Parallel Processors

Session 5

Interconnection Networks
Network Properties

- Network topology
- Node degree
- Network diameter
- Bisection width
- Data routing functions
Network Topology

• Static
  – Point to point direct connections
  – No changes when program is executed

• Dynamic
  – Switched links
    • Busses, crossbars, multistage networks
  – Dynamically configured according to communication needs as program is executed
Node Degree

• The number of links connected to the node
  – Indicates number of I/O ports $\rightarrow$ cost
  – A small node degree helps reduce cost
  – A constant node degree helps modularity and scalability
Network Diameter

• Maximum number of links between any two nodes
  – Path length is measured by the number of links traversed (hops)
    • Smaller diameter $\rightarrow$ lower communication latency
Bisection Width

• Minimum number of links between two equal parts of a network
  – Channel bisection width \((b)\): number of links
  – Wire bisection width \((B)\): number of links \(\times\) number of wires (bits)
  – \(B=bw\)
Symmetric Networks

• The topology looks the same from any node in the network
  – Easier to implement
  – Easier to program

• Homogeneous: If all nodes are the same
Data Routing Functions

• A data-routing function is used to exchange data between processing elements
• Data-routing functions are built on top of interconnection networks with different topologies such as ring, mesh, or multistage networks:
  – Shifting
  – Rotation
  – Permutation
  – Broadcast
  – Multicast
  – Personalized communication
  – Shuffle
  – Exchange
  – …
Network Topologies
Static Connection Networks

- Linear Array
- Ring and Chordal Ring
- Barrel Shifter
- Tree and Star
- Fat Tree
- Mesh and Torus
- Systolic Arrays
- Hypercubes
Linear Array

- One-dimensional network
- N nodes connected with N-1 links
- Degree
  - Internal nodes: d=2
  - Terminal nodes: d=1
- Diameter
  - D=N-1
- Bisection width
  - b=1
- Non-symmetric
- Good for small N, not good for large N
Ring

- A linear array with the two terminal nodes connected
  - Unidirectional
  - Bidirectional
- Symmetric
- Node degree $d=2$
- Diameter
  - $D=[N/2]$ if bidirectional
  - $D=N-1$ if unidirectional
Chordal Ring

- A ring with node degree larger than 2
- Extra links are added to produce chordal rings
- Network diameter is decreased as the node degree increases
- In a completely connected network node degree is \([N-1]\) and diameter is 1
Barrel Shifter

- A ring with extra links from each node to nodes having distance equal to an integer power of 2
- Node i is connected to node j if \(|j-i|=2^r|, r=0,\ldots,n-1|
- The network size is \(N=2^n\)
- \(d=2n-1\)
- \(D=n/2\)
Binary Tree

- Each node is connected to two sub nodes
- A k-level tree have $N=2^k-1$ nodes
- Maximum node degree is 3
- And the diameter is $2(k-1)$
Star

- A two-level tree with node degree of $d=N-1$
- Diameter = 2
- Good for systems with a centralized supervisor node
Fat Tree

• A binary fat tree is a binary tree in which the channel width is increased by one from each lower level to its upper level

• Solves the bottleneck problem in ordinary tree
Mesh

- Each node is connected to all neighboring nodes in a square arrangement
- A $k$-dimensional mesh with $N=n^k$ nodes has an interior node degree of $2k$ and network diameter of $k(n-1)$
Illiac Mesh

- A variant of mesh
- Edge nodes are connected with wraparound connections
- An $nxn$ illiac mesh has a diameter of $d=n-1$
Torus

- A mesh in which each row and column is connected like a ring
- Combines ring and mesh topologies
- An nxn binary torus has a node degree of 4 and a diameter of $2[n/2]$
- Torus is symmetric
Systolic Arrays

- A class of multidimensional pipelined array architectures designed for implementing fixed algorithms
- Example: A systolic array specially designed for matrix-matrix multiplication
Hypercubes

- A binary n-cube architecture consisting of $N=2^n$ nodes spanning along n dimensions with 2 nodes per dimension
- Example: A 3-cube with 8 nodes
- $D=d=n$
- Not scalable
Cube-Connected Cycles

- A k-cube-connected cycle is constructed from a k-cube with \( n=2^k \) ring with one node per dimension at the ring.
- A k-CCC can be built from a k-cube with \( k \times 2^k \) nodes.
- \( D=2k \)
- \( d=3 \) (constant)
k-ary n-Cube Networks

- $n$ is the dimension of the cube
- $k$ is the number of nodes per dimension
- $N$ total number of nodes

$N = k^n$
Network Throughput

• Network throughput:
  – Total number of messages per unit time

• Capacity:
  – Total number of simultaneous messages

• Capacity can be used to estimate the network throughput

• A hot spot:
  – A pair of nodes with a large portion of total network traffic

• Hot spot throughput:
  – Maximum rate at which messages can be sent from one specific node to another specific node

• Low dimensional networks have better throughput characteristics and lower latency as a result
Summary

• D is not critical factor but smaller D is better for resource sharing
• Number of lines l affects the cost
• Bisection width affects network bandwidth
• d affects the complexity and the cost
• Small node degree is good for implementation and scalability
• Small diameter is good for latency
• Symmetry is good for scalability and routing efficiency
Dynamic Connection Networks

- Can implement all communication patterns
- Switches or arbiters are used instead of fixed connections
  - Bus systems
  - Multistage interconnection networks
  - Crossbar switch networks
Bus Systems

- One transaction at a time!
- Arbiter needed to handle multiple requests
- Simpler, lower cost, lower throughput
Multiple Buses

• Simple way to increase bandwidth
  – use more than one bus

• Can be static or dynamic assignment to busses
  – static
    • A->B always uses bus 0
    • C-> always uses bus 1
  – dynamic
    • arbitrate for a bus, like instruction dispatch to k identical CPU resources
Crossbar network

- Each cross point can connect the row and column
- Each cross point can be set on or off dynamically
- Only one cross point can be set on in each column
Interprocessor Crossbar Network

- Only one crosspoint switch can be set on in each row and column
Multistage Networks

• Crossbar switches:
  – Have fixed latency
  – Are nonblocking
  – Are expensive

• Multistage networks:
  – Have more latency
  – Are more scalable
  – Can be blocking

• MIN: multistage interconnection network
Multistage Networks

- General structure of a multistage network is shown in the picture
- Different MINs differ in the switch modules and Interstage connections (ISC)
Switch Modules

- axb switch has a inputs and b outputs
- Usually a=b=2^k
- Switch can connect inputs to outputs
- Not more than one connection to an output at a time
- Crossbar: only one to one connections are allowed
- Binary switch: a=b=2
- 2x2 crossbar: straight or crossover
Omega Network

- Four stages of 2x2 switches for a 16x16 Omega network
- ISC is perfect shuffle
- Log n stages of 2x2 switches are required for n inputs
- n/2 switch modules in each stage, n/2 log n in total
Blocking Situation
Baseline Network

- Baseline network has a recursive structure
- $\times N$ block at first stage, $(N/2) \times (N/2)$ at second stage and so on
- Each stage is constructed using 2x2 switches
Benes Network

- A nonblocking arrangement can always be found for any permutation of inputs and outputs
Benes Structure

• Two back-to-back baseline networks
A 64x64 Benes Network
Clos Network

Clos($N$, $n$, $k$)
$N$ - inputs/outputs;

New connections between free nodes can always be made
Dilated Connections

• Have multiple outputs per logical direction
  – **Dilation**: number of outputs per direction
  – **Example**:
    • 4 outputs
    • 2 per direction
    • Dilation 2
Routing Functions

• Data routing functions are implemented on the interconnecting network to establish connections and transfer the data
• Shifting, rotation, permutation, broadcast, multicast, personalized communication, shuffle, etc.
• The functions can be implemented on ring, mesh, hypercube, or multistage networks
Permutations

• n objects can have n! permutations
• Permutations specify possible orders of the objects
• A crossbar switch can build all permutations of source/destination pairs
• A multistage switch can implement all or a subset of the permutations
Perfect Shuffle

• Perfect shuffle maps $x$ to $y$ where:

\[
x = (x_{k-1}, \ldots, x_1, x_0)
\]

\[
y = (x_{k-2}, \ldots, x_1, x_0, x_{k-1})
\]
Hypercube Routing Functions

- Data is exchanged between adjacent nodes with one bit difference
- Three routing patterns are defined:

(a) A 3-cube with nodes denoted as $C_2C_1C_0$ in binary

(b) Routing by least significant bit, $C_0$

(c) Routing by middle bit, $C_1$
Multicomputer Interconnects

- Data is exchanged through message passing
- Routers and switches are used to route messages among multiple computers