Advanced Computer Architecture (0630561)

Lecture 16
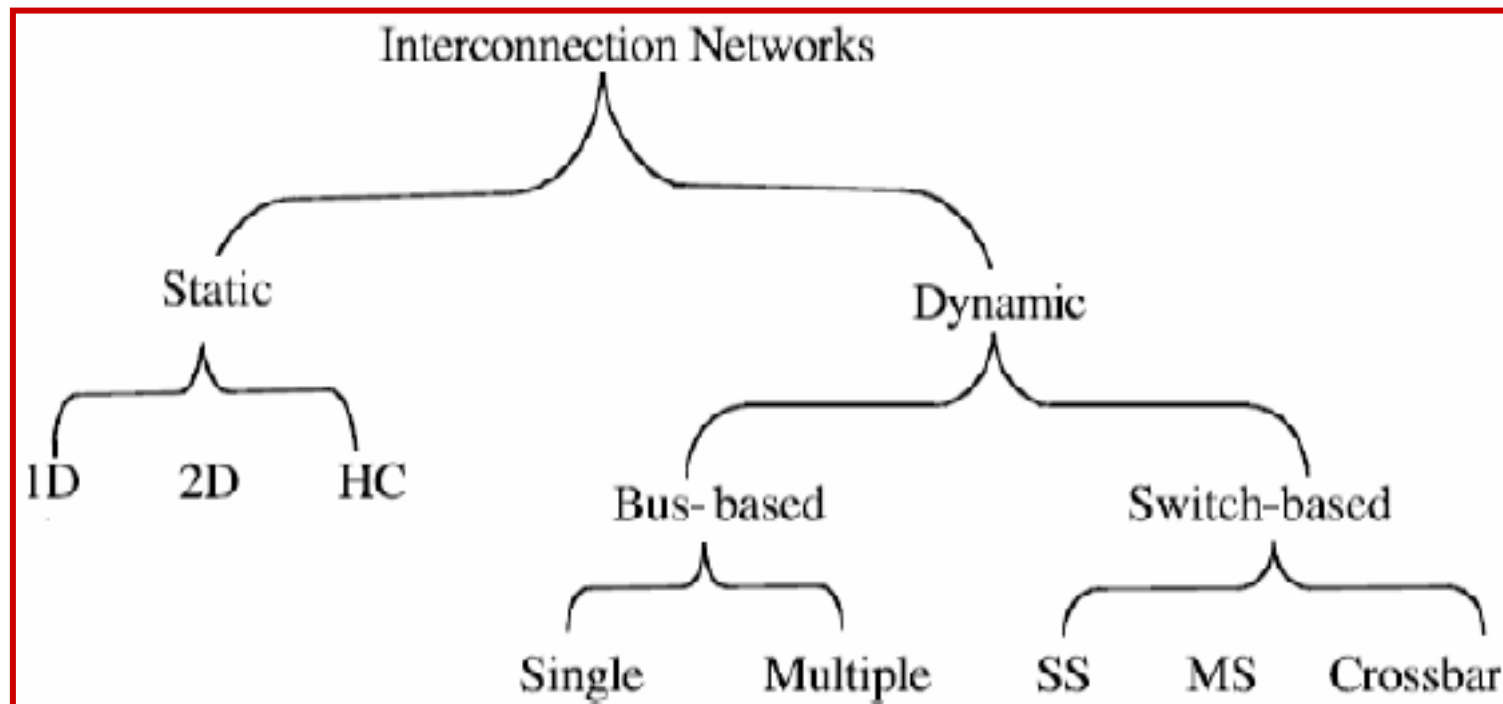
Switch-Based Interconnection Networks

Prof. Kasim M. Al-Aubidy

Computer Eng. Dept.

INs Taxonomy:

- An IN could be either static or dynamic. Connections in a static network are fixed links, while connections in a dynamic network are established on the fly as needed.
- Static networks can be further classified according to their interconnection pattern as one-dimension (1D), two-dimension (2D), or hypercube (HC).
- Dynamic networks can be classified based on interconnection scheme as bus-based versus switch-based.

Interconnection Networks

Static

1D   2D   HC

Dynamic

Bus-based

Single   Multiple

Switch-based

SS   MS   Crossbar

SWITCH-BASED INTERCONNECTION NETWORKS:

Three basic interconnection topologies exist: crossbar, single-stage, and multistage.

1. Crossbar Networks:

- The single bus can provide only a single connection, while, the crossbar can provide simultaneous connections among all its i/ps and all its o/ps.
- The crossbar contains a switching element (SE) at the intersection of any two lines extended horizontally or vertically inside the switch.
- For example, consider, the 8*8 crossbar network:
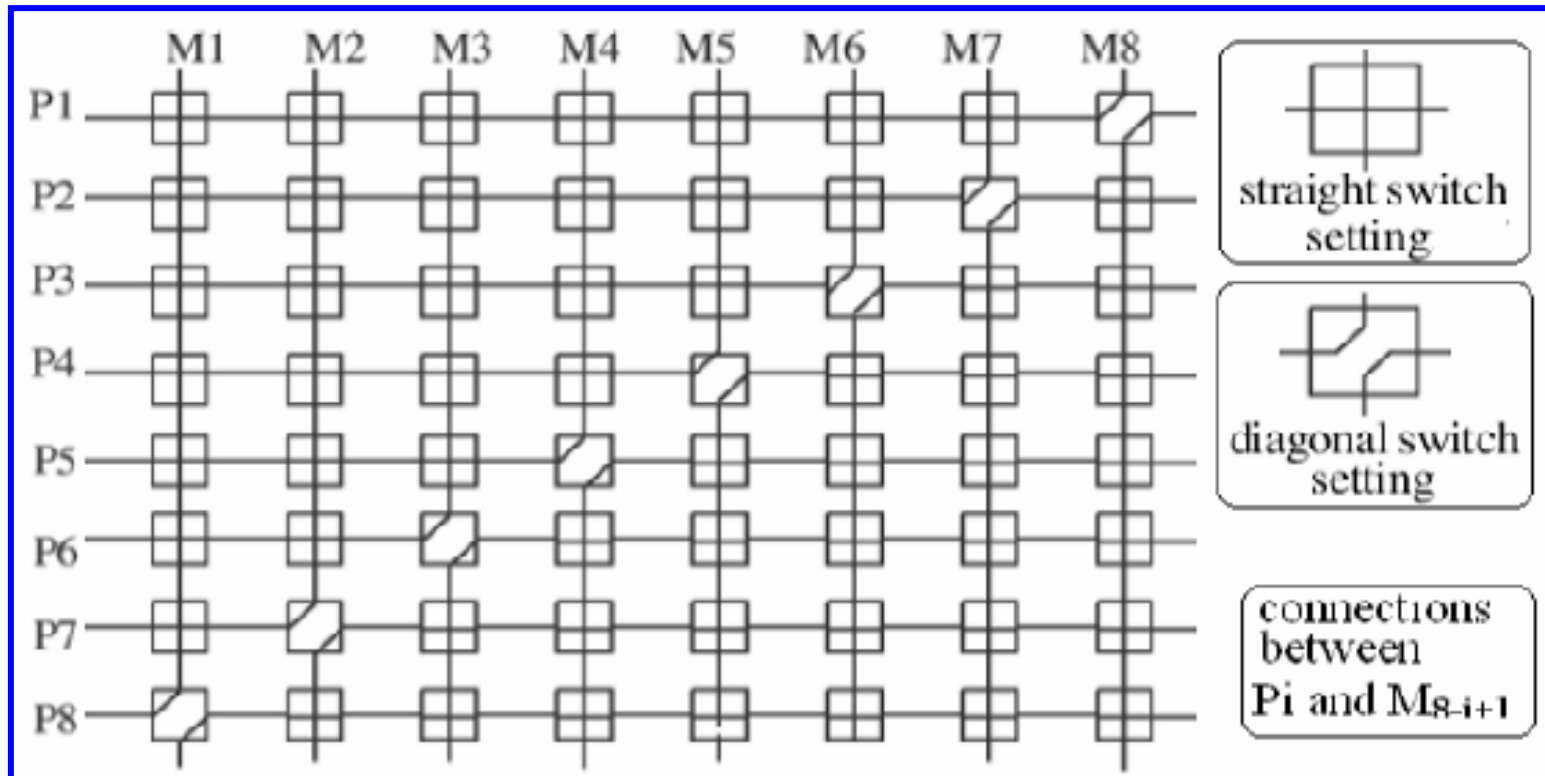
As can be seen from the figure:

- The number of SEs required is 64.

- The message delay to traverse from the input to the output is constant, regardless of which i/o are communicating.

In general for an N*N crossbar;

- The network complexity, measured in terms of the number of switching points, is O(N$\uparrow$2)

- while the time complexity, measured in terms of the input to output delay, is O(1).

- Notice that the crossbar is a nonblocking network that allows a multiple i/o connection pattern to be achieved simultaneously.

- For a large multiprocessor system the complexity of the crossbar can become a dominant financial factor.

Example: The simplest circuit for connecting n CPUs to k memories is the crossbar switch. This example shows eight CPUs connected to eight memory modules.
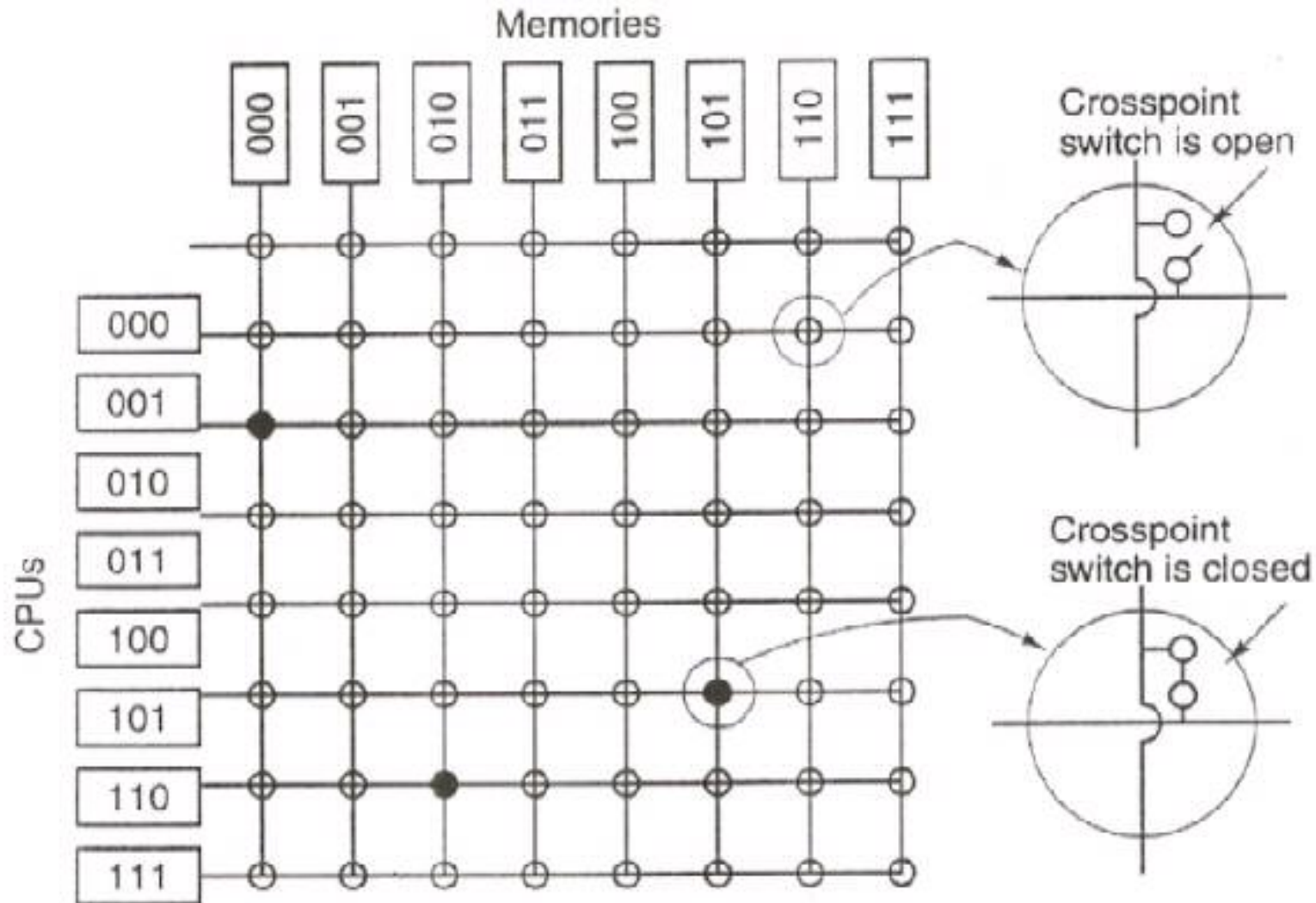
2. Single-Stage Networks:

- A single stage of SEs exists between the i/ps and the o/ps of the network.
- The simplest SE that can be used is the 2*2 SE (which is 2*2 switch).
- This switch has 2 i/ps and 2 o/ps. Messages arriving on either i/p line can be switched to either o/p line.
- Message format contains four parts:

A 2 × 2 switch.

A message format.
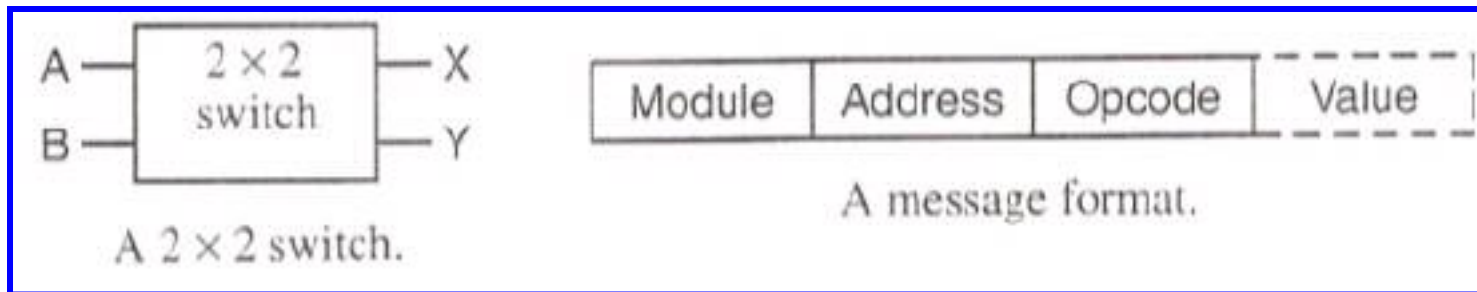
- – **Module:** tells which memory to use.
- – **Address:** specifies an address within a module.
- – **Opcode:** gives the operation, such as READ or WRITE.
- – **Value:** may contain an operand, such as 32-bit word to be written on a WRITE.

- The switch inspects the Module field and uses it to determine if the message should be sent on X or Y.

Single-Stage Networks:

- There are four possible settings;

Straight — Exchange — Upper-broadcast — Lower-broadcast
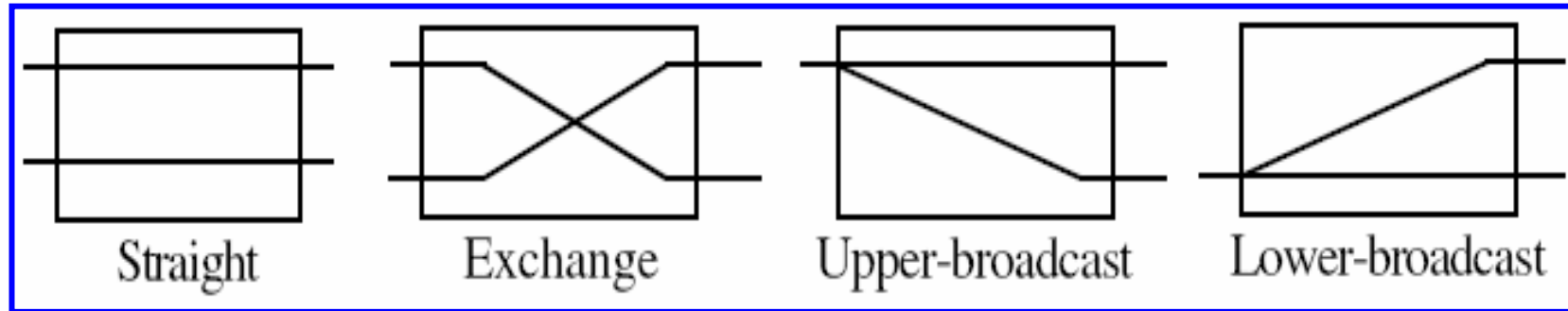
- In the **straight setting**, the upper i/p is transferred to the upper o/p and the lower i/p is transferred to the lower o/p.
- In the **exchange setting**, the upper i/p is transferred to the lower o/p and the lower i/p is transferred to the upper o/p.
- In the **upper-broadcast setting** the upper i/p is broadcast to both the upper and the lower outputs.
- In the **lower-broadcast setting** the lower i/p is broadcast to both the upper and the lower outputs.
- If the number of inputs (processors) is N and the number of outputs (memories) is N, then the number of SEs in a stage is N/2.
- The maximum length of a path from an i/p to an o/p in the network, measured by the number of SEs along the path, is $\log_2 N$.

Multistage Networks:

- These networks were introduced to improve some of the limitations of the single bus system while keeping the cost within an affordable limit.

- The most undesirable single bus limitation that multistage interconnection networks (MINs) is set to improve is the availability of only one single path between the processors and the memory modules. Such MINs provide a number of simultaneous paths between the processors and the memory modules.

- As shown in this figure, a general MIN consists of a number of stages each consisting of a set of 2*2 SEs. Stages are connected to each other using Inter-stage Connection (ISC) Pattern. These patterns may follow any of the routing functions.
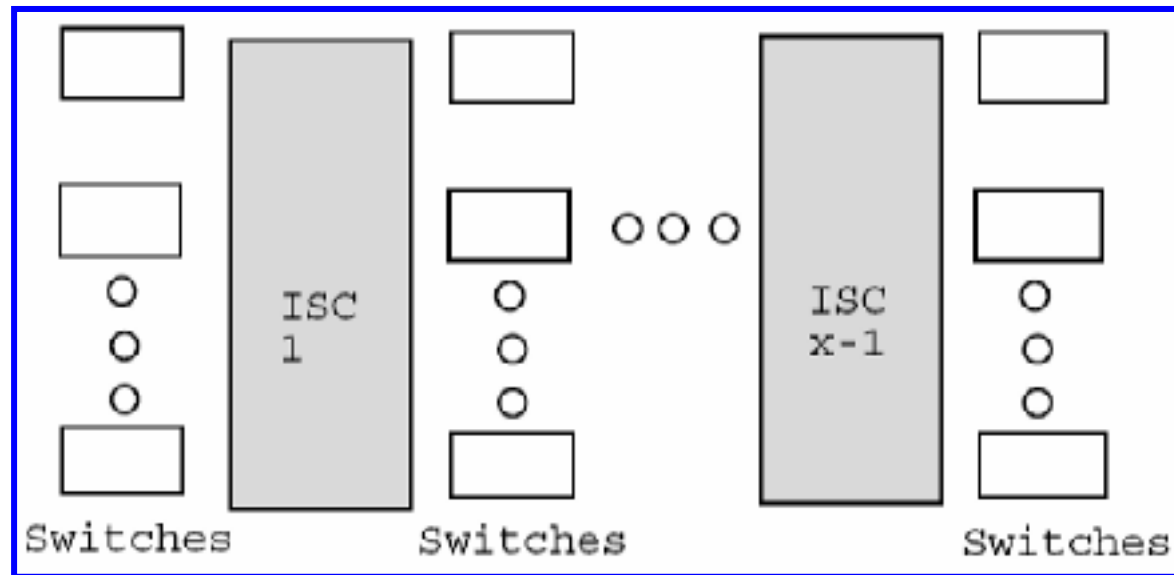
Example: An 8*8 MIN uses 2*2SEs

- This network is known as the Shuffle–Exchange network (SEN).
- The settings of the SEs illustrate how a number of paths can be established in the network.
- For example, this figure shows how three simultaneous paths connecting the three pairs of i/o: $000 \rightarrow 101$, $101 \rightarrow 011$, and $110 \rightarrow 010$ can be established.
- There exist $\log_2 N$ stages in an N*N MIN. The number of bits in any destination address in the network is $\log_2 N$.
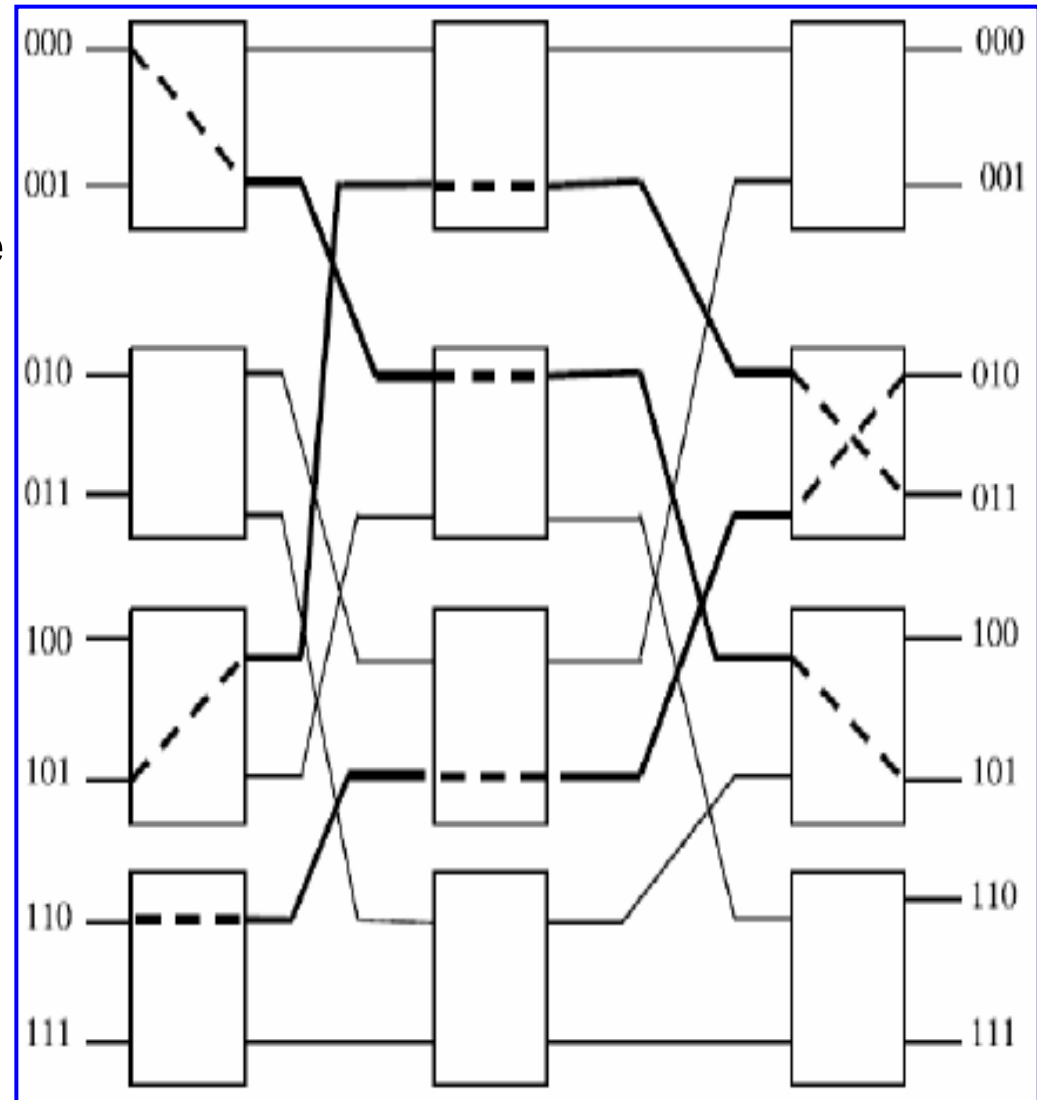
- Each bit in the destination address can be used to route the message through one stage.
- The destination address bits are scanned from left to right and the stages are traversed from left to right.
- The first MSB is used to control the routing in the first stage; the next bit is used to control the routing in the next stage, and so on.
- The convention used in routing messages is that if the bit in the destination address controlling the routing in a given stage is 0, then the message is routed to the upper output of the switch. On the other hand if the bit is 1, the message is routed to the lower output of the switch.
- Consider, for example, the routing of a message from source input 101 to destination output 011 in the 8*8 SEN;
  - Since the first bit of the destination address is 0, therefore the message is first routed to the upper output of the switch in the first (leftmost) stage.
  - Now, the next bit in the destination address is 1, thus the message is routed to the lower output of the switch in the middle stage.
  - Finally, the last bit is 1, causing the message to be routed to the lower output in the switch in the last stage.
- This sequence causes the message to arrive at the correct output.

Example: Omega Switching Network

Eight CPUs are connected to eight memory modules.

Operation: suppose CPU011 wants to read a word from memory module 110.

- The CPU sends a READ message to switch 1D containing 110 in the Module field.
- The switch takes the 1st left bit of 110 and uses it for routing. A 0 routes to the upper o/p and a 1 routes to the lower one. The message is routed via the lower o/p to 2D.
- For 2nd stage, use the 2nd bit. The message is forwarded via lower o/p to 3D.
- For 3rd stage, use the 3rd bit. The message goes out on the upper o/p and arrives at memory 110, as desired.
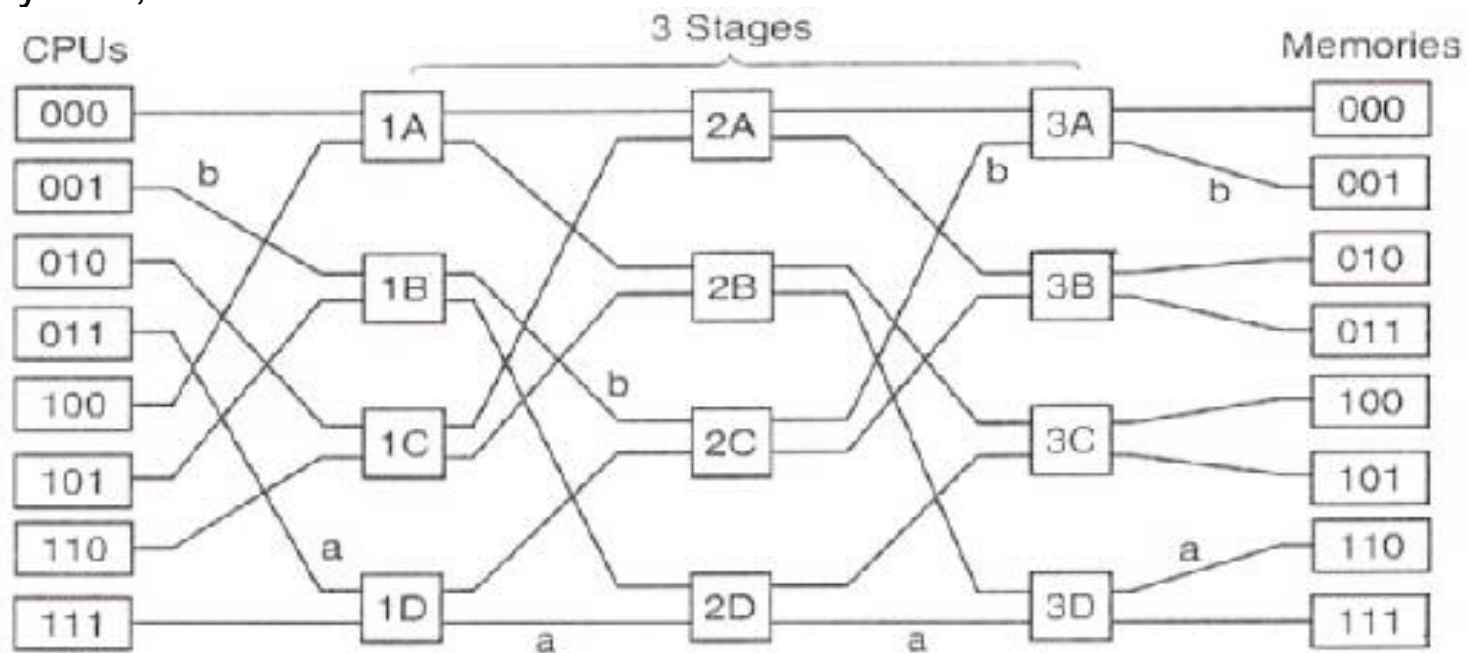
Switching Network Design:

$$\text{Number of Stages} = \log_2 n$$

$$\text{Number of Switches/Stage} = \frac{n}{2}$$

$$\text{Total number of Switches} = \frac{n}{2} \log_2 n$$

$$\text{Number of Crosspoints} = n^2$$

For the given example:
- **Number of CPUS=n= 8**
- **Number of Stages= 3**
- **Number of switches/stage= 4**
- **Total number of required switches= 12**
- **Number of Crosspoints= 16**