Introduction to Routers and LAN Switches

Session RST-101
Prerequisites

- OSI Model
- Networking Fundamentals
Agenda

- Routers and LAN Switches
- Components of an Architecture
- Summary
Routers and LAN Switches
What Are Routers and Switches

Routers Care about L3 Addresses

Switching/Forwarding Decision

Switches Care about L2 Addresses

OSI Model

Transport

Network (L3)

Data Link (L2)

Physical

OSI Model

Transport

Network (L3)

Data Link (L2)

Physical
Routers and Switches: Why Do We Need Them??

- Networks need to be connected to other networks

At L2 Same Media <> Media = **L2 Switching** (Bridging)
(No Frame Changes, Speed Mismatches Accommodated)

At L2 Different Media <> Media = **L2 Translational Switching** (Bridging)
(Frame Format Changes, Including Addresses, Flags, etc.)
Routers and Switches: Why Do We Need Them??

- However...

In Both Examples, All Connected Devices Are Part of the Same Broadcast Domains, **There Is No Layer 3 Segmentation!!**
This Becomes a Performance Issue Even in Medium Sized Networks
Routers and Switches: Why Do We Need Them??

- Networks need to be connected to other networks (LANs <> WANs, LANs<>LANs etc.)
- Networks don’t scale well at L2 and need to be segmented at L3
What Are Routers and Switches

- Routers and switches both make a decision as to how to handle packets or frames
- A frame is a L2 encapsulation
- A packet is a L3 encapsulation
Ethernet Frame and IP Packet

<table>
<thead>
<tr>
<th>Preamble</th>
<th>SFD</th>
<th>DA</th>
<th>SA</th>
<th>Length</th>
<th>Data/Payload</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 bytes</td>
<td>1 byte</td>
<td>6 bytes</td>
<td>6 bytes</td>
<td>2 bytes</td>
<td>Up to 1500 bytes</td>
<td>4 bytes</td>
</tr>
</tbody>
</table>

- **Preamble**: 6 bytes
- **SFD**: 1 byte
- **DA**: 6 bytes
- **SA**: 6 bytes
- **Length**: 2 bytes
- **Data/Payload**: Up to 1500 bytes
- **FCS**: 4 bytes

**IP Packet Structure**

- **Version**: 4 bits
- **Header Length**: 12 bits
- **TOS**: 8 bits
- **Total Length**: 16 bits
- **Identifier**: 32 bits
- **Flags**: 3 bits
- **Fragment Offset**: 13 bits
- **TTL**: 8 bits
- **Protocol**: 8 bits
- **Header Checksum**: 16 bits
- **Source Address**: 32 bits
- **Destination Address**: 32 bits
- **Options**: 0-40 bits
- **Padding**: 0-40 bits
What Switches Really Do

- L2 classification
- *Switching Table* construction
- Other activities
Switches (Operationally)

- Maintain/manipulate *Switching* information
  - Record updates in MAC Table
- Perform Layer 2 switching
  - Compare Destination Address to “Learned” MAC Table
  - Check frame for errors
- Management/billing (statistics)
  - Interface statistics—Number of frames sent, error/collision counters, utilization
Switches
(Layer 2 Frame Functionally)

- Switch frames
  Layer 2 switching based on “Switching” information
- Transmit frames
  Access outbound memory (buffers) and physical media
- Flood frames
  Flood any unknown multicast/broadcast frames out all ports of the switch
- Drop frames
  Drop any frames that contain errors
L2 Classification

1. L2 provides local link addressing and data integrity validation.

2. Look at Destination Address to determine which interface the frame will be switched through.

3. Frame stored in shared buffer.

4. Transmitted to the appropriate port based on the Switching Table.

Default behavior of a switch is to flood unknown multicast/broadcast frames.
Frames enter the switch at ingress, and are stored in the shared memory buffer. The Destination MAC Address is mapped to the appropriate port number in the Switching Table.

Switching Table Construction

<table>
<thead>
<tr>
<th>MAC Address</th>
<th>Port Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>00-10-A4-A6-FC-17</td>
<td>FastEthernet 0/14</td>
</tr>
<tr>
<td>00-10-B2-B4-FA-15</td>
<td>GigabitEthernet 0/1</td>
</tr>
</tbody>
</table>
What Routers Really Do

- L3 classification
- *Forwarding Table* construction
- *Forwarding* decision making
- Other activities
Routers (Operationally)

- Maintain/manipulate *Forwarding* information
  
  Listen for updates/update neighbors

- Classify packets for manipulation/queuing/permit-deny, etc.
  
  Compare packets to classification lists and perform control

- Perform Layer 3 switching
  
  Create outbound Layer 2 encapsulation
  
  Layer 3 checksum (IPv4 Only)
  
  TTL/hop count update

- Management/billing (statistics)
  
  Interface statistics—NetFlow export
  
  Telnet, SNMP, ping, trace route, HTTP
Routers
(Layer 3 Packet Functionally)

- **(Attempt to) switch packets**
  Layer 3 switching based on “Forwarding” information

- **(Attempt to) transmit packets**
  Access outbound memory (buffers) and media

- **Manipulate packets**
  Change contents of packet (CAR/NAT/compression/encryption)

- **Consume packets**
  Routing protocol updates etc.../services advertisements(SAP)/ICMP/SNMP

- **Generate packets**
  Routing protocol packets/SAPs/ICMP/SNMP
  Tunnels—GRE, IPSec, DLSw etc...
L3 Classification

1. L2 provides local link addressing and data integrity validation

2. Look at L3 header to determine Flow Characteristics:
   - L3 Protocol (IP, IPX, AT etc..)
   - Check to see if packet is destined for router
   - Check for any options/features (inbound)

3. Decide outbound controls, based on flow characteristics

4. L2 provides local link addressing and data integrity validation
Forwarding Table Construction

Update packets are queued for the CPU and dealt with by an appropriate software routine to build a Forwarding Table.

The router also looks at connected networks, Interface state and configured routes to complete the picture.

Forwarding Table
- 172.16.2.0/24 via 172.16.1.2
- 10.1.1.0/24 via 172.16.1.2

Updates are generated by the router and queued for transmission on interfaces configured for the given protocol.
Forwarding Table Construction

ARP Table
172.16.1.2: 50000603E...AAAA03000800
172.16.1.3: 10134567A...ECE030178654

Local ARP queries build ARP table

Requesting station IP/MAC address used to add to ARP table
L3 Forward to Next Hop

Look at L3 header to determine Flow Characteristics:
- L3 Protocol (IP, IPX, AT etc.)
- Check to see if packet is destined for router
- Check for any options/features (inbound)

L2 is just used for local link addressing and data integrity validation

Forwarding Table
1.1.0.0/16 via 172.16.2.1
10.1.1.0/24 via 172.16.1.1

ARP Table
172.16.2.1: 0F000800
172.16.1.1: 10134567A...ECE030178654

L2 is just used for local link addressing and data integrity validation

Destination
Other Activities

Check IP header Checksum (IPv4)
Check TTL
Check for ability to Fragment
Check packet against various feature lists
(I/P ACL’s, NAT, CAR, RFP etc)

Decrement TTL
Re-write IP header Checksum (IPv4)
Fragment
Check packet against various feature lists
(O/P ACL’s, NAT, Queuing)

L2 is just used for local
link addressing and data
integrity validation

L2 is just used for local
link addressing and data
integrity validation
How Do I Find PC 2?

"Where is PC 2?"
PC 1 Sends a Broadcast to See If PC 2 Is Locally Connected
Broadcast Propagation—L2 Switch

- Switch sends the broadcast frame out all the ports within the broadcast domain.
Broadcast L2 Example
Broadcast Propagation—L3 Routing

- Router terminates the broadcasts, does not propagate them everywhere

RST-101
3048_05_2001_c1 © 2001, Cisco Systems, Inc. All rights reserved.
Broadcast L3 Example

Server

PC 1
Request for PC 2

Request for PC 2

PC 2
Components of an Architecture
Components of an Architecture

- Switching fabric
- Memory/buffers
- Queuing
- Distributed vs. centralized
- Forwarding architectures
Switching Fabric

Connection Between the Slots/Ports in a Switch
Shared Memory

General Purpose CPU
(CISC older or RISC newer)

Buffers
Queues
Pointers
Headers

Forwarding Tables
IOS Image/Files
System Buffers
Processor Queues

Packet Memory

CPU Memory

Physical Media Interfaces
(Fixed or Modular)

Data/Address/Control Bus’s
Each Line card has Packet Memory, Forwarding Table Memory and a discrete CPU.

A Copy of the central forwarding table is propagated from the Central Route Processor to the Line Cards for Local switching of packets.
Cross Bar Data Path

- Multiple conflict free paths
- Typically higher bandwidth capacity
- Signaling and Scheduling more complex

ASIC X-Bar Fabric
Non-Blocking Switching Fabric

60 Gbps Fabric

10 Gbps Ports

Speed of Fabric > Ingress + Egress
Blocking Switching Fabric

Speed of Fabric < Ingress + Egress
Components of an Architecture

- Switching fabric
- Memory/buffers
- Queuing
- Distributed vs. centralized
- Forwarding architectures
Contiguous Buffer Allocation

- Buffer length fixed in size (often to MTU)
- Less expensive than particle buffering architectures
- Inefficient use of buffers
Particle Buffer Allocation

- Each buffer fixed in small increments (for example, 64 bytes each)
- Allows for efficient use of buffers
Shared Memory

Shared memory divided into ‘Pools’ of buffers and Buffer Queues.
Shared Memory

When packets arrive on interfaces they are DMA’d into appropriate buffer without interrupting CPU
CPU reads packet header information into registers and compares with forwarding table.
CPU derives next hop MAC address and loads a new header into register.

New header over-writes existing header.

Buffer ownership transferred to output interface.
Components of an Architecture

- Switching fabric
- Memory/buffers
- Queuing
- Distributed vs. centralized
- Forwarding architectures
Input Queuing

- Packets buffered at the inbound port
- Can result in head of line blocking if you have a single input queue per port
- Can reduce throughput
Output Queuing

- Buffers at the output port
- Allows for individual prioritization of traffic flows
Output Queuing/Shared Buffer

- Central pool of buffers shared between all ports
- Maximum throughput with fewest buffers
- No head of line blocking with intelligent congestion management
Multiple Queues Per Port

- Can be implemented in either output queuing or shared memory models
- Scheduling and/or congestion avoidance algorithm required
- **Note:** Number of queues affect overall number of buffers per queue
How Does Traffic Run in a Real Network?

100Mbps Port  

6 Gbps Fabric  

100Mbps Port  

WAN Port  

100Mbps Port
Head of Line Blocking (HOL)
Single Input FIFO Queue

X- Bar
(Same problem exists with Shared Memory Routers)

Single Ingress FIFO Queue

Interface

Congested Interface

A
Delayed/Dropped

B

C

CPU

Single Input FIFO Queue

C C C C B

Interface

Interface

Interface
"Lane Control"

You can only ever sit in a lane that is designated for the corresponding lane you are trying to exist the junction from.
All Destinations Have a Lane

No congested interface (Outbound) Can affect another Interface
Virtual Output Queuing

- CPU
- Interface
- FIFO
- Interface
- Congested Interface
- Interface
Virtual Output Queue

Queue Scheduler

CPU

Interface

FIFO

Interface

FIFO

Interface

B

Congested

Interface

C

A
Components of an Architecture

- Switching fabric
- Memory/buffers
- Queuing
- Centralized vs. Distributed
- Forwarding architectures
Centralized Switching

- Central forwarding table utilized
- Provides centralized control for switching and learning
- Lookup can be done in ASICs for faster processing
- Can perform a Layer 2 or Layer 3 lookup
Distributed Switching

Each Line card has Packet Memory, Forwarding Table Memory and a discrete CPU.

A Copy of the central forwarding table is propagated from the Central Route Processor to the Line Cards for Local switching of packets.
Components of an Architecture

- Switching fabric
- Memory/buffers
- Queuing
- Distributed vs. centralized
- Forwarding architectures
Serial vs. Hashed Lookup

- Sequentially look up entries in table
- Simplistic implementation
- Hash values together to obtain a value in memory
- Benefit: Faster lookup in larger tables
- VLSM causes more overhead and maintenance
Demand Generated Cache Based Switching

CPU Memory

**Forwarding Table**
- 1.1.0.0/16 via 172.16.2.1
- 10.1.1.0/24 via 172.16.1.1

**ARP Table**
- 172.16.1.1: 0F000800
- 172.16.2.1: 10134567A...ECE030178654

**Fast Cache**
- Prefix/Length: 1.1.0.0/16
  - Age: 00:00:15
  - Interface: Ethernet0
  - Next Hop: 172.16.2.1 14 00000C7EF7CF00E0B06423F60800
- Prefix/Length: 10.1.1.0/24
  - Age: 00:00:15
  - Interface: Serial1
  - Next Hop: 172.16.1.1 4 0F000800
Topography-Based Switching

- Forwarding Information Base (FIB)
- Cisco Express Forwarding (CEF)

Routing Protocols Inject Routes into the Routing Table

OSPF, IGRP, EIGRP, RIP, BGP, IS-IS

Forwarding Information Base

DA | Adjacency | Interface
--- | --------- | --------
10.1.2 | A0-B8-FE | FF 2/0

Forwarding Table

- 172.20
- Fast Ethernet 3/1
- 10.1.2
- Gigabit Ethernet 1/1
- 96.23
- VLAN 100

Switching Fabric

Distributed FIB

Distributed FIB
Topology-Based Switching

- FIB calculated based on routing table entries, not traffic flows
- FIB can be kept central or distributed
- Longest match lookup on prefix/mask
- More scalable for large enterprises and service providers
Topology-Based Switching

- Packet enters switch
- No process switching necessary
- Decision made locally or centrally irregardless of switching fabric
Summary
Summary

- LAN switches and routers use the **same** architectural features
- LAN switches give scalability at **L2**
- Routers give scalability at **L3**
- Deployment of LAN switches and routers in a hierarchical network model offers **scalability, reliability and mobility**
Network Implementation

MDF
- Centralized Server Farm
- Call Manager
- Cache Engine (Optional)

WAN Aggregation

Building A
- Local Workgroup Servers
- Workstations

Building B
- Local Workgroup Servers
- Workstations

Routed Uplinks to Core