

# 6.829 Fall 2006

Quiz 1 Review Session

October 13, 2006

Rob Beverly, Sachin Katti

{rbeverly, skatti@mit}

## Quiz Topics

- Switching
  - Circuit Switching
  - Packet Switching
  - Virtual Circuits
- LAN Interconnections
  - Switching
  - Bridges
  - Spanning Tree
- Router Hardware Design
  - HOL blocking
  - VOQ
  - Fabrics
  - Schedulers
- Wide-Area
  - CIDR
  - LPM
- Intradomain Routing
  - Distance Vector
  - Link State
- Interdomain Routing
  - BGP
  - Policy
- Wireless (Sachin)
  - MAC
  - Routing
  - Diversity
  - Coding

## Quiz 1 Review

- Format of lecture:
  - List all topics we've covered, understand what you don't know you don't know
  - Cover as many topics as possible
- Note on readings:
  - Quiz is open papers/open notes
  - Expect you to have read papers
- Previous years: students run out of time!
  - More familiar you are with papers, faster you can answer

## Switching

## Terminology: Bridging/Switching/Routing

- **Bridging vs. Switching:**
  - Bridge: connect two LANs
  - Switch: multi-port bridge
- **Shared vs. Switched LAN**
  - Competition for medium
  - Broadcast flooding
- **Routing vs. Switching**
  - Forwarding on layer 3 vs layer 2

## Switching

- Want to interconnect many computers to form larger networks (LANs)
- First, review meta-issue of switching
- Two distinct methods evolved:
  - Circuit Switching
  - Packet Switching
  - Virtual Circuits (hybrid of both)

## Circuit Switching

- Frames contain no routing information
- Setup phase
- Transfer phase
- Not work conserving
- How? ex: TDM
- Example?
  - Telephone network
  - Why? Uniform, predictable workload (CBR)

## Circuit Switching

- Disadvantages:
  - Wastes bandwidth
  - Reserve maximum bandwidth for a connection
  - Can't burst
  - Know the traffic load apriori
- Data traffic is bursty (VBR)

## Packet Switching

- Datagram Routing
- Frames contain header with forwarding information
- Implications?
  - Switches must have routing table
  - Difference between forwarding and routing table?

## Packet Switching

- What can we put in header?
  - Destination
  - Sequence of hops (source routing)
- Is source routing unattractive?
  - Hosts participate in routing protocol
  - Local path decisions vs ISP decisions
  - Paths change
  - Does it exist in the Internet?

## Virtual Circuits

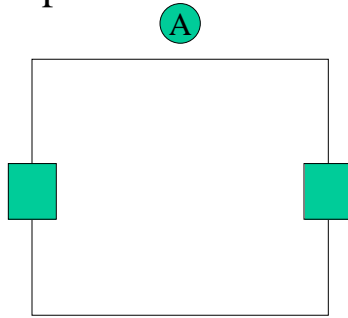
- Hybrid between circuit and packet switching
- Setup phase establishes local tags on switch
- Tags rewritten by switch
- Advantages (?):
  - Route pinning (traffic engineering)
  - Faster tag lookups (not in 2006)
  - Resource Reservation in signaling phase
- Disadvantages (?):
  - Another box in network
  - Tag overhead
  - Complexity
  - Layer 2/3 Interactions (good/bad)

## Bridges

- Interconnect hosts and LANs
- Dumb: receive and broadcast (flood)
- Transparent
- Learning bridges:
  - Promiscuous
  - What is it learning? Layer 2 MAC, e.g. Ethernet
  - Remember port-to-MAC of machines that transmit
- Mobile nodes?

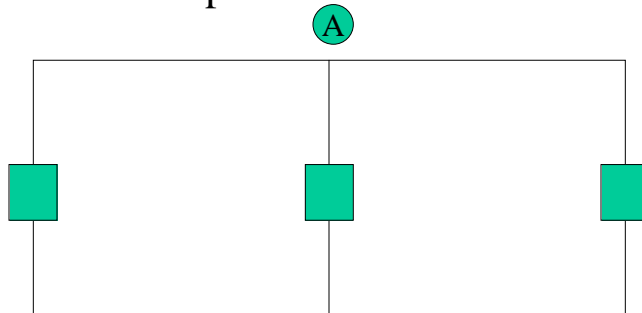
## Bridges

- What about topological loops?
- Recall examples
- Infinite Loop:



## Bridges

- What about topological loops?
- Recall examples
- Packet multiplier



# Bridges

- Solutions?
- Construct loop-free topologies
  - What about redundancy?
- Spanning-Tree:
  - Bridges have unique per-port IDs (vendor assigned)
  - Root election process
  - Form SP tree to root
  - Periodic config messages [ID][Root][d(root)]
  - Silent if hears a better metric on same LAN
- Why is Soft-state important?
  - Bridge failures
- VLANs
  - Color different segments
  - Run separate instances of spanning tree

# Interdomain Routing



## Internet Scaling

- Why not build the Internet out of learning/Spanning tree bridges?
  - Heterogeneity
  - Scaling:
    - Flooding
    - Tree messages
    - Per-host state in switches

## Area-Routing

- Why does Internet Scale?
  - Hierarchical Addressing
- How are addresses assigned?
- Classfull Addressing:
  - A = First bit 0, 7 bits network, 24 bits host
  - What's wrong with classes?
- CIDR (Classless)

## CIDR

- Stop-gap measure to prevent:
  - Address depletion
  - Route table growth
- Arbitrary network boundaries (not byte)
- Allows for proper sizing
- Allows for aggregation
- Stroke Format: prefix/mask

## CIDR

- Example:
  - 198.61.4.0 (class C)
  - 198.61.5.0 (class C)
  - Aggregate as: 198.61.4.0/23
- What about:
  - 198.61.3.0 (class C)
  - 198.61.4.0 (class C)
  - Can this be aggregated as: 198.61.3.0/23 ?
  - No! 3 = xxxxx011
  - 4 = xxxxx100

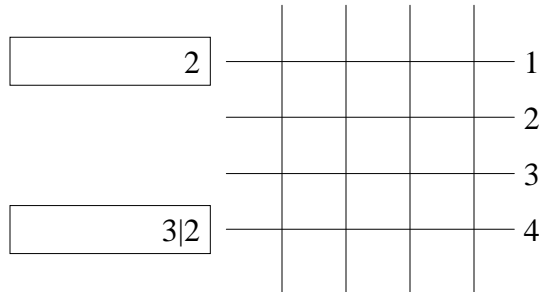
## Routing Hardware

- Let's look at iSLIP paper (assigned reading)

## Hardware Fabrics

- Shared Bus
  - Analogous to personal computer bus
- Shared Memory
  - Input writes to memory, output reads from memory
  - Pass memory pointers
- Crossbar
  - Matrix of configurable paths
- Self-routing
  - Banyan example

## Head of Line Blocking

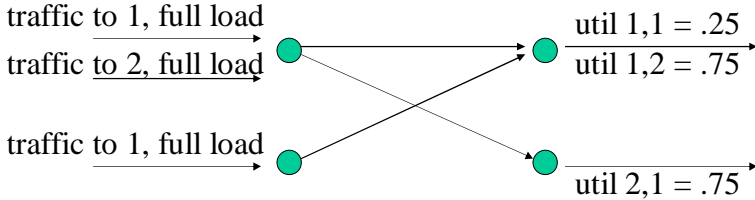


Significantly reduces efficiency of crossbar  
Solution: Virtual Output Queues (VOQ)

## Crossbar Scheduling

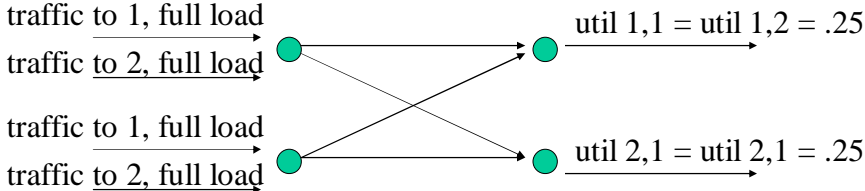
- **Parallel Iterative Matching (PIM):**
  - Step 1: request
  - Step 2: grant
  - Step 3: accept
- **Round Robin Matching (RRM):**
  - Step 1: request
  - Step 2: grant on next in fixed RR order
  - Step 3: accept on next in fixed RR order

# PIM Unfairness



Under heavy oversubscription, we get different (unfair) output rates

# RRM Problems



Under heavy oversubscription, synchronization of random arbiters leads to 50% throughput

## iSLIP

- Just like RRM
- Avoid synchronization between output arbiters
- Don't move grant pointers unless grant is accepted