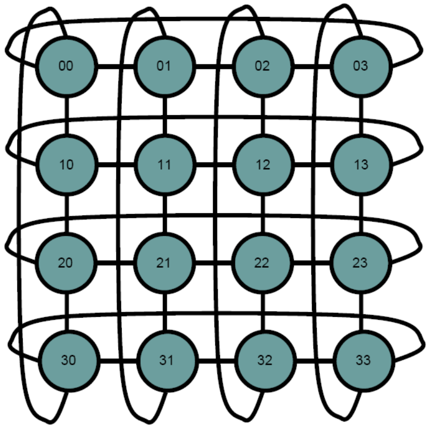
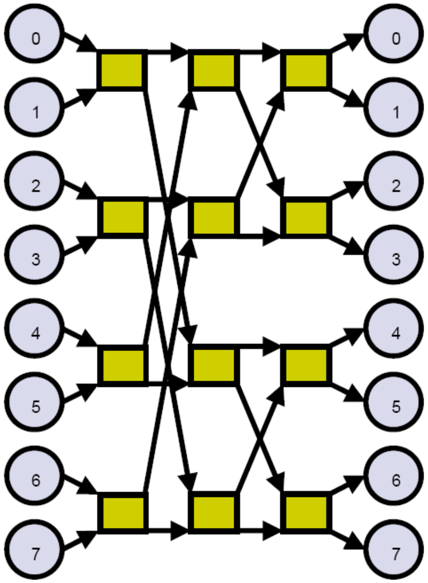
Direct and Indirect Networks

- **Direct Network**
  - ➔ Every Node in the network is both a terminal and a switch



Direct Network

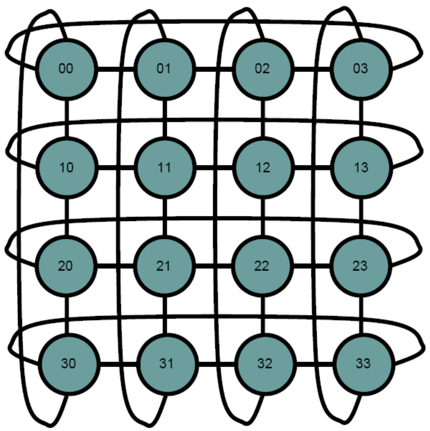
- **Indirect Network**
  - ➔ Nodes are either switches or terminal



Indirect Network

Direct Networks

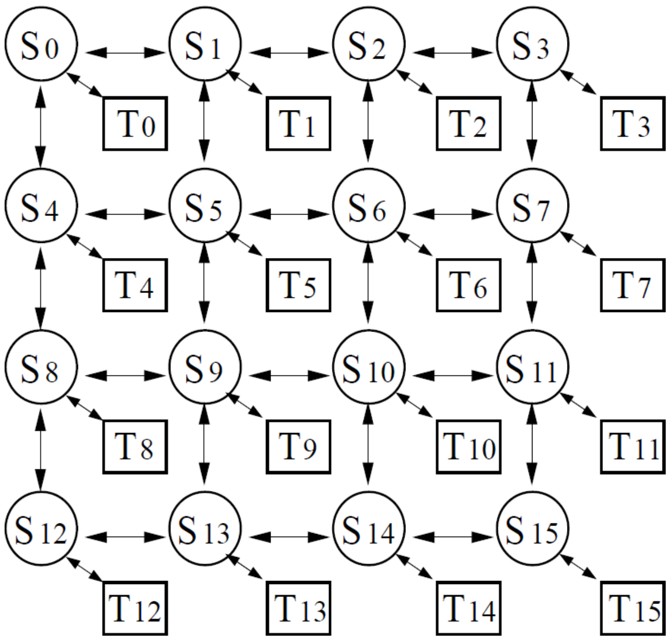
- aka point-to-point network
- Consists of a set of nodes, each one being directly connected to a (usually small) subset of other nodes in the network
  - ➔These nodes may have different functional capabilities
    - ✓ *E.g.*, vector processors, graphics processors, I/O processors, *etc.*



Direct Network

Direct Networks - Router

- A common component of the node is *the router*
  - ➔ It handles message communication among nodes
    - ✓ For this reason, direct networks are also known as router-based networks
  - ➔ Each router has direct connections to the router of its neighbors



Direct Networks - Links

- Two neighboring nodes are connected by a pair of unidirectional channels in opposite directions
- A bidirectional channel may also be used to connect two neighboring nodes

Direct Networks - Scalability

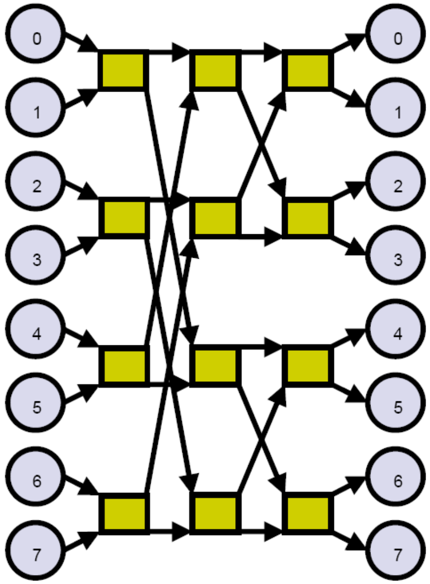
- As the number of nodes in the system increases, the total communication bandwidth also increase
  - ➔Thus, direct networks have been a popular interconnection architecture for constructing large-scale parallel computers

Direct Networks - Topologies

- Many network topologies have been proposed in terms of their graph-theoretical properties
  - ➔ Very few of them have ever been implemented
  - ➔ Most of the implemented networks have an *orthogonal topology*

Indirect Networks

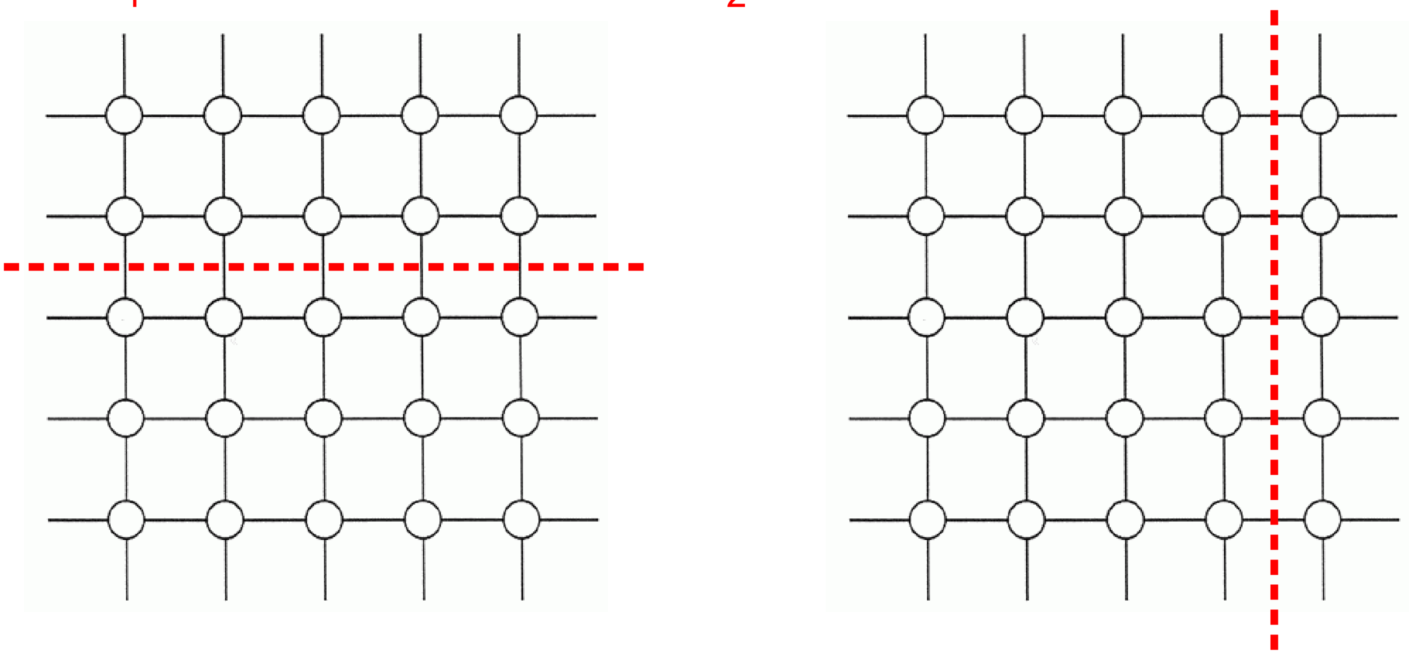
- The communication between any two nodes is carried through some switches
- Each node has a network adapter that connects to a network switch
- The interconnection of those switches defines various network topologies



Indirect Network

Cuts

- A ***cut*** of a network, $C(N_1,N_2)$, is a set of channels that partitions the set of all nodes into two disjoint sets, N_1 and N_2
  - ➔ Each element in $C(N_1,N_2)$ is a channel with a source in N_1 and destination in N_2 or vice versa



Bandwidth of the Cut

- Total ***bandwidth of the cut*** $C(N_1,N_2)$

$$B(N_1,N_2)=\sum_{c\in C(N_1,N_2)} b_c$$

Bisection

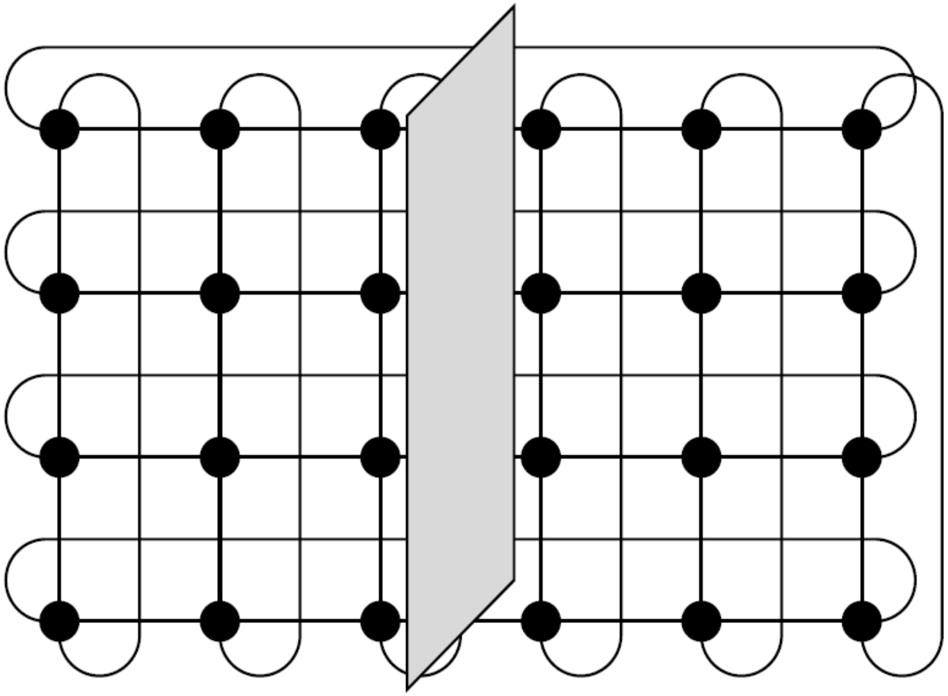
- The ***bisection*** is a cut that partitions the entire network nearly in half
- The ***channel bisection*** of a network, B_C, is the minimum channel count over all bisections

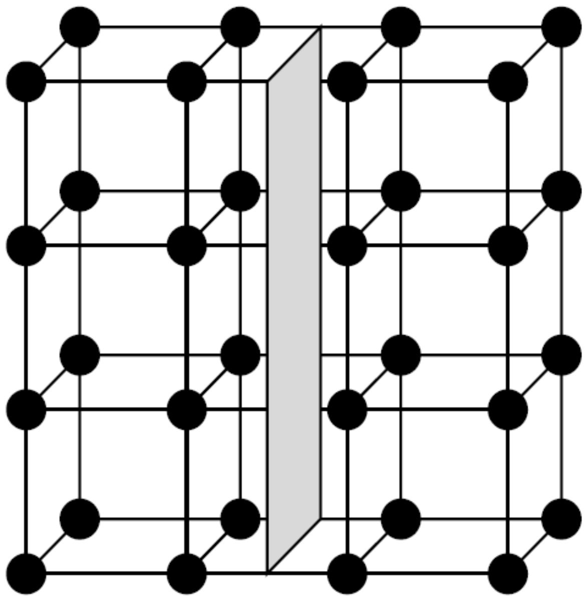
$$B_C = \min_{\text{bisections}} \left| C(N_1, N_2) \right|$$

- The ***bisection bandwidth*** of a network, B_B, is the minimum bandwidth over all bisections

$$B_B = \min_{\text{bisections}} \left| B(N_1, N_2) \right|$$

Bisection Examples





Diameter

- The ***diameter*** of a network, H_{max}, is the largest, minimal hop count over all pairs of terminal nodes

$$H_{max} = \max_{x,y \in N} |H(x, y)|$$

For a fully connected network with N terminals built from switches with out degree δ_O, H_{max} is bounded by

$$H_{max} \geq \log_{\delta_O} N \qquad (1)$$

Each terminal can reach at most δ_O other terminals after one hop

At most δ_O^2 after two hops, and at most δ_O^H after H hops

If we set $\delta_O^H = N$ and solve for H, we get (1)

Average Minimum Hop count

- The ***average minimum hop count*** of a network, H_{min}, is defined as the average hop count over all sources and destinations

$$H_{\min} = \frac{1}{N^2} \sum_{x,y \in N} H(x,y)$$