



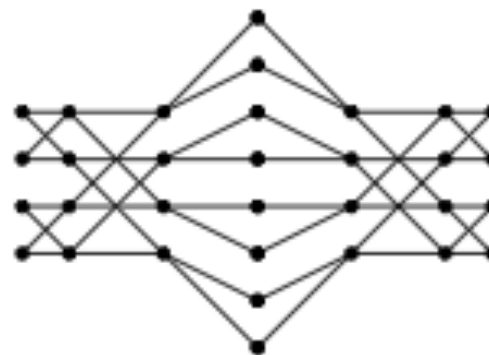
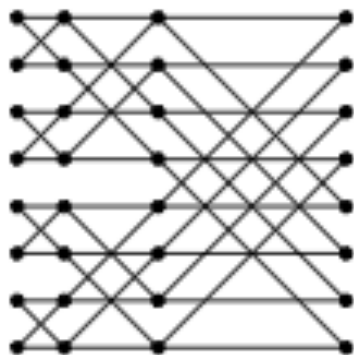
**LAYOUT DELLA BUTTERFLY DI
AREA OTTIMA**

LAYOUT DI AREA OTTIMA - IDEA (1)

- I due lavori che producono un layout di area ottima si basano sul lemma seguente:

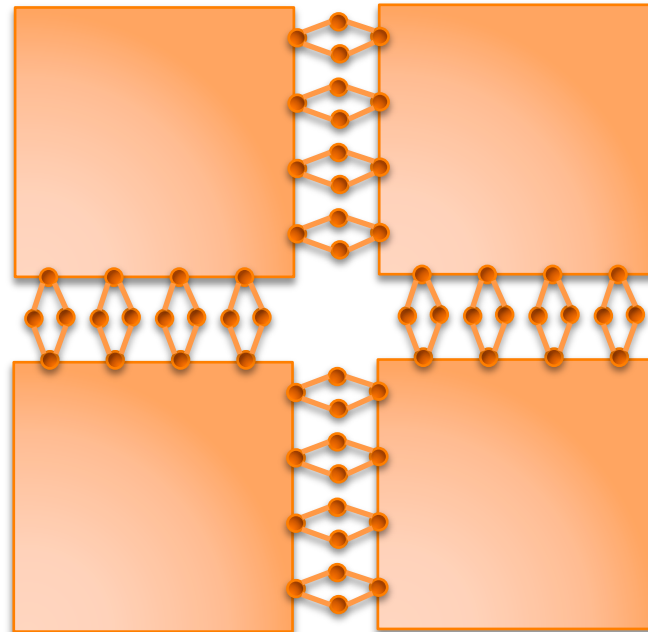
lemma: per ogni j e k interi non negativi tale che $0 \leq j \leq j+k \leq n$, il sottografo della n -Butterfly indotto dai nodi di livello $j, j+1, \dots, j+k$ è costituito da 2^{n-k} copie di k -butterflies.

- In particolare, se $j=0$ e $k=n-1$:



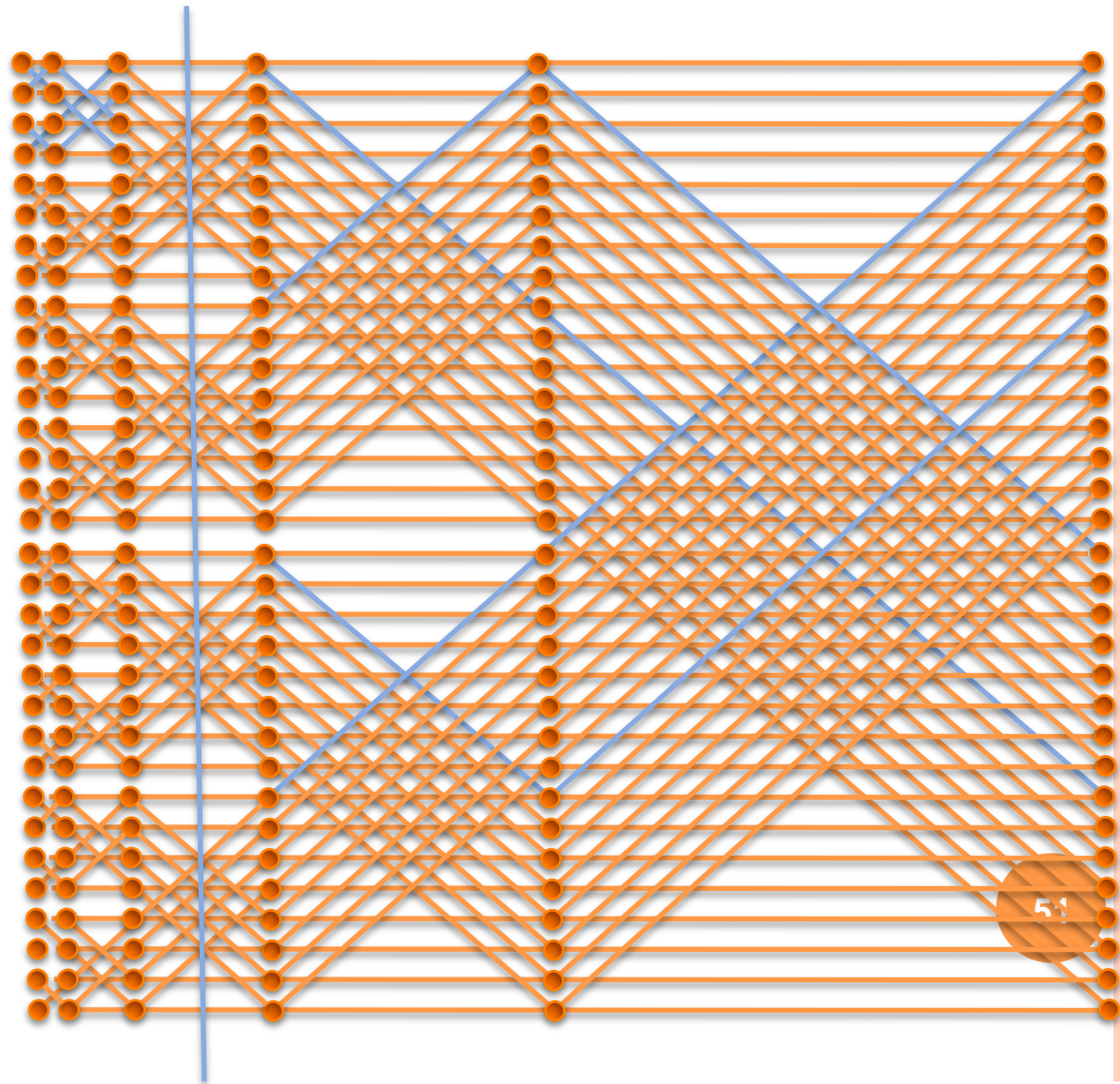
LAYOUT DI AREA OTTIMA - IDEA (2)

- A sua volta, una $(n-1)$ -Butterfly si costruisce come due $(n-2)$ -Butterflies connesse da un livello di nodi e un livello di archi.
- Tagliando dalla n -Butterfly gli input e gli output, si ottiene:



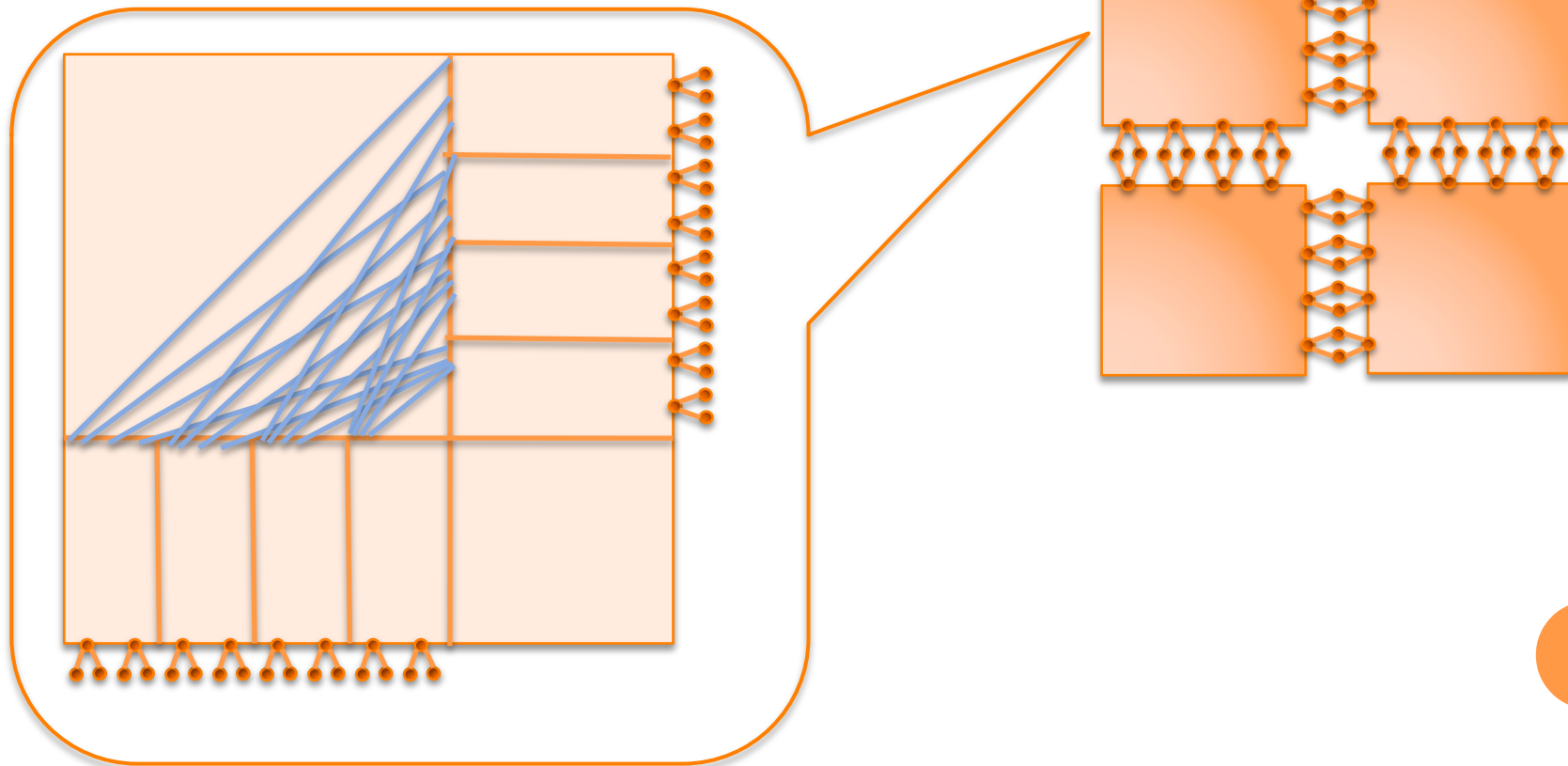
LAYOUT DI AREA OTTIMA - IDEA (3)

- Ogni $(n - 2)$ -Butterfly può, a sua volta, essere tagliata in tante Butterflies più piccole:



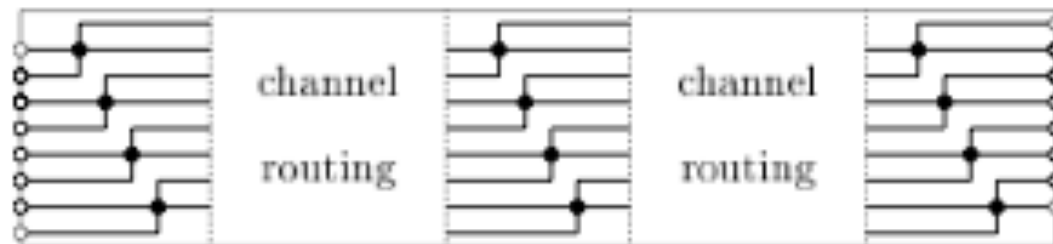
LAYOUT DI AREA OTTIMA - IDEA (4)

- Il disegno precedente si può specificare meglio:



LAYOUT DI AREA OTTIMA - IDEA (5)

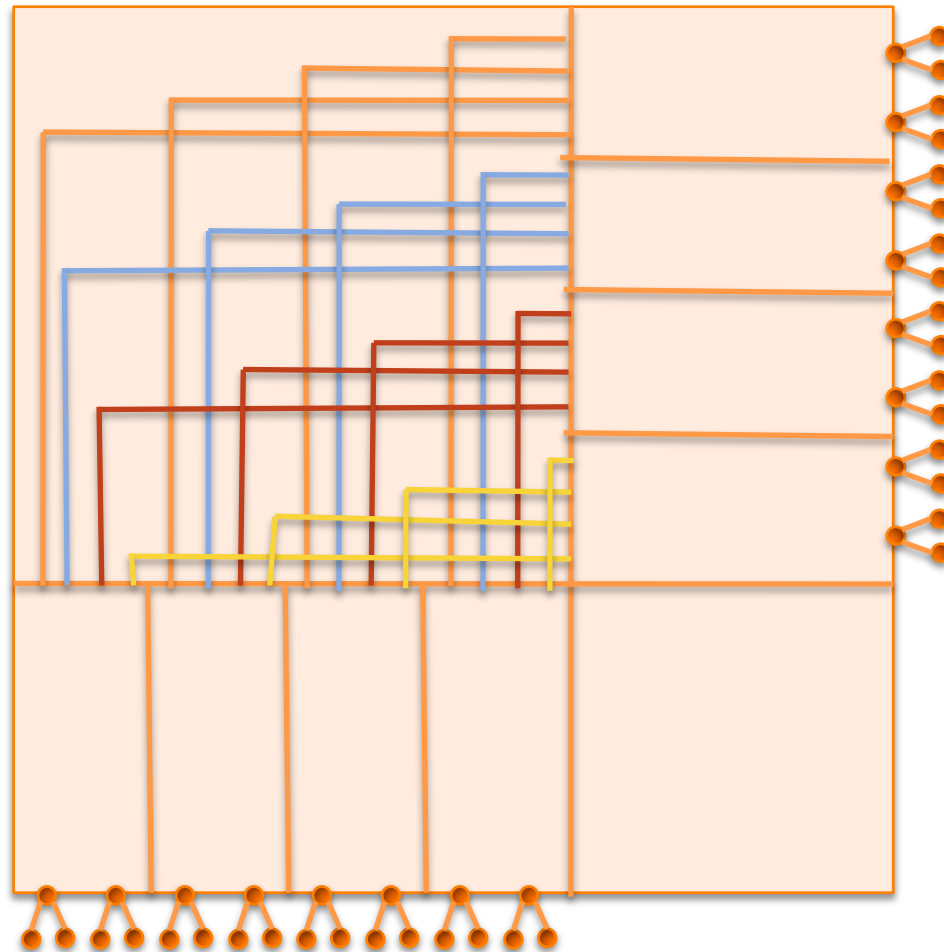
- ogni Butterfly rettangolare si disegna orizzontalmente (verticalmente) livello per livello:



- n.b. questi layout non sono ottimi ma, ciò nonostante, producono un disegno finale ottimo.

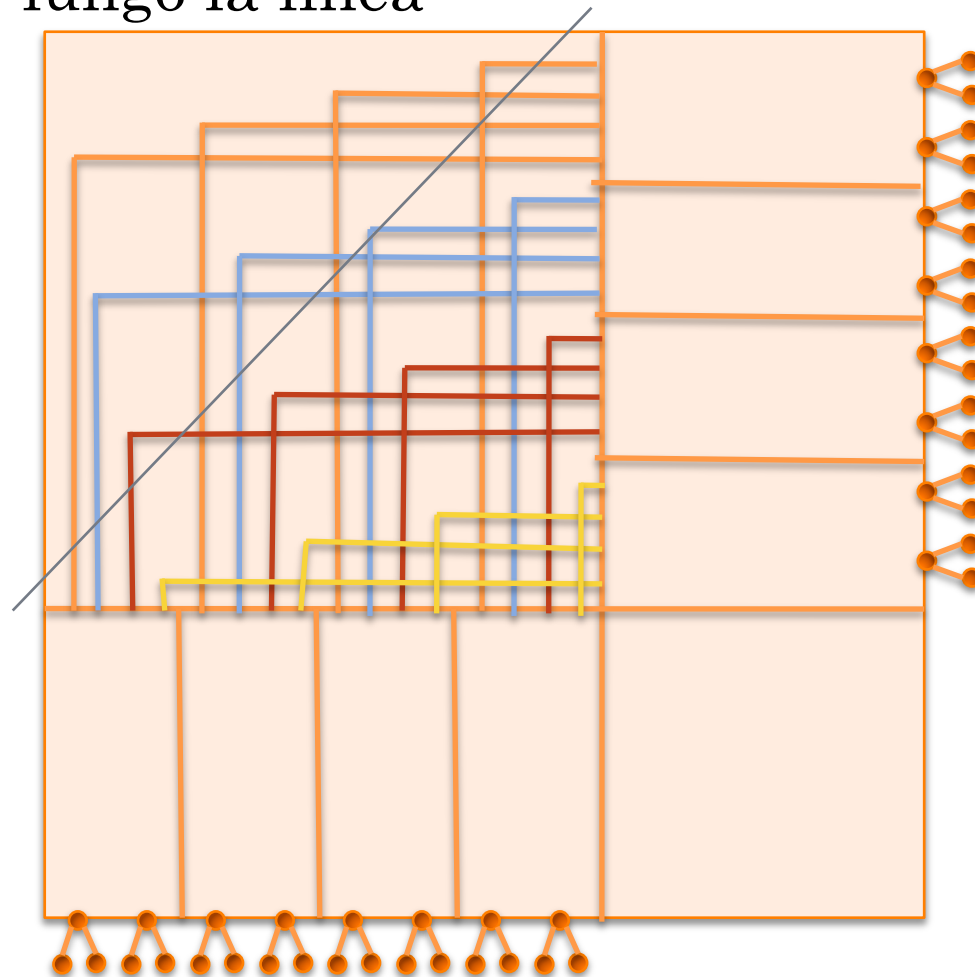
LAYOUT DI AREA OTTIMA - IDEA (6)

- rimane da connettere le piccole butterflies:



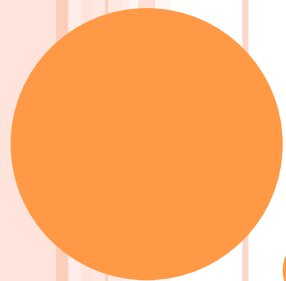
LAYOUT DI AREA OTTIMA - IDEA (7)

- nel caso di disegno con area “slanted” il disegno si “ripiega” lungo la linea



LAYOUT DI AREA OTTIMA - IDEA (8)

- E' possibile dimostrare le limitazioni inferiori per entrambi i modelli e le limitazioni superiori per entrambi i layout.
- Per chi volesse approfondire: TESINE (2 articoli)

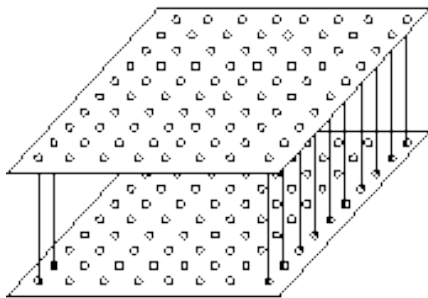


LAYOUT 3D DELL'IPERCUBO

CENNI SUL LAYOUT 3D

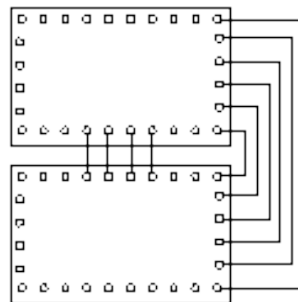
- Negli ultimi anni si è sempre più diffuso il layout 3D: la topologia è rappresentata su una serie di lastre sovrapposte
- ulteriore ottimizzazione della lunghezza dei fili e del # di svolte
- diminuzione dell'uso del silicio

3D Structure

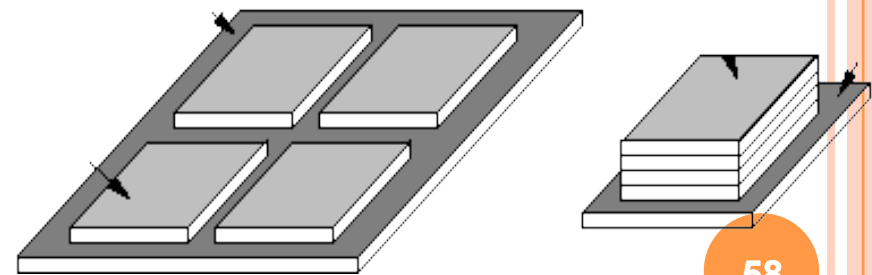


Via Hole

2D Structure



Wiring



IL PROBLEMA DEL LAYOUT 3D

Def. Un **layout 3D di una rete G** è una funzione uno a uno di G nella griglia tridimensionale tale che:

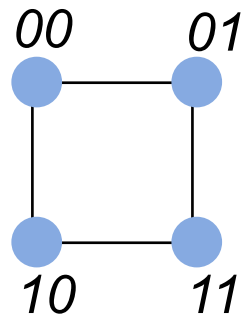
- i nodi sono mappati sui punti della griglia
 - meglio se i nodi sono sulla faccia esterna perché questo diminuisce i problemi di consumo (con conseguente generazione di calore) e semplifica la connessione ad altre parti dell'architettura
- i fili sono mappati su cammini indipendenti della griglia in modo che:
 - le mappature degli archi siano edge-disjoint;
 - non ci siano “knock-knees”
 - un arco non può passare per la mappatura di un nodo che non sia un suo estremo.

Scopo minimizzare il volume e mantenere i fili corti.

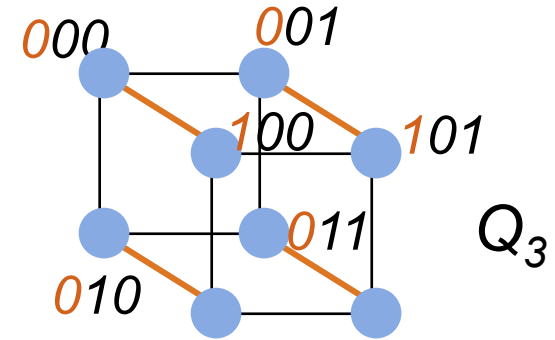
L'IPERCUBO (1)

- Largamente considerato per il calcolo parallelo per via dei suoi buoni parametri (regolarità, diametro logaritmico, fault tolerance, ...).
- **Def.** L'ipercubo di dimensione n , Q_n ha $N=2^n$ nodi, ciascuno etichettato con una stringa binaria di n bit. Due nodi sono connessi da un arco se le sue etichette differiscono in esattamente un bit.
- $Q_n: Q_{n-1} \equiv Q_{n-1}$

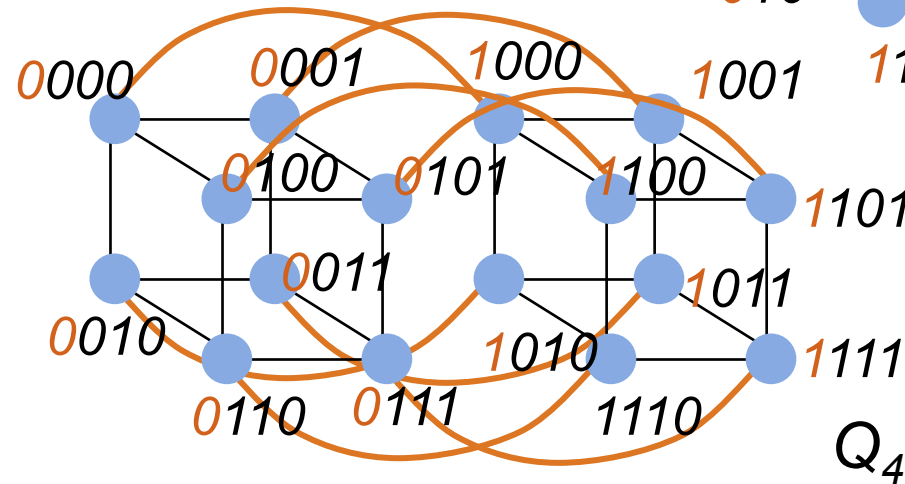
L'IPERCUBO (2)



Q_2



Q_3

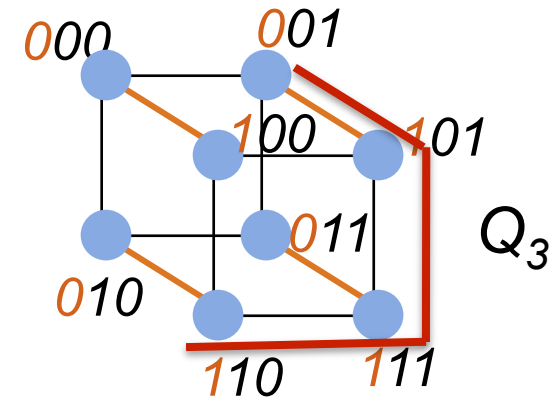


Q_4

Q_n è ottenuto connettendo nodi omonimi delle due copie di Q_{n-1} e riaggiustando le etichette.

N.B. questi archi formano un accoppiamento perfetto

L'IPERCUBO (3)



Proprietà: Q_n ha diametro $\log N$.

Dim. Ogni coppia di nodi

$u = u_1 u_2 \dots u_{\log N}$ e $v = v_1 v_2 \dots v_{\log N}$ sono connessi dal cammino:

$u_1 u_2 \dots u_{\log N} \rightarrow v_1 u_2 \dots u_{\log N} \rightarrow v_1 v_2 \dots u_{\log N} \rightarrow \dots \rightarrow v_1 v_2 \dots v_{\log N}$

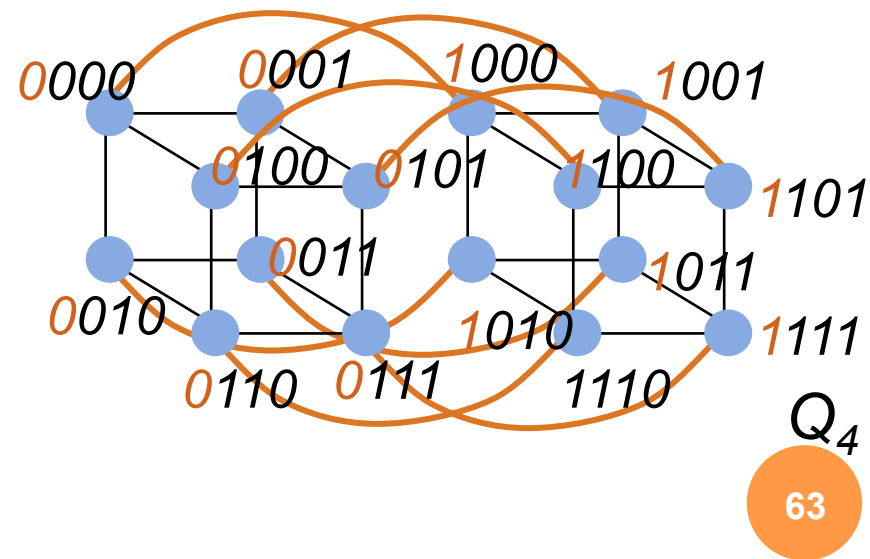
Il risultato segue banalmente.

L'IPERCUBO (4)

Def. : La **bisection width** di una rete è il minimo # di fili che bisogna tagliare per sconnettere la rete in due metà con uguale numero di nodi.

Proprietà. $BW(Q_n) = N/2$

Dim. (idea) gli archi rossi dividono l'ipercubo in due parti uguali, sono $N/2$ e non se ne possono tagliare di meno per ottenere lo stesso risultato.



L'IPERCUBO (5)

Proprietà. L'ipercubo è node symmetric ed edge symmetric, cioè rietichettando i nodi è possibile mappare ogni nodo (arco) in ogni altro.

Più precisamente, per ogni coppia di archi (u,v) e (u', v') esiste un automorfismo σ di Q_n t.c.

$$\sigma(u)=u' \text{ e } \sigma(v)=v'.$$

In effetti, esistono molti tali automorfismi.

L'IPERCUBO (6)

- Diamo un layout 3D di Q_n t.c. i nodi giacciono su due facce opposte della bounding box, equamente divisi.
- In particolare:
 - limit. inferiore sul volume: $\Omega(N^{3/2} \log^{1/2} N)$
 - limit. superiore sul volume: $O(N^{3/2} \log N)$

LOWER BOUND (1)

Th. Una limitazione inferiore per un layout 3D di Q_n è $\Omega(N^{3/2} \log^{1/2} N)$, se:

- tutti i nodi sono rappresentati come lastre rettangolari con area proporzionale al loro grado
- tutti i nodi giacciono su due piani opposti
- i nodi sono equamente divisi.

LOWER BOUND (2)

Idea della dim. ($\Omega(N^{3/2} \log^{1/2} N)$)

Considera un ipotetico layout di volume V :

di nodi = $N = 2^n$ $\Rightarrow H \times L = \Omega(N/2 \log N)$

grado = $n = \log N$

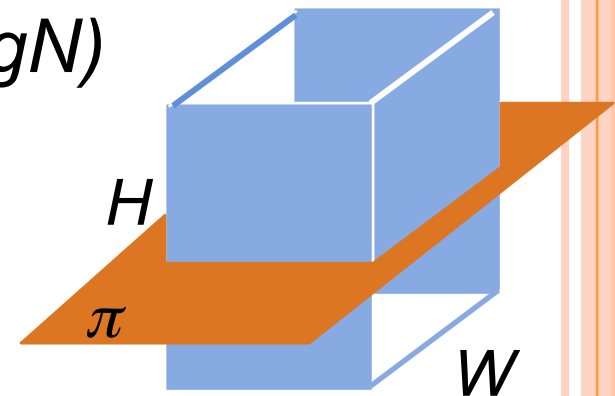
1. $W \geq H, L \Rightarrow \text{ok};$

2. $W < H \text{ o } < L$

considera π ;

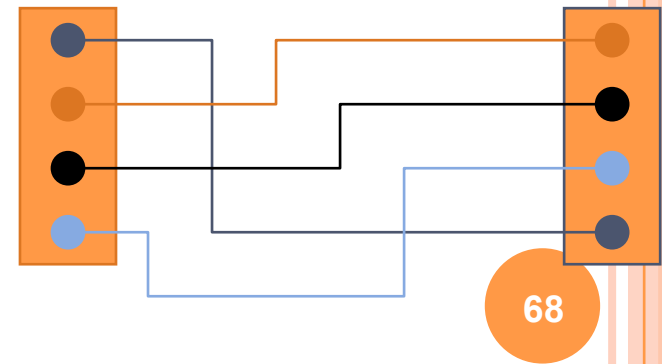
$\text{area}(\pi \cap V) = L \times W \geq BW(Q_n) \geq N/2$

$\Rightarrow W \geq N^{1/2} \log^{1/2} N \Rightarrow \text{ok}.$



UPPER BOUND (1)

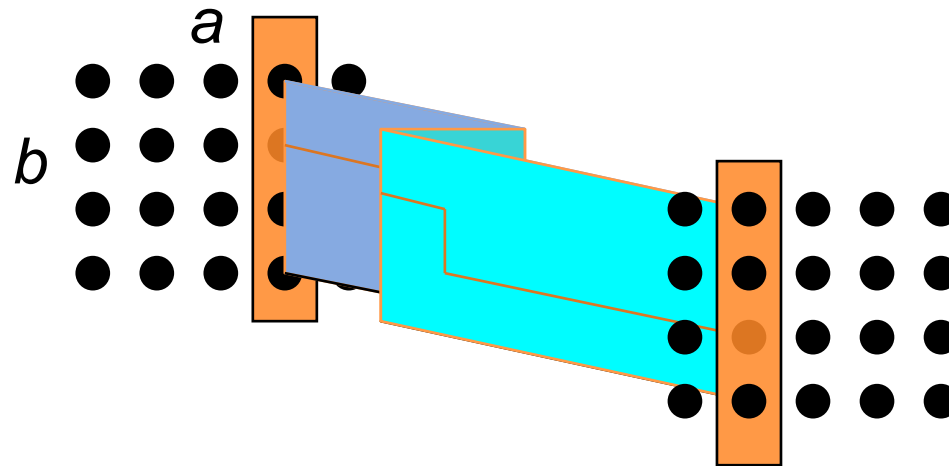
- Q_n si può costruire da 2^{n-4} copie of Q_4 , aggiungendo degli archi e riaggiustando le etichette.
- Q_n è bipartito (# di zeri pari vs. # di zeri dispari).
- **Def.** Un *k-channel routing* richiede una griglia 2D e due insiemi S and S' ciascuno di k nodi che vanno connessi da una funzione 1-1. S ed S' sono posti su lati opposti della griglia.



○ $A \leq (k+1) \times (\frac{3}{2}k + 2)$ [Pinter '82]

UPPER BOUND (2)

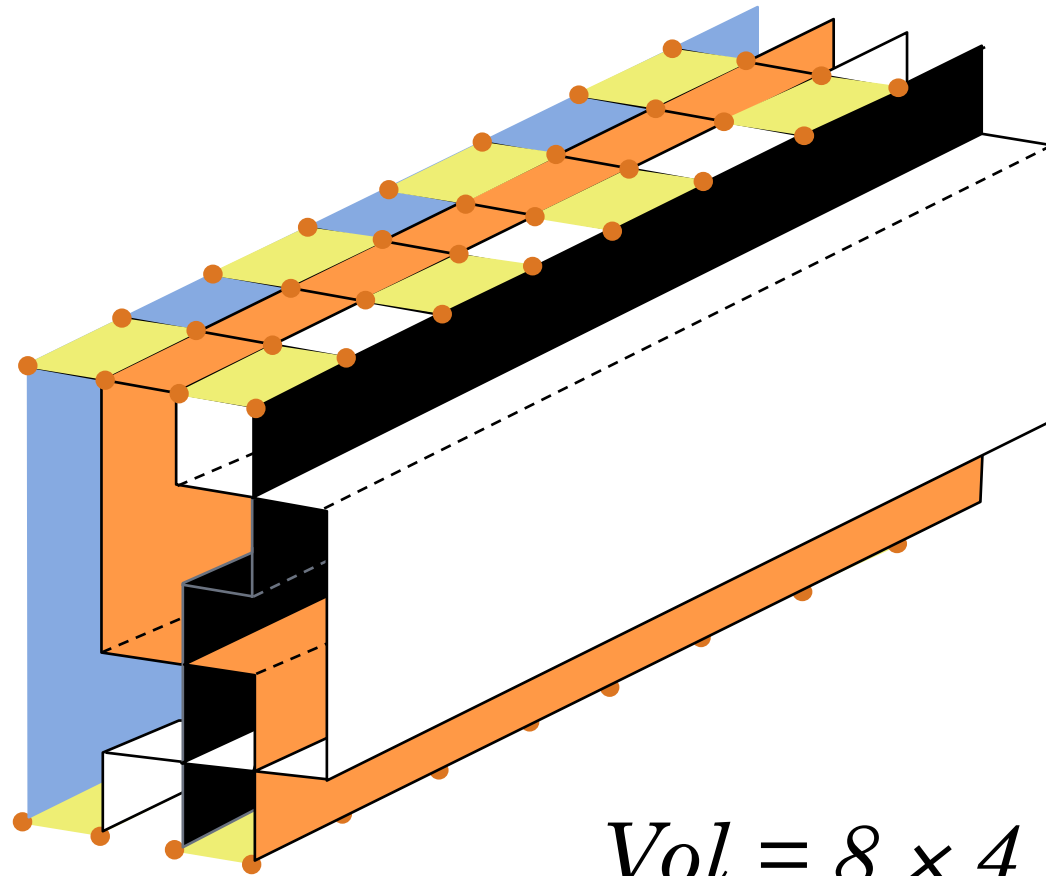
Def. $a \times b$ -3D double channel routing:



Th. $Vol \leq (a+1) \times (b+1) \times \max\{3/2 a+2, 3/2 b+2\}$

UPPER BOUND (3)

Building block: layout 3D di Q_4



$$Vol = 8 \times 4 \times 5$$

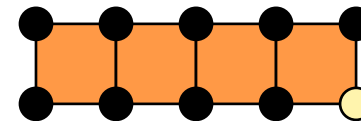
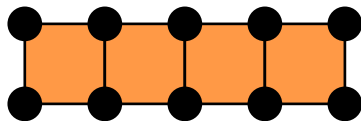
UPPER BOUND (4)

$Q_n = 2^{n-4}$ copie di Q_4 + alcuni archi.

Lemma: questi archi costituiscono un 3D double channel routing.

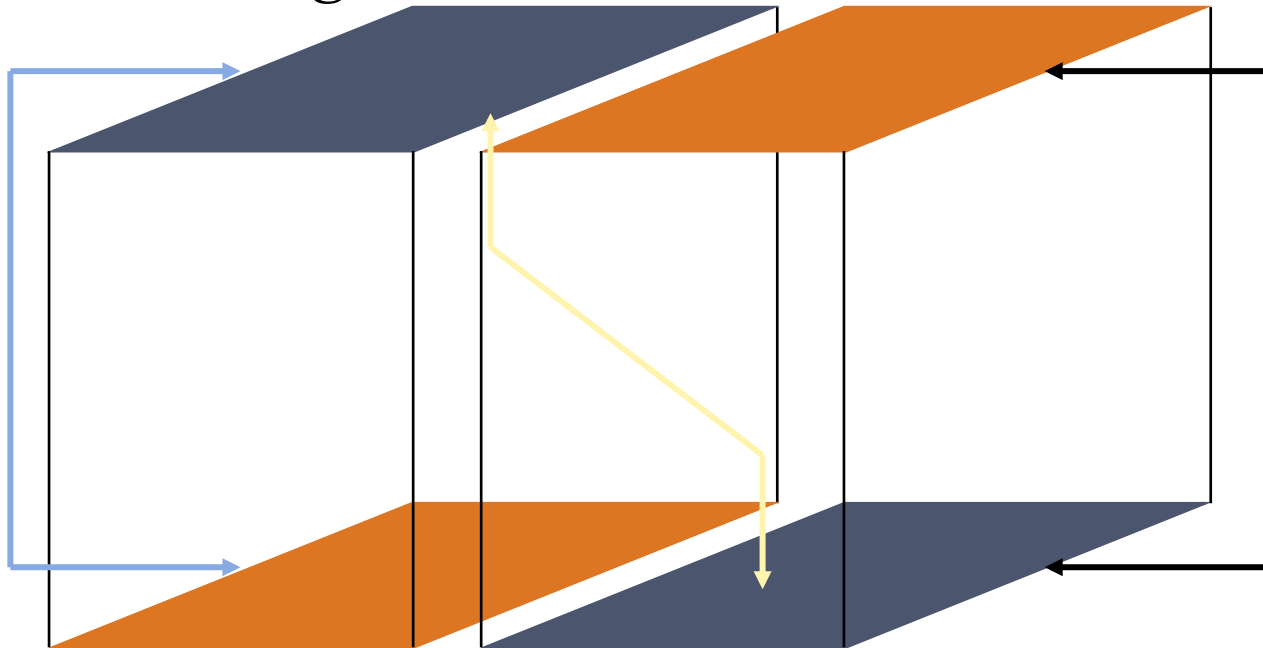
Idea della dim.: per induzione su n .

○ Nodi:



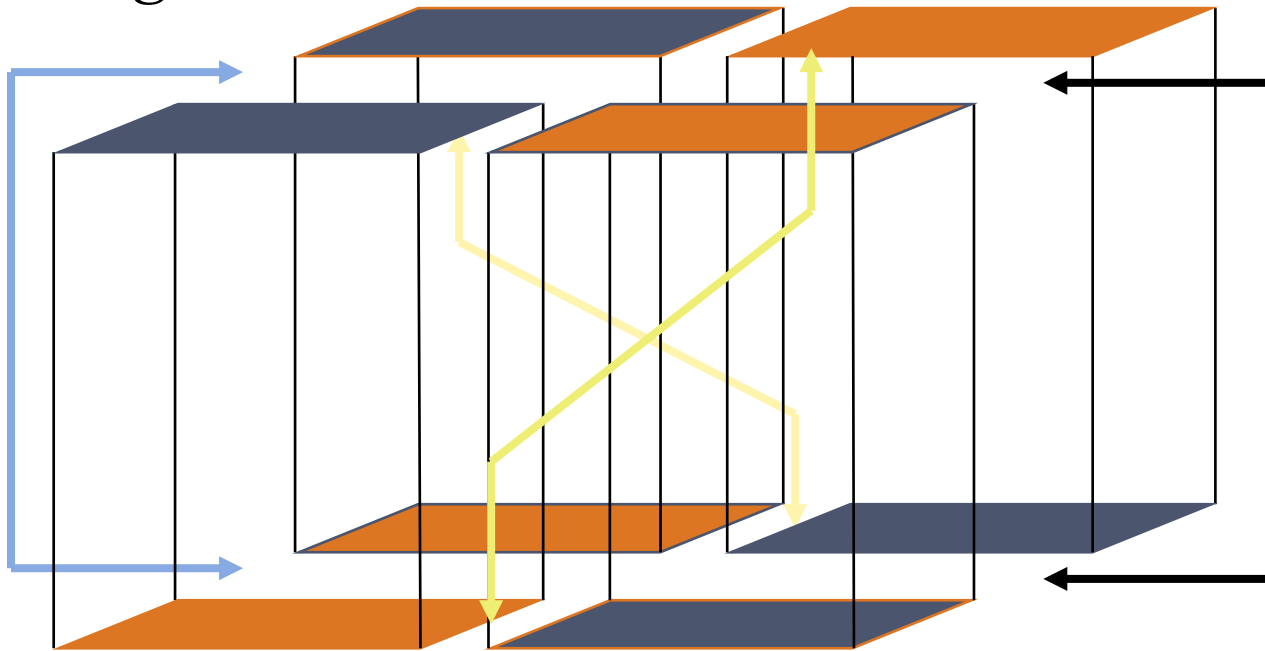
UPPER BOUND (5)

n dispari: l'insieme di archi che connettono le copie dei Q_4 è partizionato in 3 insiemi che, tutti insieme, costituiscono un 3D channel routing.



UPPER BOUND (6)

n pari: l'insieme di archi che connettono le copie dei Q_4 è partizionato in 3 insiemi che, tutti insieme, costituiscono un 3D channel routing.



UPPER BOUND (7)

Th.: Esiste un layout 3D di Q_n con volume $\Theta(N^{3/2} \log N)$.

In particolare:

$$W_n = \Theta(N^{1/2})$$

$$L_n = \Theta(N^{1/2} \log N)$$

$$H_n = \Theta(N^{1/2})$$

Per approfondire: TESINA