

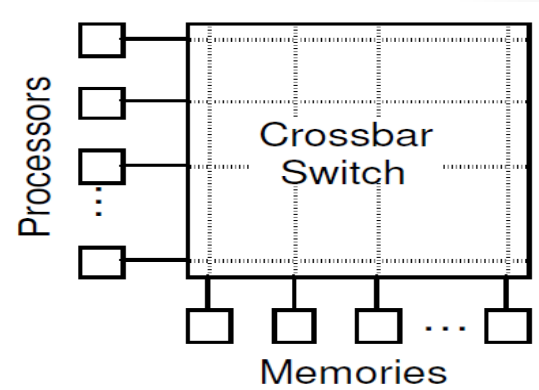
Interconnection Networks

Interconnection Networks

- When more than one processor needs to access a memory structure, interconnection networks are needed to route data
 - from processors to memories (concurrent access to a shared memory structure), or
 - from one PE (processor + memory) to another (to provide a message-passing facility).
- Inevitably, a large bandwidth is required to match the combined bandwidth of the processing elements.

Ideal Network

- For concurrent access to shared memory, the ideal structure is a **crossbar switch**, which can simultaneously connect any set of processors to any set of distinct memory modules.
- All N processors can access all M memory units with an $N \times M$ crossbar switch.
- Since there are usually about as many processors as memories, as processors are added, the complexity of a crossbar switch grows as N^2 .
- For reasonably large values of N , the crossbar switch may be more expensive than the processors and memories.



Measures of interconnection performance

- Several metrics are commonly used to describe the performance of interconnection networks:
 - **Connectivity, or degree**, the number of nodes that are immediate neighbors of a node (i.e., the number of nodes that can be reached from it in one hop).
 - **Diameter**, the maximum number of nodes through which a message must pass on its way from source to destination. Diameter measures the maximum delay in transmitting a message from one processor to another.

Measures of interconnection performance

- **Average distance**, where the distance between two nodes is defined by the number of hops in the shortest path between those nodes. Average distance is given by

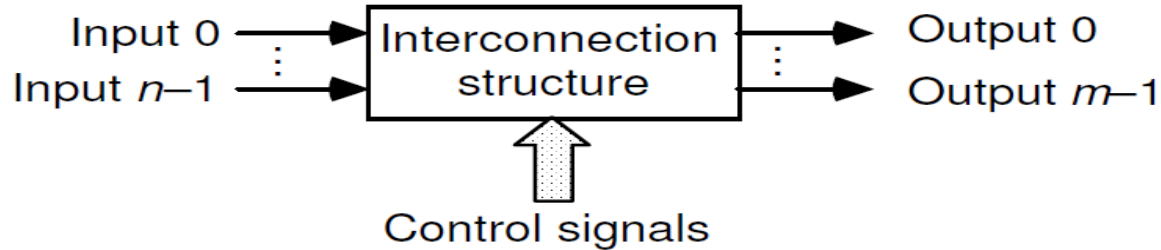
$$d_{avg} = \frac{\sum_{d=1}^r (d \cdot N_d)}{N-1}$$

where N is the number of nodes, N_d is the number of nodes at distance d apart, and r is the diameter.

- **Bisection width**, the smallest number of wires you have to cut to disconnect the network into two equal halves (± 1).

Interconnection topologies

- An idealized interconnection structure
 - takes a set of n input ports labeled $0, \dots, n-1$ and
 - sets up connections between them and a set of m output ports $0, \dots, m-1$,
 - with the connections determined by control signals.

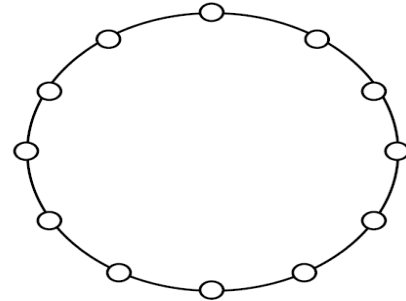


- Usually we will assume that $m = n$.

Ring

- Processor i directly connected to processors $i+1$ and $i-1$. Data can be moved from any processor to any other by a sequence of cyclic shifts.
- Motivation: Many parallel algorithms include calculations of the form

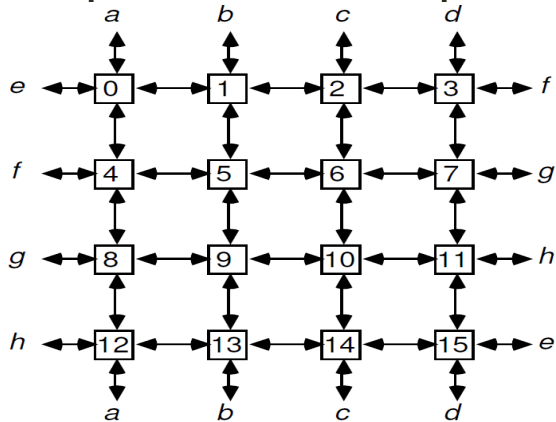
$$X[i] := \frac{X[i-1] + X[i]}{2}$$



- Usually every item of an array except the first and last is updated in this way.

Mesh interconnection network

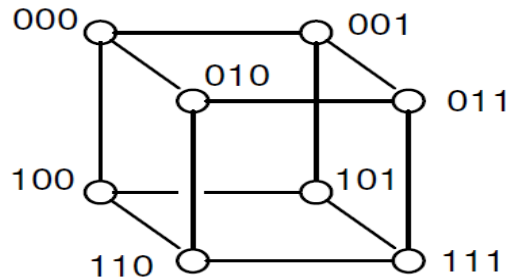
- A mesh is similar to having “row & column” cyclic shifts.
- One motivation: *Four-point iteration* is common in the solution of partial differential equations. Calculations of the f $X[i, j] := (X[i+1, j] + X[i-1, j] + X[i, j-1] + X[i, j+1]) \div 4$ are performed frequently.



Here is an example of a 16-node mesh.

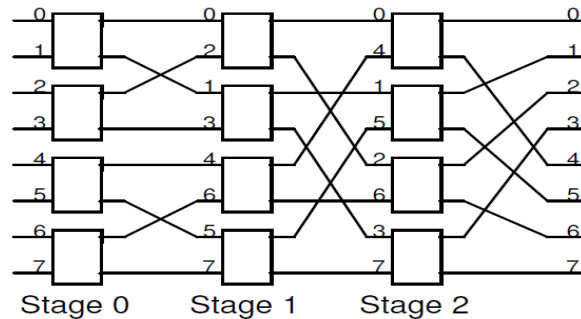
- Note that the last element in one row is connected to the first element in the next.
- If the last element in each row were connected to the first element in the same row, we would have a **torus** instead.

Hypercube



A hypercube is a generalized cube. In a hypercube, there are 2^n nodes, for some n . Each node is connected to all other nodes whose numbers differ from it in only one bit position.

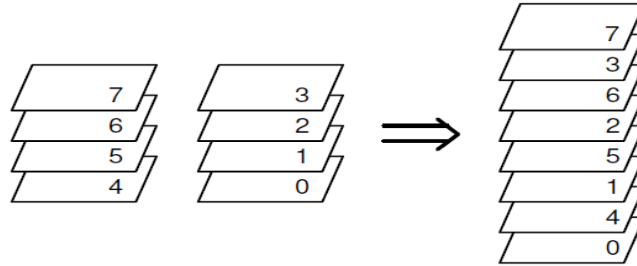
For a multistage cube network, we can diagram the paths from one cell to another like this:



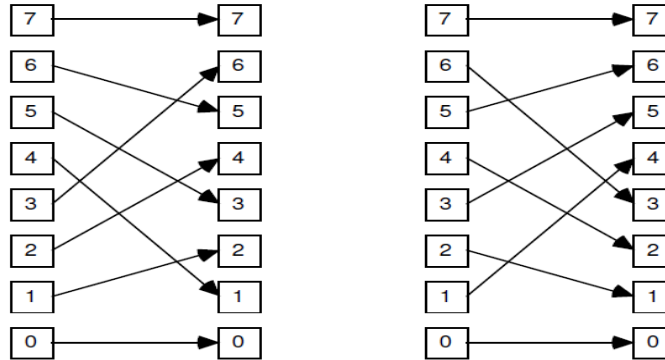
A multistage cube network is often called an *indirect binary n-cube*.

Perfect-shuffle interconnection

It describes what happens when we divide a card deck of, e.g., 8 cards into two halves and shuffle them "perfectly."



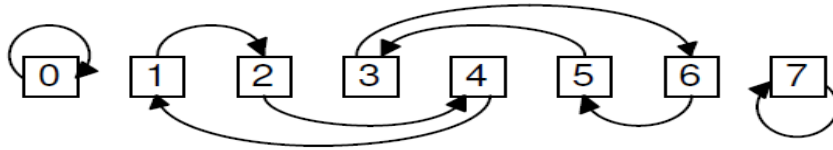
We can draw the processor interconnections required to obtain this transformation (at near right):



If the links are bidirectional, the *inverse perfect shuffle* is obtained (above, right).

Omega Network

By itself, a shuffle network is not a complete interconnection network. This can be seen by looking at what happens as data is *recirculated* through the network:



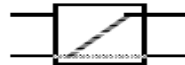
- An *exchange* permutation can be added to a shuffle network to make it into a complete interconnection structure
- A shuffle-exchange network is isomorphic to a cube network, with a suitable renumbering of boxes.
- An omega network is a $\log_2 N$ -stage shuffle-exchange interconnection, with individual cells that can perform four different operations



Pass-through



Exchange



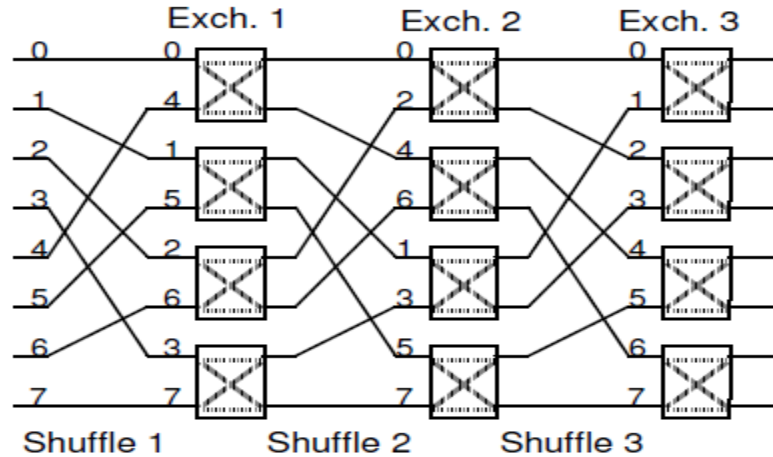
Broadcast low



Broadcast high

Omega Network

Here is a diagram of a multistage omega network for $N = 8$.



Butterfly Network

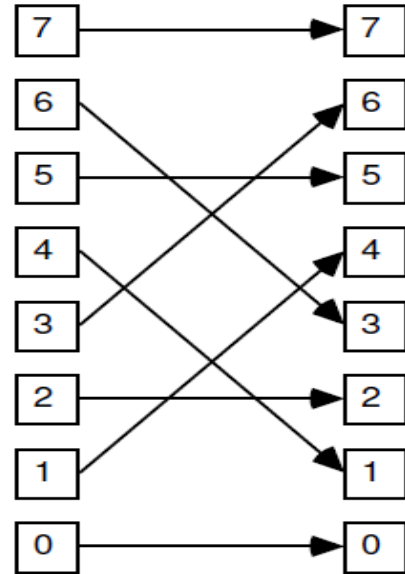
Closely related to shuffle-exchange networks.

The butterfly permutation is defined as—

$$B(a_{n-1} a_{n-2} \dots a_1 a_0) \equiv a_0 a_{n-2} \dots a_1 a_{n-1}$$

i.e., the permutation formed by interchanging the most- and least-significant bits in the binary representation of the node number.

This permutation can be diagrammed as shown at the right:

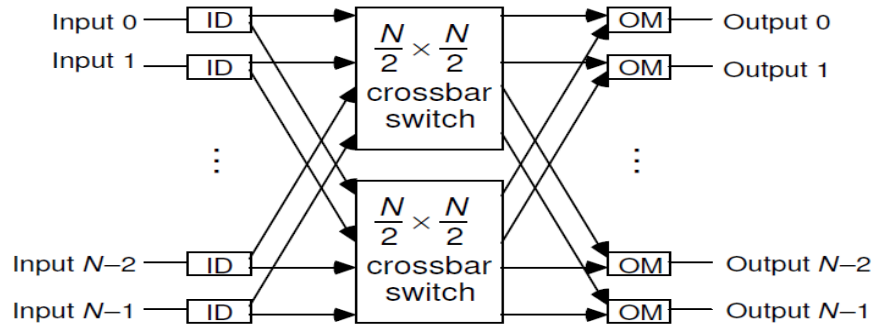


Benes network

- As we have seen, a crossbar switch is capable of connecting a set of inputs to *any* set of distinct outputs simultaneously.
- A shuffle-exchange, or multistage cube, network is not capable of doing this.
- Is it possible to achieve an *arbitrary* permutation of input-output combinations with less than a full crossbar switch?

Benes network 2

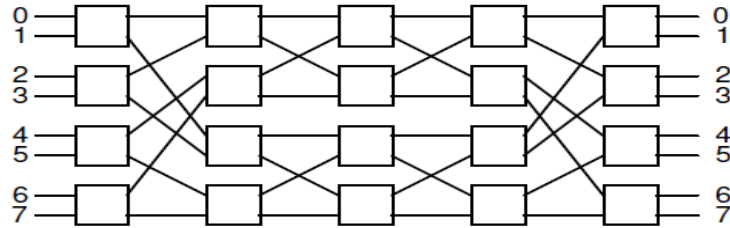
- Yes. The Benes network substitutes two $N/2 \times N/2$ crossbar switches, plus an N -input exchange switch for a full crossbar switch, as shown below



Benes network

The resulting $N/2 \times N/2$ crossbar switches can be similarly reduced.

Through this process, a full connection network can be produced from 2×2 switches at significantly lower cost than a full crossbar:



The stages of a Beneš network are connected by shuffle and inverse-shuffle permutations.

The network is called *rearrangeable*, since the switch settings can always be rearranged to accommodate any input-output mapping.

In some Beneš networks, the switches are capable of performing broadcasts, as well as pass-through or interchange.

Such Beneš networks can achieve all N^N possible input/output mappings.

Questions??