Switches

Jean Walrand
U.C. Berkeley
www.eecs.berkeley.edu/~wlr
Outline

- Switch Characteristics
- Examples
- Functions
- Output Buffer
- Input Buffer
- Virtual Output Buffer
- Shared Memory
- Modular
Switch Characteristics

- **Ports**
  - Fast Ethernet, OC-3, ATM, ...

- **Protocols**
  - ST, Link Agg., VLAN, OSPF, RIP, BGP, ...

- **Performance**
  - Throughput, 8-classes CoS, ...
Cisco 12416:

The Cisco 12416 Internet router is a 10 Gigabit, 16-slot chassis member of the Cisco 12000 series that provides a total switching capacity of 320 Gigabits per second (Gbps), with 20 Gbps (10 Gbps full duplex) capacity per slot.

With its 16-slot chassis and extensive portfolio of line cards including the new Cisco 1-port OC-192c/STM-64c and Cisco 4-port OC-48c/STM-16c POS interfaces, the Cisco 12416 Internet router supports high-density ISP aggregation and point-of-presence (POP) consolidation.
**Cisco 3600**

Voice over Frame relay (VoFR) and Voice over ATM (VoATM-AALS) on the digital voice interfaces (T1 and E1).

Gateways for the PBX and PSTN for IP telephony, enabling applications like call transfers, holds, and conferencing.
Examples (cont’d)

- Extreme Networks - Summit

- 48 10/100 ports
- 2 GE (SX, LX, or LX-70)
- 17.5Gbps non-blocking
- 10.1 Mpps
- Wire speed L2
- Wire speed L3 static or RIP
- OSPF, DVRMP, PIM, ...
Examples (cont’d)

- Foundry ServerIron

- Server Load Balancing
- Transparent Cache Switching
- Firewall Load Balancing
- Global Server Load Balancing
- Extended Layer 4-7 functionality including URL-, Cookie-, and SSL Session ID-based switching
- Secure Network Address Translation (NAT) and Port address translation (PAT)
Functions

- Data Path and Control Path:

Note: Figure from Prof. Varaiya’s notes for EE228b
Functions (cont’d)

- Data Path:
  - Extract header
  - Lookup forwarding rule
  - Queuing decision (drop, mark, queue)
  - Modify header
  - Store packet
  - Schedule transmission
  - Send packet
Functions (cont’d)

- Control Path
  - Routing table (IP, MPLS, ST, …)
  - Forwarding policies (RED, Policing, …)
  - Scheduling rules (WFQ, Priority, …)
Functions (cont’d)

- Timing:

Note: Figure from Prof. Varaiya’s notes for EE228b
Functions (cont’d)

- Table Lookup Example (routing)

Note: Figure from Prof. Varaiya’s notes for EE228b
UCB Functions (cont’d)

- CIDR Lookup

Ternary CAM

<table>
<thead>
<tr>
<th>010011**</th>
<th>00011***</th>
<th>001*****</th>
<th>11******</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = don’t care

Returns first match

Associated Data = Port

010011

00011

001

1

11
Functions (cont’d)

- Policing: Leaky Bucket

P bytes

Capacity B

Leak Rate R

Add P units

If overflow, drop packet or mark packet
Functions (cont’d)

- RED: Random Early Discard
  Objectives: Reduce synchronization of sources
  Avoid penalizing bursts
  Drop more from more active sources

Effectiveness: so so

Operations:

\[
\begin{align*}
q & \quad q_{AV} \\
0.1 & \quad L \\
H & \quad L
\end{align*}
\]
Functions (cont’d)

- RED In and Out (RIO)

IN = unmarked
OUT = marked
Priority Scheduling:

When finished transmitting a packet, serve next from first nonempty queue in order (1, 2, 3)
Functions (cont’d)

- Deficit Round Robin (DRR)

Get to new queue:
- Empty: Skip; Else: Add weight to credits
- Credits > 0 => serve one packet, subtract 1 credit per byte
- Credits < 0 => Go to next queue
Repeat

Weight

5

(5/10)×R

2

(5/10)×R

3

(5/10)×R
Functions (cont’d)

- DRR (cont’d)

```
Credits A: 0 7 -3 4 -16 -9 -2 5 2 -16
Credits B: 0 3 -11 -8 -5 -2 1 -4 ...
```

Send 10 14 20 3 18 5
functions (cont’d)

- Weighted Fair Queuing (WFQ)

Compute departure times from processor sharing queue with these weights. Serve next the queue with earliest deadline.
Functions (cont’d)

- WFQ (cont’d)

Leaves at t

- A
- B
- C

10 5
2
3

14
10

- 3/8
- 5/8

=> Serve A
Output Buffer

- Buffers at output ports
  - Limitation: Throughput < rate of shared bus

Note: Figure from Prof. Varaiya’s notes for EE228b
Input Buffer

- Buffers at input ports
  - Limitation: Head-of-Line blocking

Note: Figure from Prof. Varaiya’s notes for EE228b
Virtual Output Buffer

- OUT buffers at each input port
  - Complexity: Matching Problem

Note: Figure from Prof. Varaiya’s notes for EE228b
Virtual Output Buffer (cont’d)

- Full-Throughput Scheduling:
  Maximum Weighted Matching

\[
\begin{align*}
A &= 14 \\
B &= 11 \\
C &= 15 \\
D &= 10
\end{align*}
\]

\[B + C > A + D\]

\[\Rightarrow \text{Serve (B, C)}\]
UCB Virtual Output Buffer (cont’d)

- i-SLIP

1. Request:
   Nonempty VOBs request to outputs

2. Grant:
   Outputs grant (in RR order) VOBs

3. Accept:
   Granted VOBs accepts (in RR order)

4. Iterate:
   Repeat after removing accepts
Virtual Output Buffer (cont’d)

- i-SLIP (cont’d)

Request  Grant  Accept  Iter

Last Grant

Last Accept
Shared Buffer

- One memory pool shared by flows
  - Limitation: Memory Speed

Note: Figure from Prof. Varaiya’s notes for EE228b
Modular Switches

- Space-Division Switching
  - Blocking
Modular Switches (cont’d)

- Non-Blocking