CS6290

Fall 2009

Prof. Hyesoon Kim

Georgia Tech College of Computing

Thanks to Prof. Derek

Networks

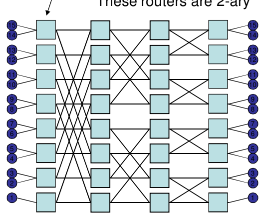
- Topology
  - Specifies way switches are wired
  - affects routing, reliability, throughput, latency, building ease
  - Important, but, with commodity switch chips, slowly becoming a non-issue
- Routing
  - How do we get from here to there?
  - Static or adaptive
- Buffering and Flow Control
  - Key resources: channels and buffers (road and parking lots)
  - What do we store within the network?
    - Entire packets, parts of packets, etc?
  - How do we throttle during oversubscription?

Terms

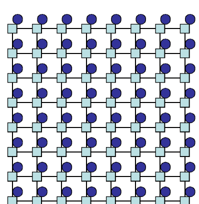
- Direct/indirect
- Direct: every node can be terminal node and switch node
- Indirect: either switch or terminal

Router (switch), Radix of 2 (2 inputs, 2 outputs)
Abbreviation: Radix-ary
These routers are 2-ary

indirect



Direct

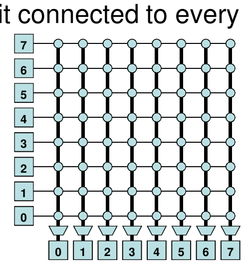


A Couple More Definitions

- Message
  - Unit of transfer for network's clients (processors, memory)
- Packet
  - Unit of transfer for network
- Flit
  - Flow control digIT
  - Unit of flow control within network
- Channel
  - A single logical connection between routers
- Node
  - A router within a network

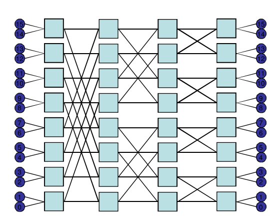
Crossbar

- Every unit connected to every other



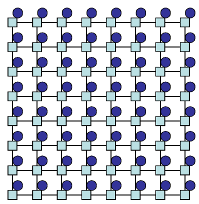
Butterfly

- Distance lg(n) to any other endpoint
- *k*-ary *n*-fly (k is radix, n is number of stages)
- 2-ary, 4-fly



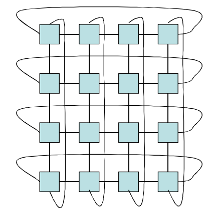
Mesh

- Easy to build
- Generally direct networks
- K-ary, n-mesh (k is side size, n is number of dimensions)
  - 8-ary 2-mesh



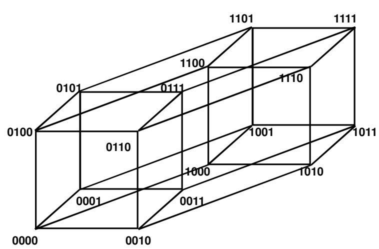
Torus

- K-ary, n-torus/cube (4-ary, 2-cube)
- Generally direct like mesh
- (3D)

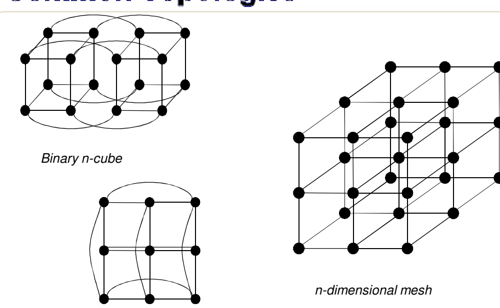


Hypercube

- High connectivity, few hops
- 4D



Common Topologies

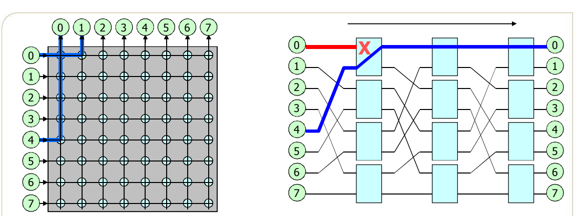


Binary n-cube

Torus (3-ary 2-cube)

n-dimensional mesh

Blocking vs. Non-blocking Networks



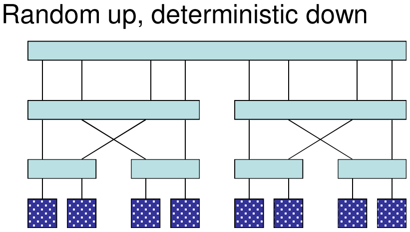
non-blocking topology

blocking topology

- Consider the permutation behavior
  - Model the input-output requests as permutations of the source addresses

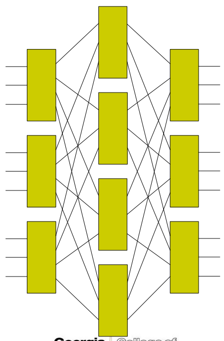
Fat Trees

- Redundancy, maintains bandwidth
- Random up, deterministic down



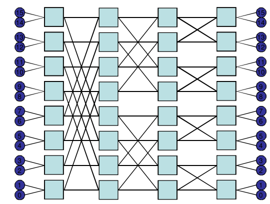
Clos Networks

- 3 stages
  - Any odd number beyond 3 can be collapsed into 3
- (m, n, r)
  - m = number of middle switches
  - n = number of input/output ports
  - r = number of input or output switches
- Each side stage completely connected to middle stage
- Benes network is Clos using radix-2 switches
- Fat tree is a Clos network as well
- and route any permutation



Speedup

- Networks are not always perfect or 100% efficient
- Often need "speedup" (overcapacity) to meet requirements



Circuit vs Packet Switching

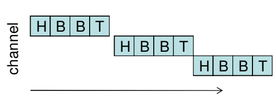
- Circuit switch sets up full path
  - no one else can use those links
  - faster and higher bandwidth
  - setting up and bringing down links slow
- Packet switching routes per packet
  - if link is free can use
  - potentially slower --- must dynamically switch
  - no setup, bring down time

Bisection Bandwidth

- Cut: $C(N_1,N_2)$: a set of channels that partitions the set of nodes N* into two disjoint sets N1, N2.
- A bisection of a network: a cut that partitions the entire network nearly half
- Channel of Bisection (Bc)
  - Minimum # of channel count over all bisections of the network
- Bisection bandwidth :
  - (Minimum # of channels spanning 2 halves) * (bandwidth of each channel)
- Often used to describe network performance
- Sometimes misleading due to switch efficiency, routing efficiency

Store and Forward Flow Control

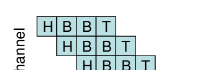
- Store and Forward
  - Message copied entirely into network node before moving to the next node
- Latency proportional to distance between networks
  - T = L/W(D+1)
    - L = length of message
    - W = width of channel
    - D = Distance



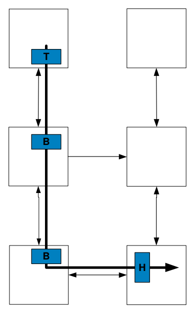
- Can we do better?

Virtual Cut Through Flow Control

- Start forwarding as soon as header is received and resources (buffer, channel, etc) allocated
  - Dramatic reduction in latency
- Still allocate buffers for full packets
- What if packets are large?
  - Internet IP packets up to 64KB



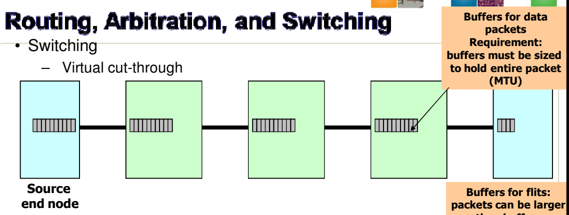
Wormhole Routing (Flow Control)



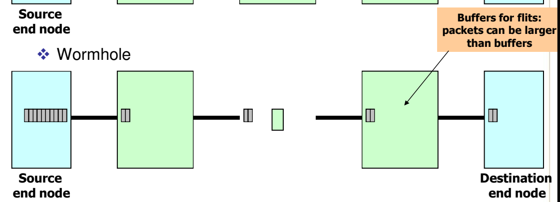
- Packets broken into (potentially) smaller flits
- Flits are sent across the fabric in a *wormhole-fashion*
  - Body follows head, tail follows body
  - Pipelined fashion
  - If head blocked, rest of packet stops
    - Ordered within fabric
  - Fabric routing information only in head

Routing, Arbitration, and Switching

- Switching
  - Virtual cut-through

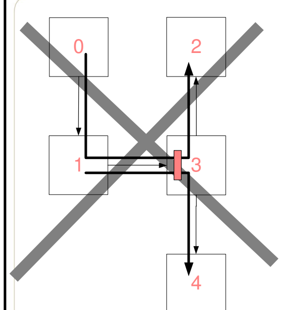


❖ Wormhole



"Virtual Cut-Through: A New Computer Communication Switching Technique," P. Kermani and L. Kleinrock, *Computer Networks*, 3, pp. 267–286, January, 1979.

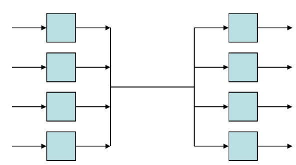
Issue: Blocking



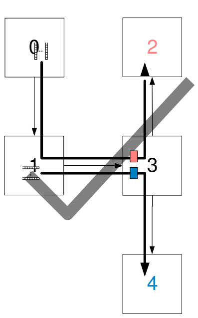
- 0 sending to 2
- 1 sending to 4
- Single fabric channel
- Flits from one packet can block flits from another packet on the same fabric channel
- Packet from 0 to 2 can block packet from 1 to 4 by consuming flit buffer in 3 and not releasing it

Virtual Channels

- Multiplex multiple channels over one real channel



One Solution: Virtual Channel Flow Control



- Separate virtual channels can eliminate blocking
  - 4 has its own set of resources independent of 2
  - If 2 is blocked, traffic to 4 is not blocked
- How many virtual channels?
- Buffers

Is Blocking a Real Issue?

- Depends on frequency and application
- For an Internet router, it is
  - Cannot let one flow block another
  - Tend to not have huge speedups
- For a parallel computer, often not
  - Memory accesses self throttling
  - Even message passing tends to be self throttling

On-Chip Interconnection Network

- Previous works are started from core-to-core, (old time computer to computer)
- Now core to core is on a chip

New challenges in NOC

- Wires are cheap on chip
- Wires are limited by pin in off-chip
- Buffers are expensive on-chip
- Power dissipation is critical
- More circuit related challenges
  - Repeaters, different delays
- Route is statically determined (when a chip is designed)

Rote Packets, Not Wires:on-chip interconnection networks, Dally and Towles, DAC 200