Sai

Physical Organization of Parallel platforms

PRAM : Ideal Model, p processors,
unbounded global Memory
All p access same memory

Simultaneous Memory Access

EREW

CREW

ERCW

CRCW

Common

Arbitrary

Priority

Sum

For p processors & m memory words,

$\Rightarrow \Theta(pm)$ interconnections

→ Very expensive

→ Not practical

$\Rightarrow$ Study of Interconnection Networks

Between Processing Nodes

Between Processors & Memory modules

Link, Switch, Network Interface $\Rightarrow$ Network

Link : Physical media connection (wires)
Capacitive coupling, Signal Attenuation

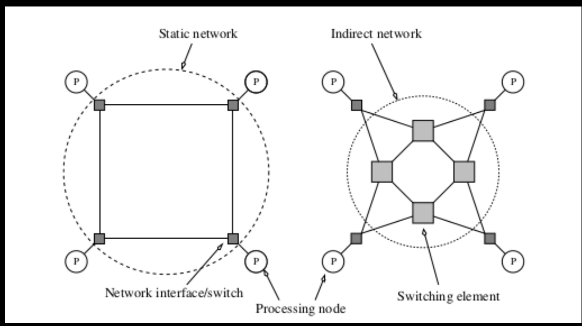
Switch : Device with set of input ports & Output ports
Capable of mapping input to Output ports
Total No. of ports = Degree of a switch
Can internally buffer, perform routing & capability to multicast.
Cost of Switch dependent on mapping hardware, peripheral hardware & packaging
($\sim$ degree^2) ($\sim$ degree^1) ($\sim$ degree^1)

Network Interface : Has input & Output ports that pipe data into and out of the network
Packetizing data
Buffer incoming and outgoing data for matching speeds of network and processing elements
Error checking
Position of placement can determine performance
(I/O Bus, Memory bus)

Network

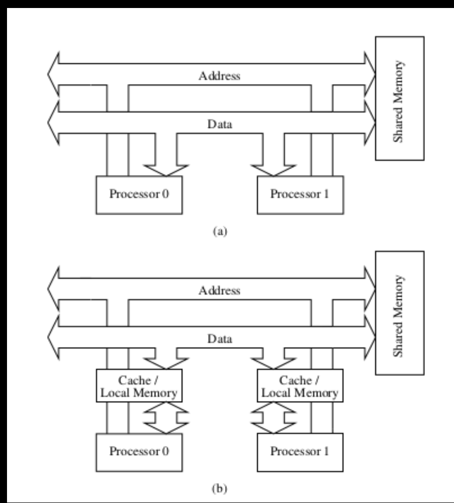
Static (Direct)

Dynamic (Indirect)



Network Topologies

1. Bus-based Networks



- Shared Medium
- Cost of Network scales linearly with P
- Distance between any two nodes is constant
- Ideal for broadcasting
- But limitations on performance as no. of nodes increase
- Add Cache at each node to reduce demands on bus bandwidth

p processors, each accessing k data items, each data access takes time t, then

Minimum time needed to complete the application

$= t\,k\,p$

Out of the k accesses, if 50% accesses are local, and if cached, then,

Minimum time needed to complete the application

$= 0.5ktp + 0.5kt$

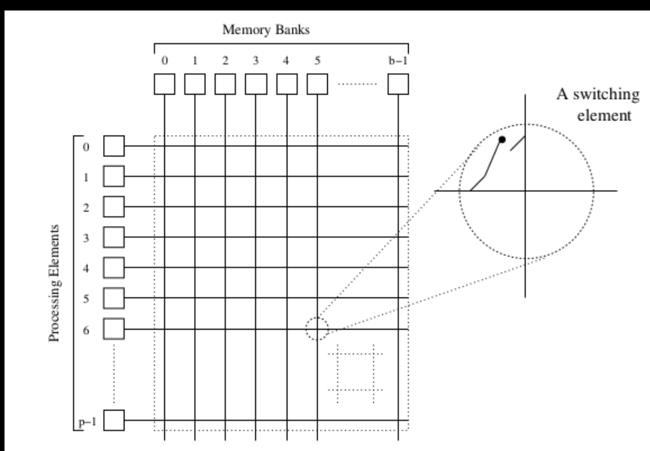
(assuming same cache access time)

and as p becomes large,

$\approx 0.5tkp \Rightarrow$ 50% improvement in the lower bound.

II. Crossbar Networks

p processors to b memory banks



- Non-Blocking network
- But total no. of switching nodes $= \Theta(pb)$
- Reasonable to assume $b \geq p$

$\Rightarrow$ No. of switching nodes $= \Omega(p^2)$

- Not scalable in terms of cost.

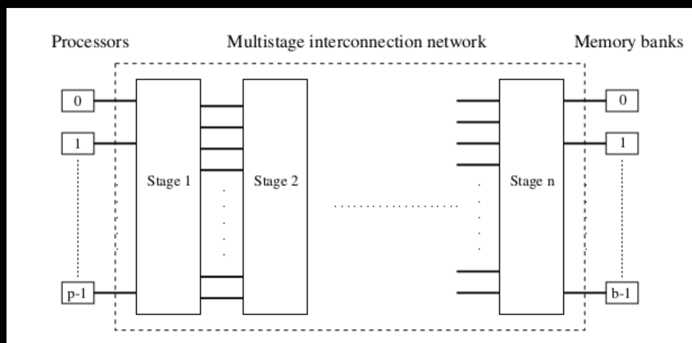
III. Multistage Networks

Crossbar - Not scalable in terms of cost

Bus - " " " " " performance

$\Rightarrow$ Intermediate solution

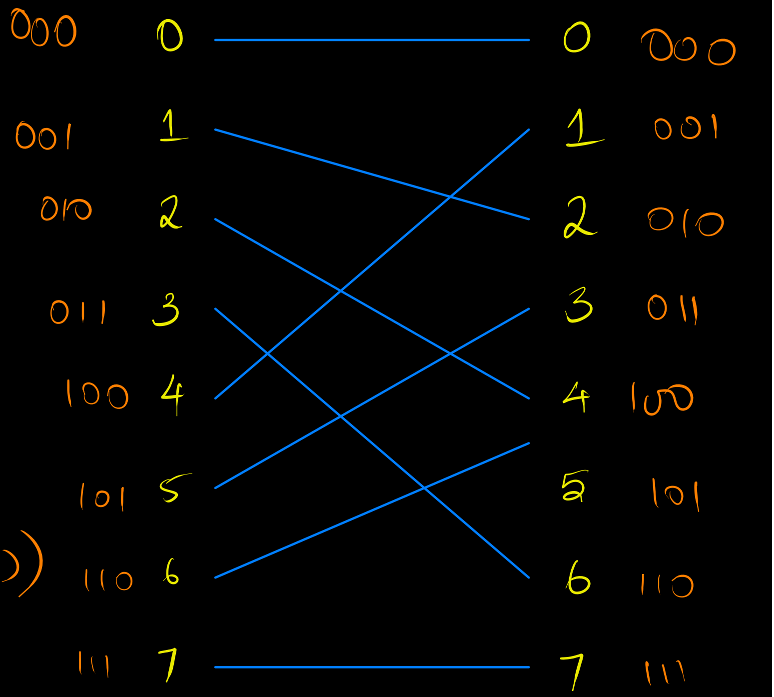
General Schematic :



Omega Network :

- Commonly used
- $p = b$
- $\log p$ stages
- $p/2$ switches at each stage
- $\Theta(p \log p)$ switches (not $\Theta(p^2)$)

Perfect Shuffle (Left Rotate)

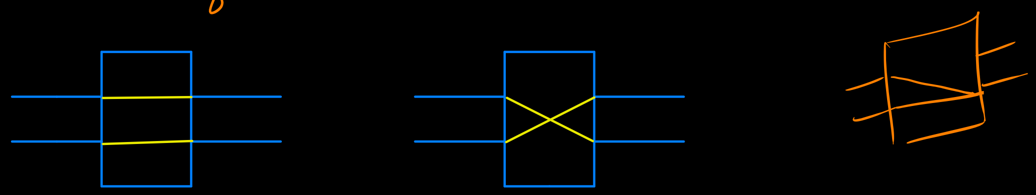


In integer arithmetic

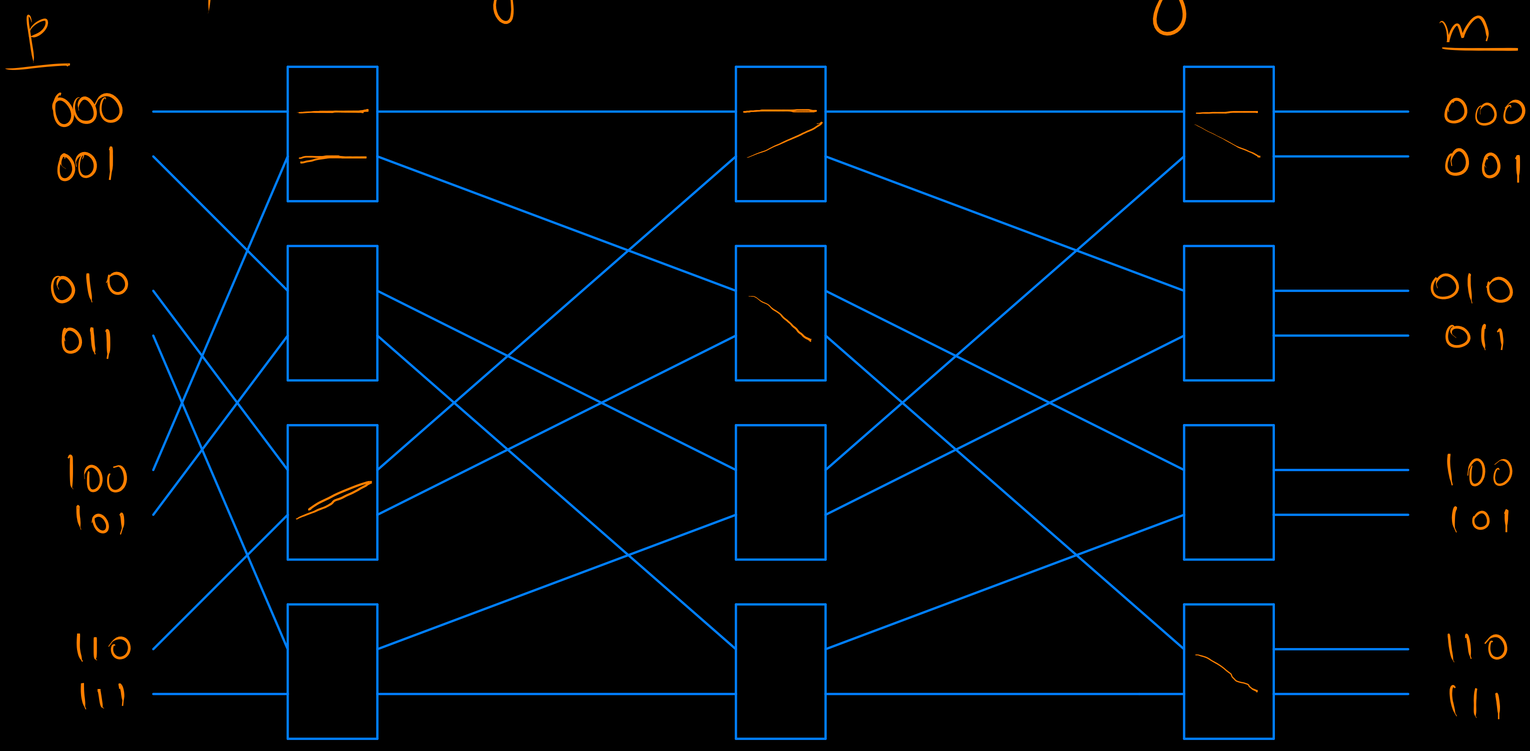
$$j = \begin{cases} 2i & 0 \le i \le p/2 - 1 \\ 2i + 1 - p & p/2 \le i \le p-1 \end{cases}$$

Switch

Pass-Through   Cross-over



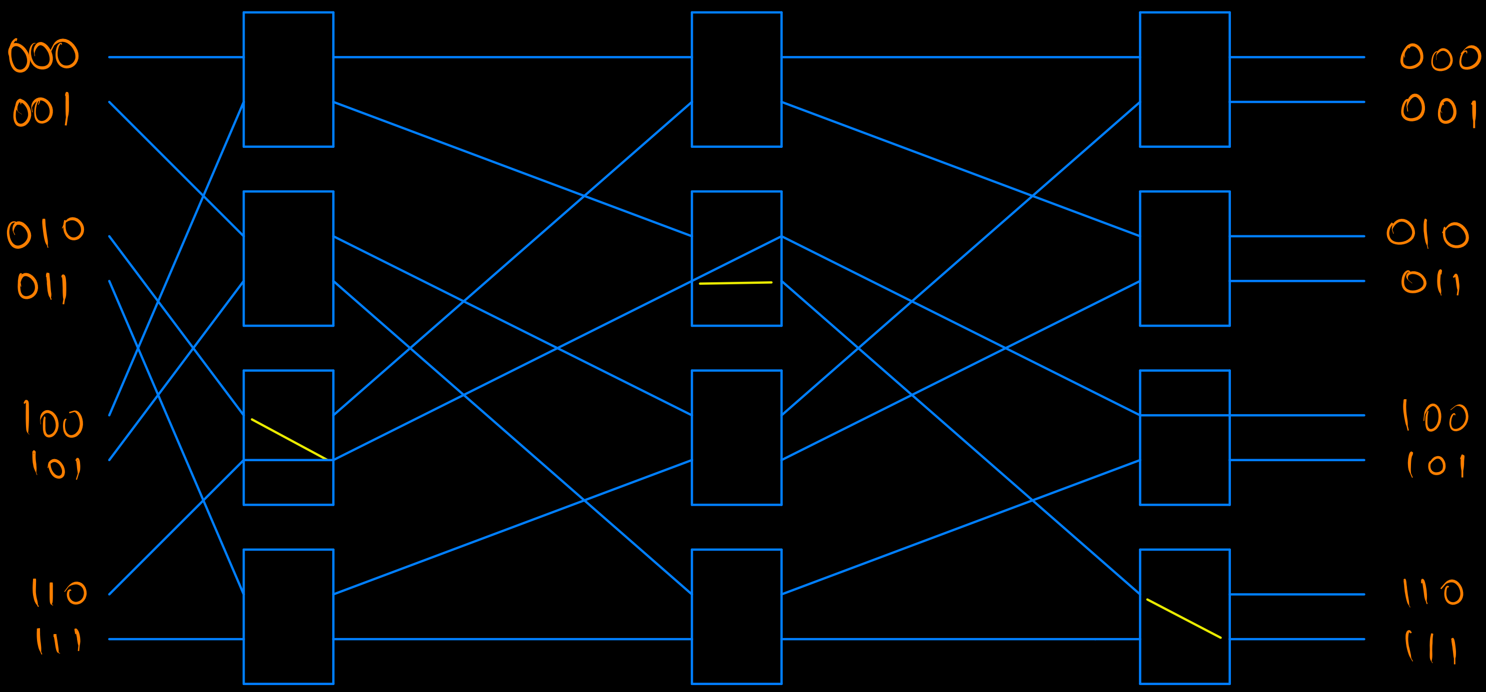
Complete Omega Network (Routing)



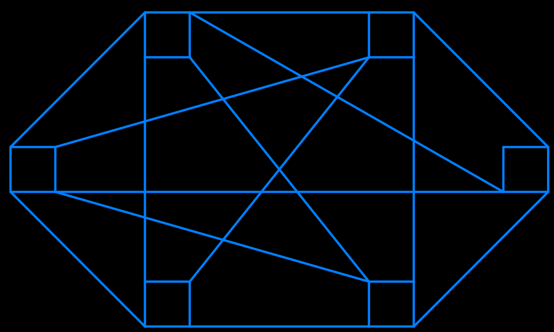
110
000
100

001
000
111

010 - 111 and 110 - 100

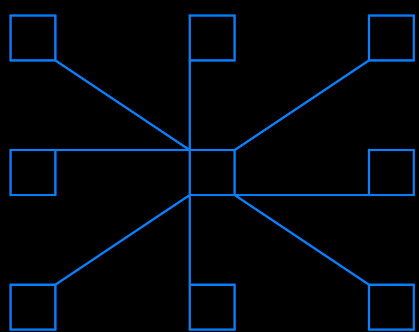


IV. Completely - Connected Network



- Direct Communication link to all other nodes
- Ideal : All messages in One Step and totally non-blocking
- Static Counterpart of Crossbar switching network

V. Star - Connected Network

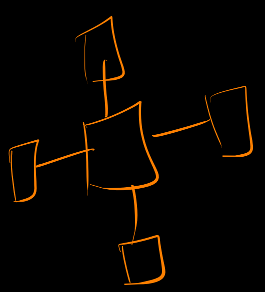


- One processor acts as the Central Processor
- All others link to the Central
- All Communication through Central processor
- Similar to bus-based networks
- Central processor can be a bottleneck

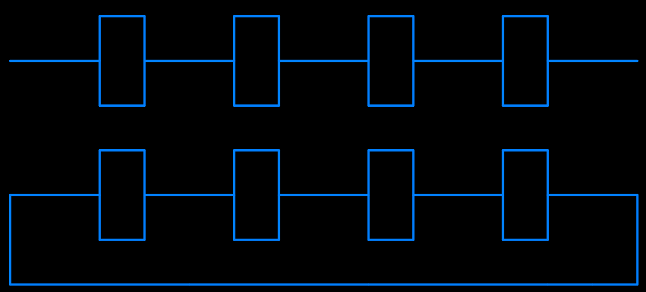
VI. Linear Arrays, Meshes, k-d Meshes

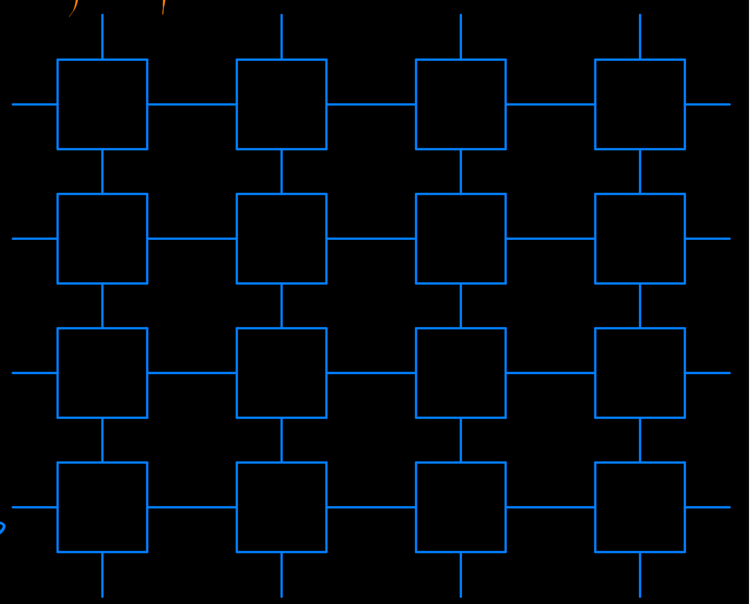
- Sparser network family
- Each node has 2 neighbours - left & right
- Linear + Wrap-around = Ring / 1-d Torus
- Linear + 2d = Mesh

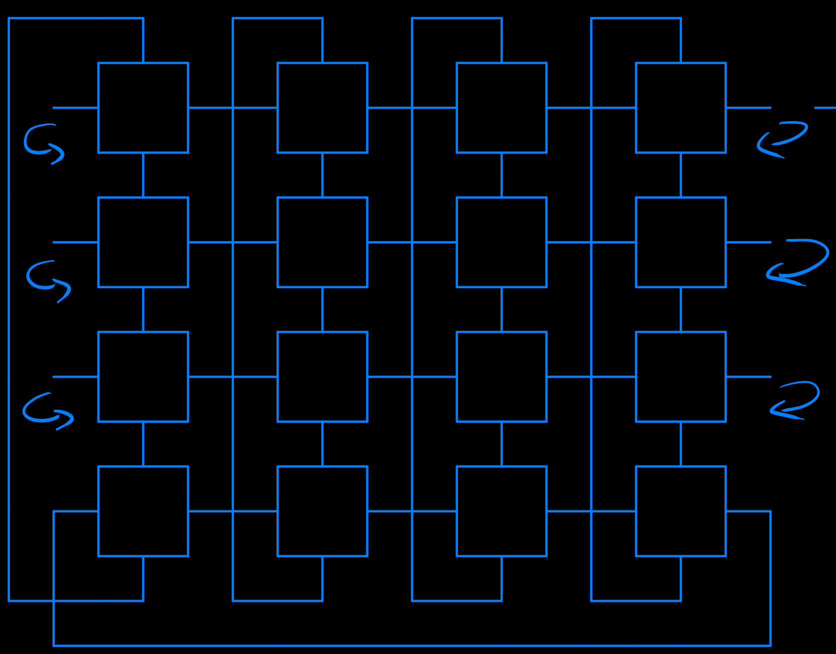
↓



- $\sqrt{p}$ nodes in each dimension
- (i, j) pair identification
- Each node connected to 4, except peripheral nodes







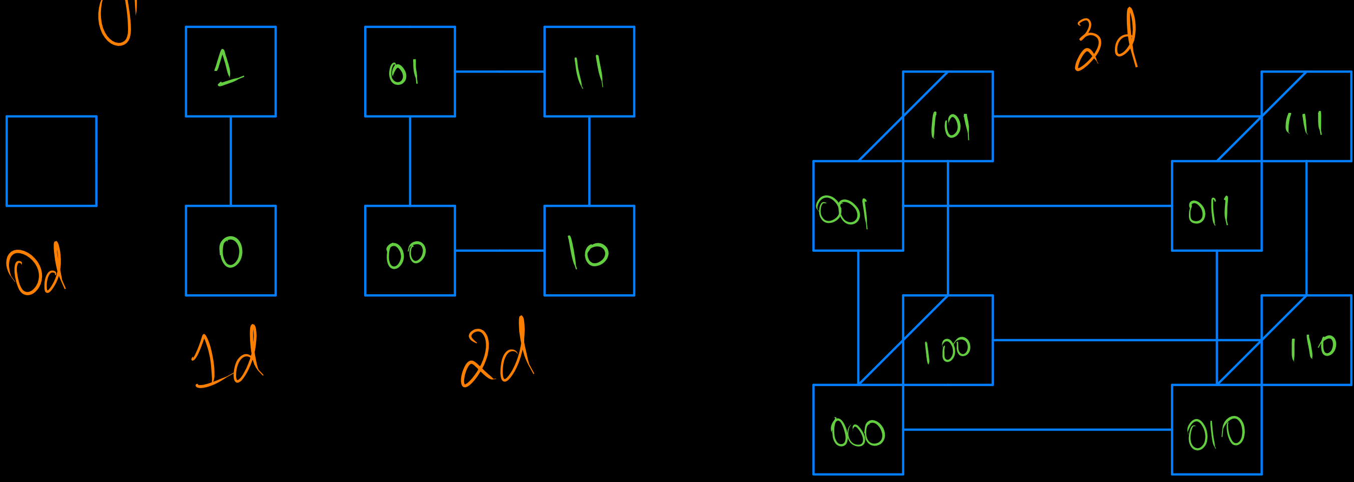
Neighbours in a 3-d Mesh ?

k-d Mesh ?

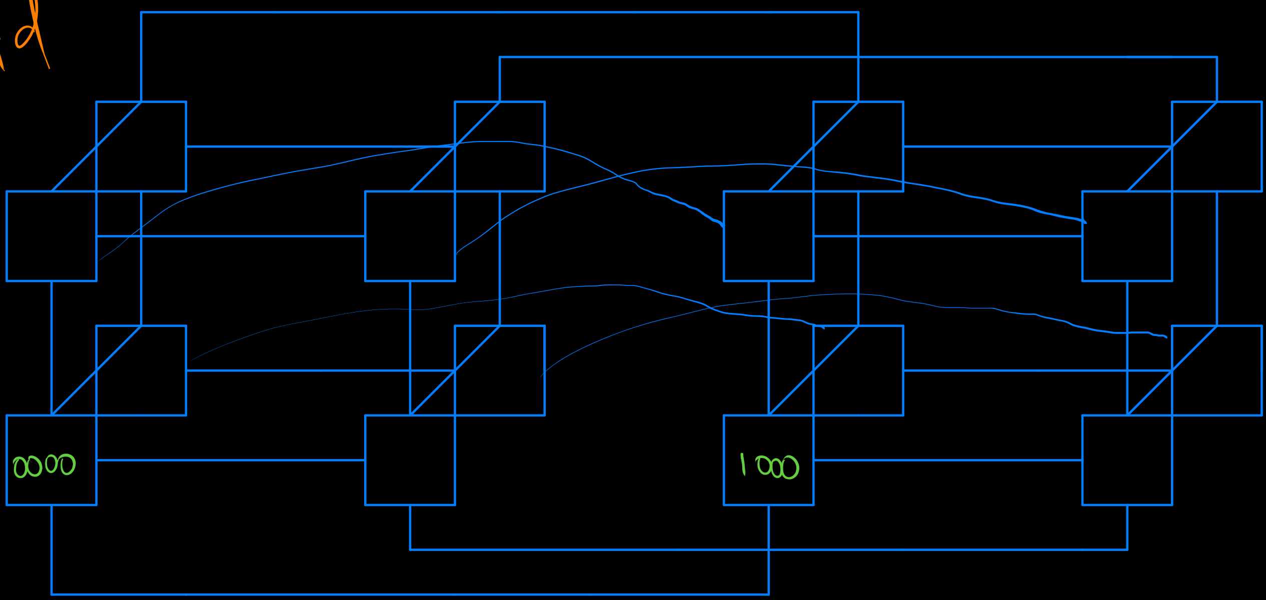
d-dimensions of $=k^d$

k-nodes along each dimension.

VII. Hypercube Network



4d



Numbering scheme for hypercube:

k-dimensional hypercube + k-dimensional hypercube $\longrightarrow$ k+1 - dimensional hypercube

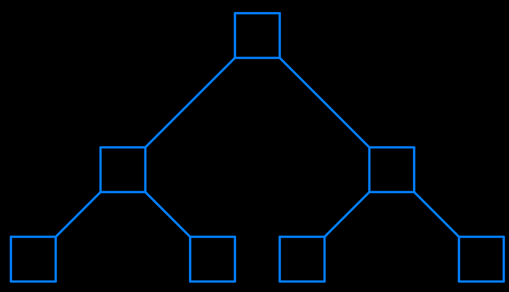
$\Rightarrow$ a) $p/2$ nodes + b) $p/2$ nodes $\longrightarrow$ $p = 2^{k+1}$ nodes

Prefix a) with 0

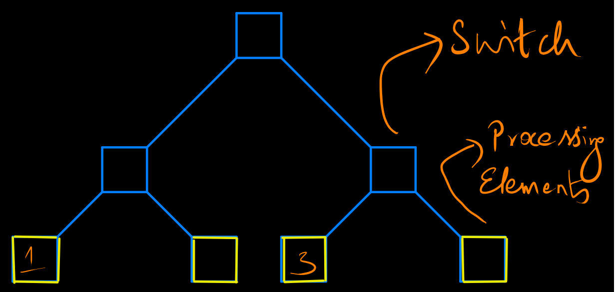
Prefix b) with 1

$\Rightarrow$ Minimum distance between two nodes = No. of bits that are different in the two labels

VIII · Tree based Networks



Static



Dynamic

- Only one path between any pair of nodes
- Star & Linear arrays are special cases
- Source node sends message up the tree until it reaches root of the smallest subtree containing both source & destination nodes.
  Then message is routed down to destination
- Bottleneck in the nodes higher up the tree
- Alleviated in dynamic tree by increasing no. of communication links and switching nodes closer to the root – Fat Tree