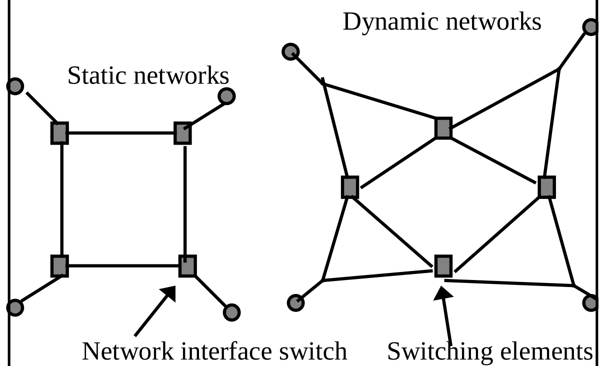
Interconnection networks

- Shared-address-space computers and message-passing computers can be constructed by connecting processors and memory units using a variety of interconnection networks.
  - Static Networks (direct networks)
    - Point-to-point communication links among processors
    - Used to construct message passing computers
  - Dynamic Networks(indirect networks)
    - Built using switches and communication links
    - Some of the switches may be connected to switches only.
    - Communication links are connected to one another dynamically by the switching elements to establish paths among processors and memory banks
    - Used to construct shared-address-space computers

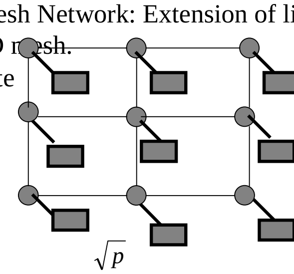


Evaluating Interconnection Networks

- Diameter
  - Largest distance between two switch nodes
- Bisection Width
  - Minimum number of edges between switch nodes that must be removed in order to divide the network into two halves.
  - Proving the bisection width of a network is difficult though it may look easy.
  - Higher bisection width desirable
- Edges per switch node
  - If the number of edges /switch is constant independent of the network size, then the processor organizes and scales more easily

Static Interconnection Network

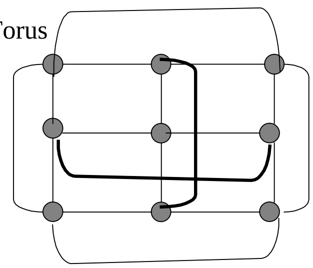
- Mesh Network: Extension of linear array to 2D mesh.



$\sqrt{p}$

Static Interconnection Network

- Torus

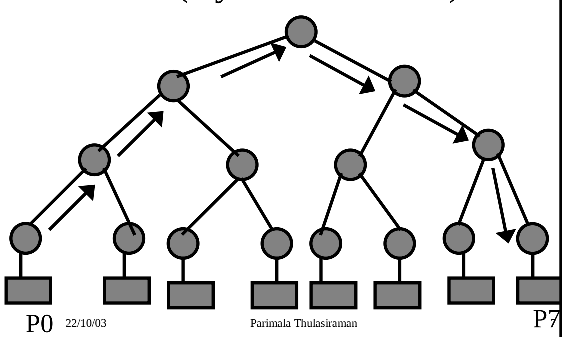


3D mesh (wraparound) mesh : TERA, Cray T3D, J-Machine

2D Mesh

- Assume n nodes with no wraparound connections
- Diameter: $2(\sqrt{p}-1)$
- Bisection width: $\sqrt{p}$
- Constant number of edges per switch

Binary Tree (Dynamic network)

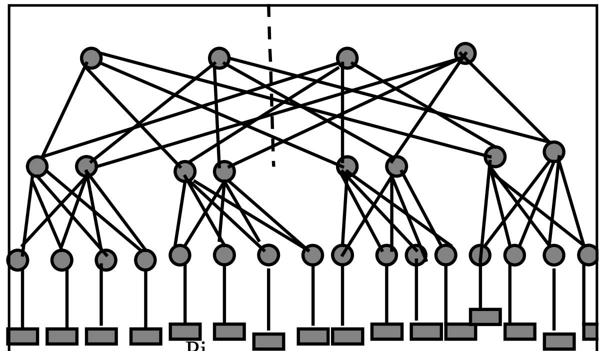


Binary Tree

- Diameter:
- Bisection width : 1
- Edges per switch node is not constant

Hypertree
(Dynamic Network)

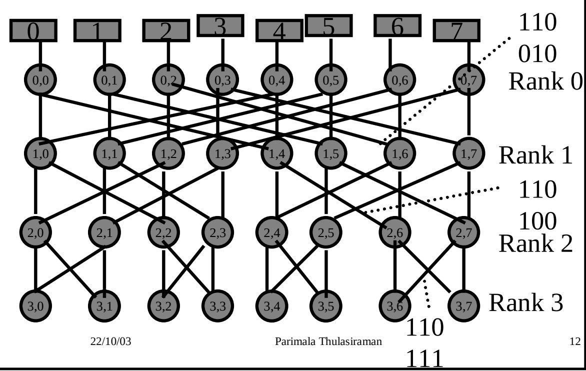
- Low diameter of a binary tree
- Improved bisection width



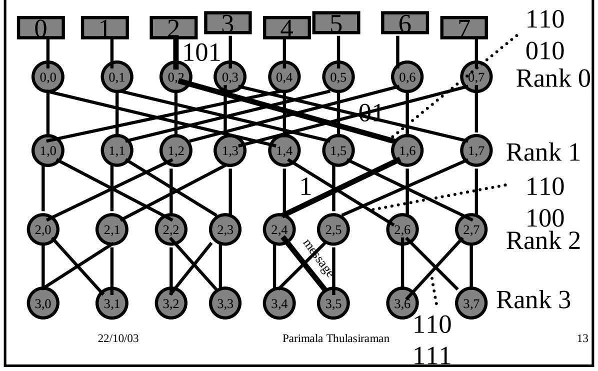
Hypertree

- Diameter: 2d
- Bisection width:

Butterfly Network
(Dynamic Network)



Butterfly Network (Dynamic Network)

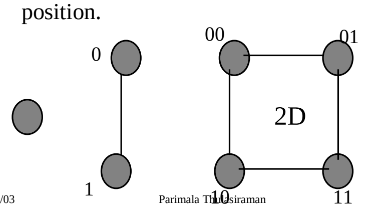


Butterfly Network

- Diameter: log n
- Bisection width: n/2

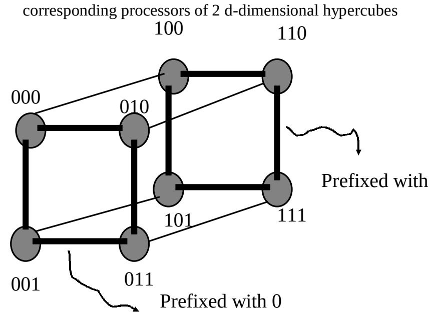
Static Interconnection Network

- Hypercube Network(binary n cube)
  - d-dimensional hypercube consists of $p=2^d$
  - 2 processors are connected by a direct link if their binary labels differ in exactly 1 bit position.



3D hypercube

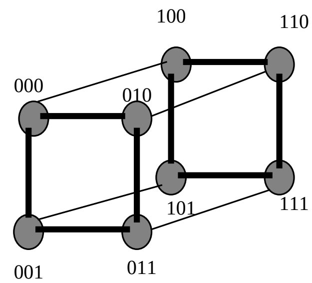
Construct (d+1) dimensional hypercube by connecting the corresponding processors of 2 d-dimensional hypercubes

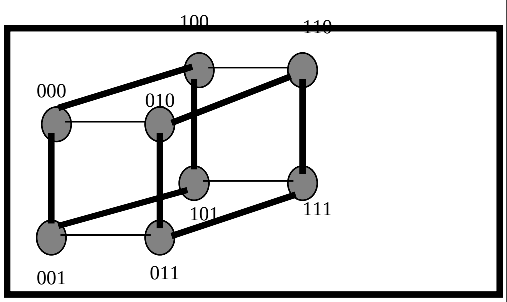


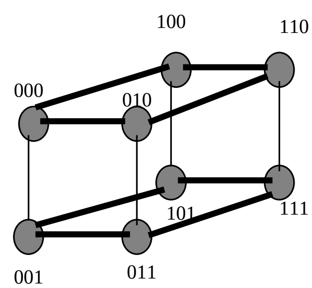
Properties of Hypercubes

- In a d-dimensional hypercube, each processor is directly connected to d other processors.
- A d-dimensional hypercube can be partitioned into 2 (d-1)-dimensional subcubes as follows:
  – Select a bit position and group together all the processors whose labels have 0 at the selected position;all of these processors make up one partition, and the remaining partition comprise the second partition. Since processor labels have d bits, d such partitions exist.

Properties of Hypercubes



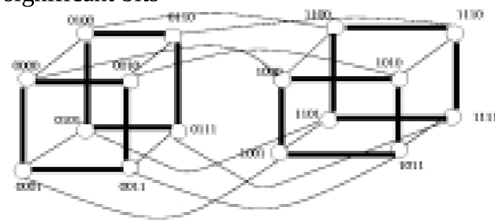




Properties of Hypercubes

- The processor labels in a d-dimensional hypercube contain d bits. Fixing any k of these bits, the processors that differ at the remaining d-k bit positions from a (d-k)-dimensional subcube composed of processors.

Subcubes formed by fixing the two most significant bits



K = 2, d = 4, , 4 subcubes of 4 processors each.

Properties of Hypercubes

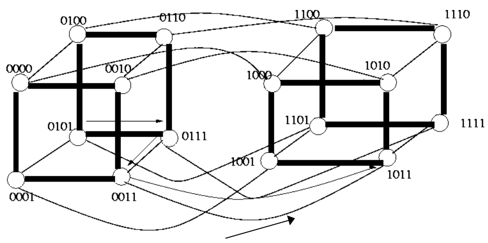
- Consider the labels s and t of 2 processors. The total number of bit positions at which these two labels differ is called the **Hamming distance** between them. Eg. Hamming distance between 011 and 101 is 2. Hamming distance between 101 and 010 is 3.
- The hamming distance between s and t is the number of bits that are 1 in the binary representation of , is the bitwise exclusive operator.

Properties of Hypercubes

- The number of communication links in the shortest path between 2 processors is the Hamming distance between their labels. A

$0101 \oplus 1011 = 1110$

nCUBE 2, Cosmic Cube

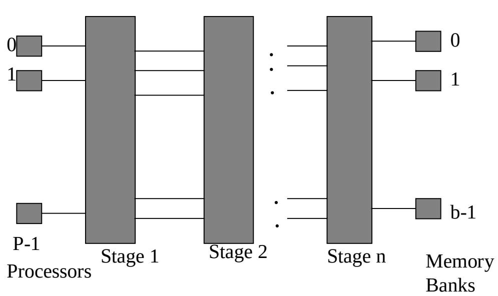


0101 – 0111 – 0011 -- 1011

Hypercube

- Diameter: logn (low diameter)
- Bisection width: n/2

Dynamic Interconnection Networks (Multistage Interconnection Networks)



Omega Network

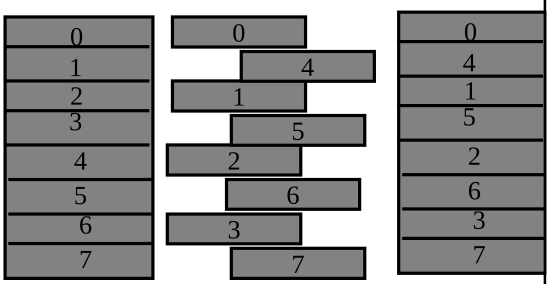
- Log(P) stages (P=B)
- Each stage consists of an interconnection pattern that connects P inputs and P outputs
- A link exists between input I and output j if:

$$j = \begin{cases} 2i, & 0 <= i <= P/2-1 \\ 2i+1-P, & P/2 <= i <= P-1 \end{cases}$$

Left-rotation operation on the binary rep. Of i to j.

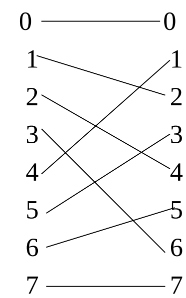
Interconnection pattern is Perfect Shuffle network

Perfect Shuffle



Perfect Shuffle

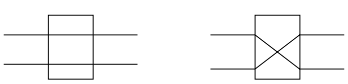
- 000
- 001
- 010
- 011
- 100
- 101
- 110
- 111



Shift every bit left one position but the leftmost bit
Wraps around the rightmost position.

Omega Network

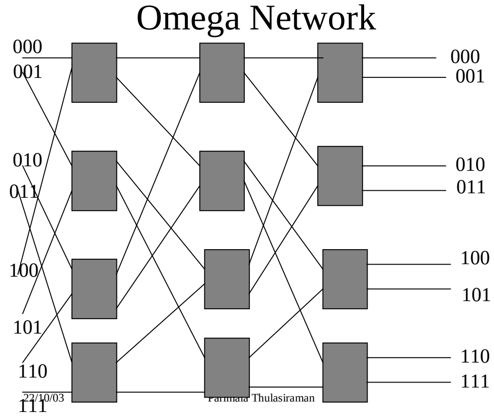
- Each stage, a perfect shuffle feeds into a set of P/2 switching elements
- Two switching configurations:



Pass through Connection Cross over connection

An omega network has P/2 x log(P) switching elements,
And the cost of such a network grows as Θ (P log P).
Less than complete cross bar switch $\Theta(p^2)$

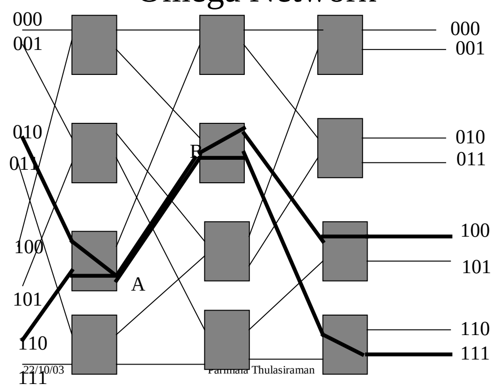
Omega Network



Omega Network (Routing Messages)

- Let s and t be the binary representations of the source and destination of the message
- The message traverses to the first switching element. If the MSBs of s and t are the same, the message is routed in pass-through mode by the switch. If different the message is routed through in crossover mode.
- This scheme is repeated at the next switching stage using the next MSB.
- Traversing log(P) stages uses all log(P) bits in the binary representations of s and t.

Omega Network

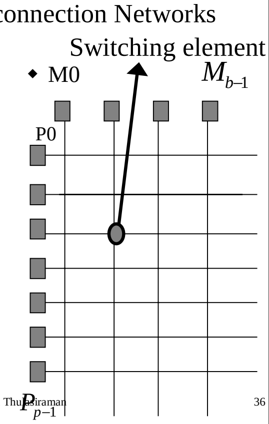


Omega Network

- AB communication link is used by both communication paths.
- Access to a memory bank by a processor may disallow access to another memory bank by another processor – Blocking Networks
- BBN Butterfly, NYU Ultracomputer

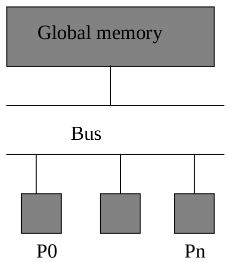
Dynamic Interconnection Networks

- Crossbar Switching Networks
  - P processors connected to b memory banks
  - A *non-blocking network*: connection of a processor to a bank does not block connection of any other processor to any other memory bank.
  - b >= P, P has atleast one memory bank to access.
  - # of switching elements required: $\Theta(Pb)$
  - As P is increased, complexity of switching network, $\Omega(P^2)$
  - Not very scalable in cost
  - Cray Y-MP, Fujitsu VP 500



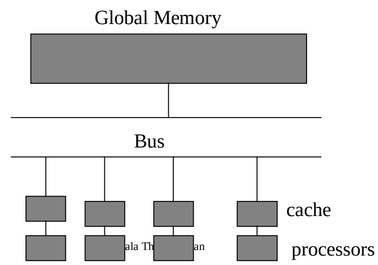
Dynamic Interconnection Networks (Bus based Networks)

- Simple to construct
- Processors connected to memory by a bus
- Processor generates request over bus; data is fetched from memory over the bus
- Uniform access to shared memory
- Bus carries only limited amount of data
- Increase processors, large amount of time waiting for memory access when bus is in use-saturation level reaches faster



Dynamic Interconnection Networks (Bus based Networks)

- Alleviate bus bottleneck—provide local cache

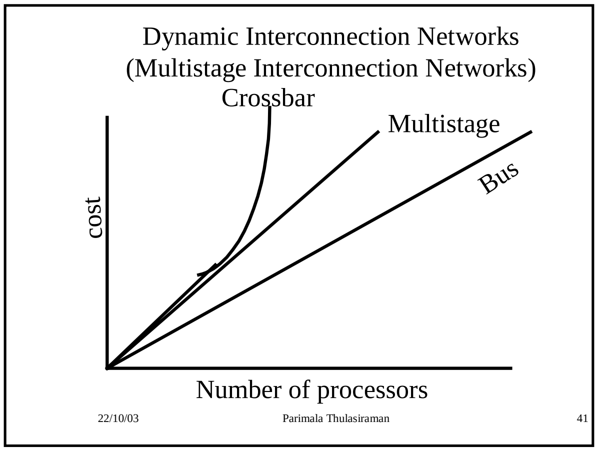


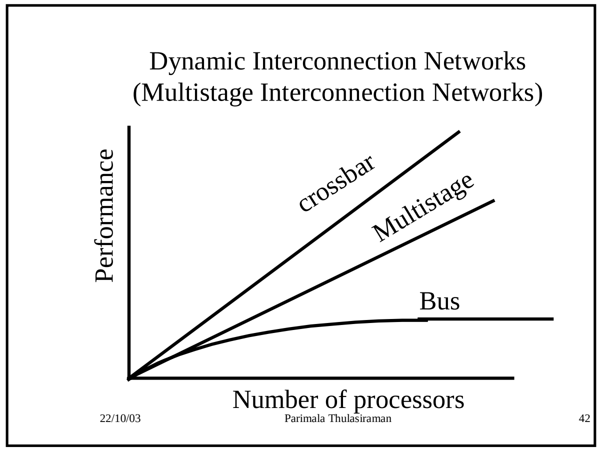
Dynamic Interconnection Networks (Bus based Networks)

- when a reference is made to a memory location, subsequent memory references are likely to be made to memory locations in the neighborhood of this location.
- Due to this locality of reference of data and instructions, once a block of data is fetched into a processor's cache memory, subsequent references will be likely be to memory words in the cache.
- In the case of a cache miss, i.e., when the word accessed is not in the cache, a block of data containing the required word is brought from the global memory across the shared bus into the local cache.
- Cache coherence problem
- Symmetry and Multimax

Dynamic Interconnection Networks (Multistage Networks)

- Crossbar scalable in terms of performance, unscalable in terms of cost.
- Bus network is scalable in terms of cost, unscalable in terms of performance.
- Multistage -- In-between bus and crossbar more scalable than the bus in terms of performance and more scalable than the crossbar in terms of cost.


Dynamic Interconnection Networks
(Multistage Interconnection Networks)
Crossbar
Multistage
Bus
cost
Number of processors
22/10/03
Parimala Thulasiraman
41


Dynamic Interconnection Networks
(Multistage Interconnection Networks)
crossbar
Multistage
Bus
Performance
Number of processors
22/10/03
Parimala Thulasiraman
42