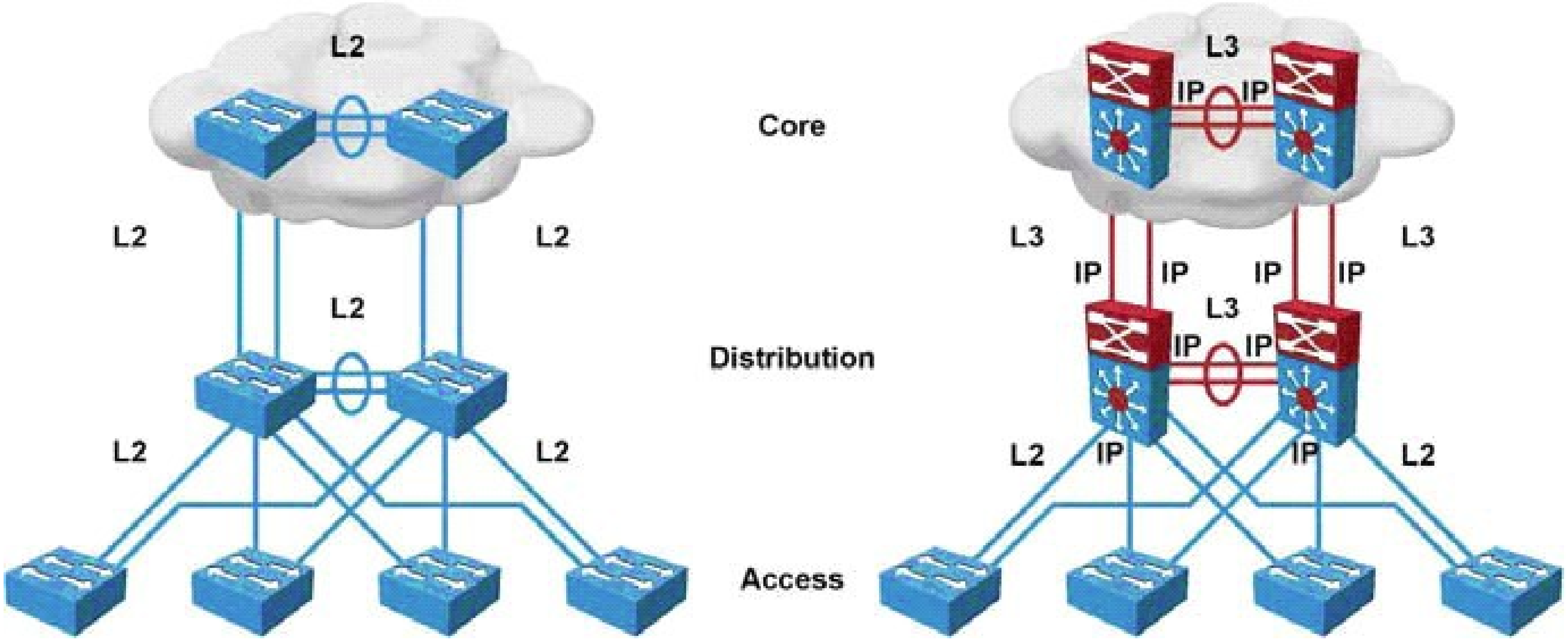
Chapter 5
L3 Functionality: Inter-VLAN Routing, DHCP
NET3011 – 17W

Inter-VLAN Routing – Basics

- Hosts in different subnets can not communicate without a router (e.g. default gateway)
- Three options exist for routing between VLANs:
  - external router / router on a stick
  - L3 switch with ports configurable as "routed"
  - L2 or L3 switch with SVIs and routing capability
- External router or L3 routed ports requires either:
  - one physical port per VLAN, or
  - trunk interface plus sub-interfaces
  - **but** in a campus network (or data centre!) there's likely *far too many* VLANs for one-port-per-VLAN
- Using SVIs with L2/L3 switches is more scaleable

L2 Switched vs L3 Routed Designs



From: FLG p. 212

External Router

- External router required if switches not L3 capable
- Advantages: simple topology, minimal ports if using a trunk, single point for troubleshooting
- Disadvantages: single path may get congested, router latency > L3 switch, single point of failure
- *<u>Not</u>* recommended for large networks
- Please ask or check notes from NET3008 for more info on external routing

```
interface FastEthernet 0/0
 no shutdown   ! Does not show in config
!
interface FastEthernet 0/0.10
 description VLAN 10
 encapsulation dot1Q 10
 ip address 172.16.10.1 255.255.255.0
```

Routed Ports on a L3 Switch

- Routed ports obtained with: `no switchport`
  Then be sure to enable routing: `ip routing`
  The port then behaves *exactly* like on a router
- Routing on a L3 switch is *no different* than on routers (manufactured in the last ~15-20 years): both forward packets in hardware at full wirespeed
- Generally can *not* configure sub-interfaces on L3 routed ports but this may depend on make & model
- A general difference between Cisco "routers" and "L3 switches" is that L3 switches have a higher port density consisting exclusively of Ethernet interfaces, making them better for aggregation

L3 LAGs

- While we're on the topic of L3 connectivity, let's confirm that LAGs can be created on L3 ports:
  - first create the port-channel logical interface
  - then add the ports into the port-channel
  - an IP address is assigned to the port-channel, not the individual links
  - if configuring on a router, omit "no switchport"

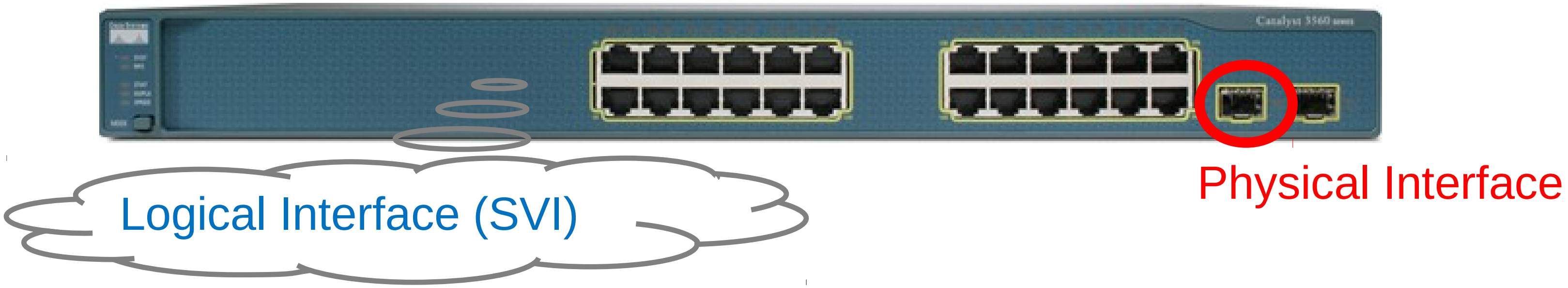
```
Sw(config)# interface port-channel 1
Sw(config-if)# no switchport                !for L3 LAG
Sw(config-if)# ip address 10.1.2.1 255.255.255.0
Sw(config)# interface range fa 0/1-4
Sw(config-if-range)# no switchport          !Must match
Sw(config-if-range)# no ip address          !Optional
Sw(config-if-range)# channel-group 1 mode active
```

If not matched: *Command rejected (Port-channel1, Fa0/1): Either port is L2 and port-channel is L3, or vice-versa*

SVI – Basics

- You used a Switch Virtual Interface in NET1002

	① VLAN Config	② SVI Config
1	VLAN 100 name DemoVLAN exit	interface VLAN100 3 ip address 1.2.3.4 255.0.0.0 4 no shutdown



- An SVI has four prerequisites to be active & reachable:

1 – the associated VLAN *must* exist

2 – there *must* be at least 1 active port/trunk in VLAN
  – can change with "autostate exclude" or "no autostate"

3 – there *must* be an address configured

4 – the SVI *must* be no-shut

SVI Advantages

- Gateway IPs for a far greater number of subnets than L3-switch routed ports
- Far greater bandwidth than external router or L3 switch routed ports (forwarding done internally)
- Less potential for congestion than external router (traffic doesn't flow both ways on same link!)
- Potential for greater bandwidth: more physical ports to create LAGs
- Reduced latency compared to external router (traffic doesn't need flow out and back)

Routing with SVIs

- To route, each VLAN needs an associated SVI
  - by default, 3560s can have multiple SVIs
  - by default, 2960s have a *single* SVI (no routing!)
    - intended for a single management interface
- Switch capabilities changed using Switch Database Manager (**SDM**), followed by a reboot
- 2960 was originally a "pure L2" access switch:
  - can route between its maximum of 8 SVIs
  - can have multiple static routes (may be <= 16?)
  - can *not* run any routing protocols
- Remember, you *must* always enable routing!

```
Sw(config)# ip routing
```

SDM Settings – Config & Verification (1)

```
2960Sw# show sdm prefer                   !IOS 12.2(55)+later
 The current template is "default" template.
 [...output omitted...]
 0 routed interfaces and 255 VLANs.
  number of unicast mac addresses:                 8K
  number of IPv4 IGMP groups:                      0.25K
  number of IPv4/MAC qos aces:                     0.125k
  number of IPv4/MAC security aces:                0.375k
2960Sw(config)# sdm prefer lanbase-routing   !now reboot
```

```
2960Sw# show sdm prefer                     ! IOS 12.2(55)
 The current template is "lanbase-routing" template.
 [...output omitted...]
 8 routed interfaces and 255 VLANs.
  number of unicast mac addresses:                 4K
  number of IPv4 IGMP groups + multicast routes:   0.25K
  number of IPv4 unicast routes:                   4.25K
    number of directly-connected IPv4 hosts:       4K
    number of indirect IPv4 routes:                0.25K
```

Possibly not correct value

SDM Settings – Config & Verification (2)

- Beware IOS 15, which gives *incorrect* info for routed interfaces but *correct* info for max routes

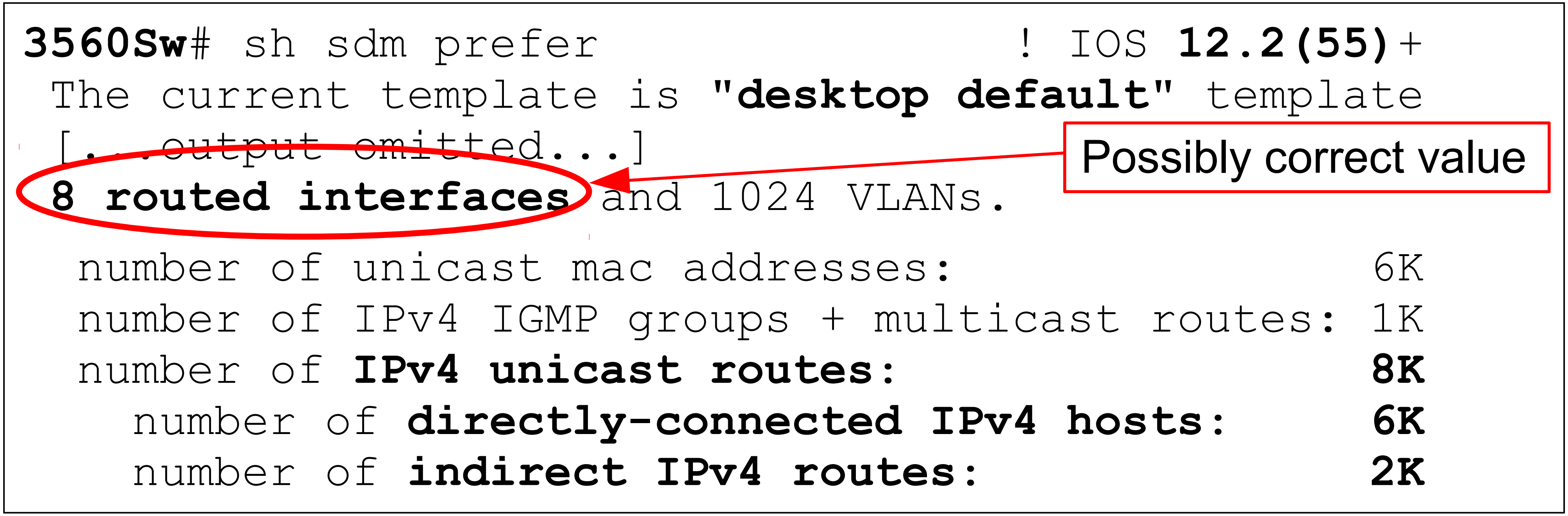
```
2960Sw# show sdm prefer                   ! IOS 15.0(2)
 The current template is "lanbase-routing" template.
 [...output omitted...]
 0 routed interfaces and 255 VLANs

   number of unicast mac addresses:                  4K
   number of IPv4 IGMP groups + multicast routes:    0.25K
   number of IPv4 unicast routes:                    4.25K
     number of directly-connected IPv4 hosts:        4K
     number of indirect IPv4 routes:                 256
   number of IPv6 multicast groups:                  0.375k
   number of IPv6 unicast routes:                    1.25K
     number of directly-connected IPv6 addresses:    0.75K
     number of indirect IPv6 unicast routes:         448
 [...output omitted...]
```

Definitely **not** correct value

SDM Settings – Config & Verification (3)

- Compare with 3560, natively a L3 switch



```
3560Sw# sh sdm prefer                     ! IOS 12.2(55)+
 The current template is "desktop default" template
 [...output omitted...]
 8 routed interfaces and 1024 VLANs.

   number of unicast mac addresses:                  6K
   number of IPv4 IGMP groups + multicast routes:    1K
   number of IPv4 unicast routes:                    8K
     number of directly-connected IPv4 hosts:        6K
     number of indirect IPv4 routes:                 2K
```

Possibly correct value

- Template availability depends entirely on IOS version:
  2960: default, dual-ipv4-and-ipv6, lanbase-routing, qos
  3560: access, default, dual-ipv4-and-ipv6, routing, vlan
- Identically named templates may have different definitions under different IOS versions

SDM Settings – Config & Verification (4)

- Some templates may disable routing on a 3560:

```
3560Sw(config)# show sdm prefer                  ! IOS 12.2(55)+
 The current template is "desktop vlan" template.
 [...output omitted...]
 8 routed interfaces and 1024 VLANs

  number of unicast mac addresses:                 12K
  number of IPv4 IGMP groups + multicast routes:   1K
  number of IPv4 unicast routes:                   0
  number of IPv4 policy based routing aces:        0
  number of IPv4/MAC qos aces:                     0.5K
  number of IPv4/MAC security aces:                1K

3560Sw(config)# ip routing
*Mar  1 00:01:52.415: %PLATFORM_UCAST-3-INV_SDM:
  IP routing configured with invalid SDM template
```

Definitely **incorrect** value!

For more info, see:

https://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst2960/software/release/12-2_55_se/configuration/guide/scg_2960/swsdm.html

https://learningnetwork.cisco.com/docs/DOC-27403

SDM Settings – Best Practices

- BEWARE that exact specs & capabilities may depend on (I)OS version!
  2960 12.2(55)SE: either routing or IPv6 at one time
  2960 15.0(2): can get routing & IPv6 simultaneously
- Coverage of SDM in the FLG textbook (pp. 364-366) may be confusing:
  - "*the default SDM is configured for optimal use of all features simultaneously*"
  - but "*IPv6 functionality is not supported with the default template*"
  - and "*By sacrificing unnecessary features in your network such as IPv6...*"
  - "*It is a best practice to change the SDM template only if you have a good reason to do so*"
  - but "*for access switches, you should only change the template for IPv6 usage unless directed by a Cisco advisor or architect*"
- "*Another common reason for changing the SDM template is because you are running out of a specific resource. … it is important to first investigate whether you can optimize the performance so that you do not need to change the SDM template.*"
- "*… always investigate the amount of systems resources being used prior to considering changes … To verify how much of the system resources are being used, use the command* **`show platform tcam utilization`**"

Inter-VLAN Routing

- Once a gateway IP is available for every VLAN, routing configuration, verification, and troubleshooting proceeds as normal
- This course does *not* add to or use anything beyond the routing covered in NET3008;
  see p. 214-220 of textbook for routing coverage
- For any routing configured on switches, use standard commands:
  – for 2960, commands for static routing only
  – for 3560, commands for static and IGP protocols

Questions??

Common Inter-VLAN Routing Problems

Problem	Possible Cause
Incorrect VLAN configuration	VLAN might not be defined across all the switches. VLAN might not be enabled on the trunk ports. Ports might not be configured in the proper VLANs
Layer 3 interface configuration	Virtual interface might have wrong IP addr or subnet mask Virtual interface might not be up Virtual interface # might not match the VLAN # (ID10T error) Routing must be enabled to route frames through routed ports or between internal VLANs via SVIs; check that it's enabled
Routing protocol misconfiguration	The interface might not be participating in the routing protocol Every interfaces or network needs to be in the routing protocol Routing (i.e. L3 reachability) must be configured if VLANs on this device must communicate with VLANs on other devices
Host misconfiguration	Host might not have the right IP or subnet mask Host might not be configured with the correct default gateway Each host has to have default gateway that is the SVI or Layer 3 interface used to communicate with other networks (VLANs)

From: FLG p. 223

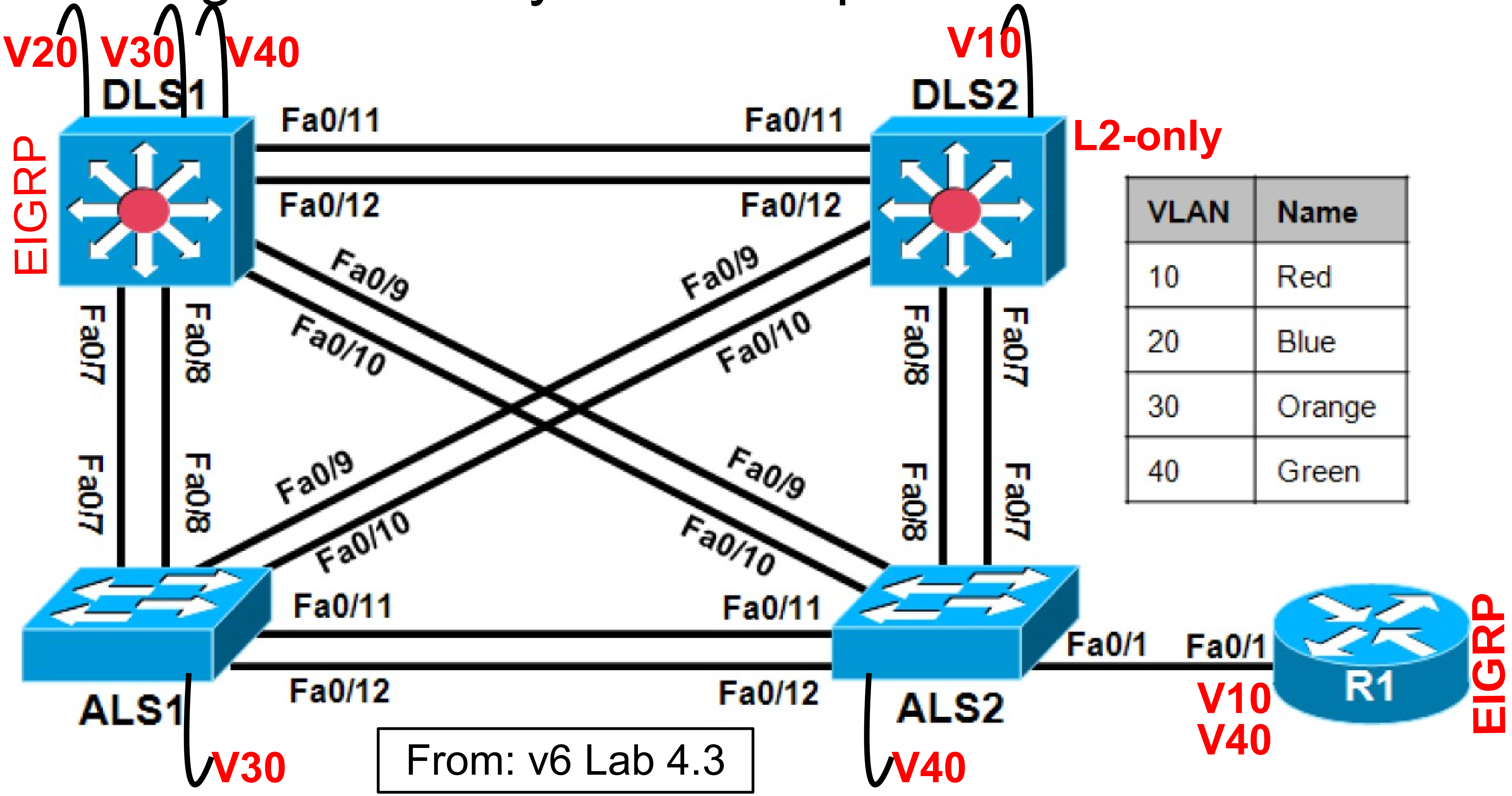
Frame Processing by a Router (reminder)

This is "Version 1.0"

1. Check CRC (if bad, drop, done); check dest MAC addr (if not me, drop, done);
   strip L2 framing
2. Check for VLAN tags and strip them (there may possibly be some processing)
   L3 Packet modified / re-written
3. Check IP header details: verify checksum and if okay then check destination IP
   if dest IP is me, punt up to Control Plane, done; else continue forwarding!
4. Decrement TTL count by 1; if now 0, punt up to Control Plane for ICMP msg
   L3 Packet modified / re-written
5. Recalculate Checksum in IP header
   L3 Packet modified / re-written
6. Check routing table for best longest(!) match to destination IP
7. Determine encap required for egress interface: Dest MAC, Src MAC, (VLAN ?)
   L2 frame modified / re-written
8. Compute CRC and add to egress frame
   L2 frame modified / re-written

L2 vs L3 Topology

- Remember: L3 is *not* necessarily the same as L2! e.g. How many routed hops between VLANs?



VLAN	Name
10	Red
20	Blue
30	Orange
40	Green

From: v6 Lab 4.3

DHCP - Basics

- Two options for configuring network settings are static (i.e. manual) and automatic / dynamic
- Almost every person uses DHCP daily (typically IPv4)
- Since DHCP allows for automatic network config, it is used almost exclusively for managing clients
- DHCPv4 provides all required network settings, and thus hides the fact that DHCP does far more than just assign a client address
- DHCP may use a combination of broadcast and unicast (IPv4), or multicast (IPv6)
- In IPv6, "auto-config" has three modes, incl DHCP
- DHCP may be distributed (localized) or centralized;
  if centralized, DHCP "local" **broadcasts will cross a router**!

DHCP – Settings Provided

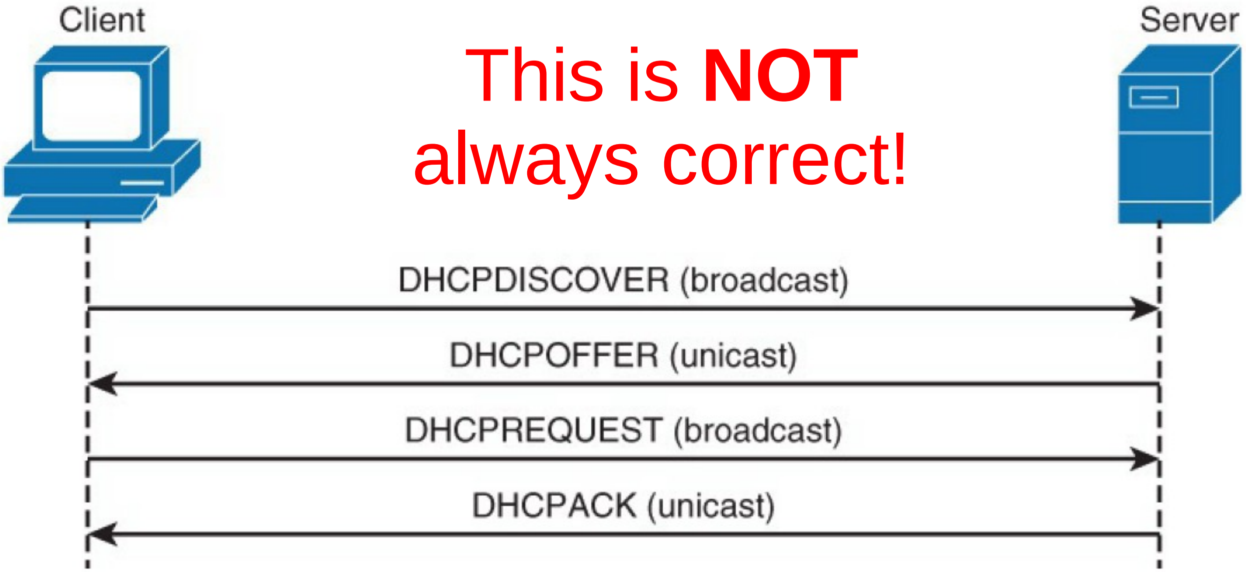
- DHCP can provide: address, mask, gateway, DNS
- It can also provide:
  - domain name
  - lease time
  - SMTP server addrs
  - POP server addrs
  - TFTP server addrs
  - MTU
  - ARP cache timeout
  - NTP server addrs
  - Wireless AP WLC addrs

… and many others as described in several RFCs (IPv4 = RFC 2132, IPv6 = RFC 3315)

Ref: https://www.cisco.com/c/en/us/td/docs/net_mgmt/network_registrar/7-1/user/guide/cnr71book/UGB_Opts.html

DHCP – Characteristics (IPv4)

- Uses the full DORA sequence for new addresses, or just Req/Ack for renewal of addresses
- Req/Ack steps required because of (possible) redundancy, to select desired Offer
- Uses UDP ports 67 (server) and 68 (client)
- Usage of bcast/ unicast depends on NIC capability



This is **NOT** always correct!

Figure 5-16 *DHCP Discovery Process*

DHCP – Configuration, Local (IPv4)

- Simply configure a pool for a subnet; automatically associated with the relevant interface
- Don't double-dip: reserve GW & server addresses

```
Srv(config)#ip dhcp pool T108  ①
Srv(config-dhcp)#network 172.16.1.0 255.255.255.0  ②
Srv(config-dhcp)#default-router 172.16.1.1  ③
Srv(config-dhcp)#dns-server 172.16.1.2  ④
Srv(config-dhcp)#netbios-name-server 172.16.1.2
Srv(config-dhcp)#domain-name t108.linux.com
Srv(config-dhcp)#lease infinite
Srv(config-dhcp)#exit

! Don't forget to ensure GW & svr addrs are unique!  ⑤
Srv(config)#ip dhcp excluded-address 172.16.1.1 172.16.1.5
```

DHCP – Configuration, Chained (IPv4)

- Settings (other than host address) can be chained

```
Main(config)#ip dhcp pool Central-Server-Pool
Main(config-dhcp)#network 10.0.0.0 255.255.255.0
Main(config-dhcp)#default-router 10.0.0.1 10.0.0.2
Main(config-dhcp)#dns-server 10.0.0.2
Main(config-dhcp)#domain-name central.com
Main(config-dhcp)#exit
Main(config)#ip dhcp excluded-address 10.0.0.1 10.0.0.5
   ! Main fa0/1 (10.0.0.1) connects to Remote fa0/0
```

```
Remote(config)#ip dhcp pool Distributed-Remote-Pool
Remote(config-dhcp)#network 20.0.0.0 255.255.255.0
Remote(config-dhcp)#default-router 20.0.0.1
Remote(config-dhcp)#import all
Remote(config-dhcp)#exit
Remote(config)#ip dhcp excluded-addr 20.0.0.1 20.0.0.5
Remote(config)#interface fa0/0
Remote(config-if)#ip address dhcp
```

DHCP – Verification (IPv4)

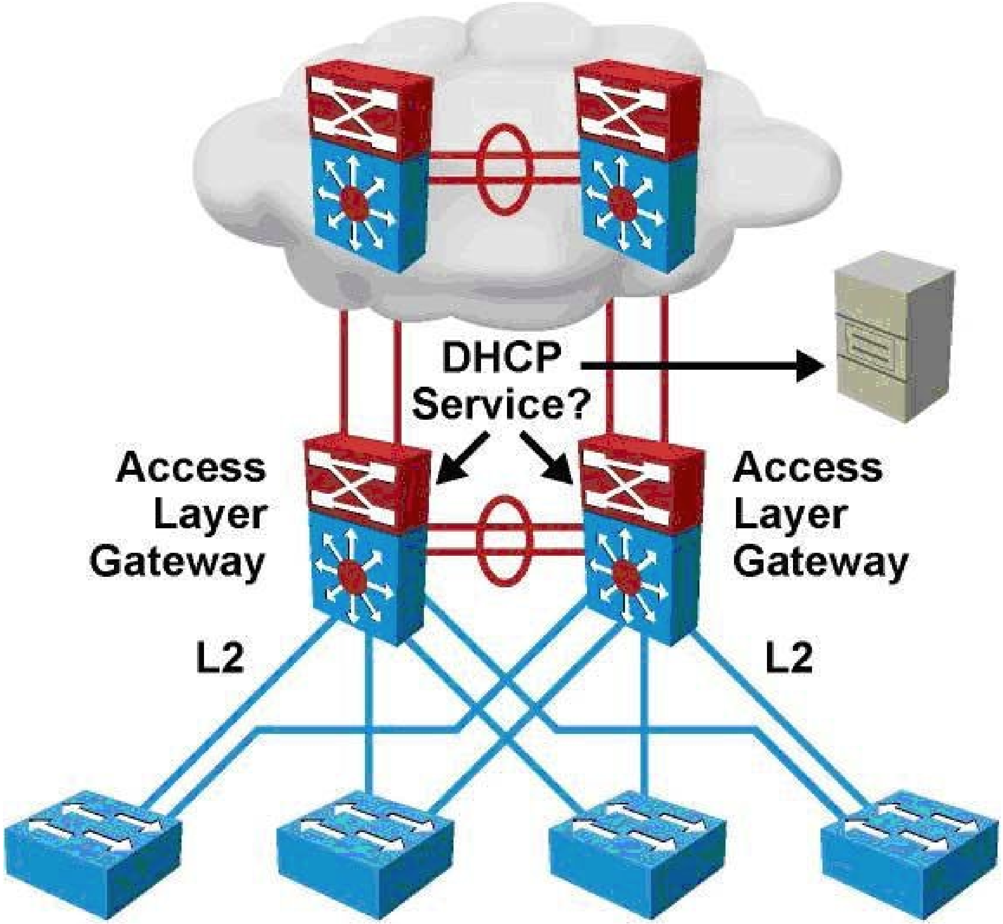
```
Srv# show ip dhcp binding
Srv# show ip dhcp conflict
Srv# show ip dhcp import
Srv# show ip dhcp database
```

```
Switch# show ip dhcp binding
Bindings from all pools not associated with VRF:
IP address  client-ID/           Lease expiration        Type
            Hardware address/
            User name
10.1.10.21  0100.1bd5.132a.d2    Jun 25 2015 06:09 AM    Automatic
10.1.10.22  0100.4096.a46a.90    Jun 25 2015 09:40 AM    Automatic

Switch# debug ip dhcp server packet
DHCPD: DHCPDISCOVER received from client 0100.1bd5.132a.d2 on
       interface Vlan6
DHCPD: Sending DHCPOFFER to client 0100.1bd5.132a.d2 (10.1.10.21)
DHCPD: broadcasting BOOTREPLY to client 001b.d513.2ad2
DHCPD: DHCPREQUEST received from client 0100.1bd5.132a.d2
DHCPD: Sending DHCPACK to client 0100.1bd5.132a.d2 (10.1.10.21)
DHCPD: broadcasting BOOTREPLY to client 001b.d513.2ad2
```

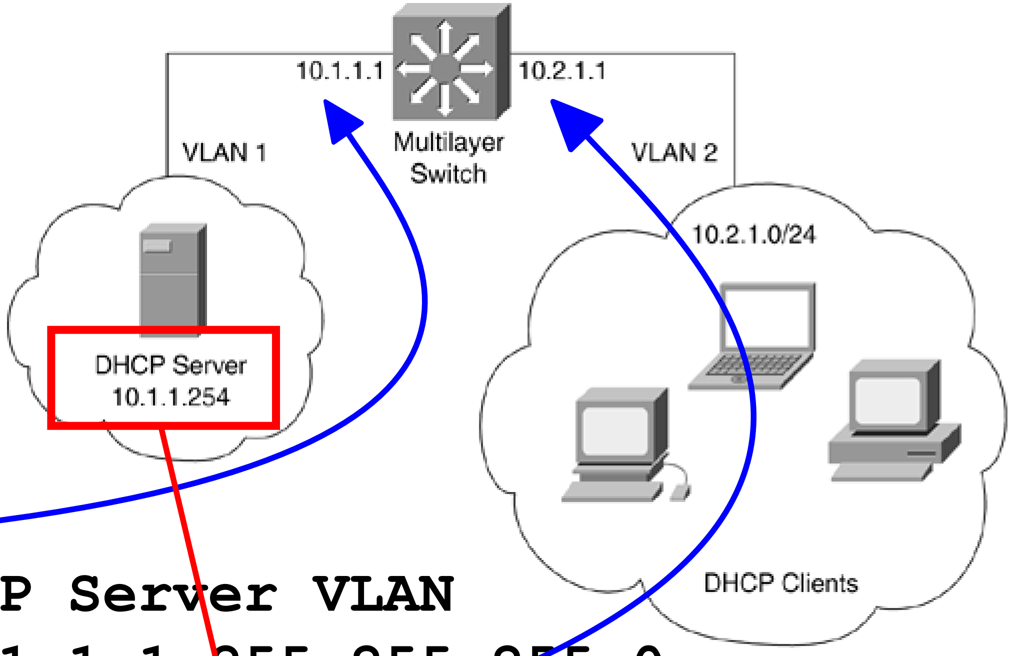
DHCP – Centralized: Basics

- Distribution layer devices often acts as L3 gateways, so it's convenient if they also provides DHCP.
- A distributed model results in higher management overhead.
- To simplify management, DHCP can be centralized to a dedicated DHCP server. This requires intermediate devices to redirect incoming DHCP requests to the central server.



DHCP – Centralized: Relay Agent (1)

- What is a broadcast domain? We've always said local broadcasts must **not** cross L3 boundaries!
- This means one DHCP server per subnet?!
  - a Relay Agent can bypass this need

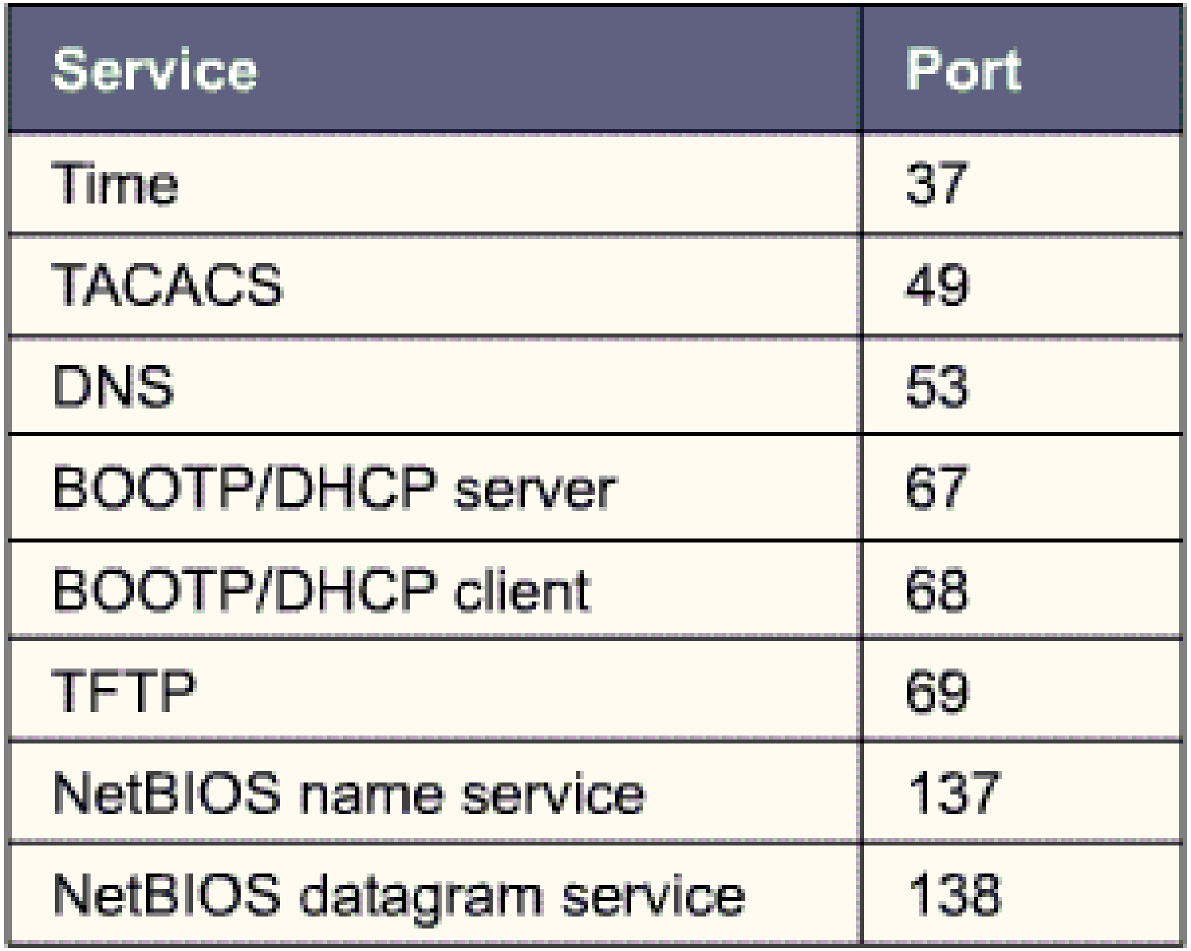


```
MLS(config)#interface vlan 1
MLS(configif)#description DHCP Server VLAN
MLS(config-if)#ip address 10.1.1.1 255.255.255.0

MLS(config)#interface vlan 2
MLS(config-ig)#description DHCP clients
MLS(config-if)#ip address 10.2.1.1 255.255.255.0
MLS(config-if)#ip helper-address 10.1.1.254
```

DHCP – Centralized: Relay Agent (2)

- Cisco devices configured as a Relay Agent will forward UDP broadcast for **any** of the following ports (they forward more than just DHCP traffic!)
- DHCP messages between server and relay agent become unicast
- RFC 1542 defines the specifications for Relay Agents



Service	Port
Time	37
TACACS	49
DNS	53
BOOTP/DHCP server	67
BOOTP/DHCP client	68
TFTP	69
NetBIOS name service	137
NetBIOS datagram service	138

DHCP – Centralized: Relay Agent (3)

- You can block, or add, the forwarding of individual UDP broadcasts with additional configuration:
  - block broadcast forwarding to TFTP server:

```
MLS(config)#no ip forward-protocol udp 69
```

  - include broadcasts destined for an NTP server:

```
MLS(config)#ip forward-protocol udp 123
```

- If redundant operation is required, you can configure multiple **`ip helper-address`** lines per interface
- *Don’t forget routing*: the central server(s) and the relay agent must have mutual IP reachability i.e. not only must server address(es) be routable from the relay agent, but the server(s) must also have a return route back to the relay agent.

IPv6 Network Addresses – Basics

- You should have covered IPv6 addressing basics in previous courses; if not, please work through the IPv6 primer posted on the course site;
  you will *need* to understand IPv6 address formats!
- With the larger number of bits available, IPv6 addressing allows for much *automatic* addressing
- In general terms, IPv6 addresses typically have a /64 mask, with the host bits often determined by the host itself (either randomly or by EUI-64)
- ICMP for IPv6 adds a set of messages, called Neighbor Discovery Protocol (**NDP** or ND), to assist with addressing; there is a strong interplay between ICMPv6 and DHCPv6

IPv6 Network Addresses – ICMPv6 (1)

Four key messages enable 10+ address functions

- **RS**: Router Solicitation (type=133; request)
  – the most relevant of the four msgs to DHCP
- **RA**: Router Advertisment
  (type=134; periodic or in response to RS)
- **NS**: Neighbour Solicitation (type=135; request)
- **NA**: Neighbour Advertisment (type=136; response)

IPv6 routers *always* generate RA messages;
L3 switches *only* generate RA when routing!
Must be configured: `Sw(config)#` **`ipv6 unicast-routing`**

IPv6 Network Addresses – ICMPv6 (2)

The four ICMPv6 messages enable 10+ functions for NDP:

- **Prefix Discovery** – Discover the prefix or prefixes assigned to that link
- **Router Discovery** – Discover the local routers without DHCP
- **Next-Hop Determination** – Determine the link-layer next hop for a destination, either as a local destination or a router to the destination
- **Address Autoconfiguration** – Determine its full address, without DHCP
- **Parameter Discovery** – Discover other params such as link MTU, hop limits for link
- **Address Resolution** – Discover link-layer addrs of other nodes on link without ARP
- **Duplicate Address Detection** – Similar to gratuitous ARP. Determine if an address it wants to use is already being used by another node on the link
- (Entire) **Subnet renumbering** (reassignment) – Swap subnet prefix for all devices on the subnet by reconfiguring only the router(s)
- **Neighbor Unreachability Detection** – Determine when a neighbor on a link, either another host or a router, is no longer reachable
- **Redirect** – Router notifies a host of a better next-hop than itself to an off-link destination. Part of basic ICMPv4 functionality, but redefined as part of NDP in IPv6

IPv6 Network Addresses – RA (1)

- The RA message is sent by all routers, with the link address used as the source address
- Flags field in RA msg has three important options:
  - **A** flag: host **A**utomatically chooses it's own addr & params based on subnet, mask, GW in RA msg
    – SLAAC = **S**tate**L**ess **A**ddress **A**uto**C**onfiguration
    – sometimes called Stateless DHCP (why?)
  - **O** flag: host obtains **O**ther info [e.g. dns, MTU, domain name, etc] from router
  - **M** flag: host's address is chosen and **M**anaged by the DHCP server (… just like IPv4)
    – also called Stateful DHCP

IPv6 Network Addresses – RA (2)

- First, consider IPv4: DHCP dictates *all* addr settings with no auto-addressing, so a DHCP pool can *only* work on a single interface, so no need to explicitly associate a DHCP pool with an interface
- Given the existence of RA msgs (and the **A** flag), SLAAC is potentially available on *every* router segment, with no distinctive feature.
- Given the existence of SLAAC plus **O**ther info, stateless DHCPv6 is potentially available on *every* router segment, with no distinctive feature.
- … So DHCPv6 pools must be explicitly associated with relevant interfaces

RA – Config and Verification (1)

- RA msg control is under the **ipv6 nd** context

```
! Enable RA msgs globally
L3Sw(config)# ipv6 unicast-routing

! Disable RA msgs entirely (per interface)
L3Sw(config-if)# ipv6 nd ra suppress

! Turn OFF the A flag
L3Sw(config-if)# [no] ipv6 nd autoconfig prefix      !IOS 12.2(55)+
L3Sw(config-if)# ipv6 nd prefix {ipv6-addr} no-autoconfig !15.0(2)

! Turn ON the O flag
L3Sw(config-if)# ipv6 nd other-config-flag

! Turn ON the M flag
L3Sw(config-if)# ipv6 nd managed-config-flag

L3Sw(config-if)# ! NO need for excluding certain host addresses!
```

The no-autoconfig parameter is **hidden**, so it will not appear in help or tab-completion!

- With RA msgs + **A** flag, SLAAC is running!

DHCPv6 – Stateless Configuration

- Stateless DHCP

```
L3Sw(config)# ! NO need for excluding certain host addresses!

L3Sw(config)# ipv6 dhcp pool MANAGEMENT_IPV6_DHCP
L3Sw(config-dhcpv6)# dns-server 2001:db8:3115:99::100
L3Sw(config-dhcpv6)# ! NO subnet address configured!!
L3Sw(config-dhcpv6)# exit

L3Sw(config)# ipv6 unicast-routing

L3Sw(config)# int vlan 100  ! Unlike IPv4, must associate with i/f
L3Sw(config-if)# ipv6 dhcp server MANAGEMENT_IPV6_DHCP
L3Sw(config-if)# ipv6 nd other-config-flag  ! Other info than Addr
L3Sw(config-if)# exit
```

DHCPv6 – Stateful Configuration

- Stateful DHCP

```
L3Sw(config)# ! NO need for excluding certain host addresses!
L3Sw(config)# ipv6 dhcp pool VLAN120-IPV6-POOL
L3Sw(config-dhcpv6)# address prefix 2001:db8:3115:120::/64 ! ADDR!
L3Sw(config-dhcpv6)# dns-server 2001:db8:3115:99::100
L3Sw(config-dhcpv6)# domain-name switch.ccnp
L3Sw(config-dhcpv6)# exit

L3Sw(config)# ipv6 unicast-routing

L3Sw(config)# int vlan 100  ! Unlike IPv4, must associate with i/f
L3Sw(config-if)# ipv6 dhcp server VLAN120-IPV6-POOL
L3Sw(config-if)# ipv6 nd managed-config-flag
L3Sw(config-if)# ipv6 nd prefix {addr} no-autoconfig      !optional
L3Sw(config-if)# exit
```

- Disabling the **A** flag is an administrative choice, but it makes sense in this context

DHCPv6 – Relay Agent

- DHCPv6 relay agent works the same as in IPv4
- Configuration is similar but not identical to IPv4:

```
L3Sw(config-if)# ipv6 dhcp relay destination {addr}

!IPv4 for comparison:
L3Sw(config-if)# ip helper-address {addr}
```

DHCPv6 – Verification

- Commands are almost identical to IPv4:

```
Srv# show ipv6 dhcp binding
Srv# show ipv6 dhcp conflict
Srv# show ipv6 dhcp database

Srv# debug ipv6 dhcp detail
Srv# debug ipv6 nd
```

Reminder

- *LOTS* of details do not appear in these slides
- You are responsible for *reading* the textbook to gain the knowledge (memorization) and understanding (apply the knowledge)
- For IPv6, you'll need to read the FLG textbook, the v7 Labs (esp 5.2), and possibly consult the recommended text "CCNP Routing and Switching SWITCH 300-115 Official Cert Guide"