Distributed-memory interconnects

- Distributed-memory interconnects are often divided into two groups:
    1. direct interconnects
    2. indirect interconnects.
- In a direct interconnect each switch is directlyconnected to a processor-memory pair, and the switches are connected to each other.
- Figure below shows a ring and a two-dimensional toroidalmesh. As before, the circlesare switches, the squares are processors, and the lines are bidirectional links.

Ring:

- Ringis superior to a simple bus since it allows multiple simultaneous communications.However, it's easy to devise communication schemes in which some of the processorsmust wait for other processors to complete their communications.
- ach node is connected to form a single closed data path
- Data from one node is passed along to the next node If that node is not the intended destination, then the data is transmitted to the next node until the destination is reached.
- A token (a special bit pattern) is circulated in the network to enable a node to capture the data.
- Used in: - Intel Larrabee/Core i7 - IBM Cell

Advantages of Ring Topology

- The ability to achieve transmission rates on the order of 10 million bits per second
- Provision of local communication via a single channel
- Cheap O(N) cost
- No central server (which reduces the cost)

Disadvantages of Ring Topology

- High latency O(N)
- Not easy to scale
- Failure of one node results in failure of the entire network
- Detection of a fault is very difficult
- Isolation of a fault is not easy

ToroidalMesh :

- Toroidal meshwill be more expensive than the ring, because the switches are more complex—theymust support five links instead of three—and if there are p processors, the number oflinks is 3p in a toroidal mesh, while it's only 2p in a ring. However, it's not difficult toconvince yourself that the number of possible simultaneous communications patternsis greater with a mesh than with a ring.

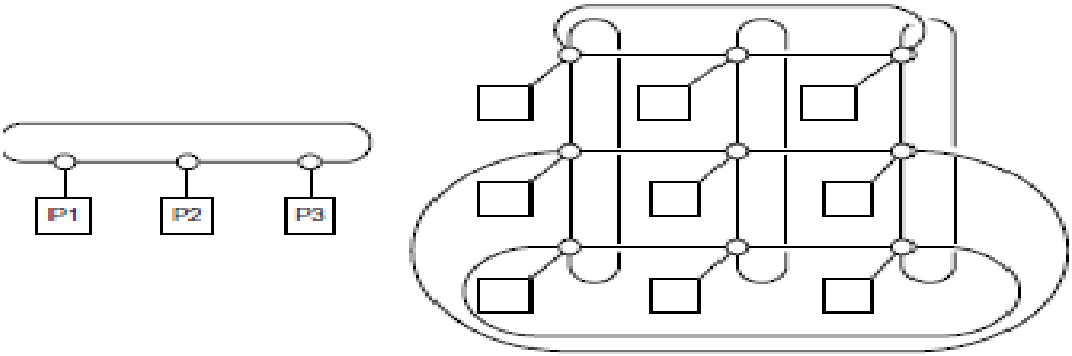

Fig: a. Ring   b. Toroidal Mesh

- One measure of “number of simultaneous communications” or “connectivity” is bisection width. To understand this measure, imagine that the parallel system is divided into two halves, and each half contains half of the processors or nodes. we’ve divided a ring with eight nodes into two groups offour nodes, and we can see that only two communications can take place betweenthe halves. Fig a below (To make the diagrams easier to read, we’ve grouped each node withits switch in this and subsequent diagrams of direct interconnects.) However, inFigure (b) we’ve divided the nodes into two parts so that four simultaneous communicationscan take placeThe bisection width issupposed to give a “worst-case” estimate, so the bisection width is two—not four.

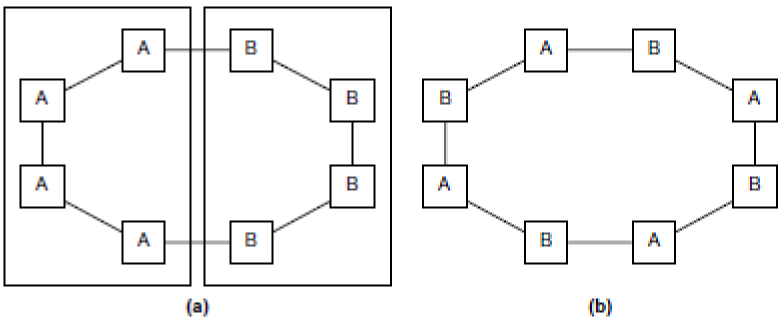

Fig: Two bisections of Ring (a) only 2 communications can takes place between the halves (b) 4 simultaneous connections can takes place.

- An alternative way of computing the bisection width is to remove the minimumnumber of links needed to split the set of nodes into two equal halves. The number oflinks removed is the bisection width. If we have a square two-dimensional toroidalmesh with $p = q^2$ nodes (where q is even), then we can split the nodes into twohalves by removing the "middle" horizontal links and the "wraparound" horizontallinks. See Figure below This suggests that the bisection width is at most $2q = 2\sqrt{p}$.

- Infact, this is the smallest possible number of links and the bisection width of a squaretwo-dimensional toroidal mesh is $2\sqrt{p}$.

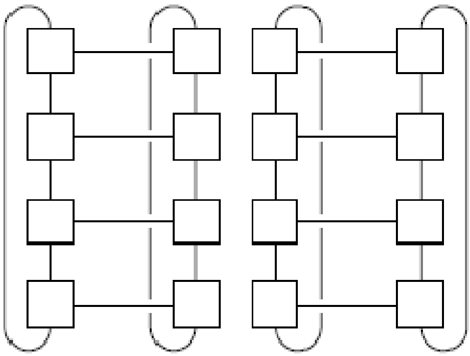

Fig: bisection of Toroidal Mesh

- The bandwidth of a link is the rate at which it can transmit data. It's usually

given in megabits or megabytes per second. Bisection bandwidth is often used as
a measure of network quality.

- The ideal direct interconnect is a fully connected network in which each switchis directly connected to every other switch. See Figure 2.11. Its bisection width is$p^2/4$.
- However, it's impractical to construct such an interconnect for systems withmore than a few nodes, since it requires a total of $p^2/2+p/2$ links, and each switch must be capable of connecting to p links.

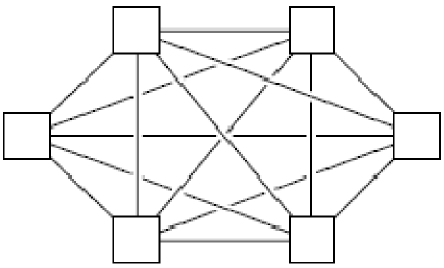

Fig :Fully connected Network

Mesh Topology:Advantages

- Each connection can carry its own data load
- It is robust
- A fault is diagnosed easily

- Provides security and privacy
- A broken node won't distract the transmission of data in a mesh network.
- Each node is connected to several other nodes which make it easier to relay data.
- Additional devices in a mesh topology will not affect its network connection
- Easy to layout on-chip: regular & equal-length links
- Path diversity: many ways to get from one node to another

Mesh Topology:Disadvantages

- There are high chances of redundancy in many of the network connections.
- Overall cost of this network is way too high as compared to other network topologies.
- Set-up and maintenance of this topology is very difficult.
- Administration of the network is tough
- O(N) cost
- Average latency: O(sqrt(N))

Hypercube:

- The hypercube is a highly connected direct interconnect that has been used in actual systems. Hypercubes are built inductively: A one-dimensional hypercube is a fully-connected system with two processors. A two-dimensional hypercube is b uiltfrom two one-dimensional hypercubes by joining "corresponding" switches. Similarly, a three-dimensional hypercube is built from two two-dimensional hypercubes

- Thus, a hypercube of dimension d has $p = 2^d$ nodes, and a switchin a d-dimensional hypercube is directly connected to a processor and d switches.
  The bisection width of a hypercube is p/2, so it has more connectivity than a ringor toroidal mesh, but the switches must be more powerful, since they must support $1+d=1+\log_2(p)$ wires, while the mesh switches only require five wires. So ahypercube with p nodes is more expensive to construct than a toroidal mesh
- Low latency O(logN)
- Hard to lay out in 2D/3D
- Used in some early message passing machines
- e.g.: - Intel iPSC, nCube

cube Interconnection:-

The hypercube (or) binary n-cube multiprocessor structure is loosely coupled system composed of $N = 2^n$ processors interconnected in an n-dimensional binary cube. Each processor forms a node of the cube. Although it is customary to refer to each node as having a processor, in effect it contains not only a CPU but also local memory and I/O interface. Each processor has direct commn paths to n other neighbor processors. These paths correspond to the edges of the cube. There are 2^n distinct n-bit binary address that can be assigned to the processors. Each processor address differ from that of its n neighbors by exactly one bit posn.

The following fig shows the hypercube structure for n = 1, 2, and 3. A one-cube structure has n = 1 and $2^n = 2$. It contains two processors interconnected by a single path. A two-cube structure has n = 2, and $2^n = 4$. It contains four nodes interconnected as a square. A three-cube structure has eight nodes interconnected as a cube. Each node is assigned a binary address in such a way that the address of two neighbors differ in exactly one bit posn.

Ex:- the three neighbours of the node with address 100 in a three-cube structure are 000, 110 and 101. Each of these binary no differ from address 100 by one bit value.

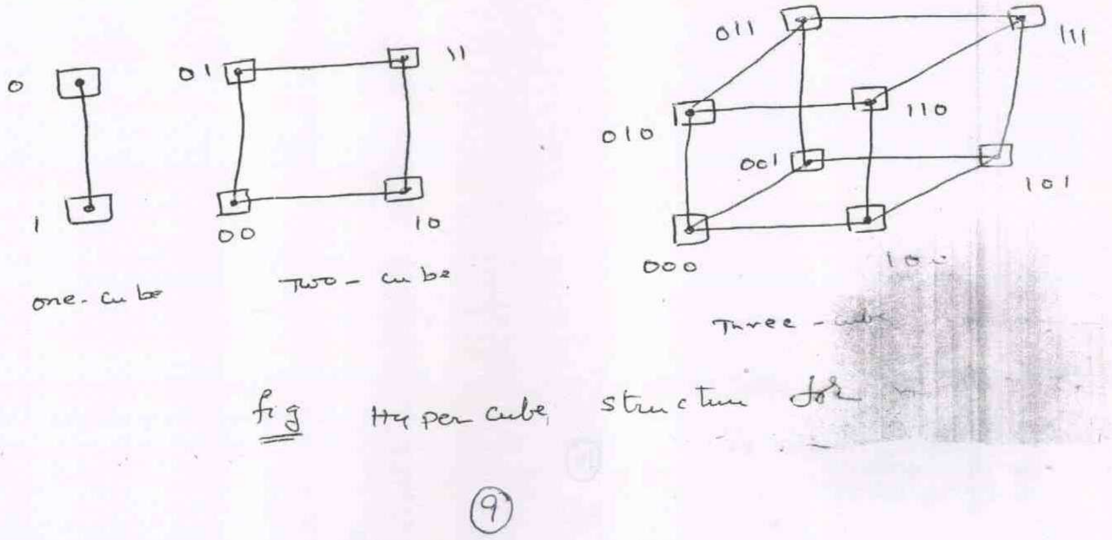

fig Hyper cube structure for

Routing messages through ... n-cube struct

from one to n links from a source node to a destination ..d

Ex:- In a three-cube structure, node 000 can communicate directly with node 001. It must cross atleast two links of communicate with 011 (from 000 to 001 to 011 or 000 to 010 to 011). It is necessary to go thru atleast three link to communicate from node 000 to node 111. A routing procedure can be computed by X-OR of source node address with destⁿ node address. The resulting binary val will have 1 bits correspo to the axes on which two nodes differ. The message is sen along any one of the axes.

Ex:- In a three cube structure, a message at 010 going to 001 produces an X-OR of the two address equal to 011 The message can be sent along the second axis to 000 and then thru the third axis to 001.

Ex:- A representative of hypercube architecture is the intel iPSC computer complex. It consists of 128 (n=7) µComp connected thru commⁿ channels.