k-ary n-cube in cui
  - ◆  $n$  è la dimensione del cubo
  - ◆  $k$  è la radice, il numero di nodi (molteplicità) lungo un'assegnata dimensione
    - $N = k^n$ ,                       $k = \sqrt[n]{N}$                        $n = \log_k N$
- Un nodo in un k-ary n-Cube può essere identificato da un indirizzo composto da  $n$  digit in base  $k$   
 $\text{ind} = a_0 a_1 a_2 \dots a_n$  essendo  $a_i$  la posizione del nodo nell' $i$ -esima dimensione del cubo

# Topologia

## Reti dirette k-ary n-cube

### Parametri delle reti k-ary n-cube

- ◆  $n$  è il numero di stadi di cui è composta la rete (la dimensione)
- ◆  $k$  è la radice, il numero di nodi di processazione afferenti al nodo di comunicazione
- ◆  $N = k^n$ ,  $k = \sqrt[n]{N}$ ,  $n = \log_k N$
- ◆  $Bd = 2N/k = 2k^{n-1}$
- ◆  $Bwd = 2NW/k$
- ◆  $d_{in} = d_{out} = n$
- ◆  $d = 2n$
- ◆  $D = nk/2$
- ◆ Ogni nodo ha un indirizzo espresso mediante un numero di  $n$  digit, base  $k$

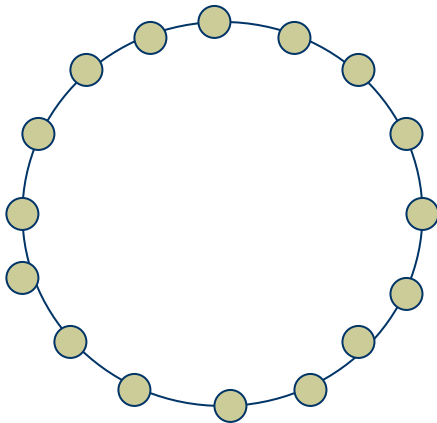
# Topologia reti dirette

- i nodi sono interconnessi ai nodi vicini che hanno indirizzi che differiscono solo per un digit,
- La famiglia delle reti k-ary n-cube include numerose tipologie di rete ampiamente utilizzate in computer message passing:
  - k=2 n-cubi binari
  - n=1 ring
  - n=2 tori
- utilizzate in numerosi sistemi commerciali di CM, Intel, MPP, Illiac nelle diverse forme di ipercubi, mesh, etc.

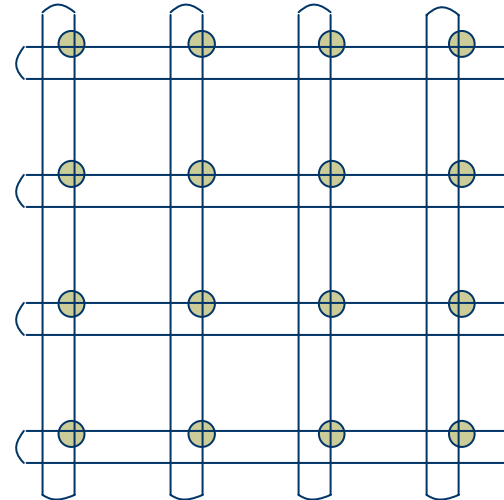
# Topologia

## Esempi di reti dirette

**Array Lineare: 8-ary 1 cube**



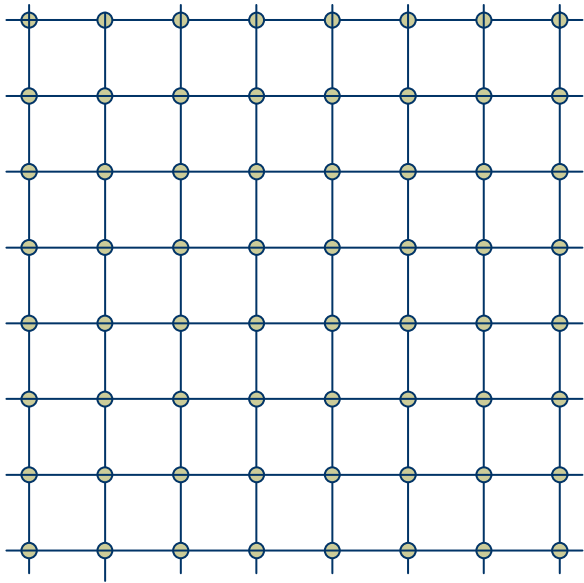
**Ring: 16-ary 1-cube**



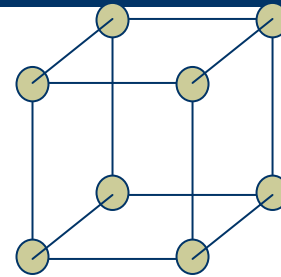
**Toro 4-ary 2-cube**

# Topologia

## Esempi di reti k-ary n-cube



Mesh: 8-ary 2-cube



2-ary 3-cube

$$N = k^n = 2^3 = 8$$

$$Bd = 2N/k = 8$$

$$Bwd = 2NW/k = 8W$$

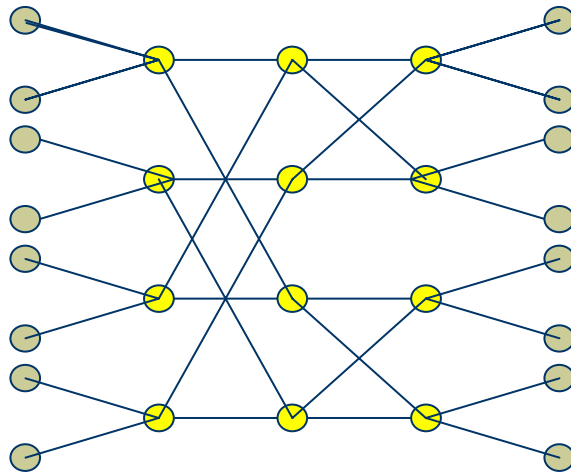
$$din = dout = n = 3$$

$$d = 2n = 6$$

$$D = nk/2 = 3$$

# Topologia

## Confronto reti indirette e dirette



2-ary 3-fly

$$n=3$$

$$k=2$$

$$N = k^n = 8$$

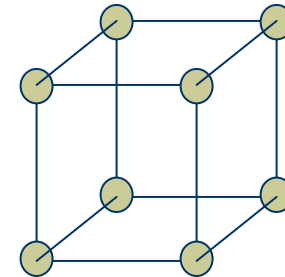
$$B_i = N/2 = 4$$

$$B_{wi} = NW/2 = 2W$$

$$d_{in} = d_{out} = k = 2$$

$$d = 2k = 4$$

$$D = n + 1 = 4$$



2-ary 3-cube

$$n=3$$

$$k=2$$

$$N = k^n = 2^3 = 8$$

$$B_d = 2N/k = 8$$

$$B_{wd} = 2NW/k = 8W$$

$$d_{in} = d_{out} = n = 3$$

$$d = 2n = 6$$

$$D = nk/2 = 3$$

# Topologia

## Caratteristiche di una rete diretta

- Per valori bassi di  $n$  le reti  $k$ -ary  $n$ -cube sono caratterizzate da un'ampiezza di bisezione  $B$  bassa e controllabile (con  $2N/k$ ), basso grado ( $2n$ ), elevato diametro ( $nk/2$ ) e connessioni corte.
- Le versioni toroidali di dette reti ( con gli estremi fra loro connessi) sono simmetriche mentre le altre sono asimmetriche.
- Le reti  $k$ -ary  $n$ -cube sono semplici per quel che concerne la realizzazione delle connessioni. La loro complessità è proporzionale ad  $N$  ( $N^2$  per quelle indirette)
- Se si sceglie un  $n \leq 3$  si ha il vantaggio aggiuntivo di avere molte connessioni corte; ciò vuol dire maggiore velocità operativa e minore potenza dissipata

# Topologia

- L'ampiezza di bisezione di una rete  $k$ -ary  $n$ -cube dipende dalla scelta di  $k$  ed  $n$  e ciò consente al progettista di poter scegliere in modo ottimo fra ampiezza di banda e diametro della rete
- Per messaggi lunghi
  - ampiezza di banda domina la latenza, ciò si ottiene scegliendo reti con  $k$  elevato ed  $n$  piccolo
- Per messaggi corti la scelta va invertita privilegiando un diametro piccolo a fronte di una banda inferiore
- La latenza è minimizzata per una scelta di  $k$  ed  $n$  che minimizzano le due componenti della latenza

# Topologia

## Confronto fra differenti reti statiche

Tipo di rete	Grado	diametro	# di link	B	Simmetr.	size
Array lineare	2	$N-1$	$N-1$	1	No	$N$ nodi
Ring	2	$\lfloor N/2 \rfloor$	$N$	2	Si	$N$ nodi
Completo. connesso	$N-1$	1	$N(N-1)/2$	$(N/2)^2$	Si	$N$ nodi
Albero binario	3	$2(h-1)$	$N-1$	1	No	$h = \log_2 N$ alt. tree
Star	$N-1$	2	$N-1$	$\lfloor N/2 \rfloor$	No	$N$ nodi
Mesh 2D	4	$2(r-1)$	$2N-2r$	$r$	No	$r \times r$ mesh $r = \sqrt{N}$
Illiac Mesh	4	$r-1$	$2N$	$2r$	No	Eq. a cord. ring $r = \sqrt{N}$
Toro 2D	4	$\lfloor r/2 \rfloor$	$2N$	$2r$	Si	$R \times r$ toro $r = \sqrt{N}$
Ipercubo	$n$	$n$	$Nn/2$	$N/2$	Si	$N$ nodi $n = \log_2 N$
CCC	3	$2k-1 + \lfloor k/2 \rfloor$	$3N/2$	$N/(2k)$	Si	$N = k \times 2^k$ Cycle $k \geq 3$
k-ary n-cube	$2n$	$n \lfloor k/2 \rfloor$	$Nn$	$2k^{n-1}$	Si	$N = k^n$ nodi

# Routing

- ◆ Routing è il metodo utilizzato per selezionare un percorso nella rete fra nodo sorgente  $N_s$  ed uno destinatario  $N_d$  di un messaggio,
- ◆ Formalmente il Routing è definito da una relazione  $R$  che identifica tutti i percorsi possibili fra  $N_s$  ed  $N_d$  ed una funzione  $\rho$  che seleziona un percorso fra quelli possibili
  - ◆  $R \subset C \times N \times C$
  - ◆  $\rho: P(C) \times \alpha \rightarrow C$
- ◆ Ad ogni passo del routing la funzione  $\rho$ , a partire dall'insieme dei successivi canali possibili  $P(C)$  e da informazioni aggiuntive sullo stato della rete  $\alpha$  seleziona un particolare canale  $C$ ,
- ◆ Le informazioni di stato di cui si tiene conto in  $\alpha$  possono essere:
  - costanti
  - casuali
  - di stato sul traffico della rete

# Routing

## ◆ Tipi di Routing:

### ■ Deterministico:

- il percorso è definito esclusivamente da  $N_s$  ed  $N_d$
- $R$  è una funzione
- $\alpha$  è una costante
- esempi: e-cube (torus chip), mesh routing chip

### ■ Oblivious:

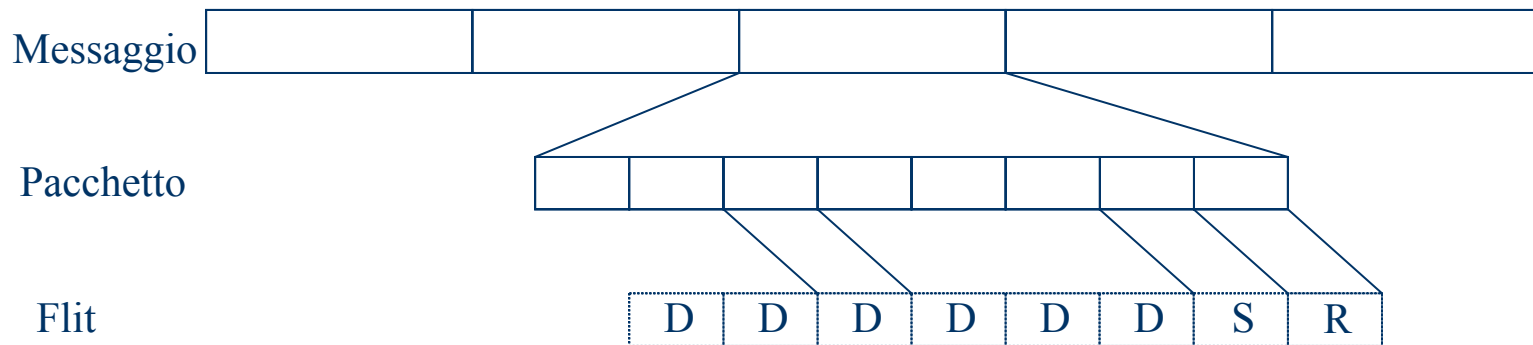
- La scelta del percorso è fatta senza alcuna informazione sullo stato della rete
- $R$  è una relazione
- $\alpha$  è una funzione random, del tempo, del contenuto di un messaggio, etc.

### ■ Adattativo:

- La scelta è fatta a partire dalle informazioni sullo stato della rete
- $R$  è una relazione
- $\alpha$  è una qualunque funzione

# Flow Control

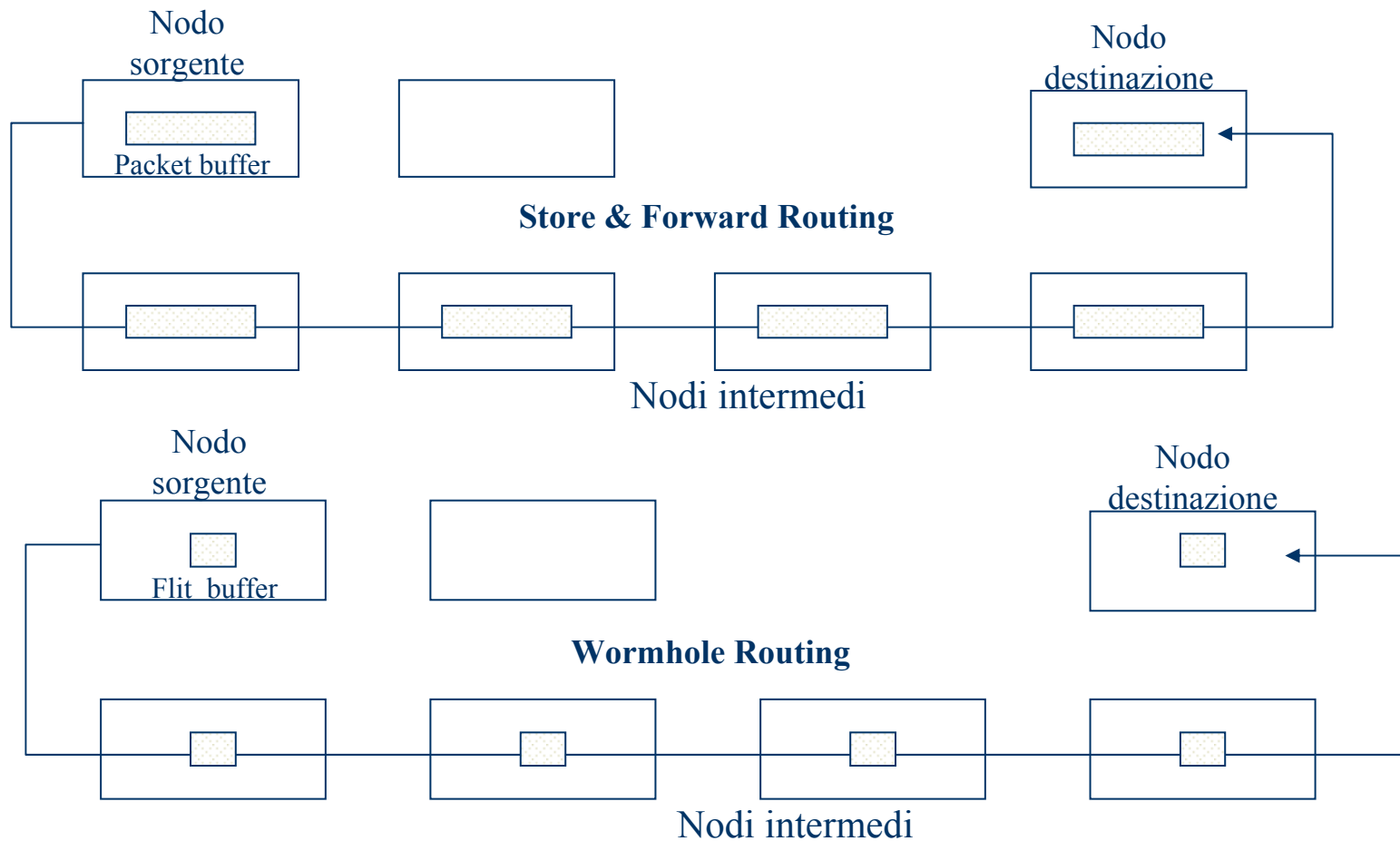
- ◆ Il Flow Control è il metodo utilizzato per regolare il traffico in una rete
- ◆ Permette di allocare le risorse di comunicazione, i link ed i buffer alle unità di informazione (messaggi, pacchetti flit)
- ◆ Formato dei messaggi, pacchetti e flit



# Flow Control

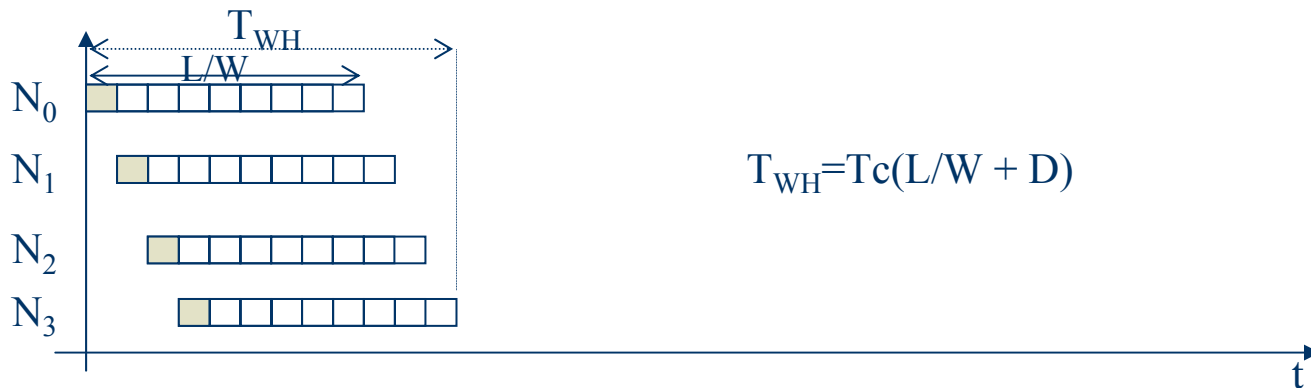
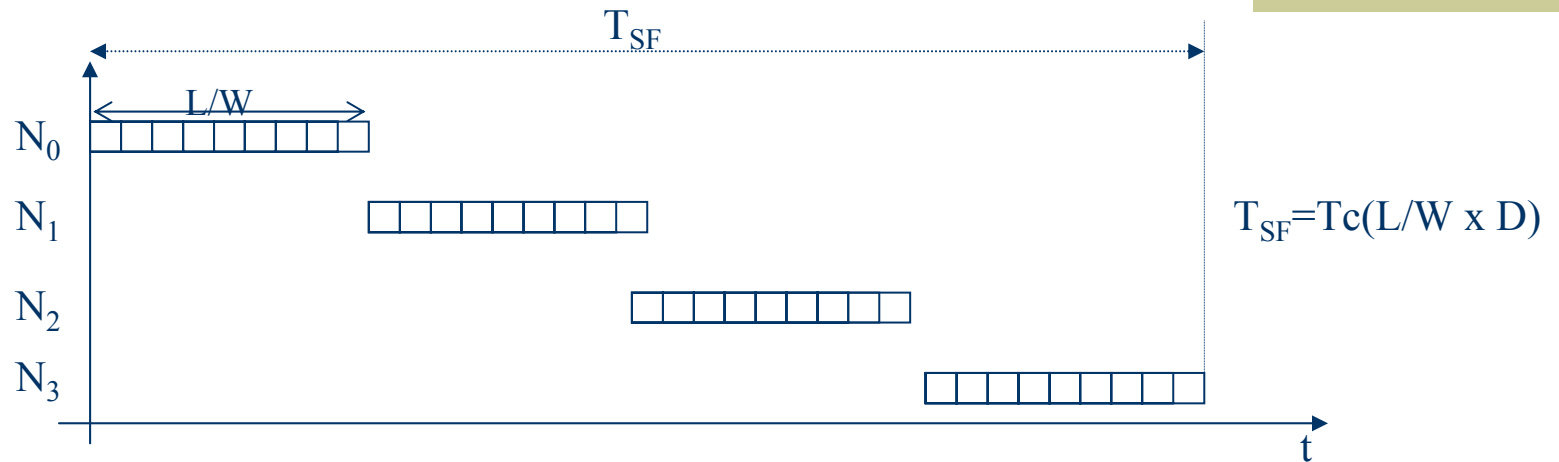
- ◆ Messaggio
  - è l'unità logica per la comunicazione fra nodi
- ◆ Pacchetto
  - è la più piccola unità che contiene informazioni di routing (ad es. dest addr); può essere inviato in modo asincrono ed essere ricevuto fuori sequenza (S è il numero di sequenza per riassemblare i pacchetti)
- ◆ Flit (flow control digit)
  - è la più piccola unità inviata e gestita dal flow control. Non contiene informazioni di routing, soltanto il pacchetto di testa di un messaggio le contiene. Le risorse di comunicazione sono allocate con riferimento ai flit
- ◆ Tipologie di routing
  - Store-and-Forward
    - l'unità minima scambiata è il pacchetto a cui deve essere allocato un opportuno buffer
  - Wormhole

# Flow Control



# Analisi della latenza

## Store-and-Forward e Whormhole



# Flow Control

- ◆ Store & Forward routing
  - è non utilizzabile per un multicomputer stante l'elevata latenza ed il consumo di risorse (buffer delle dimensioni di un pacchetto da allocare)
- ◆ Wormhole routing
  - implementa una sorta di pipeline che riduce la latenza alla somma delle due componenti e non al prodotto. Ciò rende possibile di effettuare un trade off fra le due componenti della latenza per tramite di una scelta appropriata di  $k$  ed  $n$  per le reti dirette.
- ◆ Virtual cut-through
  - strategia ibrida che alloca i buffer di memoria ai pacchetti (come nello store & forward) ma sviluppa la trasmissione come nel wormhole

# Flow Control

- ◆ I metodi di flow control consentono l'avanzamento del flit se:
  - deve essere residente in un buffer del nodo sorgente,
  - deve essere allocato un buffer vuoto nel nodo di destinazione,
  - deve essere allocato l'uso del canale fisico,
- ◆ Contesa di un canale:
  - A quale pacchetto viene allocato il canale
  - Cosa fare dell'altro pacchetto
  - 4 strategie
    - Buffering
    - Blocking
    - Dropping
    - Misrouting

# Deadlock

- ◆ La disciplina del flow control ha come obiettivo quello di allocare risorse ai pacchetti in modo tale da evitare i deadlock
- ◆ Il deadlock può manifestarsi quando c'è una dipendenza ciclica per le risorse
- ◆ Per eliminare gli stalli occorre introdurre un meccanismo per interrompere tale dipendenza
- ◆ Structured Buffer Pool
- ◆ Virtual Channel
  - è un link logico fra i due nodi  $N_s$  ed  $N_d$
  - è realizzato mediante un flit buffer nel nodo sorgente  $N_s$ , un canale fisico fra i nodi stessi, un flit buffer in  $N_d$
  - alcuni canali per la trasmissione dello stato/comando (handshaking)
  - il canale fisico è condiviso fra tutti i virtual channel che interessano i due nodi