Reti simmetriche e calcolo parallelo

Giorgio Gambosi

Dipartimento di Matematica – Universita' di Roma "Tor Vergata"

Sistemi di calcolo parallelo

- Composti da un insieme di unita' di elaborazione, di elementi di memoria e di una infrastruttura di collegamento
- I processori si ripartiscono il carico di lavoro.
- Necessitano comunque di comunicare tra loro (e con gli elementi di memoria)
- La comunicazione avviene scambiandosi dati (messaggi) mediante l'infrastruttura di collegamento

Calcolo in parallelo

- Se più processori devono accedere alla stessa memoria le reti di interconnessione sono necessarie per trasmettere i dati
  - tra processori e elementi di memoria condivisa
  - tra processori diversi
- Molti dati devono essere scambiati contemporaneamente.
- Reti di interconnessione diverse determinano macchine e algoritmi diversi

Importanza dell'infrastruttura di comunicazione

- Le esigenze di comunicazione tra processori dipendono dal procedimento da eseguire (e dal problema da risolvere)
- Decomponibilita' del problema
- In generale, le esigenze di comunicazione sono considerevoli
- La comunicazione rappresenta spesso un collo di bottiglia per il numero di passi eseguiti

Importanza dell'infrastruttura di comunicazione

- Considerazioni sull'infrastruttura di comunicazione determinano valutazioni (inferiori) sul numero di passi da eseguire
- Le considerazioni sull'infrastrutture sono di carattere topologico

- Infrastruttura come grafo
  - Nodi: processori
  - Archi: collegamenti tra coppie di nodi

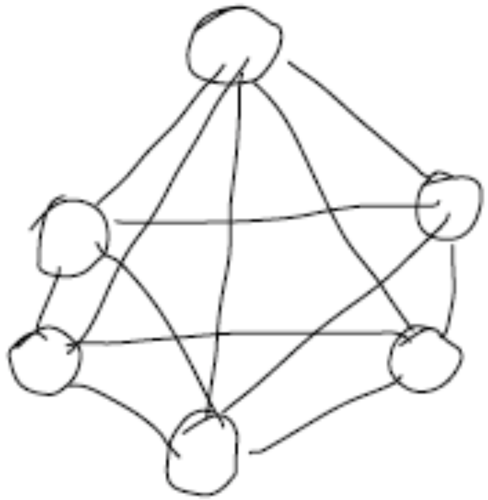
Caso estremo

- Canale (bus) condiviso

- Nessun parallelismo: una sola comunicazione alla volta
- La banda trasmissiva sul canale deve crescere con il numero di processori

Caso ideale

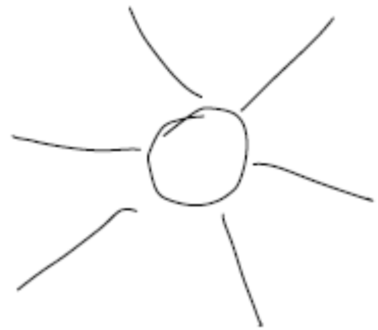
- Rete completa



- Parallelismo totale: tutti possono comunicare contemporaneamente
- Numero quadratico di collegamenti

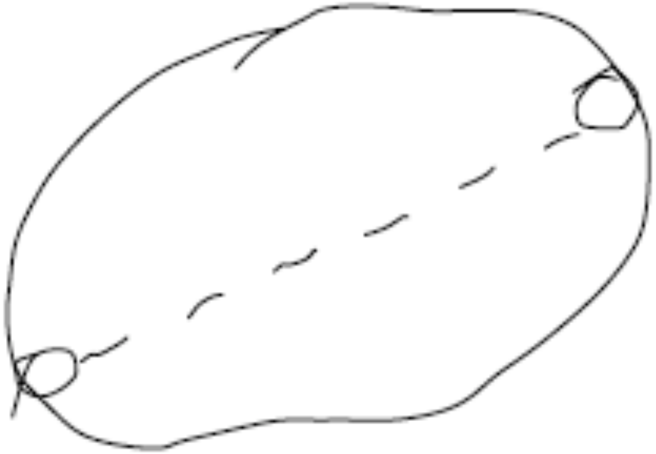
Misure importanti

- Grado (numero di vicini collegati ad un processore).
- Per la rete: grado massimo
- Desiderabile: grado indipendente da numero processori (o crescente in modo lento)



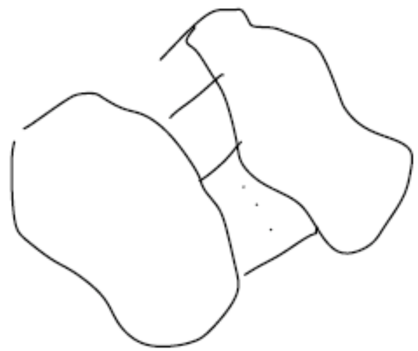
Misure importanti

- Diametro (massima distanza tra due nodi)
- Desiderabile basso: determina il massimo tempo necessario per la comunicazione tra due processori
- Fornisce un limite inferiore al numero di passi di algoritmi che richiedono che processori qualunque possano interagire



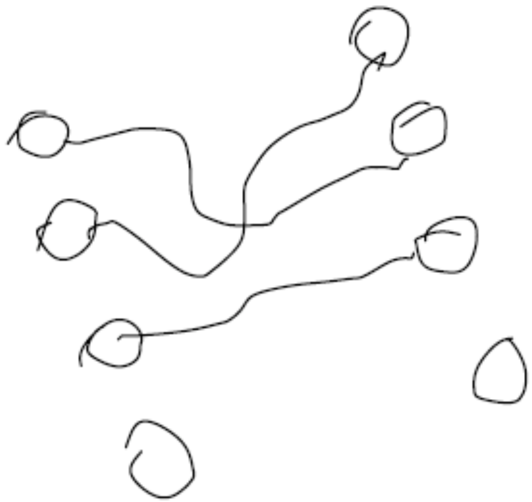
Misure importanti

- Connettivita' (minimo numero di collegamenti da rimuovere per disconnettere la rete)
- Desiderabile alta
- Correlata al livello di parallelismo raggiungibile



Misure importanti

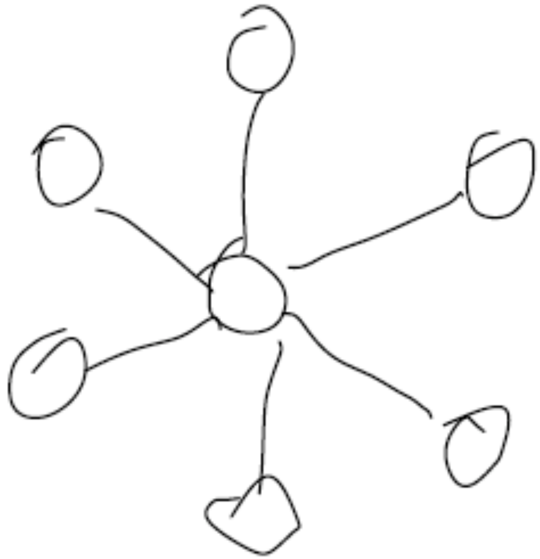
- Larghezza di bisezione (massimo numero di comunicazioni attivabili contemporanemante tra n/2 e n/2 nodi)
- O anche: numero minimo di collegamenti da rimuovere per disconnetere due insiemi di n/2 nodi
- Desiderabile alta



- Per algoritmi che spostano molti dati: almeno un numero di passi pari a quantita' di dati / bisezione

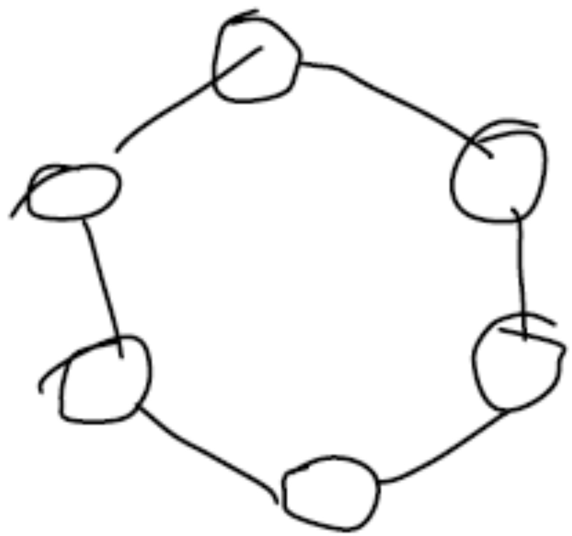
Rete a stella

- Come il canale comune
- n-1 collegamenti
- Grado: n-1 ☹
- Diametro: 2 ☺
- Connettivita': 1 ☹
- Bisezione: n/2 ☺



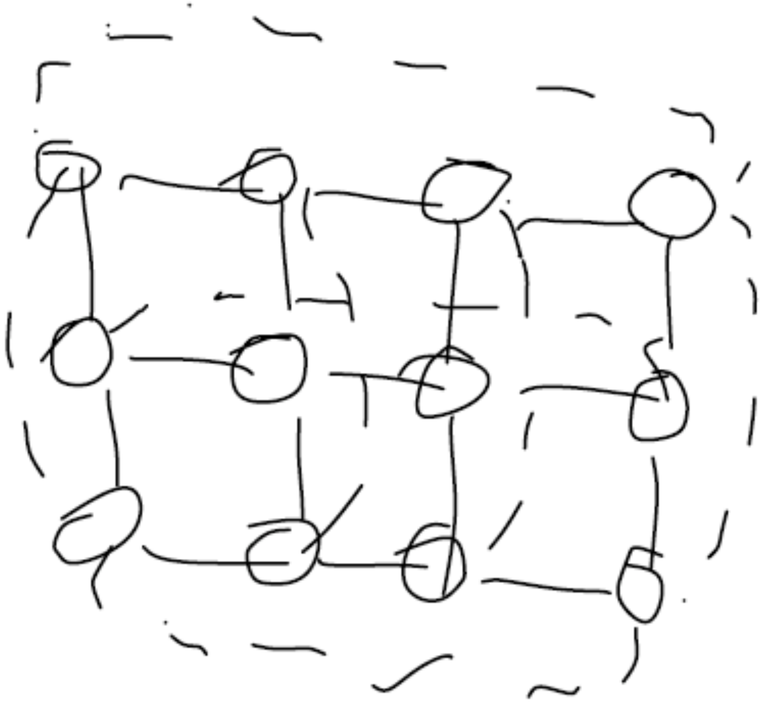
Anello

- n-1 collegamenti
- Grado: 2 ☺
- Diametro: n/2 ☹
- Connettivita': 2 ☹
- Bisezione: 2 ☹

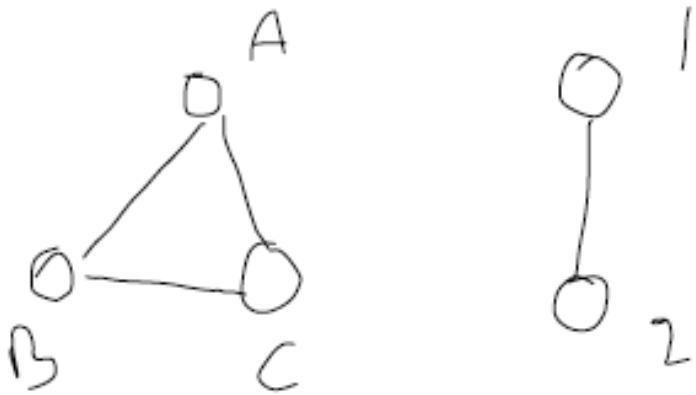


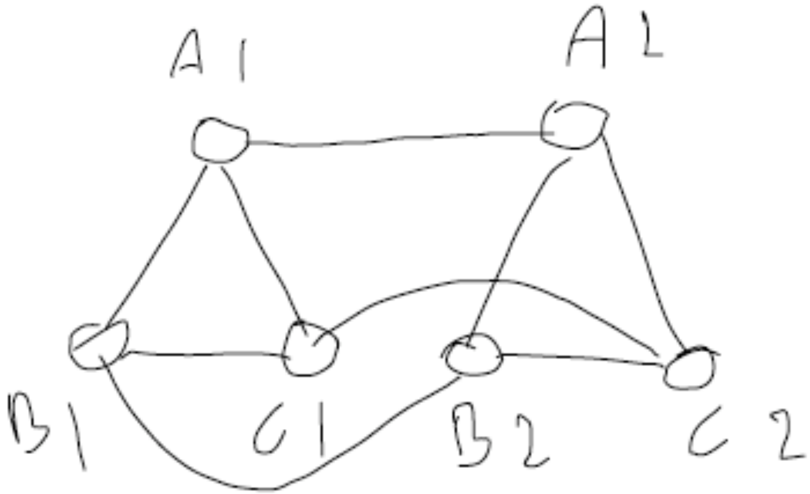
Mesh/Toro

- 4n collegamenti
- Grado: 4 ☺
- Diametro: $\sqrt{n}$ 😐
- Connettivita': 4 ☹
- Bisezione: $\sqrt{n}$ ☺

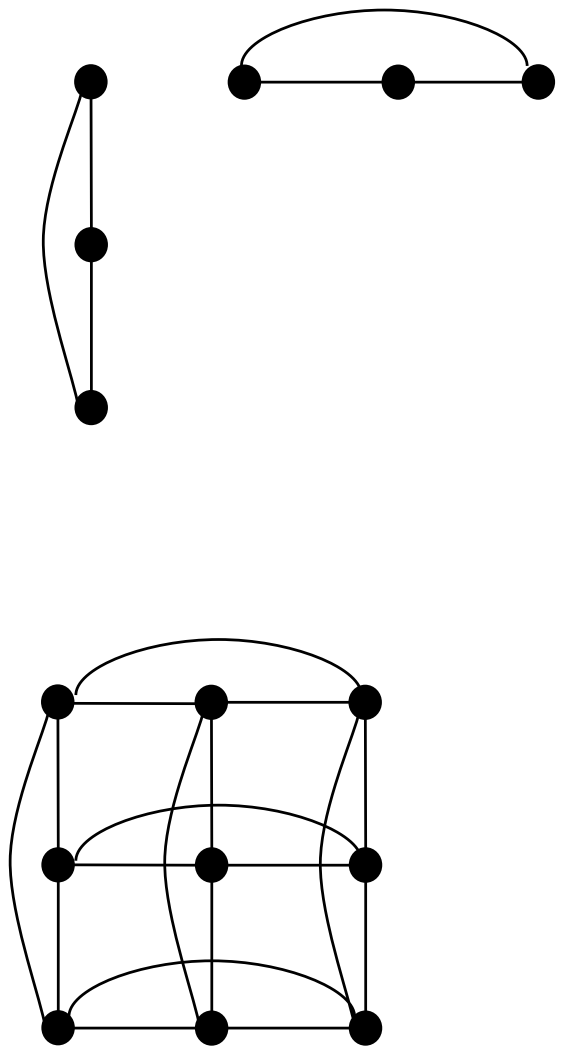


Prodotto cartesiano di grafi

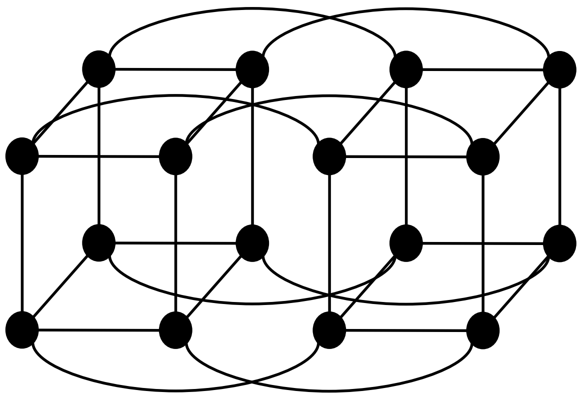


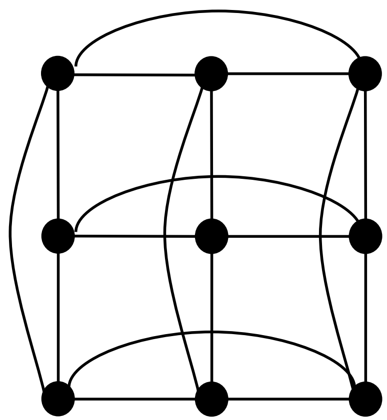


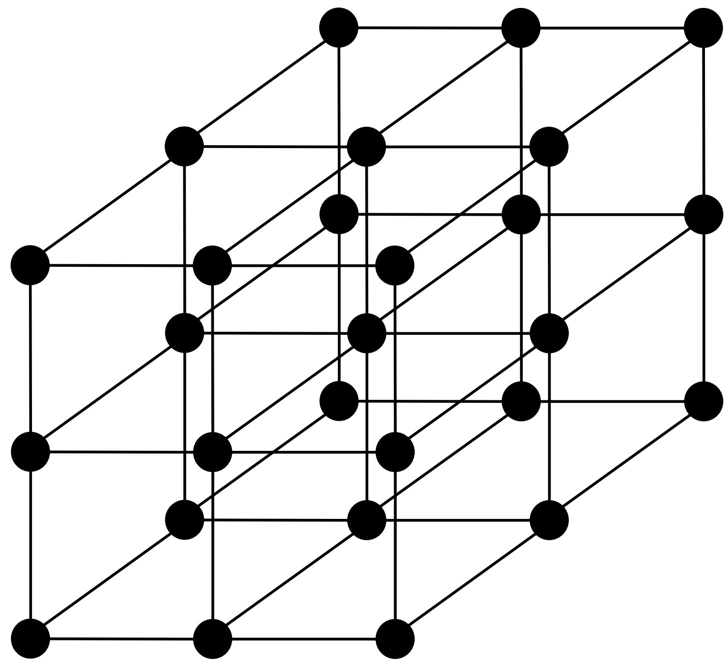
Mesh prodotto di anelli



Reti prodotto

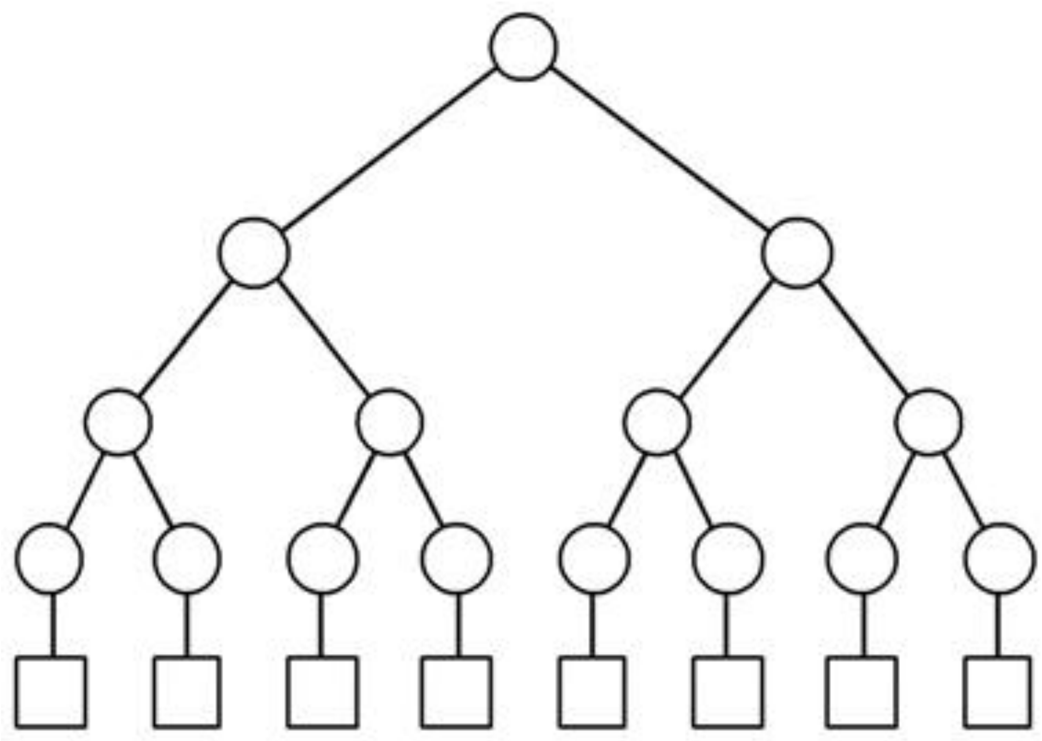






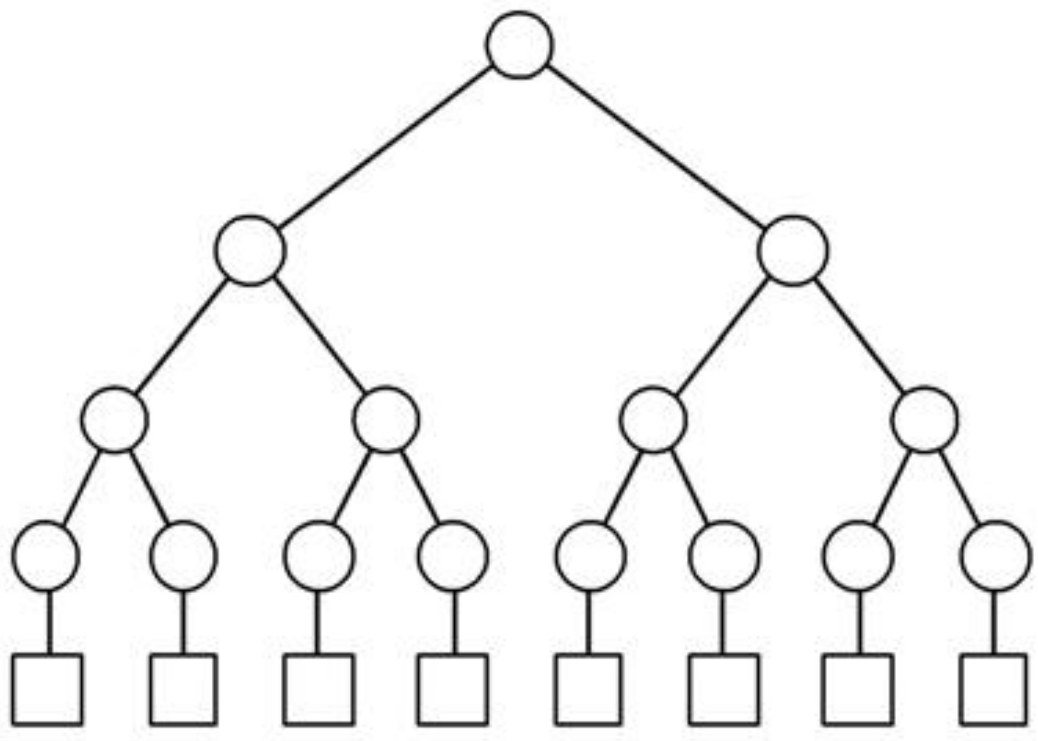
Albero binario

- $n = 2^d$ processori, $2n-1$ nodi aggiuntivi (switch)
- d= numero di livelli

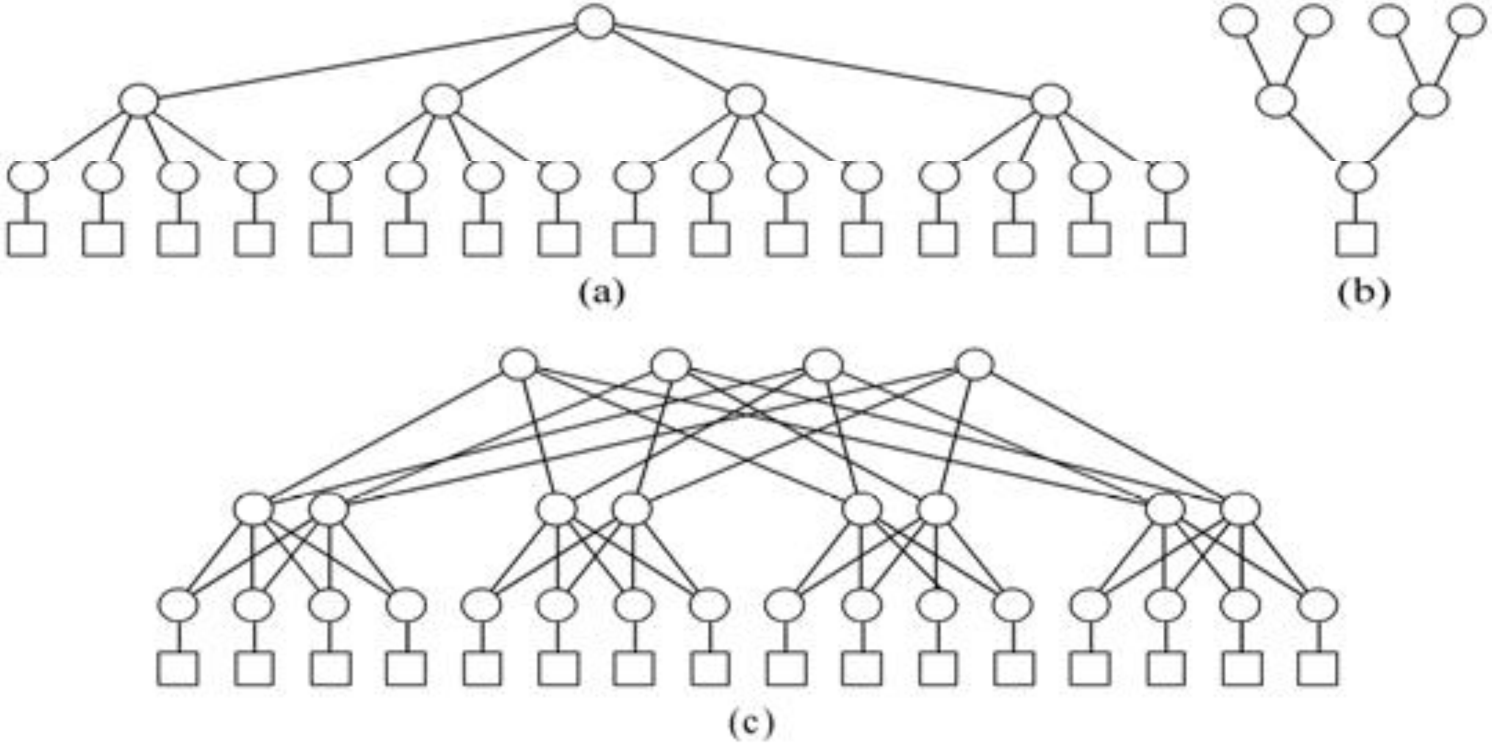


Alberi binari

- Diametro: $O(\lg n)$ ☺
- Bisezione: 1 ☹
- Connettivita': 1 ☹
- Grado: 3 ☺



Hypertree (grado4, altezza 2)



(a) Di fronte: albero di grado 4 e altezza 2

(b) Di lato: albero binario di altezza 2 a testa in giu'

Hypertree

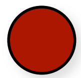
- $n = 4^d$ processori
  - In generale, $n = k^d$ per hypertree di grado k
- Diametro: $2d = 2 \log n$ ☺
- Connettivita': 1 ☹
- Bisezione: 2^{d+1} ☺
- Grado: 6 ☺

Hypercube

- d dimensioni
- $N=2^d$ processori

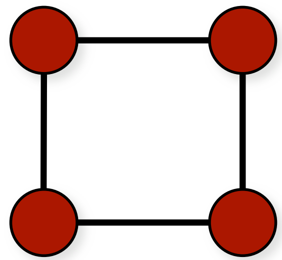
- Ogni nodo: etichetta binaria di d bit
- Collegamento tra nodi con etichette diverse per un solo bit: distanza di Hamming =1

Hypercube

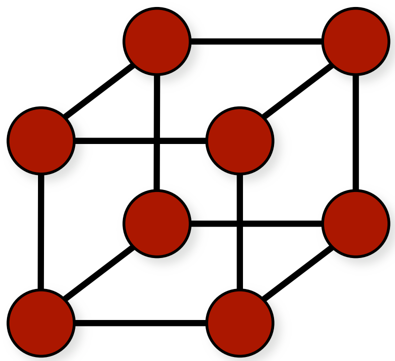


$d = 0$
$N = 1$

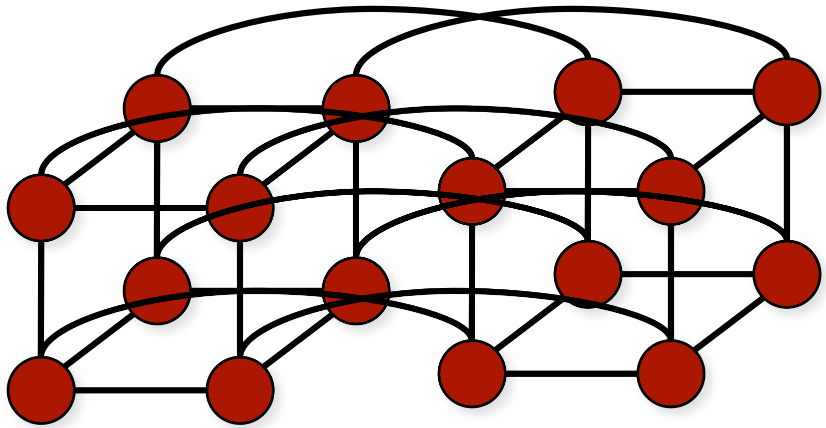
$d = 1$
$N = 2$



$d = 2$
$N = 4$



$d = 3$
$N = 8$

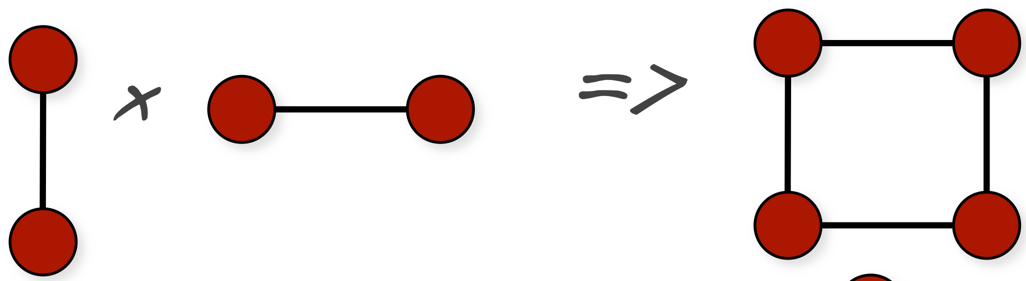


$d = 4$
$N = 16$

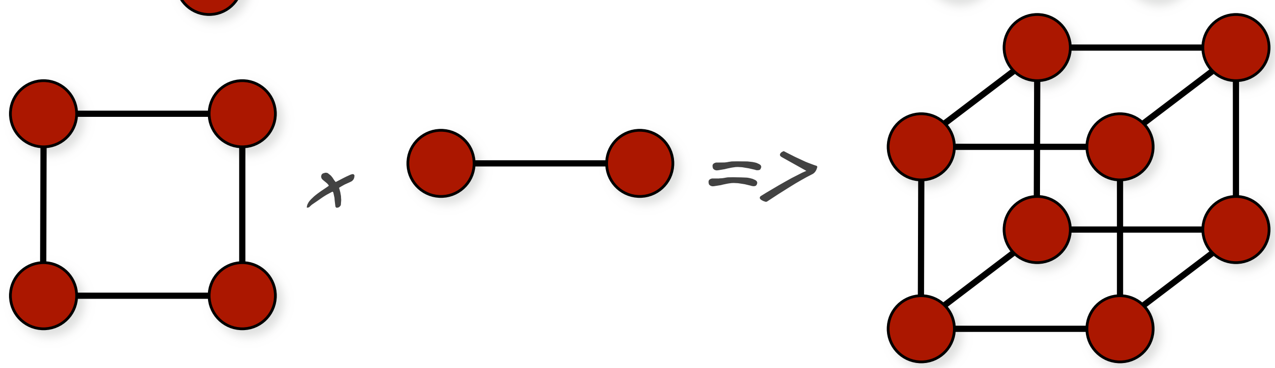
Hypercube

- Definizione ricorsiva di ipercubi Q_r

- $Q_1 = K_2$
- $Q_r = Q_{r-1} \times Q_1$
- $Q_2 = Q_1 \times Q_1$



- $Q_3 = Q_2 \times Q_1$



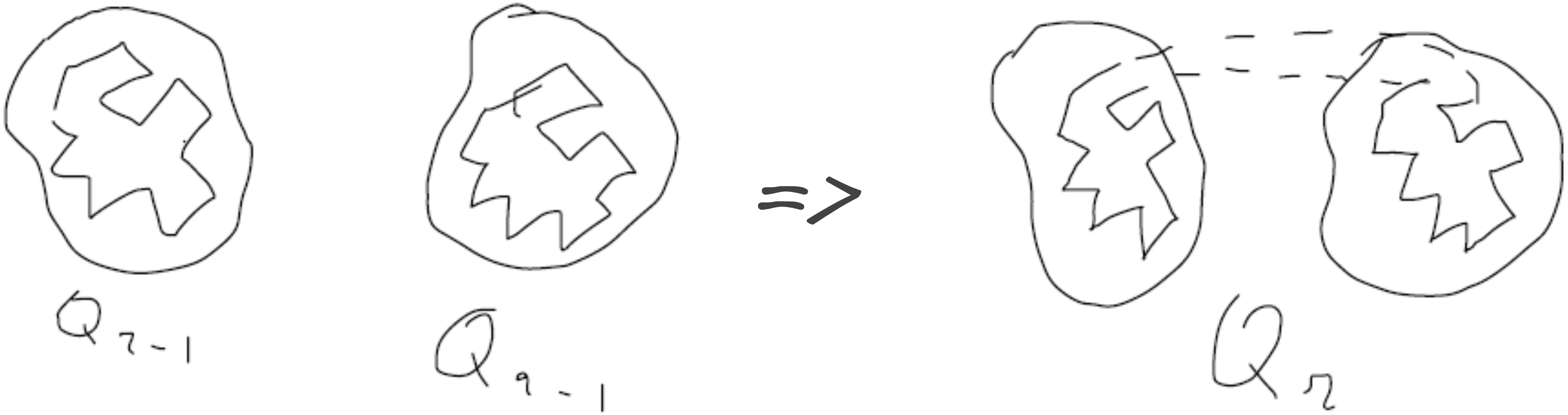
Proprieta' di Q_r

- $n=2^r$ processori, $r2^{r-1}$ collegamenti
- Regolare di grado $r=O(\log n)$ ☺
- Bipartito
- Hamiltoniano
- Diametro: $r=O(\log n)$ ☺
- Connettivita': $r=O(\log n)$ ☺
- Bisezione: $n/2$ ☺

Cammino minimo su Q_r

- Distanza tra due nodi= distanza di Hamming tra le etichette (numero di bit diversi)
- Cammino minimo da x a y:
  - Sia $k<=r$ la distanza da x a y
  - Siano $b_{i1}, b_{i2}, \ldots, b_{ik}$ i bit diversi
  - Selezionare uno tra i bit diversi,
    - Invertire il bit nell'etichetta di x
    - Muoversi sul nodo con 'etichetta ottenuta (collegato a x)
  - Iterare

Simulare reti sull'ipercubo

- Q_r contiene una rete lineare di $n=2^r$ processori (hamiltoniano)
- vero per $r=1,2,3$
- Per $r>3$, per induzione
- 

Hamiltoniano e codici di Gray

- Codice di Gray G_r a r bit
  - $G_1=(0,1)$
  - $G_r=(0G_{r-1}, 1G_{r-1}R)$, costruzione a specchio

G2=(00,01,11,10)

G3=(000,001,011,010,110,111,101,100)

- Hamiltoniano = codice di Gray

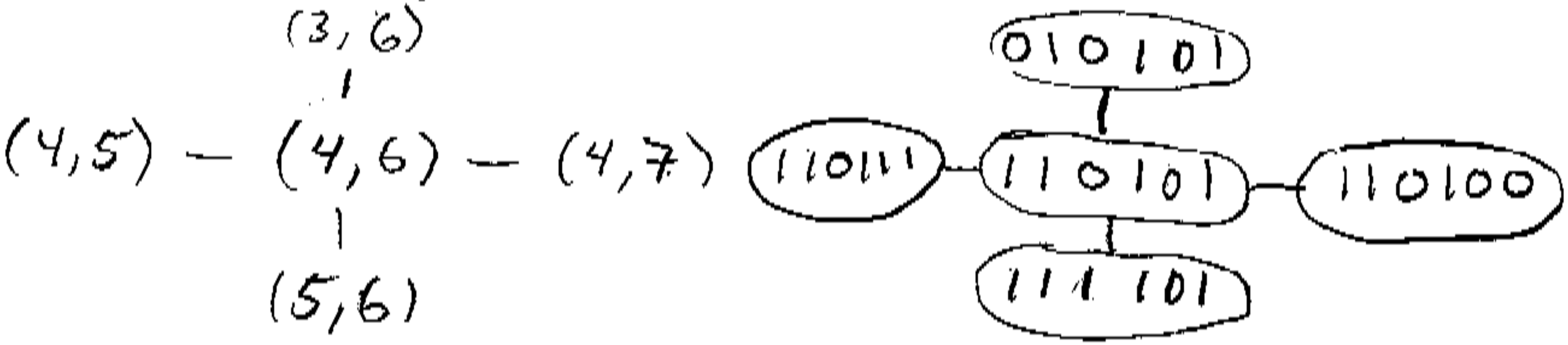
Embedding di mesh su ipercubo

- Una mesh $d_1 x d_2$ puo' essere immersa (embedded) in Q_d, dove $d=d_1+d_2$

- Nodo (x,y) della mesh
  - $G_{d1}(x)$ codice di Gray di x
  - $G_{d2}(y)$ codice di Gray di y
  - Mappa (x,y) sul nodo $G_{d1}(x) \cdot G_{d2}(y)$ dell'ipercubo

Embedding di mesh su ipercubo

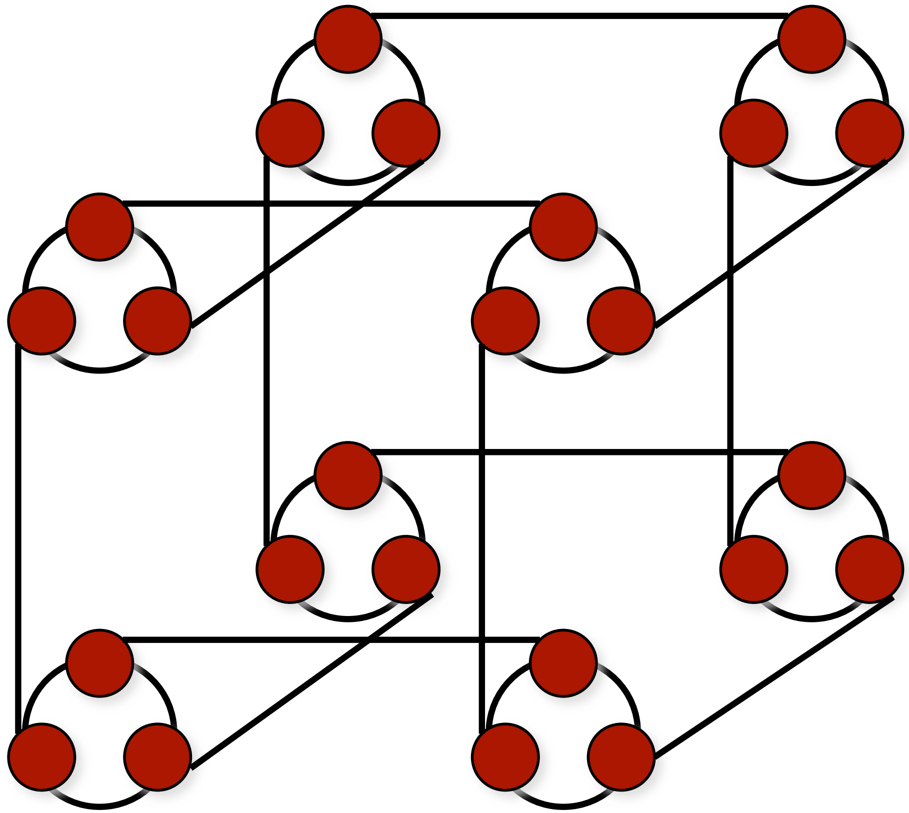
Ex. 8 × 8 mesh.



Ottenere grado costante

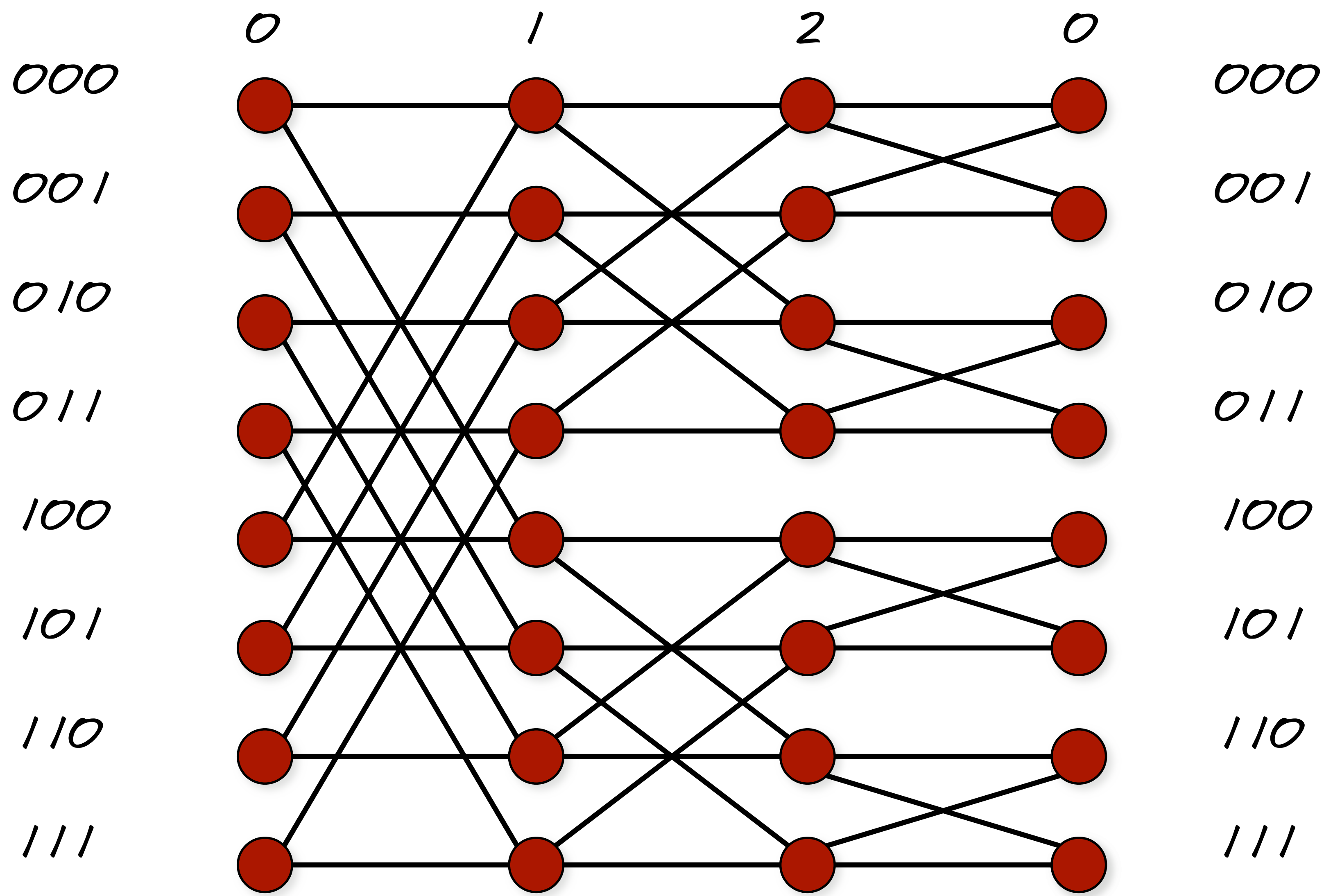
- Sostuire ad ogni nodo dell'ipercubo un anello di r nodi, ognuno collegato ad un diverso (tra gli r) processore collegato
- Effetto:
  - Grado: 3 ☺
  - Numero di nodi: $r2^r$ ☹

Cube connected cycles



Butterfly

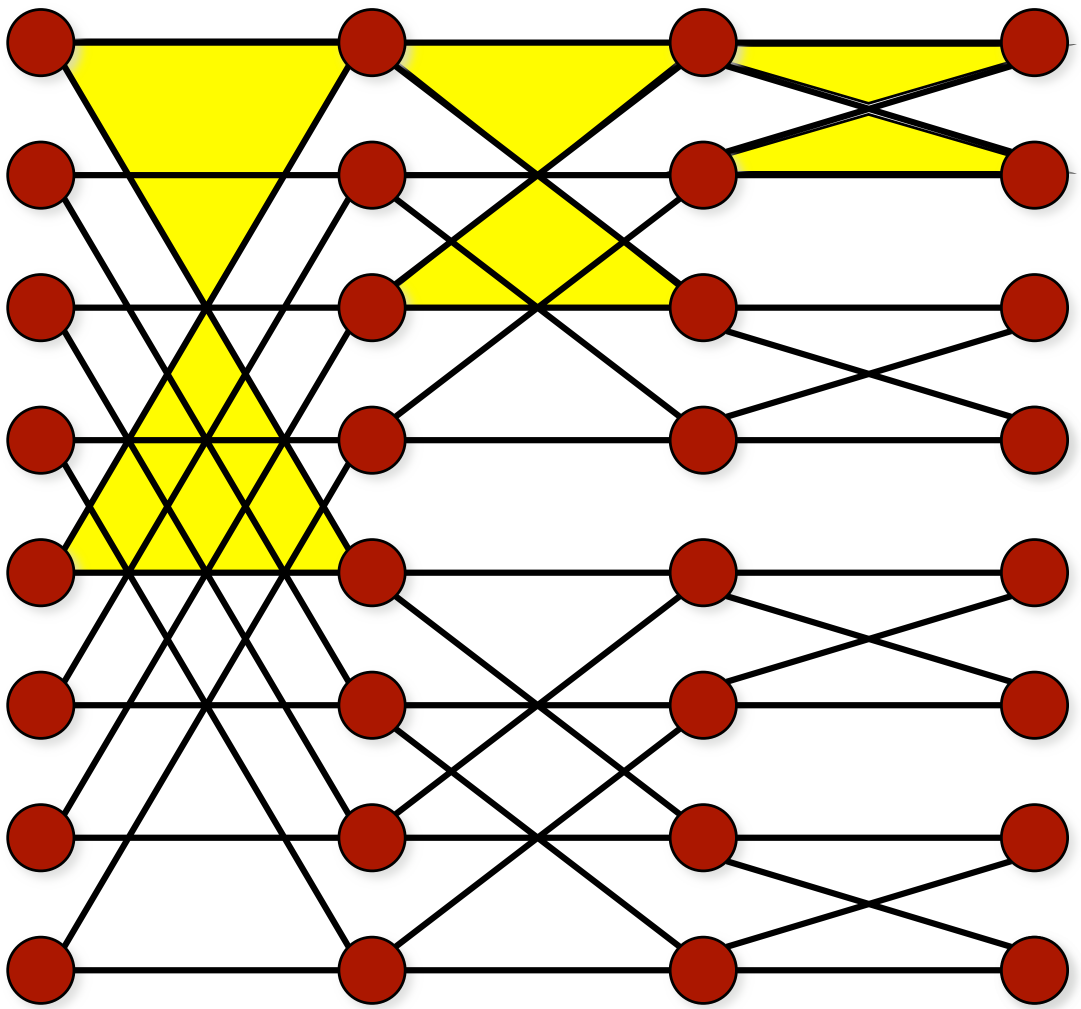
- $n = 2^d$ processori collegati da $n(\log n + 1)$ switch



Costruzione della 2^3 Butterfly

- 8 processori
- 4 colonne di 8 switch ciascuna
- Collegamenti:
  - Il nodo (i,j), for $i > 0$, e' collegato a due nodi della colonna on $i-1$: $(i-1,j)$ e $(i-1,m)$
    - m e' l'intero ottenuto invertendo l'i-esimo bit di j
      - Ad esempio, per $i = 2$ e $j = 3$, il nodo $(2,3)$ e' collegato a $(1,3)$ e a $(1,1)$ in quanto $3 = 011_2$ e si inverte il 2 bit di 011 ottenendo 001_2
- Nodi collegati da un collegamento trasversale tra le colonne I e i+1 hanno etichette che differiscono solo per il bit i+1.

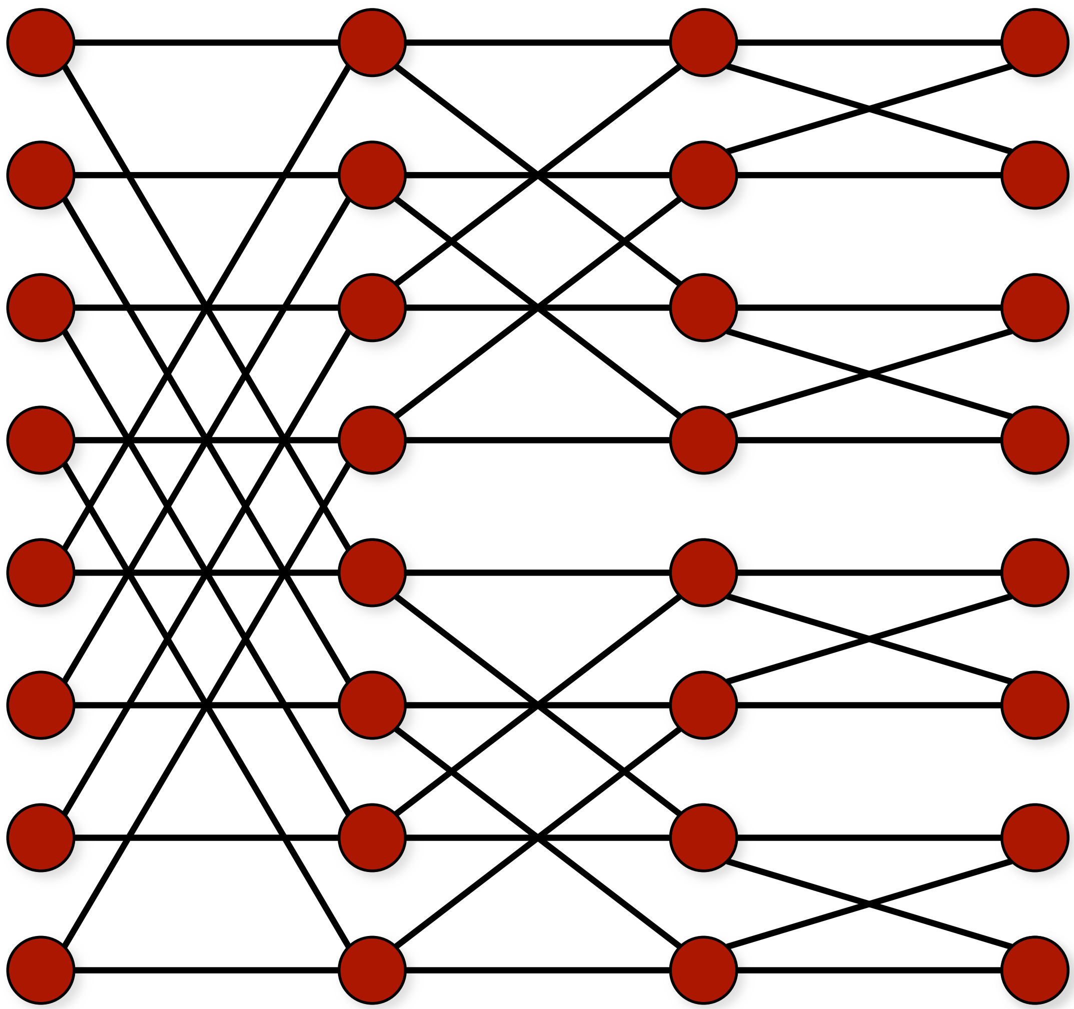
Perche' questo nome



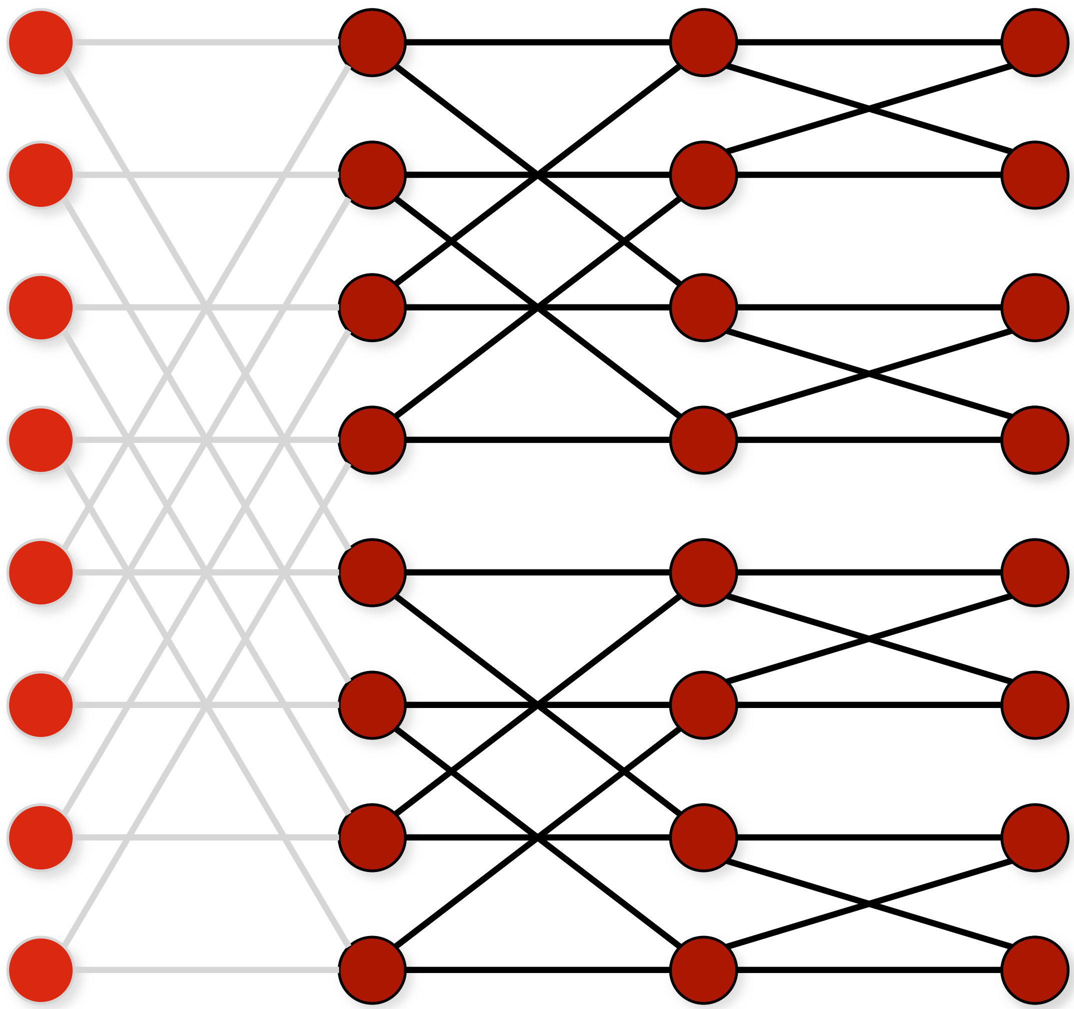
Caratteristiche della Butterfly

- n input, n output
- $n\log n$ nodi 😐
- Grado: 4 ☺
- Diametro: $O(\log n)$ ☺
- Bisezione: n ☺

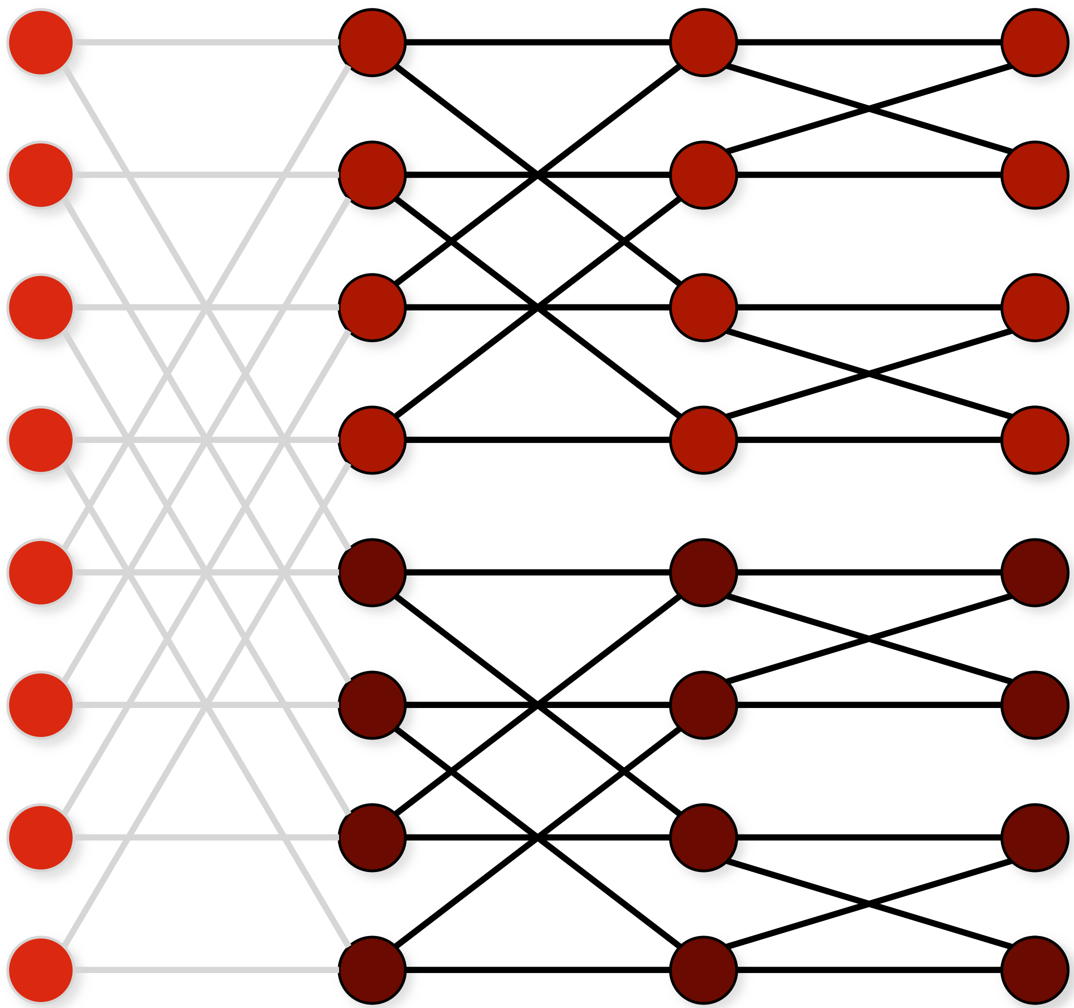
Decomposizione di una butterfly



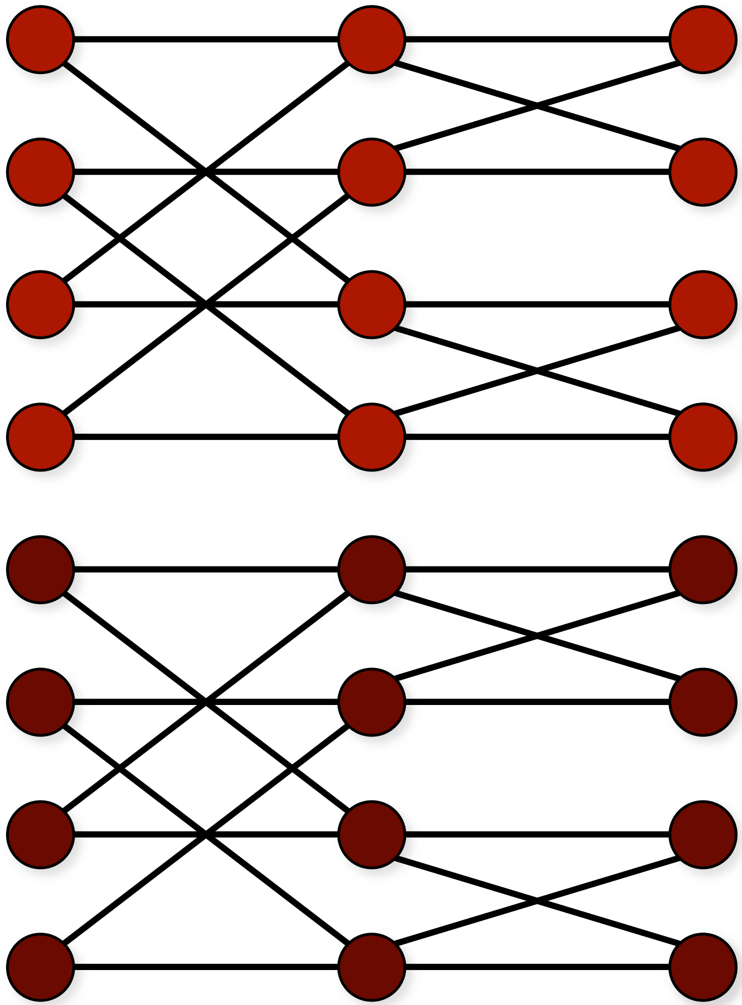
Decomposizione di una butterfly



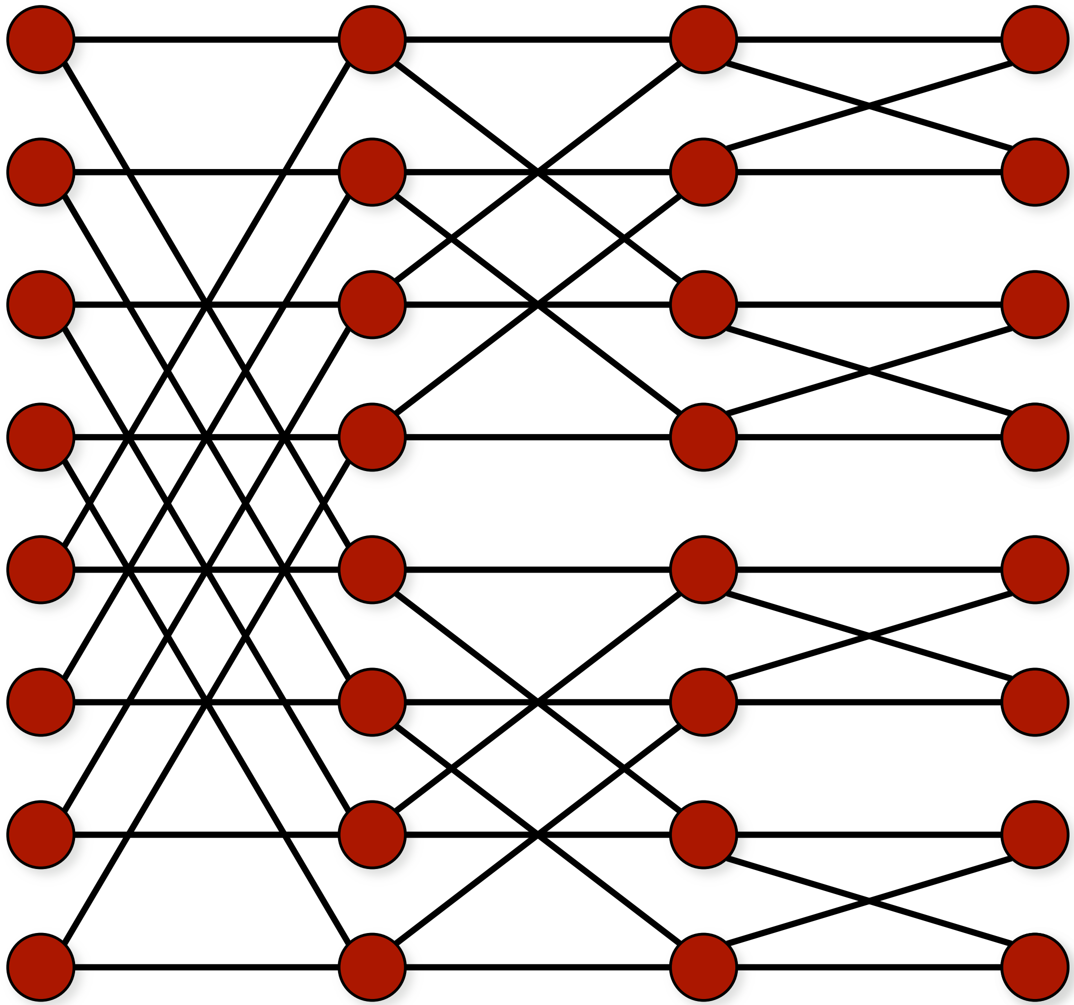
Decomposizione di una butterfly



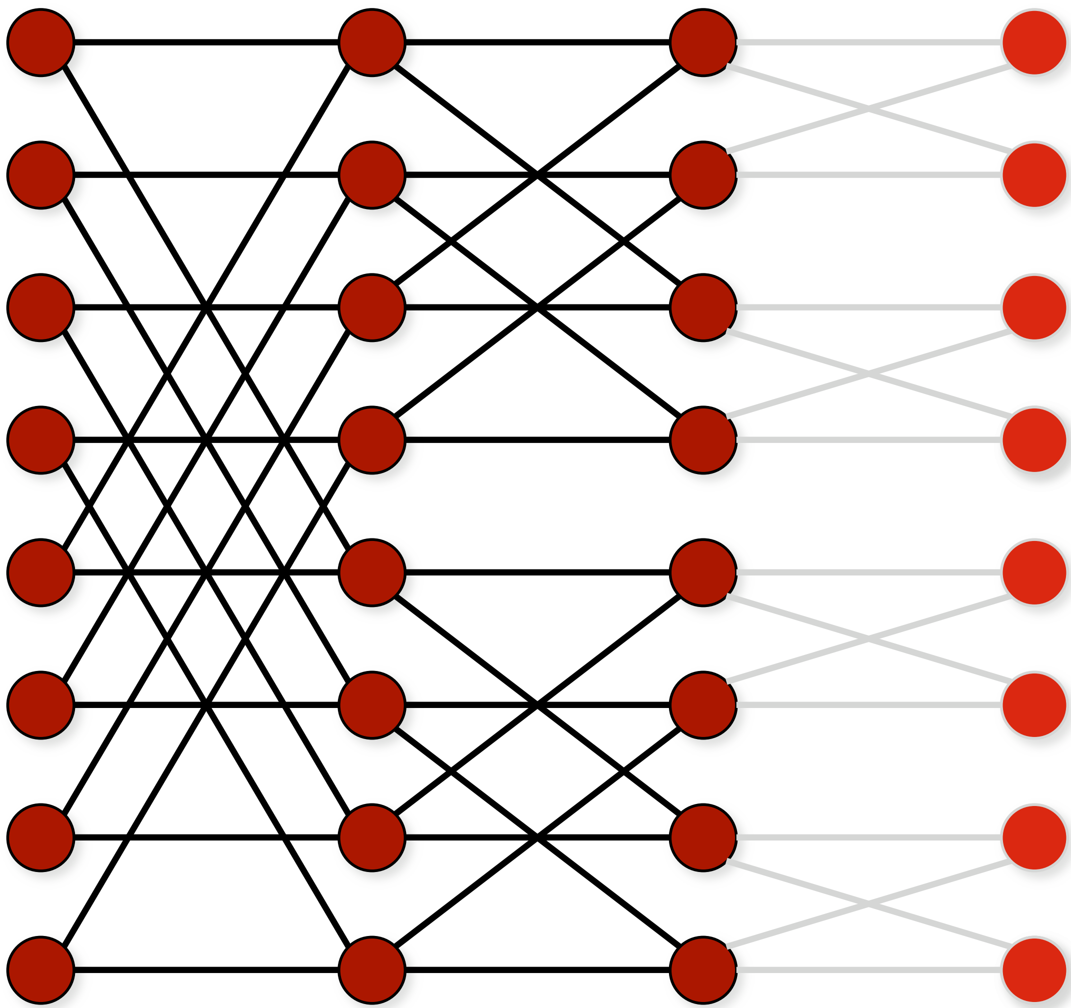
Decomposizione di una butterfly



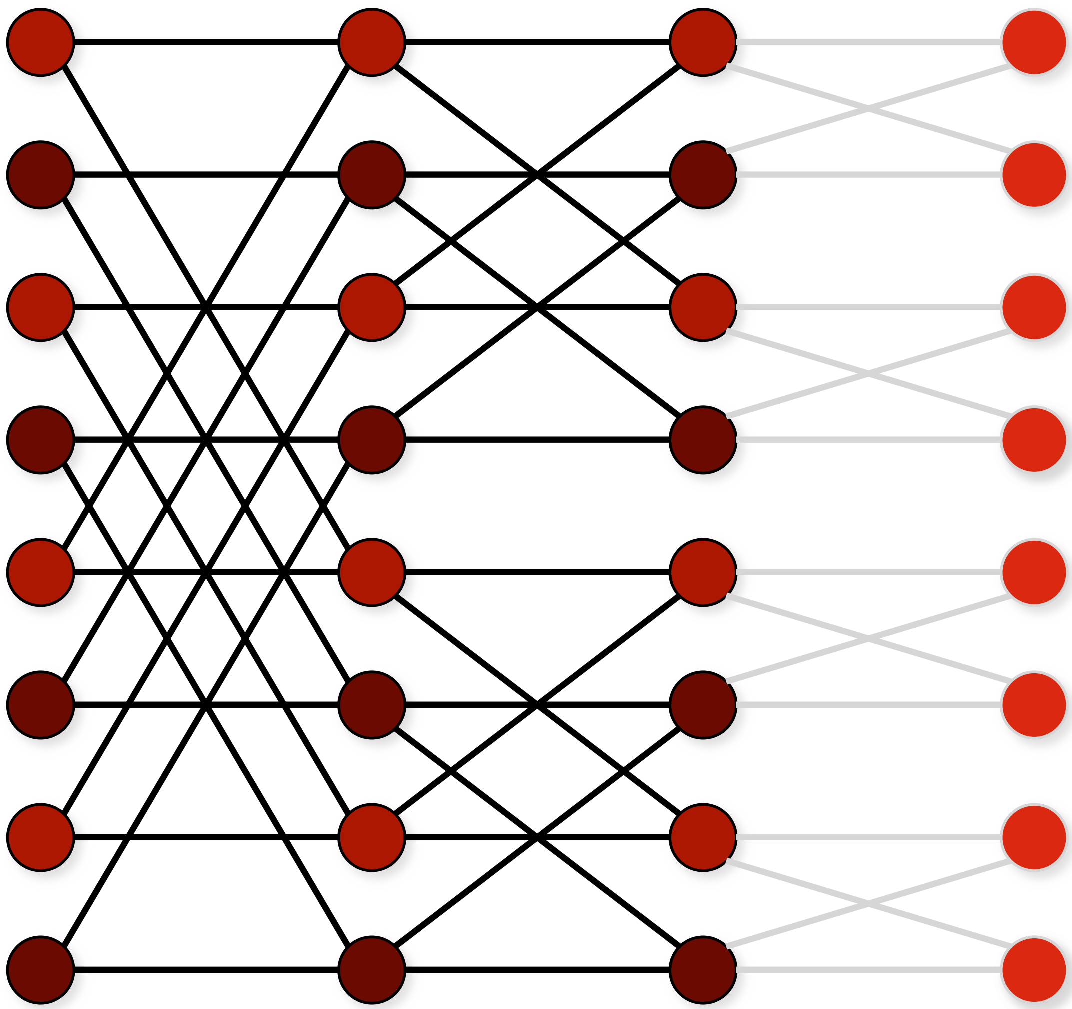
Decomposizione di una butterfly



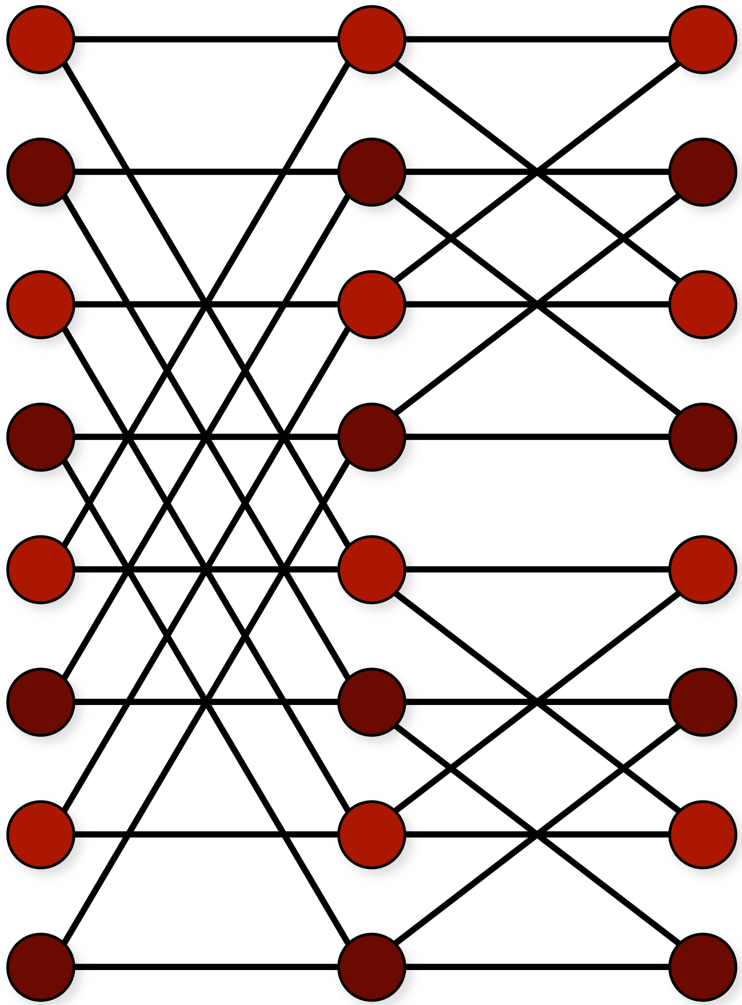
Decomposizione di una butterfly



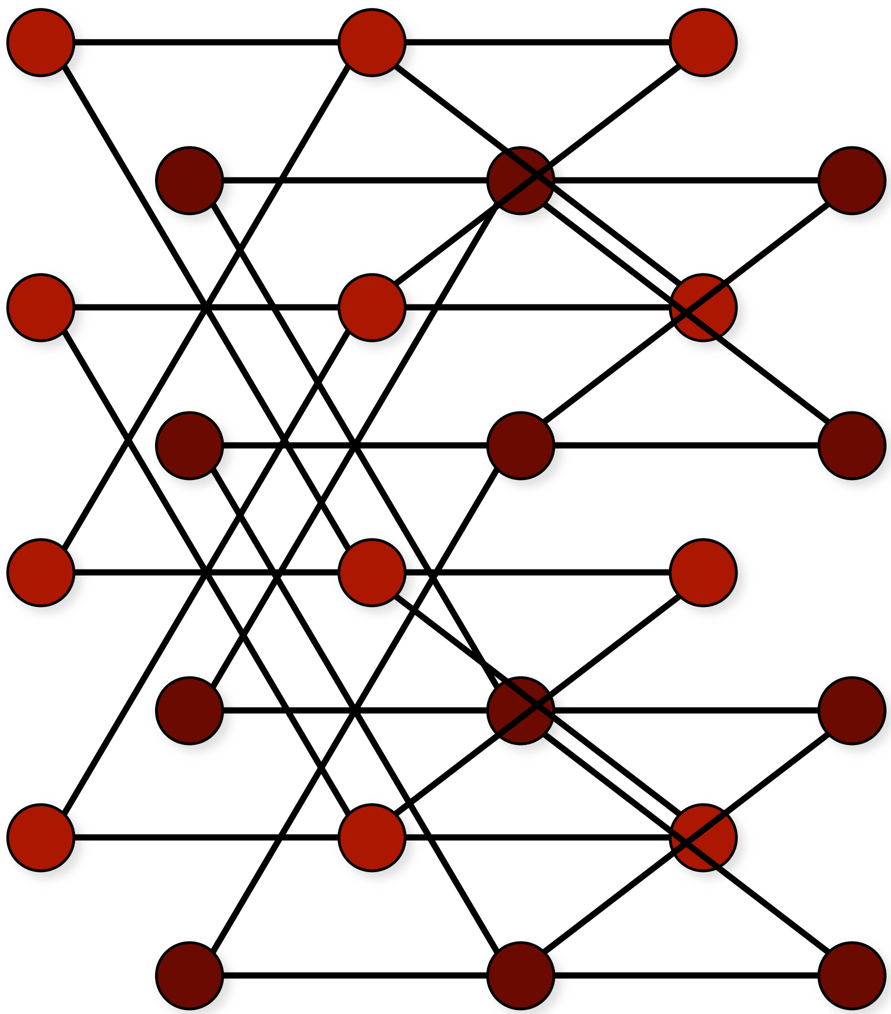
Decomposizione di una butterfly



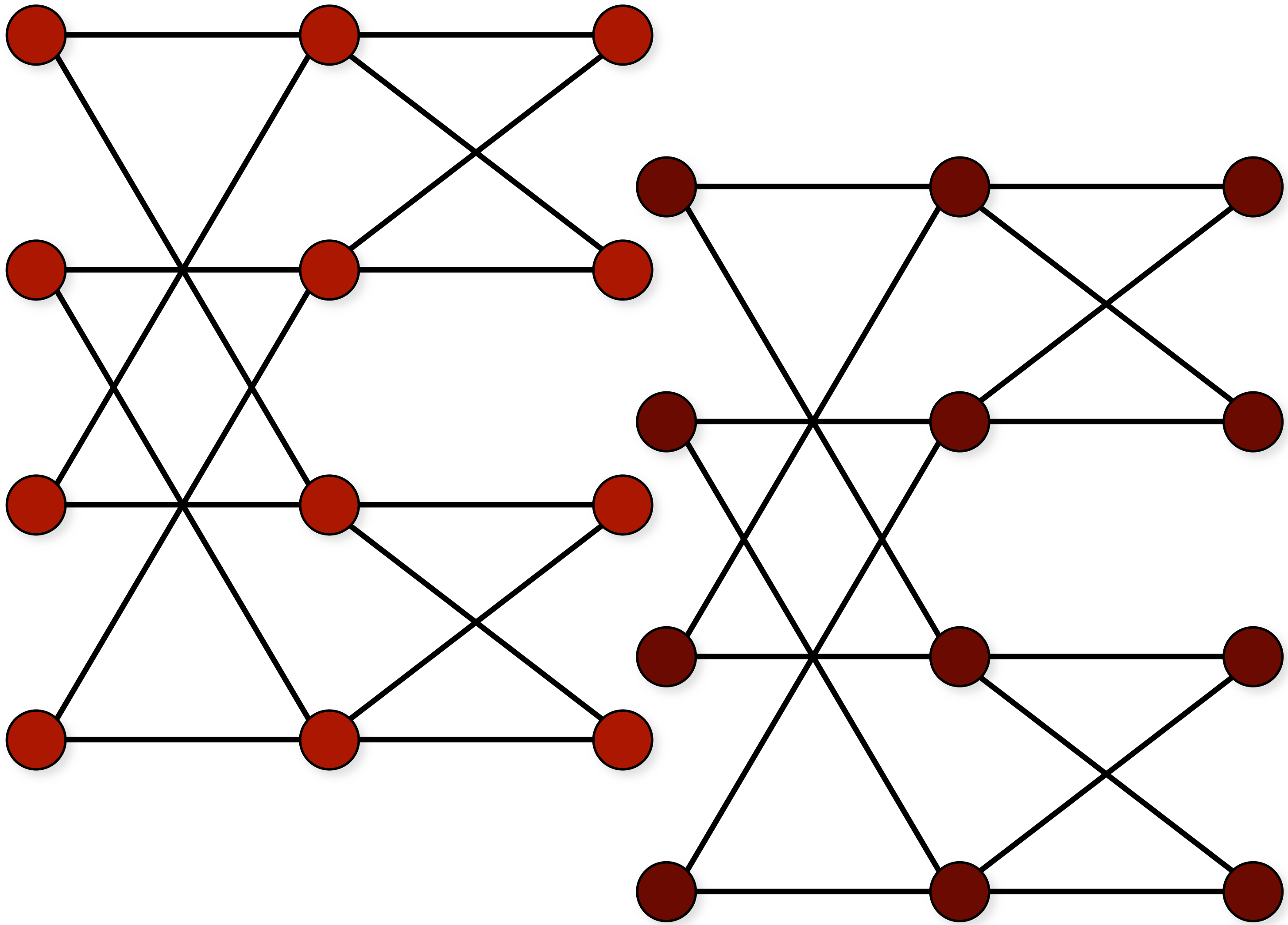
Decomposizione di una butterfly



Decomposizione di una butterfly



Decomposizione di una butterfly



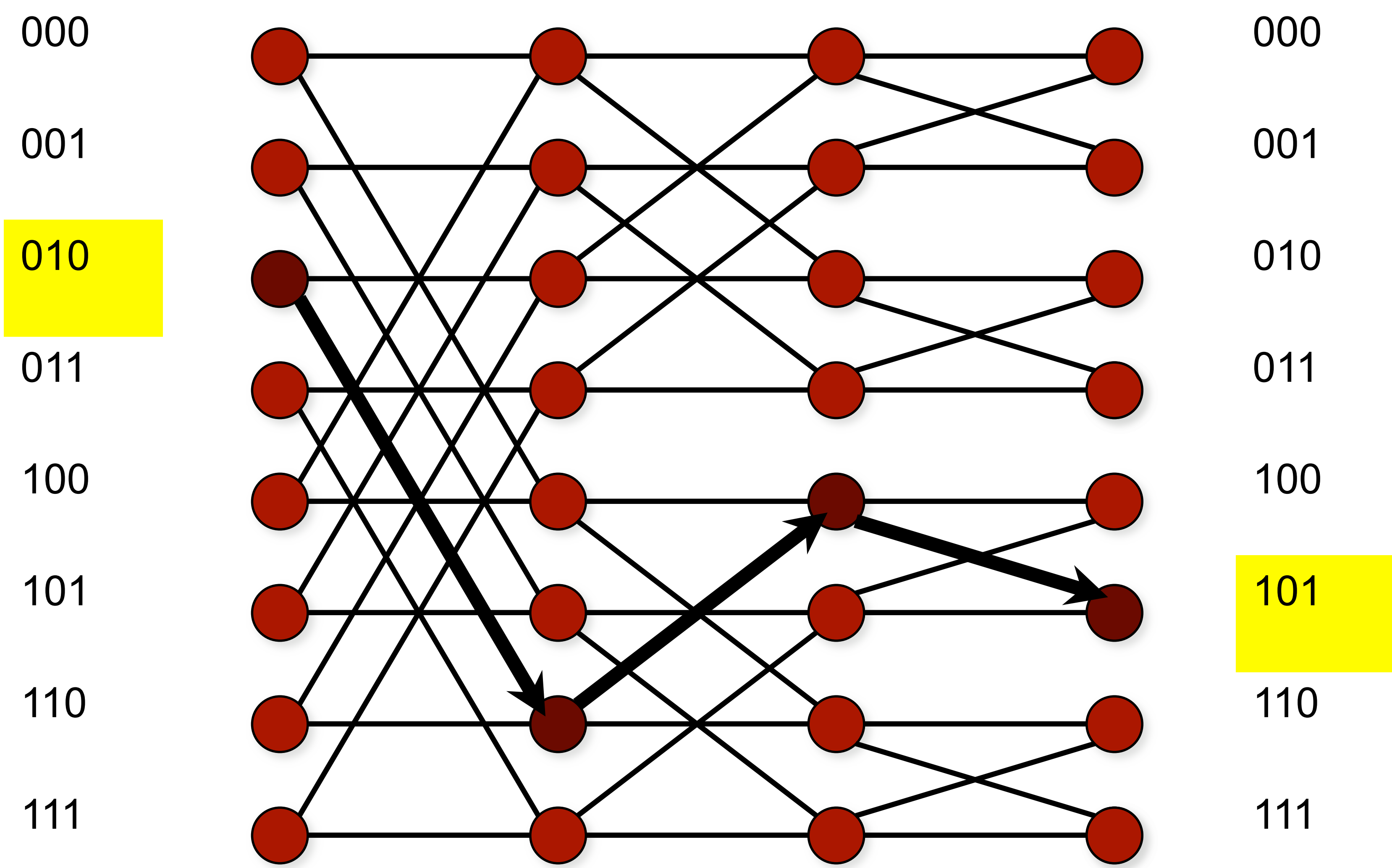
Routing su Butterfly

<u>Algoritmo</u>:
1 bit della destinazione alla volta
0 : vai a sx;
1 : vai a dx.

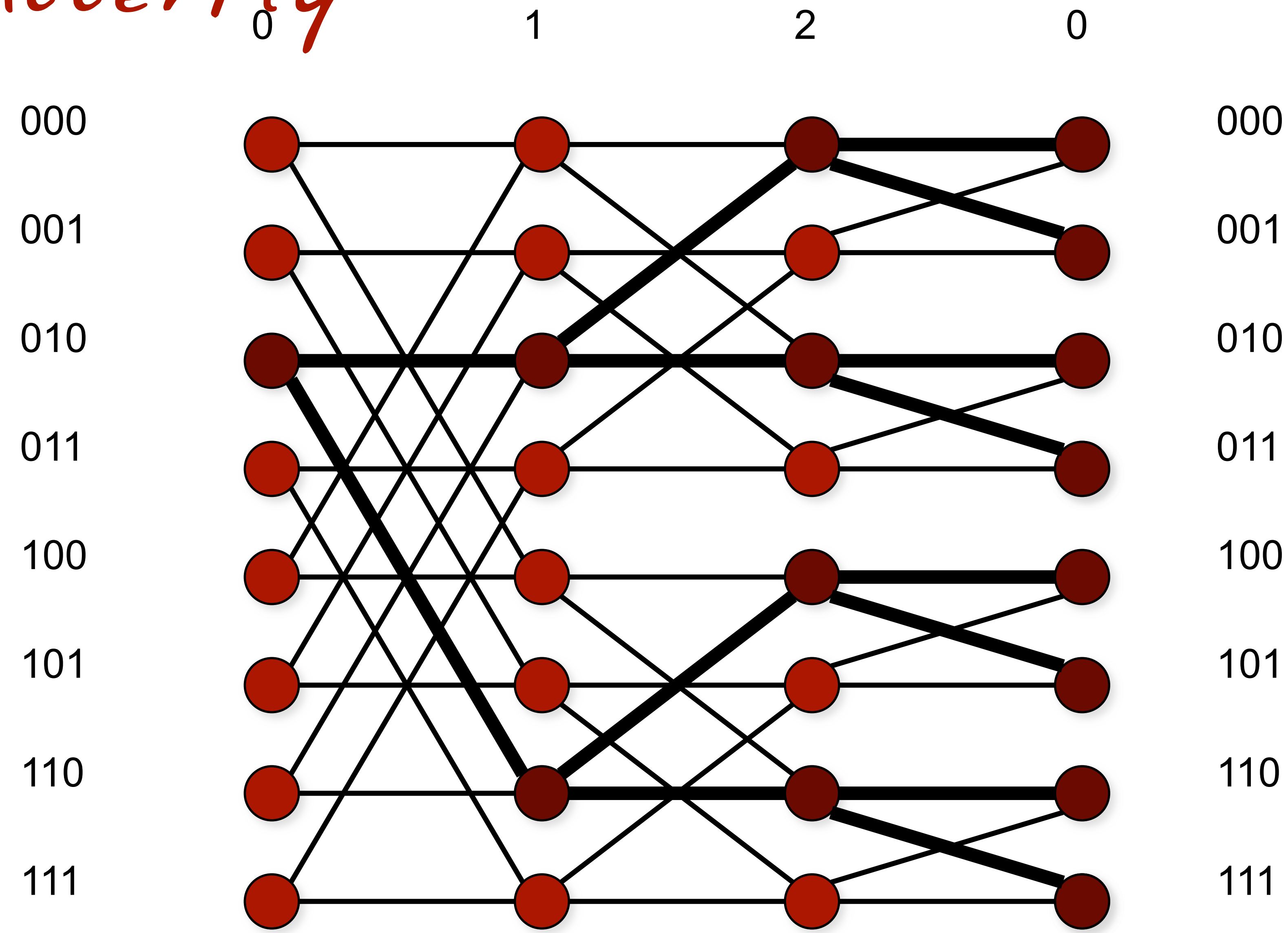
Messaggio da 2 a 5

1) 5 = 101
2) Primo bit (a sx) = 1 vai a dx (resta 01).
3) Primo bit (a sx) = 0 vai a sx (resta 1).
4) Primo bit (a sx) = 1 vai a dx.

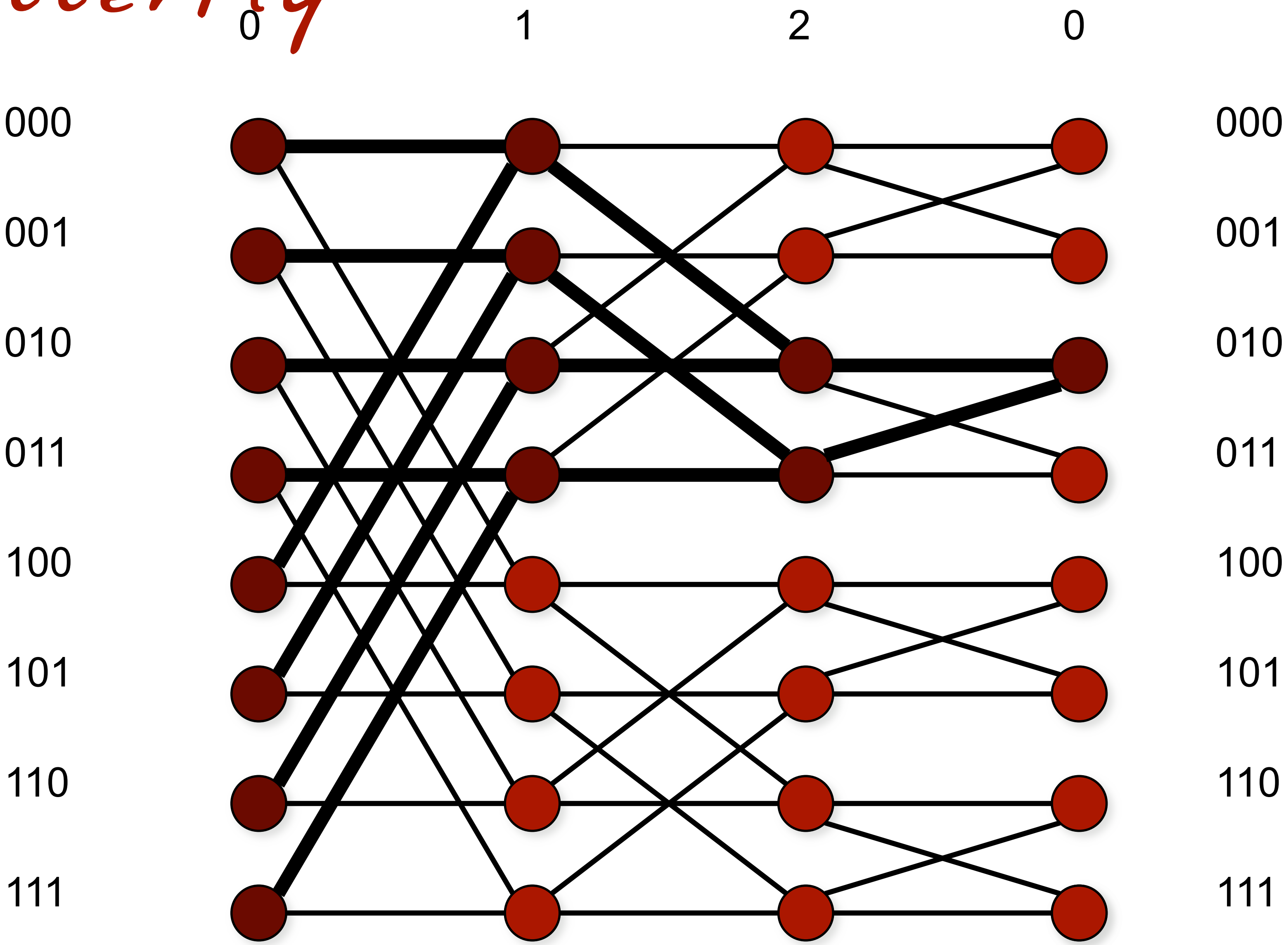
Routing su Butterfly



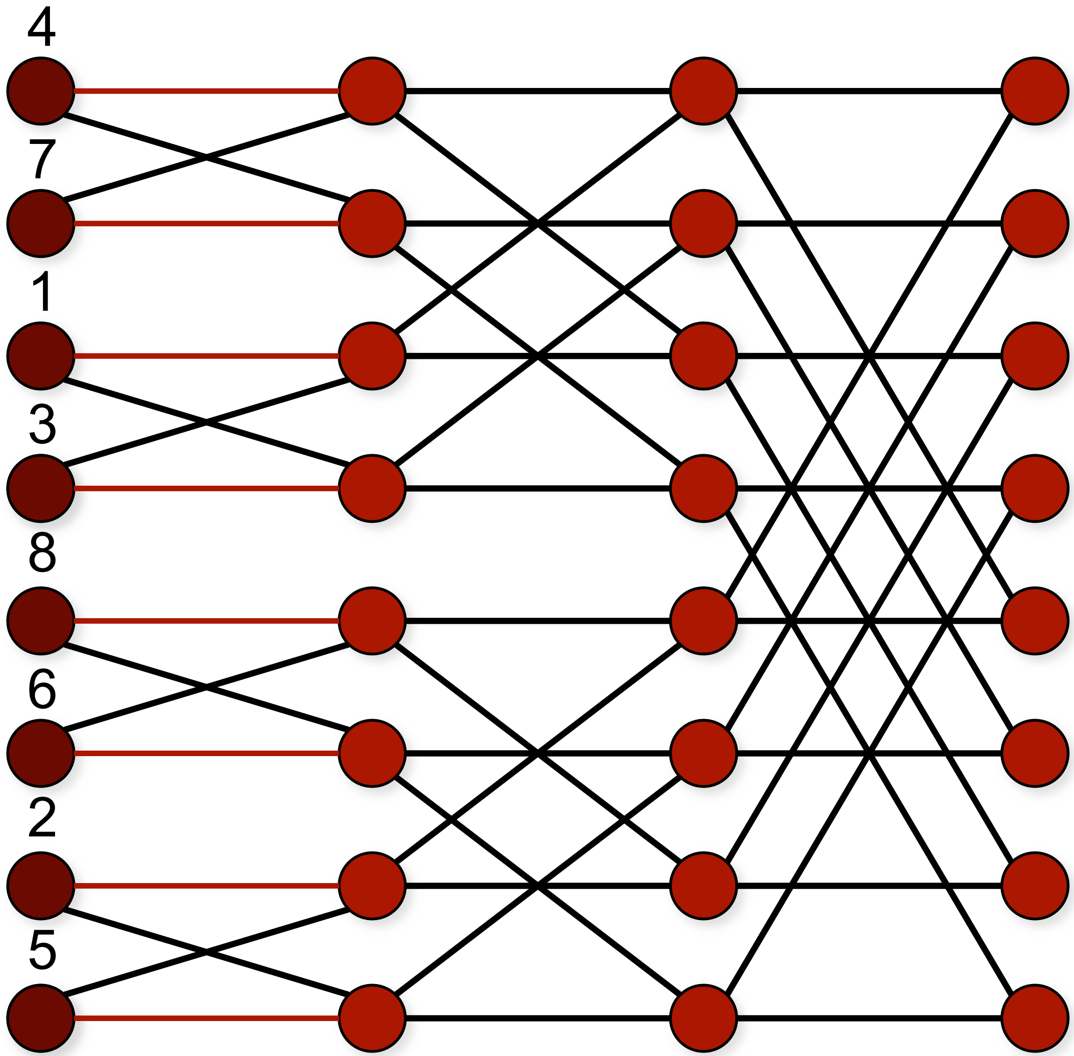
Alberi dei cammini su Butterfly

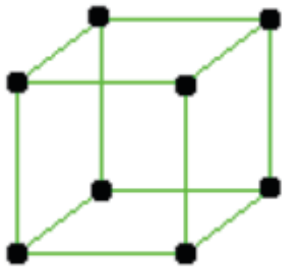


Alberi dei cammini su Butterfly

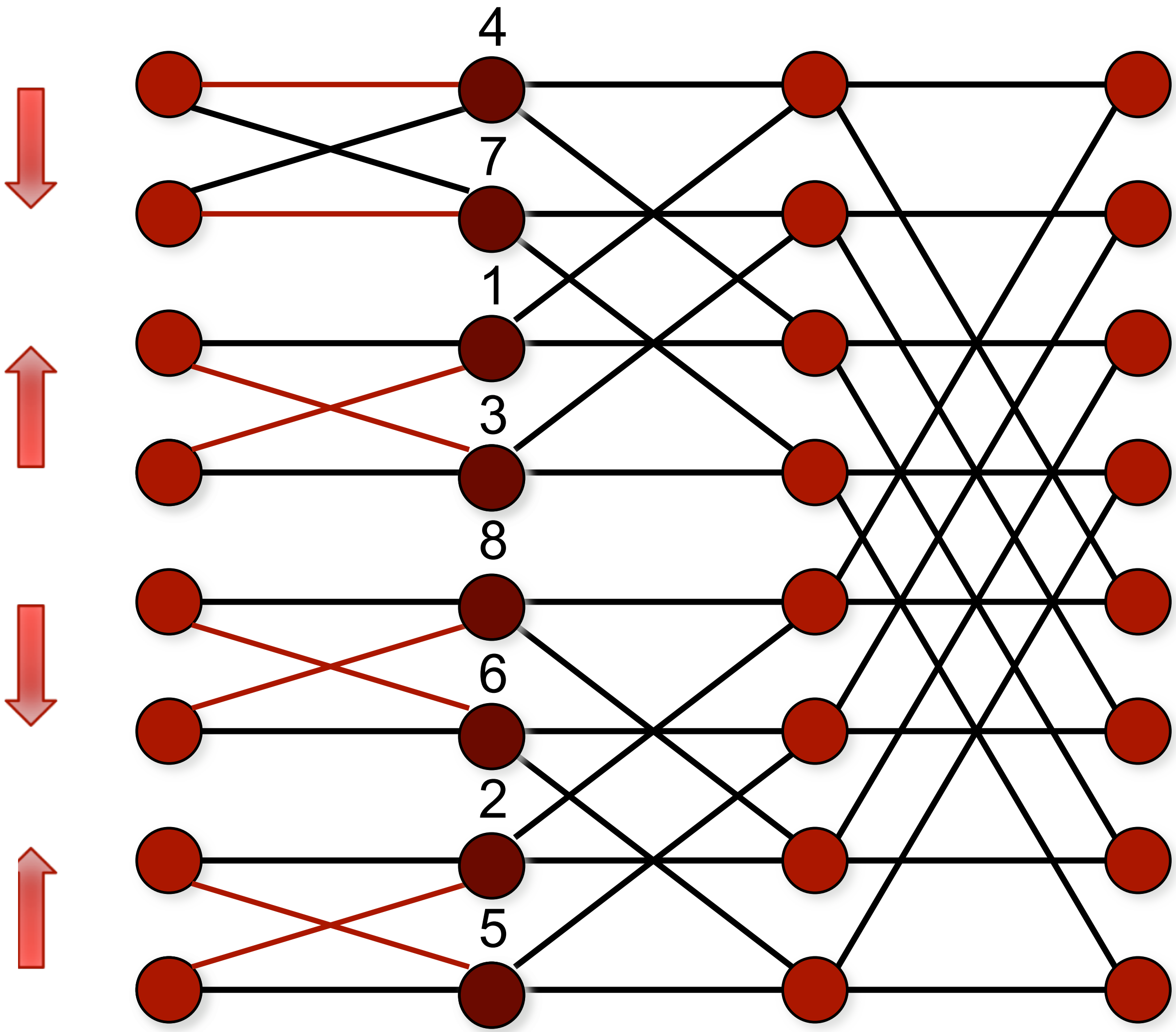


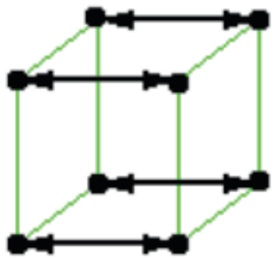
Bitonic sort su Butterfly



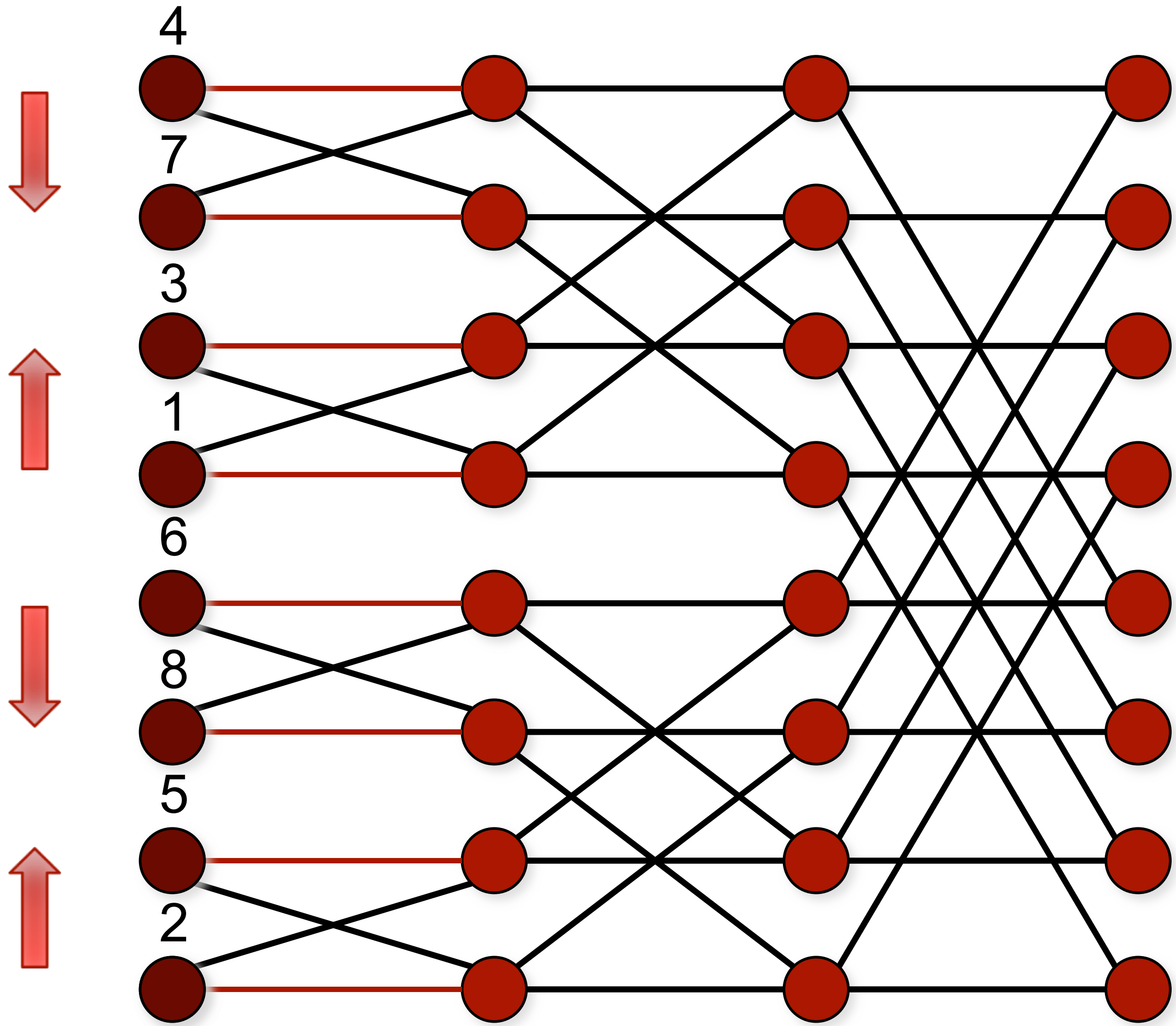


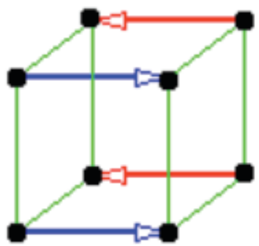
Bitonic sort su Butterfly



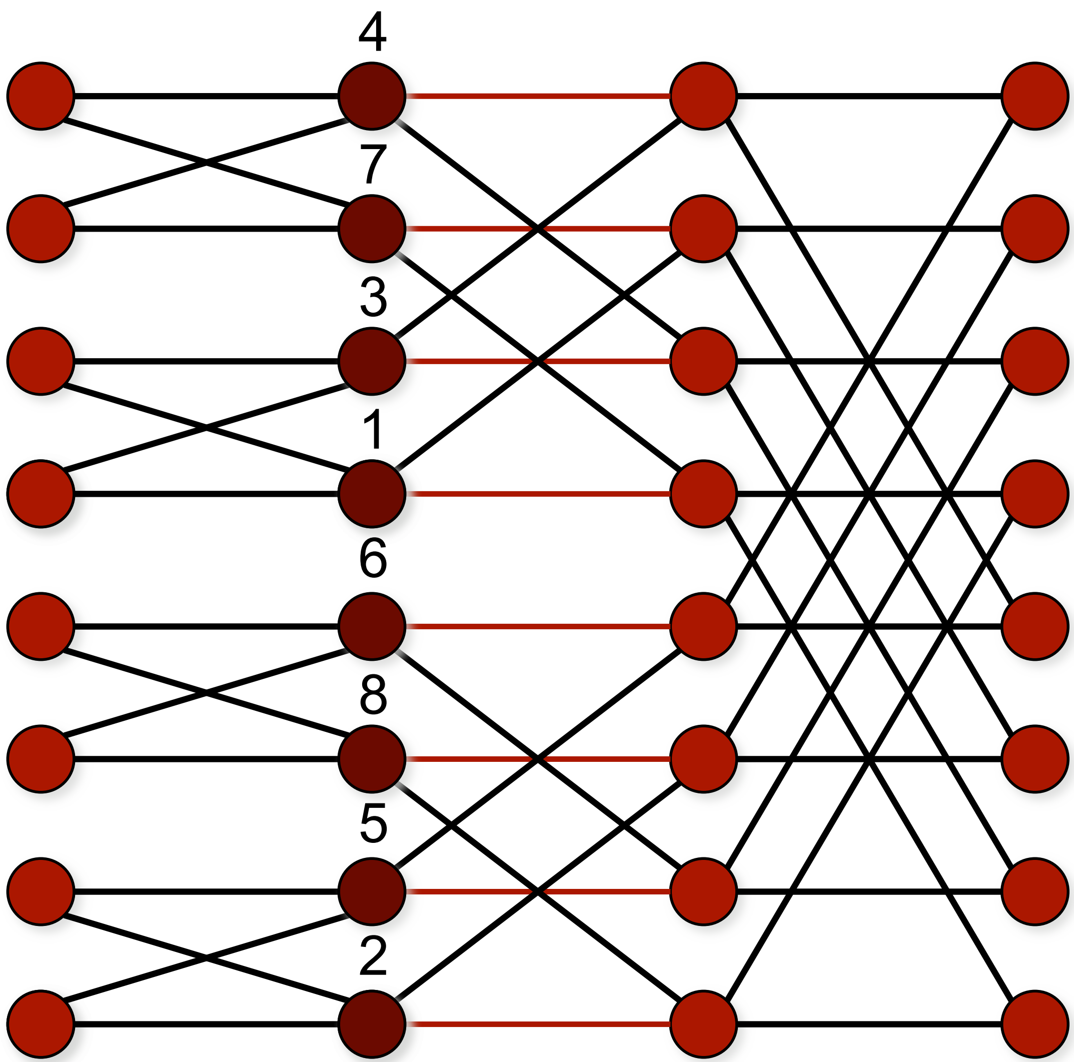


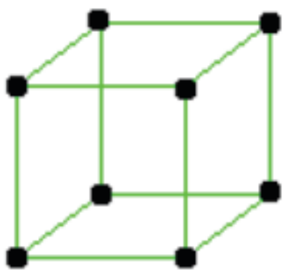
Bitonic sort su Butterfly



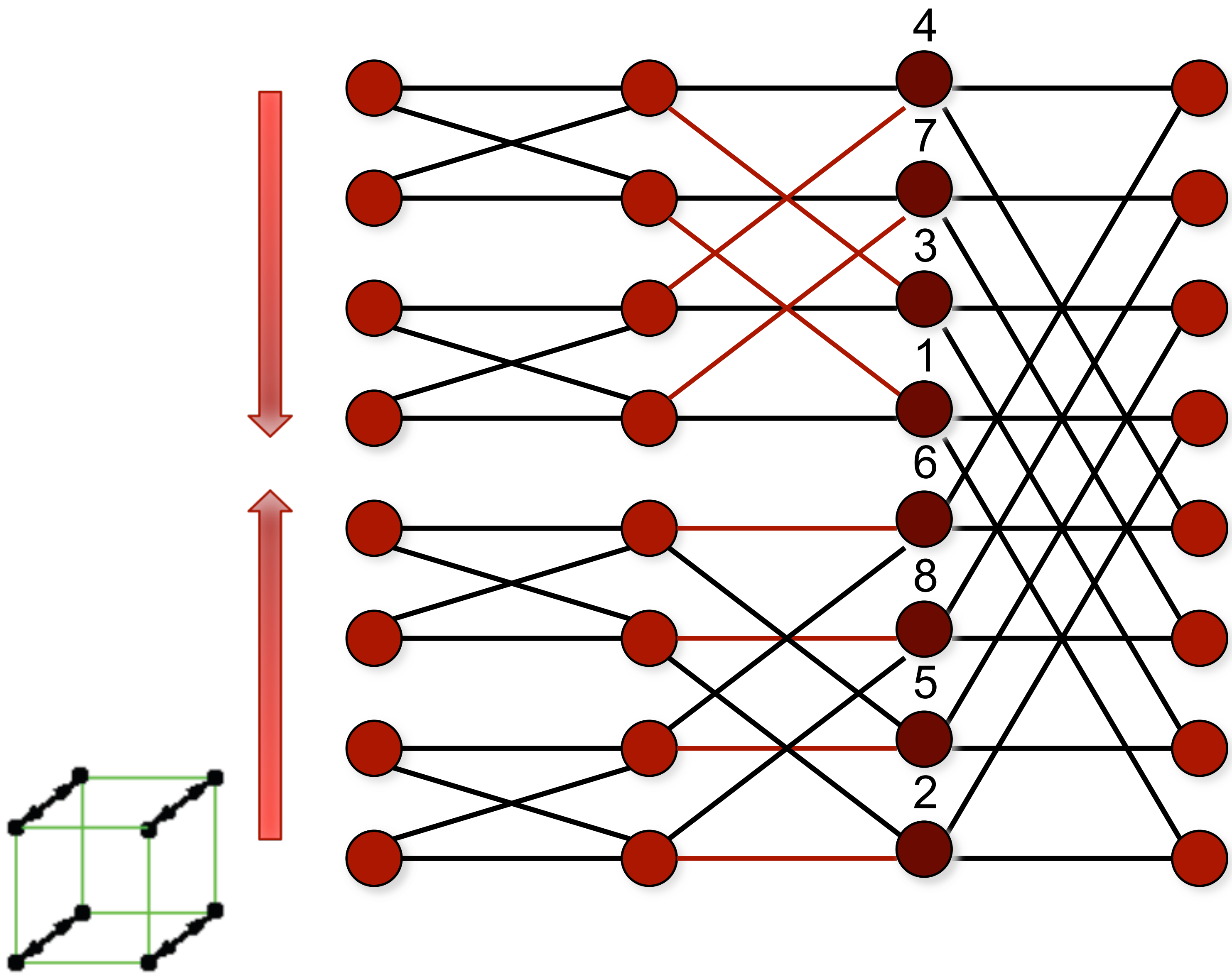


Bitonic sort su Butterfly

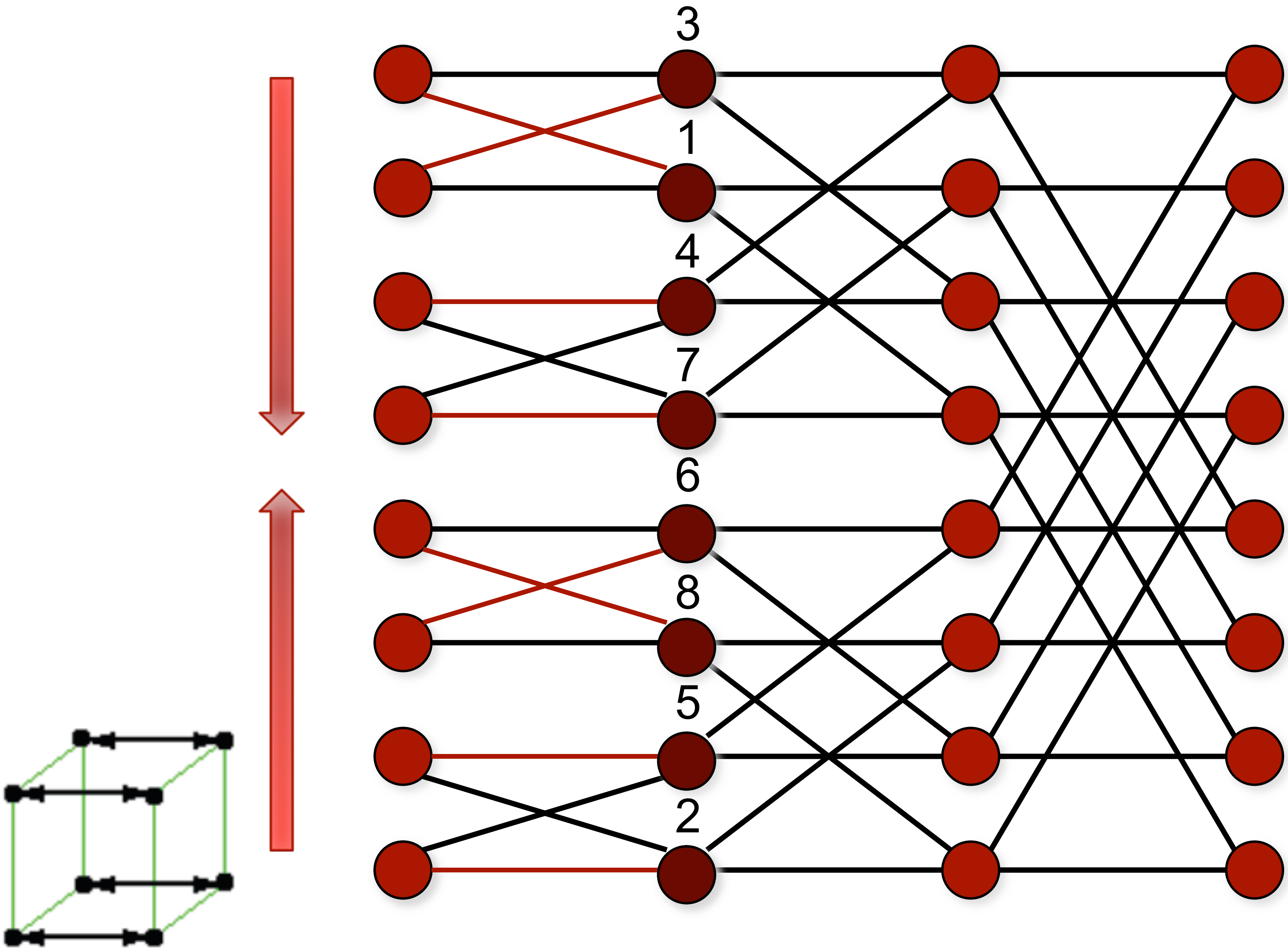




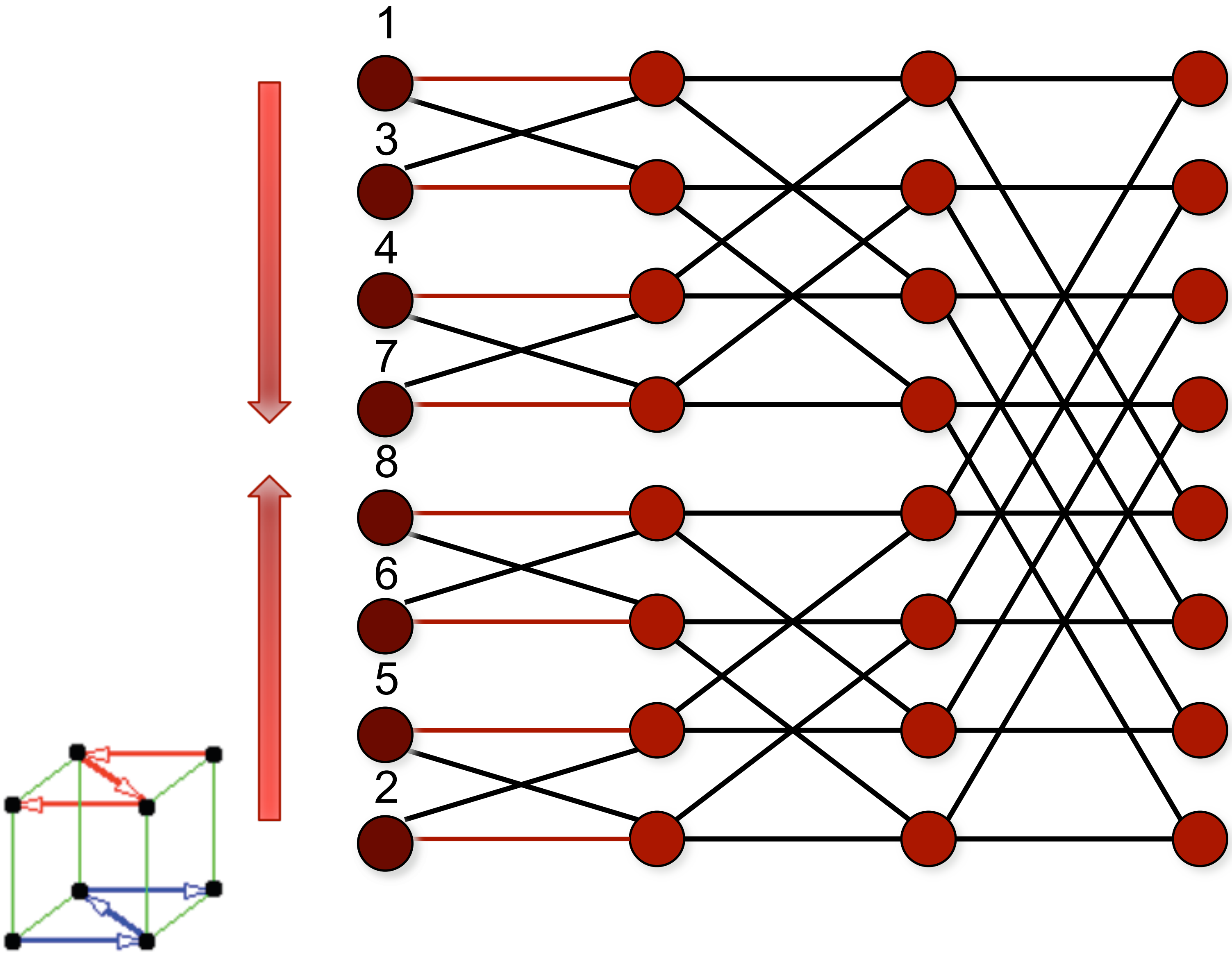
Bitonic sort su Butterfly



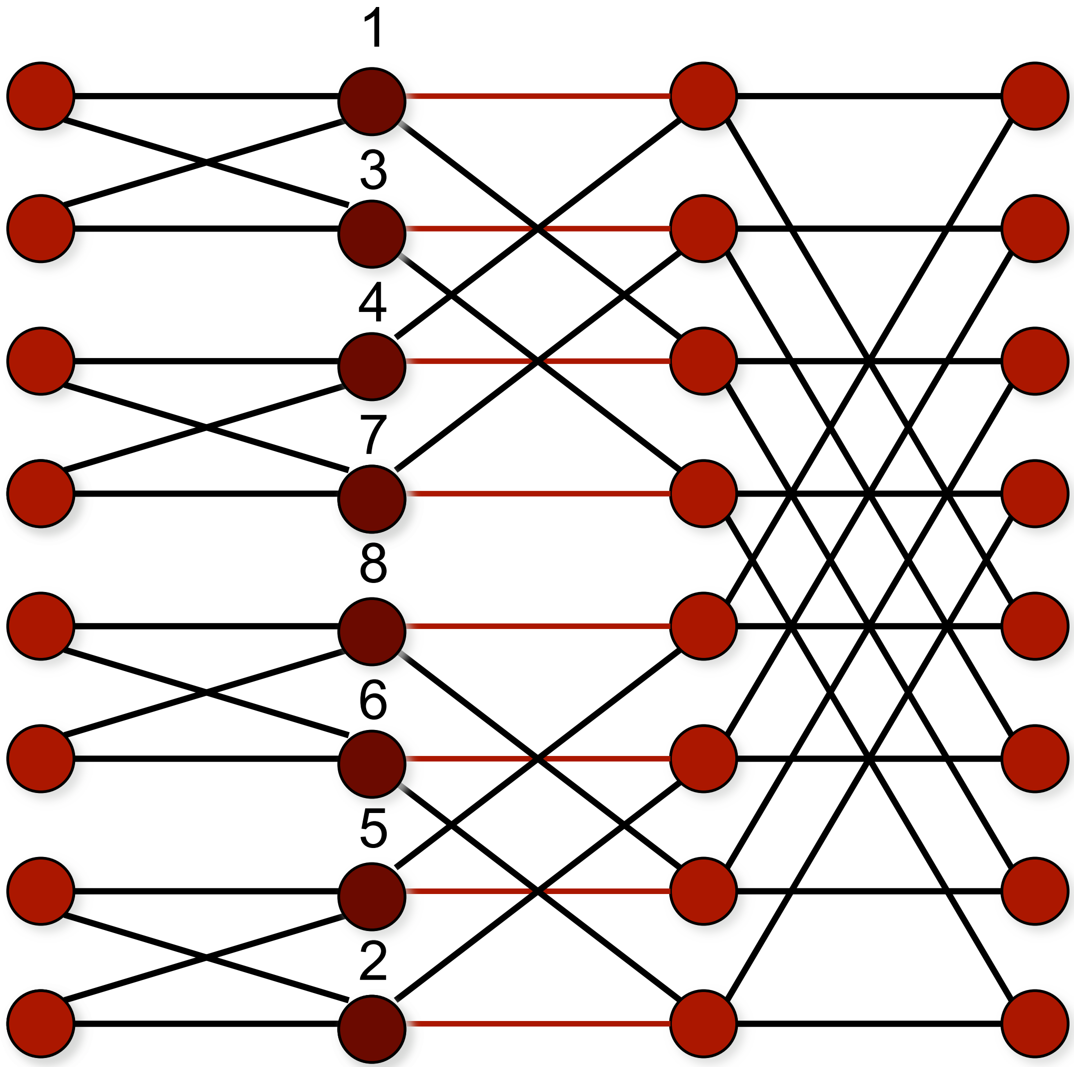
Bitonic sort su Butterfly

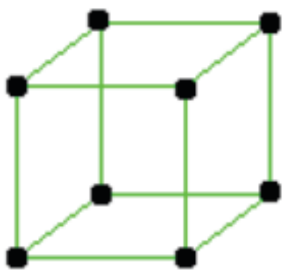


Bitonic sort su Butterfly

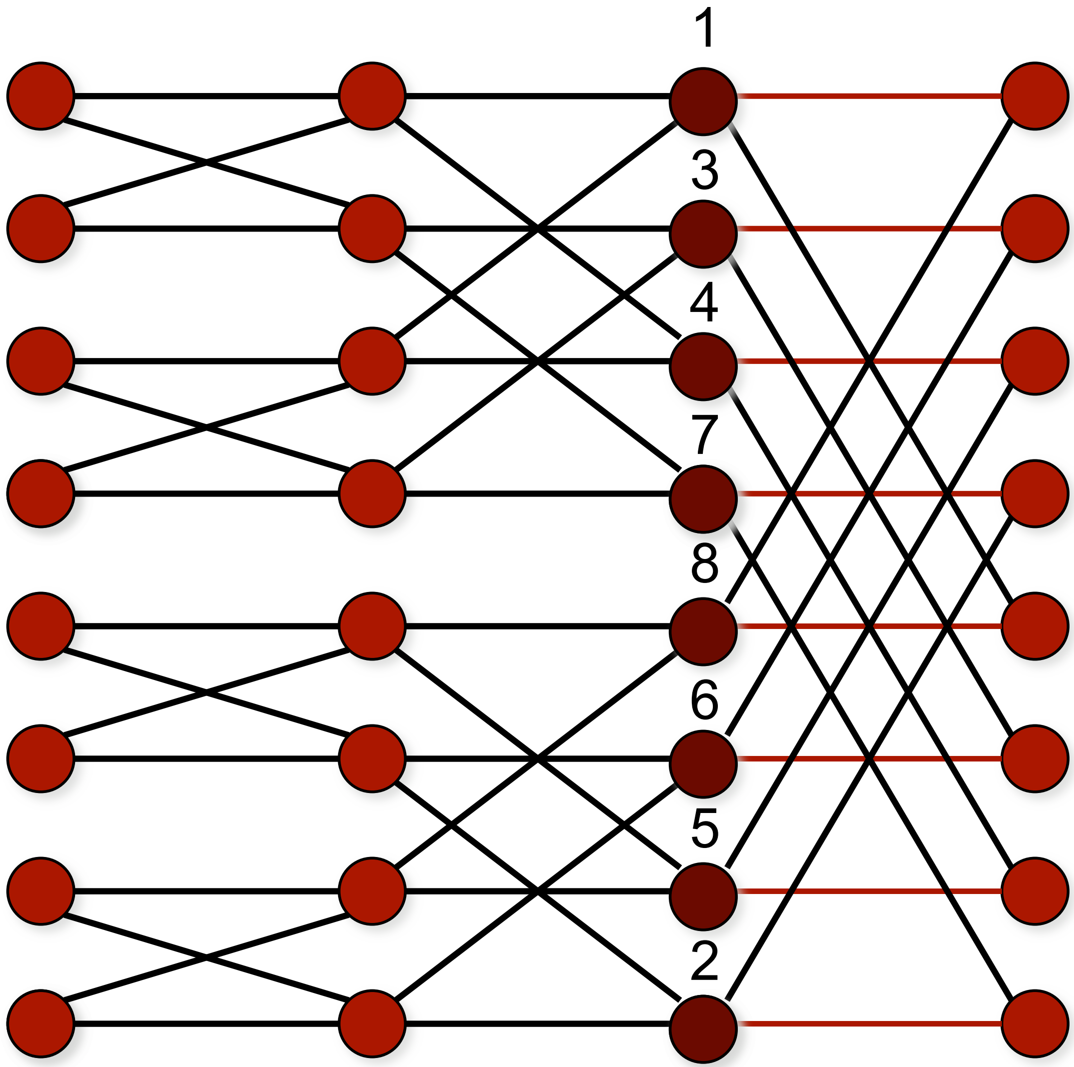


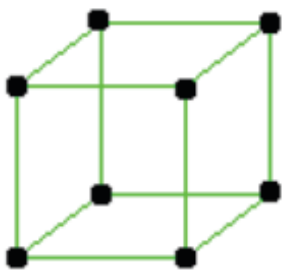
Bitonic sort su Butterfly



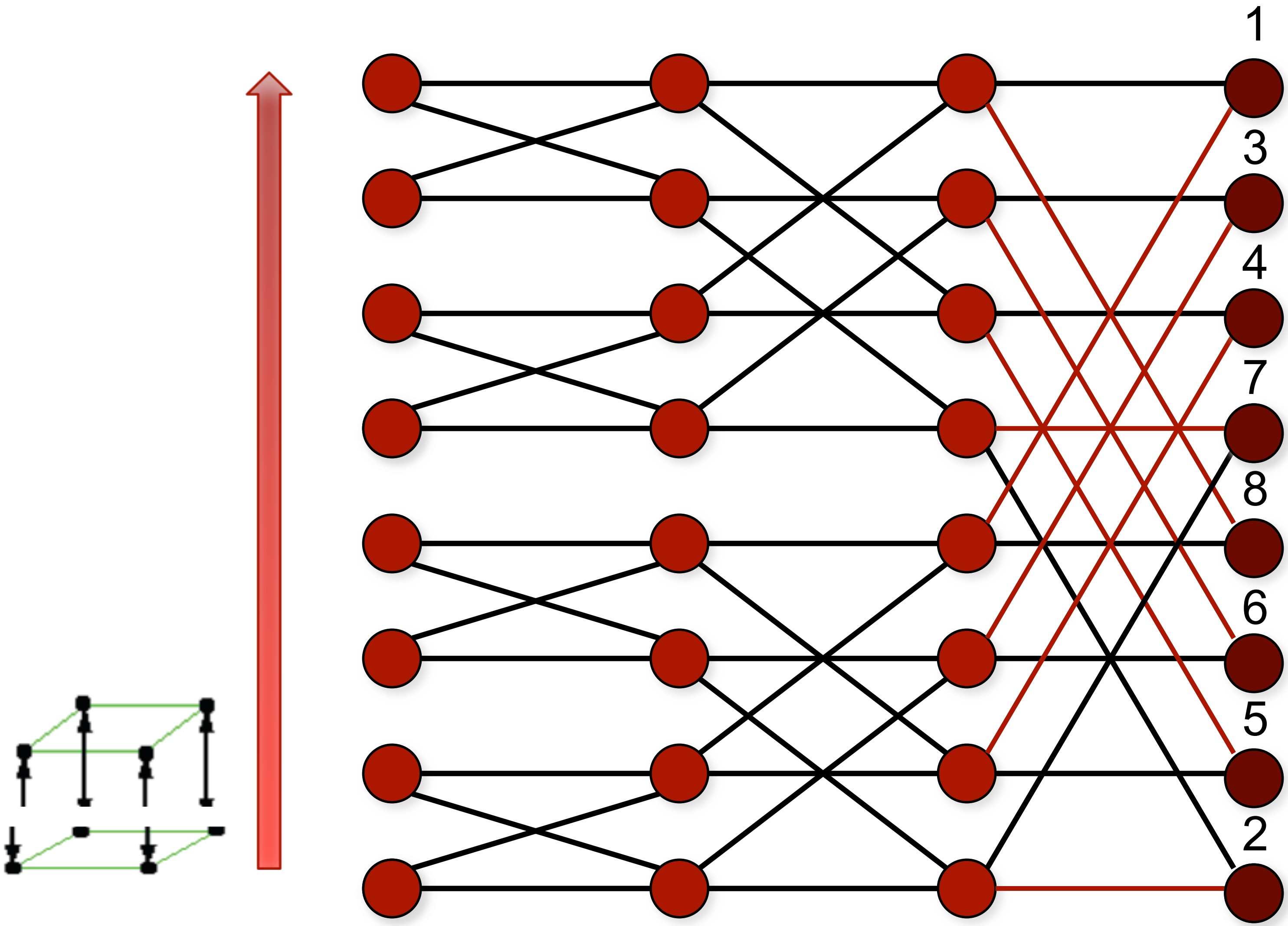


Bitonic sort su Butterfly

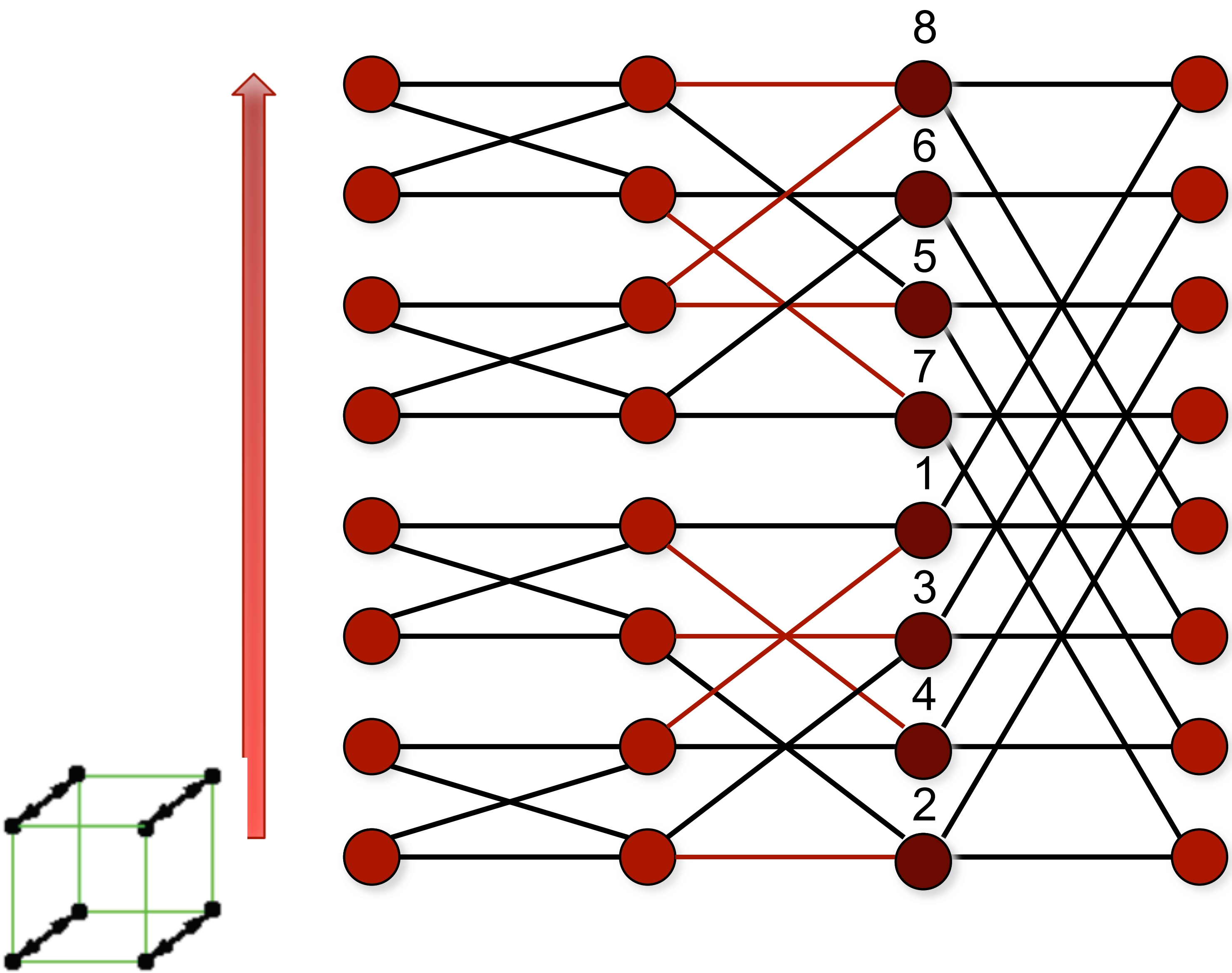




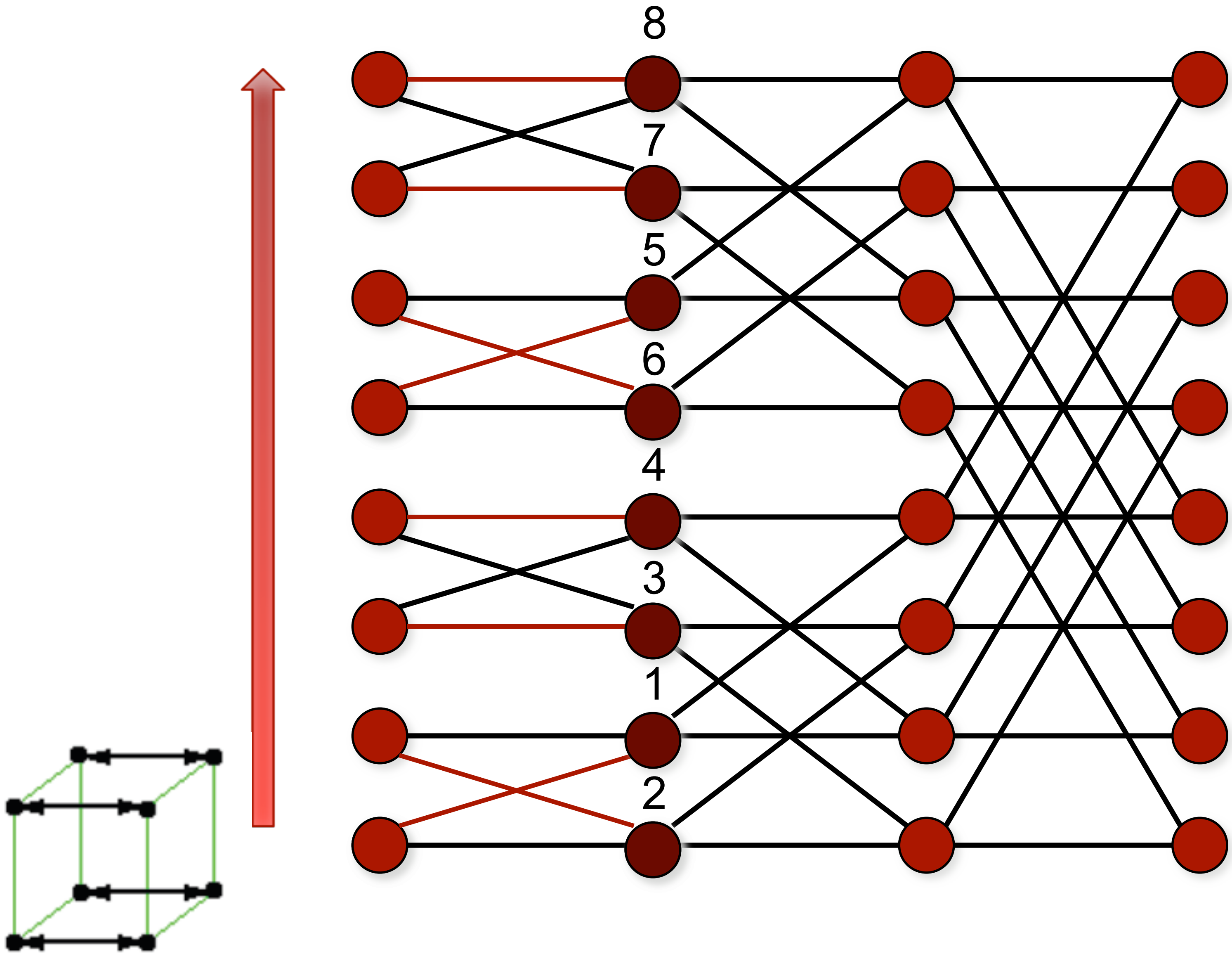
Bitonic sort su Butterfly



Bitonic sort su Butterfly



Bitonic sort su Butterfly



Bitonic sort su Butterfly