CSPP 57130
Router VS Switch

- Protocols
- Hardware
Router HW

- Architectures
- Capabilities
Switching

- Shared memory, shared CPU switches
- Optimal for data moves
- Limited by memory bandwidth
- Shared Bus, multiple CPUs
- Shared Bus, per-card CPUs
- Forwarding decision handled off-card
Crossbar switches

- Take-a-ticket
- \(N^2\) crosspoints
- One output per destination
  - No more than one input per output at a time
- Request, grant, connect
Take-A-Ticket

- Ticket is small integer
- Control bus
- Watch those before
  - Self-schedule asynchronously
  - Hunt groups
  - Hash to ensure order
  - Dec GigaSwitch(multicast problems)
Head-of-Line Blocking

- Head blocking limits parallelism
- About half of bandwidth
- No help if all traffic destined for one port
- Ideal for other orders
- What is traffic distribution?
Output queuing

- Run switch fabric $N$ times faster
- Fails for lots of ports
- Only need $k$ times fast
- Use cells
- $N=2k$, $N>k$, fairness
Parallel Iterative Matching

- Input queue without HOL blocking
- Schedule more than just the input head
- Only $N$ possible destinations
- Only schedule first cell to each port
- $N^2$ bit bitmap for queue
- Input port contention resolution
PIM Input Resolution

- Random
- Might be suboptimal (all choose A)
- Iterate
  - Mask out matches, run again
- Worst case N matches
- AN-2 30 Port switch used 3 levels
iSLIP

Avoid randomization

Use round robin pointer

Initial sync problem, but disappears

Keep grant pointer

Avoids starvation

Some commercial versions use one iteration
Priorities

Cisco VoIP

- Modify iSlip to keep grant pointer for each priority, $k$
- Each output grants only for highest priority it receives
- Each input accepts for highest priority it receives
Multicast

- Take-A-Ticket handles multicast poorly
- Lower performances and high waste
- Crossbar hardware is ideal for multicast
- How to schedule it?
- iSLIP for multicast, ESLIP
  - One additional multicast queue per input
  - Some HOL on multicast port
Multicast vs Priority

What is the priority of multicast?

ESLIP uses every other slot for handling multicast, if present

Might take two slots to finish a single multicast fan-out

Cisco implementation has 4 priorities and one multicast
Large Switches

- Crossbar limit is chip pins
- Up 200 can fit in a chip
- Interconnect chips
- Clos network (1955-2002 Juniper T-Series)
  - Nonblocking – connection if possible
  - Clos nonblocking for $k \geq 2n - 1$
  - $5.6N \times \sqrt{N}$
Speed Scaling

- Short Links
- 1D, 2D and 3D torus
- Memory Scaling
- Buffer one RTT for congestion-control
Product Specifics

- Cisco Catalyst 3750 switch
- Architecture
- IOS
Cisco  5500/6500
Cisco 3750 Traffic Shaping

- Weighted Tail Drop
- three Thresholds
  - CoS 5 100%
    - No packets dropped unless full
  - CoS 0-4 30% drop threshold
  - CoS 6-7 70% drop threshold
Cisco 3750 Cont.

- Four hardware egress queues
- Shaped Round Robin
- Aggregate rate shaping
  - Specify maximum combined transmission rate for all four queues
- Individual rate shaping
  - Queues not shaped are shared by SRR
Cisco 3750 Cont.

- Traffic Policing
  - Like shaping, but no buffer
  - Traffic beyond limit is dropped
  - Non-blocking switch fabric
  - Two ingress queues
  - Strict priority traffic, other traffic
VLANS

Define broadcast domain virtually, rather than by hardware connection.

VLAN tags inserted into packet sent over "Trunk" connections--new CRC

4 bytes (802.1Q)

2 Byte, length, Type

3 bit priority, 1 bit unused, 12 bit VLAN
VLANs

- Sometimes called “layer 2 routing”
- Useful for security configurations
- Only as strong as security on switches
- Useful for traffic and administration segregation
- Can support bad habits