Packet Forwarding and Switching basics review Module 1:

Content:

- Network Device Communication - How switches forward traffic from L2 & routers via L3
- Forwarding Architectures - Mechanisms used in routers and switches to forward traffic
- High Availability - Stateful switchover (SSO), Nonstop Forwarding (NSF)

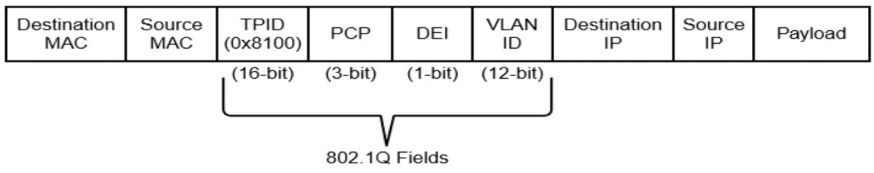
Collision Domains on a Hub vs a switch

- Unknown unicast flooding occurs when a packet containing a destination MAC Address is not in the switch's MAC address table. The switch forwards the packet out of every switch port.
- Broadcast traffic is network traffic intended for every host on the LAN and is forwarded out of every switch port interface
- Network broadcasts do not cross Layer 3 boundaries (different subnets)

Virtual LANs

- Adding a router between LAN segments helps shrink broadcast domains
- Provide logical segmentation by creating multiple broadcast domains on the same network switch. VLANS provide higher utilization of switch ports because a port can be associated to the necessary broadcast domain, and multiple broadcast domains can reside on the same switch.
- VLANs are defined in the IEEE 802.1Q standard that states that the 4 bytes are added to the packet header with the following fields: tag Protocol identifier (TPID), Priority code point (PCP) 3 bits, drop eligible indicator (DEI), VLAN Identifier (VLAN ID) 12 bits

Figure 1-4 displays the VLAN packet structure.



- **Priority code point (PCP): This 3-bit field indicates a class of service (CoS) as part of Layer 2 quality of service (QoS) between switches.**

Native VLANs

↳ In 802.1Q, any traffic that is advertised or received on a trunk port without the 802.1Q VLAN tag is associated to the native VLAN
↳ Default = 1
↳ Has to match on both trunk ports, or traffic can change VLANs unintentionally. Will be a headache if not done properly.
↳ Port specific configuration

MAC Address Table

↳ Resides in the Content addressable memory (CAM) table.

CAM uses high-speed memory that is faster than typical computer RAM due to its search techniques.
↳ Provides a binary result for any query of 0 (true) or 1 (false)

Forwarding traffic from L2 uses destination MAC Address
— — - L3 uses destination IP Address

Layer 3 Forwarding and Layer 2 Frame Forwarding "Local Network Forwarding"

↳ Some Layer 3 forwarding logic occurs before Layer 2 forwarding.
1) Forwarding traffic to the same subnet
2) — — — different subnet

Local Network forwarding (on the same subnet)

Two devices that reside on the same subnet communicate locally. As the data is encapsulated with its IP address, the device detects the destination is on the same network. However, the device still needs to encapsulate the Layer 2 information to the packet. It knows its own MAC address but not the destination MAC address initially.
↳ Solved by using Address Resolution Protocol (ARP) to map L3 ip addresses to L2 MAC addresses

Forwarding Architectures

Process Switching:

↳ AKA software switching (slow)
↳ Process that uses the general purpose CPU on a router for packet switching

The types of packets that use this process are the following:
1. Packets sourced or destined to the router
2. Fragmented Packets
3. Encrypted Packets
4. Heavily complex for hardware to handle (IP packets with IP options)

↳ Routing table (RIB) is built from information obtained from dynamic routing protocols and directly connected + static routes.
↳ ARP table is built from info gathered by the ARP protocol
↳ RIB & ARP table both reside in the device control plane.

Switching Database Management (SDM) Templates

- The number of MAC addresses a switch needs, compared to the number of routes it holds, depends on where it is deployed.
- The memory for Tenary content Addressable Memory (TCAM) tables is statically allocated during the boot up sequence of the switch.
- Memory allocation ratios between the various TCAM tables are stored and can be modified with SDM templates
  - Switch (config)# sdm prefer {vlan | advanced}
    # write
    then reload switch

Centralized Forwarding and Distributed Forwarding

- When a route processor (RP) engine is equipped with a forwarding engine so that it can make all the packet switching decisions, this is known as centralized forwarding architecture
  - When a packet is received on the ingress line card, it is transmitted to the forwarding engine.
    - Examines packet headers to determine that the packet will be sent out a port on the egress line card and forwards the packet
- If the line cards are equipped with forwarding engines so that they can make packet switching decisions without intervention of the RP, this is known as a Distributed forwarding architecture
  - When a packet is received on the ingress line card, it is transmitted to the local forwarding engine.
  - Forwarding engine performs a packet lookup, and if it determines that the outbound int is local, it forwards the packet out a local interface
  - If outbound interface is located on a different line card, the packet is sent across the switch fabric directly to the egress line card, bypassing RP

CEF and TCAM

Cisco Express Forwarding:

- Cisco proprietary switching mechanism
- Used by default on all Cisco platforms
- uses specialized application-specific integrated circuits (ASICs) and Network Processing Units (NPU) for high packet throughput

Ternary Content Addressable Memory:

- Allows matching and evaluation of a packet on more than one field
- Entries are stored in Value, Mask, and Result (VMR) format
  - value = Indicates fields that should be searched
  - mask = Indicates field of interest and should be queried
  - result : Action taken that should be taken with a match on the value & mask
- Operates in hardware, providing faster processing and scalability than processing and scalability than process switching.

Components part of TCAM operation:

1. Feature manager (FM) - after an access list, QoS or routing table has been created or configured, the feature manager software complies, or merges, entries into the TCAM table. It can then be consulted at full frame forwarding speed.

2. Switching Database Manager (SDM) - Can partition the TCAM table memory allocations on Catalyst switches into areas for different functions.

Forwarding Architectures - CEF

Primary data structures:

1- Forwarding Information Base (FIB) - built directly from the routing table and contains the next-hop IP address for each destination
   ↳ "show ip cef" to display FIB content

2- Adjacency table - Contains both directly connected next-hop IP address and MAC address, as well as egress interface's MAC address
   ↳ "Show adjacency detail" to display CEF adjacency content

(Centralized) CEF

The Supervisor Engine (SUP)

Most important component in Catalyst switches that is needed to forward traffic

- Two SUPs can be installed on a single chassis removing the single point of failure
  - First module to successfully boot becomes the active supervisor
  - The other remains in a standby role in case the supervisor fails

Stateful Switchover (SSO)

- Provides minimal Layer 2 traffic disruption during Supervisor switchover
- Redundant Supervisor starts up in fully initialized state and synchronizes with start + running configuration of active supervisor
- Standby supervisor in SSO mode keeps in sync with active for all changes in hardware and software states

Configuration:

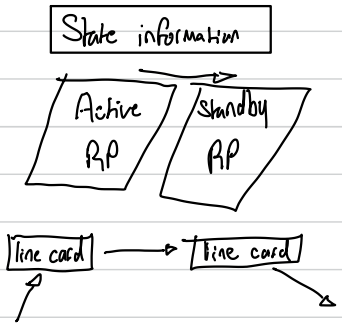
```
Switch(config)# redundancy
Switch(config-red)# mode sso
```

SSO can synchronize most **Layer 2 protocols** between the Primary and the Backup Supervisor Engines for both control plane and date plane

- 802.3x (Flow control)
- 802.3ad (LACP) and PagP
- 802.1X (Authentication and port security)
- 802.3af (Inline power)
- VTP
- Dynamic ARP inspection / DHCP snooping / IP Source Guard
- IGMP snooping (Versions 1 and 2) DTP
- (802.1Q and ISL) MST/PVST+/Rapid-
- PVST
- PortFast / UplinkFast / BackboneFast / BPDU Guard & Filtering
- Voice VLAN
- Unicast MAC filtering
- Access control lists (ACLs; VLAN ACLs, port ACLs
- Multicast storm control / broadcast storm control

NonStop Forwarding (NSF) with Stateful Switchover (SSO)

- NSF is a Layer 3 function that works with SSO to provide nonstop forwarding for L3 traffic in the event of failure of one of the member supervisor engines
- Focuses on quickly rebuilding all RIBs:
  - All major routing protocols (BGP, OSPF, ISIS, EIGRP) have extensions via NSF to continue sending IP traffic upon supervisor failure



Configuring and verifying NSF with SSO

- NSF is an additional configuration option for configuring SSO.
- Use nsf command to configure NSF for OSPF, EIGRP, and IS-IS
- Use bgp graceful-restart to configure BGP for nsf support.

```
Switch# configure terminal
Enter configuration commands, one per line. End with
CNTL/Z.
Switch(config)# router ospf 200
Switch(config-router)# nsf
Switch(config-router)# exit
Switch(config)# router bgp 100
Switch(config-router)# bgp graceful-restart
Switch(config-router)# end
Switch#
```

Forwarding Architecture

Process switching

↳ Software based
↳ Uses general purpose CPU for packet switching

Packet types that use this type of forwarding architecture:
1. Packets sourced or destined to the router
2. Fragmented Packets
3. Encrypted Packets
4. Heavily complex for hardware to handle (IP packets with IP options)

Switching Database Management Template (SDM)

↳ Number of MAC addresses a switch needs (based on routes) depends on where it is deployed.
↳ Memory for Ternary content Addressable memory (TCAM) tables is allocated statically during boot up sequence
  ↳ Overflow data goes to CPU negatively hampering performance

Configuration:
Switch(config)# sdm prefer {vlan | advanced} - enable sdm Switch# show sdm prefer - view current SDM template
Switch# write
reload

Centralized Forwarding

When a RP engine is equipped with a forwarding engine to make all packet switching decisions
↳ When a packet is received on the ingress line card, it is forwarded to the engine
↳ Engine examines the packet header to determine what port it leaves on the egress line card

Distributed Forwarding

If line cards are equipped with forwarding engines to make packet switching decisions without intervention from the RP
↳ When a packet is received, it is transmitted to the local forwarding engine
↳ Engine performs a packet lookup, if the outbound int is local, the packet is forwarded out a local int
  ↳ If the outbound int is located on a different line card, packet is sent across the switch fabric directly to the egress line card bypassing the RP

Ternary content Addressable Memory (TCAM)

Allows for matching and evaluation of a packet on more than one field and is operated in hardware
↳ entries are stored in value, mask and Result (VMR)
  Value: field to be searched (IP address & Protocol fields)
  Mask: field of interest that should be queried
  result: Action to be taken with a match on value & mask

2 components in operation:
1. Feature Manager: After an access-list, QoS or routing table has been created or configured, it compiles or merges the entries into the TCAM table. It can then be consulted at full speed
2. Switching Database Manager: Can partition the TCAM table memory allocation for different functions or tuning partitions

Cisco Express Forwarding (CEF)

Cisco proprietary switching mechanism that is default on all cisco platforms utilizing application-specific integrated circuits (ASICs) and network processing units (NPU) for high packet throughput

CEF data structures

1. Forwarding Information Base (FIB)
   - Built directly from the routing table to make IP destination prefix-based switching decisions
   - Contains next-hop IP address for each destination in the network
   - Changes made to routing table are directly reflected onto FIB
   - "# show ip cef" → displays FIB content

2. Adjacency Table / Adjacency Information Base (AIB)
   - Contains directly connected next-hop IP addresses + MAC Address + egress int
   - Populated with data from ARP table or other Layer 2 protocol tables
   - "# show adjacency detail" → display AIB

Centralized CEF

Upon receiving an IP packet, FIB is checked for a valid entry
- If entry is missing → "glean" adjacency which sends packet to CPU since CEF cannot handle it
- Valid FIB entries continue processing by checking adjacency table for each destination IP address
- Missing adjacency entries invoke ARP process. When resolved, CEF entry is created

Distributed CEF

- ASIC allow high packet rates, but limited functionality because they are hardwired to perform specific tasks.
- Routers have network processing units (NPUs) that are designed to overcome ASIC limitations
- Packet switching is done via dCEF
- CEF data structures are downloaded to forwarding ASICs and CPUs of all line cards to participate in packet switching