Traceroute / Ping

Friday, December 11, 2015
8:49 PM

Traceroute

- Traceroute has default TTL of 30.
- Steps performed when traceroute is executed:
  - Sends 3 packets with TTL=1 to first-hop router. FH router responds with time-exceeded (ICMP Type-11).
  - In response sends 3 packets with TTL=2 to FH router, second-hop router responds with TTL message.
  - Continues until packets arrive at destination, last-hop router responds with unreachable (ICMP Type-3).
  - The LH router sends back a unreachable message because the destination is an unreachable port.

Traceroute Output

The * means that ICMP rate limit is enabled at the last-hop router. The default timeout is 500 msec.

- The reason only the LH router shows this is because intermediate routers send a time exceeded TTL message.
- The second traceroute packet usually times out because that one is within the 500 msec interval, the third packet is not.
- The same applies to ping with U.U.U output, the 1st message is sent back as unreachable by the LH router. The 2nd times out because it is within the 500 msec interval, the 3rd is unreachable again and so on...

ip icmp rate-limit unreachable 500

show ip icmp rate-limit

Traceroute Responses

*	The probe timed out
A	Administratively prohibited (ACL)
U	Port unreachable
H	Host unreachable
N	Network unreachable

Ping

- Time exceeded on ping means TTL expired (ICMP Type-11). This is used in traceroute.
- Specify how often ICMP unreachable messages are sent to neighbors with the ip icmp rate-limit unreachable command.
- ICMP redirect messages are used to notify hosts that a better route (other router) is available for a particular destination.
- The kernel is configured to send redirects by default. Disable with the interface command no ip redirects.

Cisco routers send ICMP redirects when all of these conditions are met:

- The ingress interface is the same as the egress interface of the packet.
- The source is on the same subnet as the better next-hop.
- The source does not use source-routing.

ICMP types

0	Echo Reply
3	Destination Unreachable
5	Redirect
8	Echo

11	Time Exceeded (TTL)

ICMP Responses

!	Reply
.	Timed Out
U	Destination Unreachable

AAA

Authentication, Authorization and Accounting (AAA)

- Place local login before group authentication so that the specified usernames are authenticated first.

```
aaa authentication login default local-case group radius
aaa authentication login default local group tacacs+
```

AAA Auto-Command

- Automatically logout particular users (PPP usernames for example) and prevent them from managing the router.
- The autocommand function will only work if authorization is configured using AAA.
- The console is treated differently by default, and requires additional commands in order to automatically logout users.
- The aaa authorization console applies AAA rules to the console as well.

```
username R1 autocommand logout
aaa authorization exec default if-authenticated

aaa authorization console
line con 0
 authorization exec default
```

Lines / SSH

Wednesday, December 2, 2015
7:09 PM

Login

- Log both successful and failed login attempts.
- 3-second delay between successive login attempts.
- Do not allow login attempts for 10 seconds if two tries fail within 15 seconds.
- During login block allow hosts in ACL.

```
login on-failure log
login on-success log
login delay 3
login block-for 10 attempts 2 within 15
login quiet-mode access-class QUIET

ip access-list standard QUIET
 permit host 192.168.0.1
```

Secure Shell (SSH)

- Enable SSH without ip domain-name by using the label keyword.
- Normally the first generated RSA key is linked to SSH. Override this with the keypair-name command.

```
crypto key generate rsa modulus 768 label R1.lab.local
ip ssh version 2
ip ssh rsa keypair-name R1.lab.local
username admin privilege 15 password cisco
line vty 0 4
 login local

show crypto key mypubkey rsa
ssh -v 2 -l admin 192.168.0.1
```

Telnet

- Default ToS is 192 (C0). Change the Telnet ToS with the ip telnet tos command.

Hide IP / hostname information when establishing Telnet sessions:

```
ip telnet hidden addresses
ip telnet hidden hostnames
```

Or:

```
service hide-telnet-addresses
```

Show the line connected to when Telnet establishes:

```
service linenumber

show users
show line
```

Privilege

Wednesday, December 2, 2015
8:32 AM

Privilege Access Control

- The all keyword specifies the configuration and all underlying sub-configurations.
- This will give access to all interfaces and all configurations under the interfaces.
- Every level has access to all the commands available underneath it. 8 has access to privilege 1-7 and 8
- In IOS there are 3 default privilege levels:
    - Privilege 0.
    - Privilege 1. User exec.
    - Privilege 15. Privilege exec.

Create custom commands for level 10:
privilege exec level 10 [commands]
privilege configure level 10 [commands]
privilege interface level 10 [commands]

Radius

Radius

- By default all defined radius servers will be used if you configure aaa...default group radius.
- Servers are consulted in the order in which they were configured.
- Either set default key, timeout and retransmit settings, or per host.
- Host specific setting will override the default. Default timeout is 5 sec, default retransmit amount is 3 times.
- Default radius ports are 1645 for authentication and 1646 for accounting (newer ports are 1812 and 1813).

```
aaa new-model
radius-server host 1.1.1.1 auth-port 1645 acct-port 1646 timeout 5 retransmit 3 key cisco

radius server RADIUS
 address ipv6 1::1 auth-port 1812 acct-port 1813
 timeout 3
 retransmit 0
 key cisco

aaa authentication login default group radius local
aaa authorization exec default group radius local

line vty 0 4
 login authentication default
 authorization exec default

show radius server-group all
debug radius
```

AAA Server Groups

- Group servers to use for a single purpose. For example a group used only for PPP authentication.
- Grouped private servers do not have to exist in the global config. And are used for a single purpose by a single AAA server group.
- Grouped public servers have to exist in the global config.
    - Public servers configured with the tacacs server command, have to be linked with the server name command.
    - Public servers configured with the tacacs-server host command, have to be linked with the server command.

```
aaa new-model
radius-server host 1.1.1.1 auth-port 1645 acct-port 1646 timeout 5 retransmit 3 key cisco

radius server RADIUS
 address ipv6 1::1 auth-port 1812 acct-port 1813
 timeout 3
 retransmit 0
 key cisco

aaa group server radius MYRADIUS
 server name RADIUS
 server 1.1.1.1
 server-private 2.2.2.2 timeout 5 retransmit 0 key cisco

aaa authentication login default group MYRADIUS local
aaa authorization exec default group MYRADIUS local

show aaa servers
```

Radius Change of Authorization (CoA)

- The CoA feature provides a mechanism to change the attributes of a AAA session after it is authenticated.
- The change is initiated on the radius server and 'pushed' to the router.
- The default port for packet of disconnect is 1700, ACS uses 3379.
- The client is the radius server that sends the CoA request.
- Requests allow for:
    - Session identification
    - Host re-authentication
    - Session termination

```
aaa new-model
aaa server radius dynamic-author
 client 1.1.1.1 server-key cisco
 port 3799
 auth-type all

show aaa clients
```

RBAC

Wednesday, December 2, 2015
8:31 AM

Role-Based Access Control (Parser Views)

- Requires AAA enabled.
- AAA authorization is recommended.
- Requires enable 'root' password configured.
- A superview combines two more more parser views.

```
enable secret cisco
enable view root

username bpin view VIEW privilege 15 password cisco

aaa new-model
aaa authentication login default local
aaa authorization exec default local

parser view VIEW
 secret cisco
 commands exec include configure terminal
 commands configure ...

show parser view
enable view VIEW

parser view SUPERVIEW superview
 secret cisco
 view VIEW1
 view VIEW2
```

Tacacs+

Wednesday, December 2, 2015
8:32 AM

Tacacs+

- The tacacs-server host command will be replaced by tacacs server. The old-style command has no option for IPv6.
- This new form will only show up after aaa new-model has been configured.
- The single-connection option will reuse the TCP session. This is more efficient because the tcp session doesn't have to be rebuild.
- Either set default key and timeout settings, or per host. Host specific setting will override the default. Default timeout is 5 sec.
- By default all defined tacacs+ servers will be used if you configure aaa...default group tacacs+.
- Servers are consulted in the order in which they were configured.

```
aaa new-model
tacacs-server host 1.1.1.1 single-connection port 49 timeout 5 key cisco

tacacs server TACACS
 address ipv6 1::1
 key cisco
 timeout 5
 single-connection

aaa authentication login default group tacacs+ local
aaa authorization exec default group tacacs+ local
aaa accounting commands 15 default start-stop group tacacs+

line vty 0 4
 login authentication default
 authorization exec default

show tacacs
debug tacacs
debug tacacs events
debug tacacs packets
```

AAA Server Groups

- Group servers to use for a single purpose. For example a group used only for PPP authentication.
- Grouped private servers do not have to exist in the global config. And are used for a single purpose by a single AAA server group.
- Grouped public servers have to exist in the global config.
    - Public servers configured with the tacacs server command, have to be linked with the server name command.
    - Public servers configured with the tacacs-server host command, have to be linked with the server command.

```
tacacs-server host 1.1.1.1 single-connection port 49 timeout 5 key cisco
tacacs server TACACS
 address ipv6 1::1
 key cisco
 timeout 5
 single-connection

aaa group server tacacs+ MYTACACS
 server name TACACS
 server 1.1.1.1
 server-private 2.2.2.2 single-connection port 49 timeout 5 key cisco
```

```
aaa authentication login default group MYTACACS local
aaa authorization exec default group MYTACACS local
aaa accounting commands 15 default start-stop group MYTACACS

show aaa servers
```

Access-Lists

Sunday, December 13, 2015
10:23 PM

Compiled Access-Lists

- The turbo ACL feature is designed in order to process ACLs more efficiently.
- Applied on all access-lists.

```
access-list compiled
show access-list compiled
```

VTY Access-Lists

- Outbound ACL on interface only filters transit traffic, not locally generated traffic.
- Outbound ACL on VTY lines only applies to connections generated from the existing VTY session.
- Inbound ACL on VTY is applied to outside connections coming into the router.

```
ip access-list standard VTY_IN
 10 permit host 192.168.0.1
 99 deny any log

ip access-list standard VTY_OUT
 10 permit host 192.168.0.2
 99 deny any log

ip access-list logging interval 1
ip access-list log-update threshold 1

line vty 0 4
 access-class VTY_OUT out
 access-class VTY_IN in
```

Dynamic

Sunday, December 13, 2015
10:23 PM

Dynamic Access-Lists (Lock-and-Key)

- Blocks traffic until users telnet into the router and are authenticated.
- A single time-based dynamic entry is added to the existing ACL.
- Idle timeouts are configured with the autocommand. Absolute timeouts are configured in the ACL.
- The absolute timeout value must be greater than the idle timeout value, if using both.
- If using none, the access-list entry will remain indefinitely.
- It is also possible to set the autocommand access-enable host timeout directly on the VTY line.
- Absolute timers can be extended by 6 minutes using the access-list dynamic-extended command. Requires re-authentication.

```
username bpin password cisco
username bpin autocommand access-enable host timeout 4

ip access-list extended DYNAMIC
 10 permit ospf any any
 20 permit tcp host 10.0.12.1 host 10.0.12.2 eq telnet
 30 dynamic ICMP timeout 8 permit icmp host 10.0.12.1 any
 99 deny ip any any log-input

 access-list dynamic-extended

int fa0/0
 ip access-group DYNAMIC in
```

IPv6

Sunday, December 13, 2015
10:23 PM

IPv6 Access-Lists

- By default IPv6 access-lists add (hidden) permit statements that are necessary for IPv6 to function.
- IPv6 only supports named and extended ACLs.
- If using an explicit deny, make sure to manually include the ND and RA/RS statements.

```
ipv6 access-list ALLOW_TELNET_OSPF
 sequence 10 permit 89 host FE80::2 any
 sequence 20 permit tcp host 12::2 any eq telnet
 sequence 30 permit icmp any any nd-na
 sequence 40 permit icmp any any nd-ns
 sequence 50 permit icmp any any router-advertisement
 sequence 60 permit icmp any any router-solicitation
 sequence 99 deny ipv6 any any log-input

int fa0/0
 ipv6 traffic-filter ALLOW_TELNET_OSPF in
```

Reflexive

Sunday, December 13, 2015
10:23 PM

Reflexive Access-Lists

- Only allow return traffic if inside source initiated the traffic.
- The reflect keyword links ACLs together, the evaluate entry is dynamically created based on this reflect entry.
- By default the dynamic entry will timeout after 60 seconds.

```
ip access-list extended TRAFFIC_FROM_R3
 10 permit ospf any any
 20 permit icmp host 192.168.0.3 any reflect ICMP

ip access-list extended TRAFFIC_TO_R3
 10 evaluate ICMP
 99 deny ip any any log-input

ip access-list logging interval 1
ip access-list log-update threshold 1
ip reflexive-list timeout 60

int fa0/0
 description LINK_TO_R3
 ip access-group TRAFFIC_FROM_R3 in
 ip access-group TRAFFIC_TO_R3 out
```

Time-Based

Sunday, December 13, 2015
10:23 PM

Time-Based Access-Lists

- Specify time ranges that will allow traffic during specific periods.
- Based on local system time.

```
time-range DAILY
 periodic daily 09:00 to 17:00

ip access-list extended ALLOW_ICMP_TIME
 10 permit ospf any any
 20 permit icmp host 192.168.0.3 any time-range DAILY
 30 deny ip any any log-input

int fa0/0
 description LINK_TO_R3
 ip access-group ALLOW_ICMP_TIME in
```

BGP

Wednesday, January 13, 2016
5:04 PM

AS Range

- Private AS range is 64512 - 65534, 65535 is reserved for special use.

asplain	2-byte: 1 to 65535 4-byte: 65536 to 4294967295
asdot	2-byte: 1 to 65535 4-byte: 1.0 to 65535.65535

BGP Best Path Selection

- A peer Autonomous System (AS) is what defines an external or internal route.
- A route learned from an external peer will be external, until the point where the route is forwarded (or redistributed) into own AS.

First criteria is that the NEXT_HOP is reachable, meaning that it is present in the routing table.
Which protocol was responsible for generating the route is irrelevant.

- WEIGHT (highest)
- LOCAL_PREF (highest)
- Locally injected (network, aggregate) over remotely learned
- AS_PATH (shortest)
- ORIGIN (lowest) (0 over 1 over 2) 0 is internal, 1 is external (absent), 2 is redistributed (incomplete)
- MUTLI_EXIT-DISC / MED (lowest)
- eBGP over iBGP learned routes
- Lowest IGP cost to the NEXT_HOP

Best-Path Tie Breakers (No Multipath)

- If both paths are external, prefer the older one
- If both paths are internal, prefer the lowest ROUTER_ID
    - If ORIGINATOR_ID is the same, prefer one with the shorter CLUSTER_LIST
    - Finally, prefer the one with the lowest neighbor's IP-address

Multipath is enabled

- Weight, Local preference, AS_PATH length, Origin, MED, AS PATH content must match + eBGP and iBGP specific additional rules.

BGP Synchronization

- Synchronization is only relevant in iBGP. Disabled by default.
- Routes will only be passed on if they are synchronized.
- All iBGP learned prefixes must have an identical route in the routing table learned from an IGP.
- The originator of the route in the routing table has a router-id (RID).
- This IGP RID has to match the BGP neighbor that the prefix was received from otherwise its not synchronized.
- To fix synchronization problems change either the originators IGP RID or the neighbors BGP RID to the same value.

BGP Attributes

- BGP uses Network Layer Reachability Information (NLRI), which is the route and prefix and certain attributes.

Well known, Mandatory	Must appear in every UPDATE message. Must be supported by all BGP software implementations. If missing, a NOTIFICATION message must be sent to the peer.	NEXT_HOP AS_PATH ORIGIN
Well-Known, Discretionary	May or may not appear in an UPDATE message. Must be supported by any BGP software implementation.	LOCAL_PREF AT_AGGREGATE

Optional, Transitive	May or may not be supported in all BGP implementations. Will be passed on if not recognized by the receiver.	AGGREGATOR COMMUNITY
Optional, Non-Transitive	May or may not be supported in all BGP implementations. Not required to pass on, may be safely ignored.	MED ORIGINATOR_ID CLUSTER_LIST

BGP Attributes Scope

- Transitive, these attributes are across AS boundaries.
- Non-transitive, these attributes are restricted to the same AS.
- Local, these attributes are local to the router only (weight).

Transitive attributes can be altered to be local only, non-transitive attributes cannot be altered.
MED can be transitive but only to one neighbor AS, not more.

NEXT_HOP Attribute

This is a well-know, mandatory, transitive attribute that must be present in all updates.

- Is the peers IP-address if remotely learned.
- Is 0.0.0.0 for routes advertised using the network or aggregate commands.
- Is the IP next-hop for redistributed routes.
- The next-hop must be reachable, meaning that it must be present in the routing table.
- Remains unchanged in the same AS by default, but can (or should) be modified.
- Is changed by default when forwarded between different AS. Will become the IP address of the router that passed on the route.

AS_PATH Attribute

This is a well-know, mandatory, transitive attribute that must be present in all updates.

- Ordered List (AS_SEQUENCE).
- Read from right to left.
- Can have an unordered component AS_SET.
- Used for loop prevention only.
- Shorter count is better (hop count).
- Local ASN is added when advertised to an external peer.
- Can be modified using route-maps.
- AS Path Prepending.
- In special cases can be shortened using neighbor "neighbor-ip-address" remove-private-as. This will remove the private AS that is connected to a non-private as from appearing in the AS_PATH. Neighboring AS will assume that the route originated from the non-private AS.

ORIGIN Attribute

This is a well-know, mandatory, transitive attribute that must be present in all updates. Origin is not part of AS_PATH.
Three possible values:

- IGP (0, shown in IOS as "i").
- EGP (1, shown in IOS as "e").
- Incomplete (2, shown in IOS as "?").

Set at the injection point:

- Using "network" or "aggregate" will result in IGP origin code.
- Using redistribution will result in incomplete origin code.

Can be modified using route-maps using the neighbor statement or at the insertion point.

WEIGHT Attribute

This is a proprietary, optional, non-transitive attribute that is Cisco only and is only of local significance.
Higher value preferred, default values are:

- 32768 for locally inserted.
- 0 for remotely learned.

Can be changed using neighbor statement (directly) or using route-maps (only in the inbound direction)

LOCAL_PREF Attribute

This is a well-known, discretionary, non-transitive attribute that must be supported by all vendors, but may not appear in every UPDATE message. The local-preference is only of significance to the same AS.

- Higher value preferred, default value is 100. Not always visible using show commands.
- Takes effect in the entire autonomous system, even in confederations (sub-autonomous systems)
- Most effective way to influence path preference for incoming routes, meaning outgoing traffic paths.

Can be modified using route-maps.

ATOMIC_AGGREGATE Attribute

- Well-known discretionary attribute
- Must be recognized by all BGP implementations, but does not have to appear in all UPDATES.
- This attribute is set when routes are aggregated (summarized) at a BGP speaker and forwarded to another AS.
- Alerts BGP peers that information may have been lost in the aggregation and that it might not be the best path to the destination.
- For example, if 10.0.1.0/24 and 10.0.2.0/24 are summarized to 10.0.0.0/22 it will include the 10.0.0.0/24 and 10.0.3.0/24.
- Aggregated routes will inherit the attributes of the component routes.
- When routes are aggregated, the aggregator attaches its RID to the route into the AGGREGATOR_ID attribute, unless the AS_PATH is set using the AS_SET statement.
- It is not possible to disable the discard route in BGP.

Communities Attribute

- This is an optional, transitive attribute that is comparable to route-tags.
- Not used for best-path selection.
- Large number (32 bits) that are conventionally displayed as ASN:ID (for example 64512:100).
- This notation however must be enabled using the ip bgp-community new-format command.
- This community "tag" is stripped by neighbors by default (even within the same AS). Use neighbor statement with route-maps.
- Routes can be in multiple communities.

MULTI EXIT_DISC/MED/METRIC Attribute

This is an optional, non-transitive attribute meaning that its not required to pass on when received from another AS.
MED can actually be transitive but only to one neighbor AS, not more.

- Lower value preferred, but the default value is missing in IOS (treated as 0 which is equal and thus best). This can be modified so that 0 is treated as worst. Use the bgp bestpath med missing-as-worst command.
- Influences preferred "exit point" when peering with the same AS at multiple locations.
- Requirements are that WEIGHT, LOCAL_PREF, AS_PATH length and peer AS must be the same. These requirements can also be modified, using the bgp always-compare-med command.

Can be modified (increased from 0) using route-maps or redistribution.

- Setting the MED outbound will influence the remote AS.
- Setting the MED inbound will influence own AS.

ORIGINATOR_ID

Route reflectors use the ORIGINATOR_ID and CLUSTER_LIST attributes for loop prevention.

- Each routing table entry will have the originating routers RID inserted by the RR.
- The RR will pass along routes to the RR-Clients, but it is these clients themselves who check the update for the presence of the RID (ORIGINATOR_ID). If this ORIGINATOR_ID matches their own, the update is denied.

CLUSTER_LIST

Set by RR by default to the value of the RID. The CLUSTER_ID identifies a group, or a single RR.

- RR can use this information to discover other RR present on the network.
- Prevents the installation of multiple routes in the BGP table that were reflected by RR neighbor.

Advertise-Map

Wednesday, January 13, 2016
5:04 PM

BGP Advertise-Map

Advertise prefixes based on the existence of other prefixes in the BGP table.

- Network in the non-exist map has to be present in the BGP table.
- Possible to advertise default route if another route is present.
- This does not work with the default-information originate command.

```
route-map NON_EXIST permit 10
 match ip address prefix-list Lo2
route-map ADV_MAP permit 10
 match ip address prefix-list ADV_PREFIX

ip prefix-list ADV_PREFIX permit 192.168.0.1/32
ip prefix-list Lo2 permit 192.168.0.2/32

router bgp 1
 address-family ipv4
  network 192.168.0.1 mask 255.255.255.255
  neighbor 10.0.13.3 advertise-map ADV_MAP non-exist-map NON_EXIST

show ip bgp neighbors 10.0.13.3 | i Condition
```

Advertise default:

```
route-map EXIST permit 10
 match ip address prefix-list Lo2
route-map ADV_MAP permit 10
 match ip address prefix-list ADV_DEFAULT

ip prefix-list ADV_DEFAULT permit 0.0.0.0/0
ip prefix-list Lo2 permit 192.168.0.2/32

ip route 0.0.0.0 0.0.0.0 null0

router bgp 1
 address-family ipv4
  network 0.0.0.0 mask 0.0.0.0
  neighbor 10.0.13.3 advertise-map ADV_MAP exist-map EXIST
```

Aggregation

Wednesday, January 13, 2016
5:04 PM

ATOMIC_AGGREGATE Attribute

- Well-known discretionary attribute
- Must be recognized by all BGP implementations, but does not have to appear in all UPDATES.
- This attribute is set when routes are aggregated (summarized) at a BGP speaker and forwarded to another AS.
- Alerts BGP peers that information may have been lost in the aggregation and that it might not be the best path to the destination.
- For example, if 10.0.1.0/24 and 10.0.2.0/24 are summarized to 10.0.0.0/22 it will include the 10.0.0.0/24 and 10.0.3.0/24.
- Aggregated routes will inherit the attributes of the component routes.
- When routes are aggregated, the aggregator attaches its RID to the route into the AGGREGATOR_ID attribute, unless the AS_PATH is set using the AS_SET statement.
- It is not possible to disable the discard route in BGP.

The as_set keyword preserves the AS_PATH information, meaning that the AS information is not overwritten by the aggregator.

- In this case the AT_AGGREGATE is not set and the original AS (or multiple AS) is still present in the aggregated route.
- The as-confed-set keyword is the same as as-set, but applies to confederations.

By default the more specific routes are still advertised to peers in addition to the summary.

- This behavior can be altered with the summary-only command to not advertise the more specific routes.

BGP Aggregation Suppress-Map

- Accomplishes the same as summary-only, except only subnets matched in the suppress-map will not be advertised.

```
ip prefix-list 1 permit 3.0.0.1/32
ip prefix-list 3 permit 3.0.0.3/32

route-map SUPPRESS permit 10
 match ip address prefix-list 1
 match ip address prefix-list 3

router bgp 1
 address-family ipv4
  aggregate-address 3.0.0.0 255.255.255.252 as-set suppress-map SUPPRESS
```

BGP Aggregation Unsuppress-Map

- Subnets matched will be advertised alongside the suppressed summary.
- Applied on neighbor statement instead of aggregate statement.
- Will not inherit attributes set by other neighbor statements such as community values.
- In order to send attributes, neighbor route-map configurations have to be copied to the unsuppress route-map.

```
ip prefix-list 1 permit 3.0.0.1/32
ip prefix-list 3 permit 3.0.0.3/32

route-map UNSUPPRESS permit 10
 match ip address prefix-list 1
 match ip address prefix-list 3
 set community 1:1 additive

router bgp 1
 address-family ipv4
  aggregate-address 3.0.0.0 255.255.255.252 as-set summary-only
  neighbor 10.0.12.2 unsuppress-map UNSUPPRESS
```

BGP Inject-Map

- Routers can conditionally inject a more specific route based on the presence of an aggregate in the BGP table.
- The injected subnets must fall within the range of the aggregate.
- Copy-attributes will also transfer AS information, otherwise origin will be incomplete.
- Route-source must match the neighbor the aggregate was received from, NOT the originator.

```
ip prefix-list R3 permit 10.0.13.3/32
ip prefix-list AGG permit 3.0.0.0/29
ip prefix-list INJECT permit 3.0.0.5/32
ip prefix-list INJECT permit 3.0.0.6/32
ip prefix-list INJECT permit 3.0.0.7/32

route-map R3_AGG
 match ip address prefix-list AGG
 match ip route-source prefix-list R3
route-map INJECT
 set ip address prefix-list INJECT

router bgp 1
 address-family ipv4
  bgp inject-map INJECT exist-map R3_AGG copy-attributes

show ip bgp injected-paths
```

BGP Aggregation Advertise-map

- This function works alongside AS_SET and summary-only.
- Used when multiple AS share the same prefixes and are both included in the aggregation.
- AS that are filtered do not become unreachable, they are just hidden in the aggregate.

```
show ip bgp regexp _4$
ip as-path access-list 4 permit _4$

route-map AGG_HIDE_AS4 deny 10
 match as-path 4
route-map AGG_HIDE_AS4 permit 99

router bgp 1
 add ipv4
 aggregate-address 34.0.0.0 255.255.255.248 summary-only as-set advertise-map AGG_HIDE_AS4
```

BGP Aggregation Attribute-Map

- Optionally add additional attributes to the aggregate with the attribute-map keyword.
- Using a route-map instead will provide the same results, and will be converted to attribute-map in the config.

```
route-map AGG_ATTRIBUTE permit 10
 set metric 500
 set local-preference 200
 set origin igp
 set community 3:1
 etc..

router bgp 1
 address-family ipv4
  aggregate-address 3.0.0.0 255.255.255.252 attribute-map AGG_ATTRIBUTE
  aggregate-address 3.0.0.0 255.255.255.252 route-map AGG_ATTRIBUTE
```

AS_PATH

Wednesday, January 13, 2016
5:04 PM

BGP Prepend AS_PATH

Prepend AS 3 five times to the AS_PATH:

```
route-map PREPEND_AS permit 10
 set as-path prepend 3 3 3 3 3

router bgp 1
 neighbor 10.0.12.2 route-map PREPEND_AS out
```

Prepend the last AS in the path 4 times. This will lead to 5 entries on neighbors (1 original + 4 prepend)

```
route-map PREPEND_LAST_AS permit 10
 set as-path prepend last-as 4

router bgp 1
 neighbor 10.0.12.2 route-map PREPEND_LAST_AS out
```

BGP AS_PATH Tagging

- The AS_PATH is converted to a tag by default when redistributing BGP into IGP.
- Convert this tag back when redistributing the IGP back into another BGP process on another router with set as-path tag.
- Configure set automatic-tag on the original redistributing router (BGP->OSPF) to preserve the origin code as well (i instead of ?).
- The automatic tag has to match a specific AS in the route-map.
- It is not possible to prepend these redistributed routes using the same route-map.

AS_PATH tag:

```
route-map AS_PATH_TAG permit 10
 set as-path tag

router bgp 1
 add ipv4
  redistribute ospf 1 route-map AS_PATH_TAG
```

Automatic Tag:

```
ip as-path access-list 1 permit .*
route-map AS_ORIGIN_TABLE_MAP permit 10
 match as-path 1
 set automatic-tag

router bgp 2
 add ipv4
  table-map AS_ORIGIN_TABLE_MAP

clear ip bgp ipv4 unicast table-map
```

BGP Local-AS

- Use a different AS then is configured when neighboring with peers.

By default R2 will see AS 2 prepended before the AS 1 on routes received from R1.

- AS 2 will also be prepended to all R1 routes passed on to other BGP peers.
- To override this behavior, configure no-prepend on R2.

By default R1 will see AS 2 prepended before the actual AS of 64512.

- To override this behavior, configure replace-as on R2.

- R1 will see prefixes from R2 with AS2, other peers will see AS 64512.

The dual-as keyword will allow R1 to peer with either the correct AS 64512 or the local-as 2.

```
router bgp 1
 neighbor 10.0.12.2 remote-as 2

router bgp 64512
 neighbor 10.0.12.1 remote-as 1
 neighbor 10.0.12.1 local-as 2 no-prepend replace-as dual-as
```

Communities

Wednesday, January 13, 2016
5:04 PM

Communities Attribute

- This is an optional, transitive attribute that is comparable to route-tags.
- Not used for best-path selection.
- Large number (32 bits) that are conventionally displayed as ASN:ID (for example 64512:100).
- This notation however must be enabled using the ip bgp-community new-format command.
- This community "tag" is stripped by neighbors by default (even within the same AS). Use neighbor statement with route-maps.
- Routes can be in multiple communities.

BGP Filtering using Communities

There are three well-known BGP communities:

- No-advertise. Do not advertise route to any peers. CxFFFFFF02
- No-export. Do not advertise route to external peers. Advertise to internal and confederation peers only. OxFFFFF01
- Local-as. Do not advertise route to external and confederation peers. Advertise to internal peers only. OxFFFFFF03

It is possible to add the communities at the network statement or redistribution point.

- This will allow the router to advertise routes to specific neighbors only.
- With the neighbor statement the routes are advertised to that neighbor but the community is applied after reception.
- With the network statement the community is applied immediately.

Advertise routes with incomplete origin (alternative to redistribution):

```
route-map BGP_ORIGIN permit 10
 set origin incomplete

router bgp 1
 address-family ipv4
  network 1.0.1.0 mask 255.255.255.0 route-map BGP_ORIGIN
  network 1.0.2.0 mask 255.255.255.0 route-map BGP_ORIGIN
  network 1.0.3.0 mask 255.255.255.0 route-map BGP_ORIGIN
  network 1.0.4.0 mask 255.255.255.0 route-map BGP_ORIGIN

ip prefix-list Lo1 permit 1.0.1.0/24
ip prefix-list Lo2 permit 1.0.2.0/24
ip prefix-list Lo3 permit 1.0.3.0/24
ip prefix-list Lo4 permit 1.0.4.0/24

route-map BGP_COMMUNITY permit 10
 match ip address prefix-list Lo1
 set community no-advertise 1:1

route-map BGP_COMMUNITY permit 20
 match ip address prefix-list Lo2
 set community local-as 1:1

route-map BGP_COMMUNITY permit 30
 match ip address prefix-list Lo3
 set community no-export 1:1

route-map BGP_COMMUNITY permit 99

router bgp 1
 address-family ipv4
  neighbor 10.0.12.2 route-map BGP_COMMUNITY out
```

```
show ip bgp community 1:1
```

Match community values and modify on other routers:

- The set comm-list 1 delete keyword will only delete the community matched in the list.
- The additive keyword will add the 'internet' community to the existing communities, and not overwrite them.

```
ip community-list 1 permit no-export
ip prefix-list Lo4 permit 1.0.4.0/24

route-map MODIFY_COMMUNITY permit 10
 match ip address prefix-list Lo4
 match community 1
 set comm-list 1 delete
 set community internet additive
route-map RM1 permit 100

router bgp 2
 address-family ipv4
  neighbor 10.0.12.1 route-map MODIFY_COMMUNITY in
```

Confederations

Wednesday, January 13, 2016
5:04 PM

iBGP Confederations

- Divides autonomous system into smaller sub-autonomous systems.
- Large (container) AS is a confederation, smaller systems are "members".
- Inside the member, regular iBGP rules apply using Full-mesh and/or RR.
- Perceived externally only as the confederation AS
- Local Preference (LOCAL_PREF) Applies to the entire confederation.

Configuration Considerations

- All routers must be configured to be aware of the confederation identifier.
- All routers should be configured to be aware of all the confederation peers (members).
- Can be configured alongside RR.

```
router bgp 1
 bgp confederation identifier 123
 bgp confederation peers 2
 bgp confederation peers 3
neighbor 192.168.0.2 remote-as 2
neighbor 192.168.0.3 remote-as 3
neighbor 192.168.0.2 update-source Lo0
neighbor 192.168.0.3 update-source Lo0
neighbor 192.168.0.2 ebgp-multihop
neighbor 192.168.0.3 ebgp-multihop
 address-family ipv4
  neighbor 192.168.0.2 next-hop-self
  neighbor 192.168.0.3 next-hop-self
```

The peering between AS1 and AS2, AS3 is an internal peering using different AS.

- Meaning that routers R3 and R2 are unaware of the routes injected by R1, unless a full-mesh is configured.
- Another option is to use next-hop-self for all peers.

When peering with loopbacks between confederation peers the peering is basically the same as eBGP.

- Meaning that ebgp-multihop has to be configured.

eBGP Peering

Wednesday, January 13, 2016
5:04 PM

eBGP Peering Rules

- Neighbor must be in a different AS.
- Neighbor should be directly connected. Override with ebgp-multihop which changes the default TTL.
- If not, an underlying routing protocol must provide reachability (static, IGP). Beware of recursive routing.
- NEXT_HOP is changed when routes are advertised (by the advertising router).
- Routes with own AS are rejected if observed in incoming AS_PATH updates (AS_SEQUENCE and / or AS_SET).
- All routes are advertised and accepted.

If using an IGP to advertise loopbacks for the external peering (multi-hop). It is important not to advertise the same loopbacks into BGP.

- BGP has a lower AD, meaning that the IGP routes used for the peering will become BGP routes. This will lead to recursive routing.
- Solve by increasing the AD of eBGP (not recommended), decreasing the AD of the IGP (possible solution) or use backdoor routes.

The backdoor route will set the iBGP administrative distance (200) for the route instead of the eBGP distance.

- This ensures that it is higher than the IGP AD.
- The backdoor statement is specified when advertising the loopback of the neighbor into BGP.

```
int se1/0
 description EBGP_LINK_TO_R2
 ip add 10.0.12.1 255.255.255.0
interface lo0
 description EBGP_PEERING_LOOPBACK
 ip address 192.168.0.1 255.255.255.255

router eigrp 1
 network 10.0.12.0 0.0.0.255
 network 192.168.0.1 0.0.0.0

ip bgp-community new-format
router bgp 1
 bgp router-id 192.168.0.1
 neighbor 192.168.0.2 remote-as 2
 neighbor 192.168.0.2 ebgp-multihop 2
 neighbor 192.168.0.2 update-source Loopback0
 address-family ipv4
  network 192.168.0.1 mask 255.255.255.255
  network 192.168.0.2 mask 255.255.255.255 backdoor
  neighbor 192.168.0.2 activate
  neighbor 192.168.0.2 send-community
```

eBGP Peering using Default Route

BGP will not actively open (initiate) BGP session with a peer if the only route to it is the default route.

- BGP will initiate a session if the active BGP peer has a more specific route to its peer.
- The passive BGP peer is allowed to only have a default route.
- One side has to have a specific route for this to work.

Force active/passive peerings between peers:

```
router bgp 1
 neighbor 192.168.0.2 transport connection-mode active

router bgp 2
```

```
 neighbor 192.168.0.1 transport connection-mode passive
```

If Peer 1 (being the active peer by force) only has a default route to Peer 2, the session will not form.

```
ip route 0.0.0.0 0.0.0.0 10.0.12.2

ip bgp-community new-format
router bgp 1
 bgp router-id 192.168.0.1
 neighbor 192.168.0.2 remote-as 2
 neighbor 192.168.0.2 ebgp-multihop
 neighbor 192.168.0.2 update-source Loopback0
 neighbor 192.168.0.2 transport connection-mode passive
 address-family ipv4
  neighbor 192.168.0.2 activate
  neighbor 192.168.0.2 send-community
```

BGP Connected-Check

- Allows directly connected neighbors to peer with loopback addresses without configuring multi-hop.

```
router bgp 1
 neighbor 10.0.12.2 remote-as 2
 neighbor 10.0.12.2 disable-connected-check
```

iBGP Peering

Wednesday, January 13, 2016
5:04 PM

BGP Internal Routing

- Does not replace IGP. No direct connectivity required.
- Works alongside IGPs, this is why direct connectivity is not required.
- Carry external prefixes throughout the AS.
- "Split-horizon" as the loop-prevention. iBGP routes are not forwarded to other iBGP peers.
- Does not re-advertise internally-learned prefixes to other iBGP peers, because the AS_PATH is not modified inside the AS.
- Multiple sessions to same neighbor are not permitted, use loopback destinations instead.

Peering Rules

- Neighbor must be in the same AS.
- Neighbors do not have to be directly connected, if not a underlying routing protocol must provide reachability.

Default Policy Behavior

- NEXT_HOP is not changed when routes are advertised.
- External routes are advertised.
- Routes learned from other internal peers are not advertised to other iBGP neighbors. But are advertised to eBGP neighbors.

iBGP Next-Hop-Self

- Because NEXT_HOP is not changed, IGPs are needed to reach neighbors.
- Another solution is to change the NEXT_HOP to the local router address.
- The address that will be chosen for this is the loopback address (in case of next-hop-self) or the internal address used to peer with the neighbors.
- Make sure the networks next_hop_self is set to are advertised into iBGP. Another solution is to redistribute the IGP.

```
router bgp 1
 neighbor 10.0.12.2 remote-as 2
 address-family ipv4
  neighbor 10.0.12.2 next-hop-self all
  network 10.0.12.0 mask 255.255.255.0
  network 10.0.13.0 mask 255.255.255.0
  redistribute ospf 1 metric 2 match internal
```

The all keyword enables next-hop-self for both eBGP and iBGP received paths. Default is only for iBGP received paths.

- BGP only redistributes ebgp routes into an IGP
- BGP does not redistribute ospf external routes by default
- If no metric is specified it will be equal to the number of hops a peer needs to reach the networks.
- If a metric is specified, it will be applied to all routes advertised, meaning that all routes in the OSPF domain will receive the same metric.
- The default is to only match OSPF internal routes, redistributed routes into OSPF are not included.

iBGP Peer-Groups

```
router bgp 123
 bgp listen range 10.0.123.0/24 peer-group PEERS
 bgp listen limit 2
 neighbor PEERS peer-group
 neighbor PEERS remote-as 123
 address-family ipv4
  neighbor PEERS activate
  neighbor PEERS send-community
  neighbor PEERS etc..
```

IPv6

Wednesday, January 13, 2016
5:06 PM

IPv6 Networks over IPv4

```
router bgp 1
 neighbor 10.0.12.2 remote-as 2
 address-family ipv4
  neighbor 10.0.12.2 activate
 address-family ipv6
  neighbor 10.0.12.2 activate
  neighbor 10.0.12.2 route-map IPV6_NEXT_HOP out
  network 1::1/128

route-map IPV6_NEXT_HOP permit 10
 set ipv6 next-hop 2001:10:0:12::1
```

IPv6 Peer Link-Local

```
router bgp 1
  neighbor FE80::2%Serial1/0 remote-as 2
 address-family ipv6
  neighbor FE80::2%Serial1/0 activate
  network 1::1/128
```

IPv6 Peer Loopback

- Same as IPv4, IPv6 neighbors are not automatically activated.
- Specify router-id if no IPv4 addresses are used on the router.

```
ipv6 route 2::2/128 2001:10:0:12::2

router bgp 1
 bgp router-id 192.168.0.1
 neighbor 2::2 remote-as 2
 neighbor 2::2 update-source Loopback 0
 neighbor 2::2 disable-connected-check
 address-family ipv6
  neighbor 2::2 activate
```

IPv4 Networks over IPv6

- Peer with directly connected interfaces (NOT loopbacks) when advertising IPv4 prefixes over an IPv6 connection.

```
router bgp 1
 neighbor 2001:10:0:12::2 remote-as 2
 address-family ipv4
  neighbor 2001:10:0:12::2 activate
  neighbor 2001:10:0:12::2 route-map IPV4_NEXT_HOP in
  network 192.168.0.1 mask 255.255.255.255
 address-family ipv6
  neighbor 2001:10:0:12::2 activate

route—map IPV4_NEXT_HOP permit 10
 set ip next—hop 10.0.12.2
```

Filtering AS

Wednesday, January 13, 2016
5:07 PM

AS-Path Filters

_	White space, start of string or end of string.
^	Start of string, PEER AS.
$	End of string, ORIGINATING AS.
.	Matches any character.
+	repeats the previous character one or more times.
*	repeats the previous character zero or many times
?	repeats the previous character one or zero times.
I	A logical "or" statement.
\	Removes special meanings.
()	Matches a group of characters.
[]	Matches a range of characters.

Definition and Use of AS-Path Access-Lists

- AS-Path access-lists match a single character, every number in an AS is a single character.
  - Example. AS 2456 consists of the single characters 2,4,5 and 6.
- Can match a range of characters using brackets [].
  - Example. [0-3] consists of 0 or 1 or 2 or 3
- Can match a group of characters.
  - Example. (123) matches only 123 when the characters follow each other. However they can be part of a greater AS such as 55123 or 12300.
  - Example. (123_) matches only 123 when its at the end of the AS, such as 55123 but not 12300.
  - Example. (_123_) matches only AS123 and it cannot be part of a greater AS.

AS-Path AS-Sequence

- In IOS the AS-Sequence is the AS_PATH that a route travels through.
- The AS that the router receives the route from is called the PEER AS.
- The router that originated the route is called the ORIGIN AS, others are called TRANSIT AS.
  - Example. ^500_400_300_200_100$. 500 is the PEER AS, 100 is the ORIGIN AS, others are TRANSIT AS.

AS-Path Filtering

Match prefixes that originated in the connected AS:

```
ip as-path access-list 1 permit ^[0-9]+$
route-map BGP_CONNECTED_AS permit 10
 match as-path 1

router bgp 1
 address-family ipv4
  neighbor 10.0.12.2 route-map BGP_CONNECTED_AS in
  neighbor 10.0.12.2 filter-list 1 in
```

This matches all numbers but does not allow blank spaces. Meaning that there can only be one AS in the path, which is the neighbors AS. The + means that the pattern must appear, so blank AS (our own) will not match.

^$	Local AS
.*	All AS
^5_	Directly connected AS 5
_5_	Transferring AS 5
_5$	Originated in AS 5
^[0-9]+$	Originated in directly connected AS
^[0-9]*$	Originated in directly connected AS + empty AS (local AS)
^([0-9]+)_5	AS 5 which passed through directly connected AS
^2_([0-9]+)	Directly connected to AS 2
^\(1234\)	Confederation peer 1234.
_(4|5)$	Originated in AS 4 or AS 5
^(2|3)$	Originated in AS 2 or AS 3 that are directly connected
_2_(4|5)$	Originated in AS 4 or AS 5 that passed through AS 2
^[0-9]+([0-9]+)?$	Originated in directly connected AS or directly connected to our directly connected AS

? Basically means true or false, the secondary AS that are being matched can appear or not.

Filtering PL / ACL

Wednesday, January 13, 2016
5:07 PM

BGP Extended Access-Lists Filtering

```
ip access-list extended PREFIXES
 deny ip 2.0.0.0 0.0.0.255 host 255.255.255.252
 deny ip 2.0.0.0 0.0.0.255 host 255.255.255.254
 permit ip any any

router bgp 1
 address-family ipv4
  neighbor 10.0.12.2 distribute-list PREFIXES in
```

BGP Prefix-List Filtering

```
ip prefix-list PREFIXES deny 1.0.0.2/31
ip prefix-list PREFIXES deny 1.0.0.4/30
ip prefix-list PREFIXES permit 0.0.0.0/0 le 32

router bgp 2
 address-family ipv4
  neighbor 10.0.12.1 prefix-list PREFIXES in
```

BGP Outbound Route Filtering (ORF)

- Normally when a router filters incoming routes the peer will still advertise them.
- To signal the peer that some routes do not need to be advertised, configure an ORF based on the same prefix-list that is filtering.

```
ip prefix-list PREFIXES deny 1.0.0.1/32
ip prefix-list PREFIXES deny 1.0.0.2/32
ip prefix-list PREFIXES deny 1.0.0.3/32
ip prefix-list PREFIXES deny 1.0.0.4/32
ip prefix-list PREFIXES permit 0.0.0.0/0 le 32

router bgp 2
 neighbor 10.0.12.1 remote-as 1
 address-family ipv4
  network 192.168.0.2 m 255.255.255.255
  neighbor 10.0.12.1 activate
  neighbor 10.0.12.1 capability orf prefix-list send
  neighbor 10.0.12.1 prefix-list PREFIXES in

router bgp 1
 neighbor 10.0.12.2 remote-as 2
 address-family ipv4
  neighbor 10.0.12.2 activate
  neighbor 10.0.12.2 capability orf prefix-list receive
  network 192.168.0.1 m 255.255.255.255
  network 1.0.0.1 m 255.255.255.255
  network 1.0.0.2 m 255.255.255.255
  network 1.0.0.3 m 255.255.255.255
  network 1.0.0.4 m 255.255.255.255

show ip bgp neighbors 10.0.12.1 | s capabilities
show ip bgp neighbors 10.0.12.2 advertised-routes
```

Misc

Wednesday, January 13, 2016
5:07 PM

BGP Fall-over

- Holdtime does not have to expire in order to tear down the session.
- The fall-over will still allow bgp to form neighborships using a default or summary route.
- Configure a route-map to make sure that the specific route has to be present in the routing table.

```
ip route 192.168.0.2 255.255.255.255 10.0.12.2

ip prefix-list R2_LOOPBACK permit 192.168.0.2/32
route-map R2_FALLOVER permit 10
 match ip address prefix-list R2_LOOPBACK

router bgp 1
 neighbor 192.168.0.2 fall-over route-map R2_FALLOVER
```

BGP Next Hop Tracking (NHT)

- On-by-default feature that notifies BGP to a change in routing for BGP prefix next-hops.
- NHT makes the process of dealing with next-hop changes event-driven, instead of using the 60 second scanner process.
- The bgp nexthop trigger delay defines how long for the NHT process to delay updating BGP.

```
router bgp 1
 bgp nexthop trigger enable
 bgp nexthop trigger delay 5
```

BGP Selective Address Tracking (SAT)

- The route-map determines what prefixes can be seen as valid prefixes for next-hops.
- Allows for specific addresses as viable next-hops, and will pull prefixes from the bgp table if the next-hop does not match.
- Re-use the same route-map in fall-over to also tear down the session.

```
ip prefix-list LOOPBACKS permit 0.0.0.0/0 ge 29
route-map SAT permit 10
 match ip address prefix-list LOOPBACKS

router bgp 1
 address-family ipv4
  bgp nexthop route-map SAT
  neighbor 10.0.13.3 fall-over route-map SAT
  neighbor 10.0.12.2 fall-over route-map SAT
```

BGP Multi-Session

Use a different TCP session for each address-family:

```
router bgp 1
 neighbor 192.168.0.2 transport multi-session (must agree on both sides)
```

BGP TTL-Security

- Configuring TTL-Security automatically allows peering over multiple hops (without explicitly configuring multi-hop).
- Configuring a TTL-Security hop-count of 2 will cause the router to expect incoming packets have a TTL value of 253.
- If packets arrive with a TTL of lower than 253, they will be discarded.

```
router bgp 1
 bgp router-id 192.168.0.1
 neighbor 192.168.0.2 remote-as 2
```

```
neighbor 192.168.0.2 update-source Loopback0
neighbor 192.168.0.2 ttl-security hops 2
address-family ipv4
 neighbor 192.168.0.2 activate
```

Route-Reflector

Wednesday, January 13, 2016
5:06 PM

iBGP Route-Reflector

- Because routes learned from other internal peers are not advertised to other iBGP neighbors, iBGP needs a full-mesh topology.
- This approach has drawbacks however when a lot of routers are present in the AS.
- Another solution is to use a Route-Reflector (RR) to basically re-advertise the routes to internal peers.

BGP Route-Reflectors Terminology:

- Route-Reflector (RR), the router that has to feature enabled on some or all interfaces.
- Route-Reflector Client, router peering with the RR for which the RR has the feature enabled.
- Non-Client, router that is in the same AS but is not enabled for the feature on the RR.

A Route Reflector reflects routes considered as best routes only.

- If more than one update is received for the same destination, only the BGP best route is reflected.
- A Route Reflector is not allowed to change attributes of the reflected routes including the next-hop attribute.

Route Target Constraint

With Route Target Constraint (RTC) the Route Reflector sends only wanted VPN4/6 prefixes to the PE.

- 'Wanted' means that the PE has a VRF importing the specific prefixes.

ORIGINATOR_ID

Route reflectors use the ORIGINATOR_ID and CLUSTER_LIST attributes for loop prevention.

- Each routing table entry will have the originating routers RID inserted by the RR.
- The RR will pass along routes to the RR-Clients, but it is these clients themselves who check the update for the presence of the RID (ORIGINATOR_ID). If this ORIGINATOR_ID matches their own, the update is denied.

CLUSTER_LIST

Set by RR by default to the value of the RID. The CLUSTER_ID identifies a group, or a single RR.

- RR can use this information to discover other RR present on the network.
- Prevents the installation of multiple routes in the BGP table that were reflected by RR neighbors.

Set the CLUSTER_ID to be equal on all reflectors:

```
router bgp 1
 bgp cluster-id 13.13.13.13
```

Advertisement Rules

- Routes received from iBGP are advertised to external peers.
- So there is no need for RR in this case.

Non-client -> iBGP = Forwarded to eBGP and RR-Clients.

- When a NC receives a route from an eBGP peer it will be forwarded to the RR, which forwards it to all eBGP peers and iBGP RR-Clients.

External route from RR-Client -> iBGP = Forwarded to eBGP, RR-Clients and non-clients

- When a RR-Client receives a route from an eBGP peer it will be forwarded to the RR, which forwards it to all eBGP peers and iBGP peers.

External route from RR -> iBGP = Forwarded to eBGP, RR-Clients and non-clients

- When a RR receives a route from an eBGP peer it will be forwarded to all eBGP and iBGP peers.

Session

Wednesday, January 13, 2016
5:07 PM

BGP Session

- BGP uses TCP port 179 for destination and a random source port.
- Commands that affect the BGP session are: shutdown, password, update-source and modification of timers.
- When using the network statement, BGP looks in the routing table and verifies that the route exists.
- The subnet used in the network statement must match exactly in order to inserted into BGP.

If two BGP peers initiate a session, the router that initiated the session will use a random TCP port to set up the TCP connection, the router that received the session request will use TCP port 179.

- The initiating router message will be OPEN_ACTIVE.
- The receiving router message will be OPEN_PASSIVE.
- The receiving router will also send the OPEN_ACTIVE message, but this will fail because it is no longer needed.

BGP Messages

OPEN	First message sent between peers after a TCP connection has been established. Confirmed by KEEPALIVE message.
KEEPALIVE	Sent periodically to ensure that the connection is live.
NOTIFICATION	Sent in response to errors or special conditions. If a connection encounters an error, a NOTIFICATION message is sent and the connection is closed.
UPDATE	Advertises routes between BGP speakers. Multiple routes that have the same path attributes can be advertised in a single message. A single UPDATE message may simultaneously advertise a feasible route and withdraw multiple unfeasible routes from service.

When BGP fast peering session deactivation is enabled, BGP will monitor the peering session with the specified neighbor.

- Adjacency changes are detected, and terminated peering sessions are deactivated in between the default or configured BGP scanning interval.

BGP States

IDLE	No peering, router is looking for neighbor. Connections refused. Idle (admin) means that the neighbor relationship has been administratively shut down.
CONNECT	Establishing TCP session. Transitions to ACTIVE if TCP session has failed.
ACTIVE	An OPEN message is periodically sent to try to establish the TCP session peering.
OPENSENT	TCP session connection is established. Router is OPEN and expects an OPEN message from peer.
OPENCONFIRM	Expecting either KEEPALIVE (meaning that session is approved), or NOTIFICATION.
ESTABLISHED	Routers have a BGP peering session. UPDATES are sent between BGP speakers.

Table-Map

Wednesday, January 13, 2016
5:07 PM

BGP Table-Map

- Sits between the BGP table and the RIB. Does not filter bgp prefixes received.
- Can be used to match prefixes and apply tags or communities.
- Can also be used to filter BGP prefixes from the local RIB with the filter keyword.
- Typically used to apply QoS settings and tags/communities.
- Applies to all prefixes received from all peers.

Filter specific prefixes from the RIB:

```
ip prefix-list DENY_192 permit 192.168.0.2/32
route-map TABLE_MAP_DENY deny 10
 match ip add prefix DENY_192

router bgp 1
 add ipv4
  table-map TABLE_MAP_DENY filter

clear ip bgp ipv4 unicast table-map
```

Apply QoS to specific communities:

```
ip community-list standard 2:2 permit 2:2
ip community-list standard 3:3 permit 3:3

route-map TABLE_MAP_QOS permit 10
 match community 2:2
 set ip precedence 2
route-map TABLE_MAP_QOS permit 20
 match community 3:3
 set ip precedence 3
route-map TABLE_MAP_QOS permit 99

router bgp 1
 add ipv4
  table-map TABLE_MAP_QOS
```

Verify QoS application to routes:

```
show ip cef 192.168.0.1/32
```

Configure the ingress interface to perform classification based on IPP:

```
int fa0/0
 bgp-policy source ip-prec-map
 bgp-policy destination ip-prec-map
```

BGP Table-Map Automatic-Tag

```
ip as-path access-list 1 permit .*
route-map AS_ORIGIN_TABLE_MAP permit 10
 match as-path 1
 set automatic-tag

router bgp 2
 add ipv4
  table-map AS_ORIGIN_TABLE_MAP

clear ip bgp ipv4 unicast table-map
```

Cryptography

Sunday, December 13, 2015
11:31 AM

Internet Key Exchange (IKE)

- IKEv1 uses UDP port 500. UDP port 4500 is used for NAT-Traversal in IKEv2.
- Encapsulation Header (ESP) is UDP port 50. Authentication Header (AH) is UDP port 51.

IKE Phase 1 (P1)

- IKE authenticates IPsec peers.
- Negotiates IKE Security Associations (SAs).
- Sets up a secure channel for negotiating IPsec SAs in P2.
- Protects the identities of IPsec peers.
- Main mode hides the identities of the two sides, but takes more time to negotiate. Default mode.
- Aggressive mode takes less time to negotiate at the cost of less security.
- MM and AG states differ, aggressive mode skips the SA_SETUP state. Both lead to QM_IDLE.

IKE Phase 2 (P2)

- IKE negotiates IPSec SA parameters.
- Sets up matching IPSec SAs in the peers.
- Sets up protection suite using ESP or AH.
- Two SAs are set up. One for sending and one for receiving traffic.
- Multiple P2 SAs can be established over the same P1 SA.
- Only one state, QM_IDLE.

Main Mode States

- MM_NO_STATE.
- MM_SA_SETUP. The peers have agreed on parameters for the ISAKMP SA.
- MM_KEY_EXCH. Exchanged DH information but remains unauthenticated.
- MM_KEY_AUTH. Authenticated.

Aggressive Mode States

- AG_NO_STATE.
- AG_INIT_EXCH. Exchanged DH information but remains unauthenticated.
- AG_AUTH. Authenticated.

Crypto Maps

Sunday, December 13, 2015
11:25 AM

Crypto Maps

- Requires specification of traffic (ACL) in order to function.
- Applied on physical interface.
- The VPN will only be initiated when traffic is generated that matches the ACL.

```
show crypto map
show crypto isakmp profile

debug crypto isakmp
debug crypto ipsec
```

Main Mode

```
crypto isakmp policy 10
 encr aes 256
 hash sha256
 authentication pre-share
 group 14
 life time 3600

crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode tunnel

crypto isakmp key cisco address 10.0.12.2

crypto map CMAP 10 ipsec-isakmp
 set peer 10.0.12.2
 set transform-set TS
 match address VPN
 qos pre-classify

ip access-list extended VPN
 permit ip host 192.168.0.1 host 192.168.0.2

ip route 192.168.0.2 255.255.255.255 10.0.12.2

int fa0/0
 crypto map CMAP
```

Aggressive Mode

```
crypto isakmp policy 10
 encr aes 256
 hash sha256
 authentication pre-share
 group 14
 life time 3600

crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode tunnel

crypto keyring KEYRING
 pre-shared-key address 10.0.12.2 key cisco

crypto isakmp profile ISAKMP
```

```
 keyring KEYRING
 initiate mode aggressive
 match identity address 10.0.12.2

crypto map CMAP 10 ipsec-isakmp
 set peer 10.0.12.2
 set transform-set TS
 match address VPN
 qos pre-classify
 set isakmp ISAKMP

ip access-list extended VPN
 permit ip host 192.168.0.1 host 192.168.0.2

ip route 192.168.0.2 255.255.255.255 10.0.12.2

int fa0/0
 crypto map CMAP
```

Dynamic VTI

Sunday, December 13, 2015
11:25 AM

Dynamic Virtual Tunnel Interfaces (VTI)

- Provides a separate virtual interface for each VPN session cloned from virtual template.
- When using EIGRP the split-horizon and next-hop-self configuration needs to be placed on the virtual-template.
- The VPN will be initiated even if no traffic is generated.

Hub configuration:

```
crypto isakmp policy 10
 encryption aes 256
 hash sha256
 authentication pre-share
 group 14
 lifetime 3600

crypto keyring KEYRING
 pre-shared-key address 123.0.0.2 key cisco
 pre-shared-key address 123.0.0.3 key cisco

crypto isakmp profile ISAKMP
   keyring KEYRING
   match identity address 123.0.0.2
   match identity address 123.0.0.3
   virtual-template 123

crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode tunnel

crypto ipsec profile IPSEC
 set transform-set TS
 set isakmp-profile ISAKMP

int lo1
 ip add 10.0.123.1 255.255.255.0

int virtual-template 123 type tunnel
 ip unnumbered loopback 1
 tunnel mode ipsec ipv4
 tunnel protection ipsec profile IPSEC
 no ip next-hop-self eigrp 123
 no ip split-horizon eigrp 123

router eigrp 123
 network 10.0.123.0 0.0.0.255
 network 192.168.0.1 0.0.0.0
```

Spoke configuration:

```
crypto isakmp policy 10
 encr aes 256
 hash sha256
 authentication pre-share
 group 14
 lifetime 3600

crypto isakmp key cisco address 123.0.0.1
```

```
crypto isakmp key cisco address 123.0.0.3

crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode tunnel

crypto ipsec profile IPSEC
 set transform-set TS

interface tun0
 ip add 10.0.123.2 255.255.255.0
 tunnel source fa0/0
 tunnel destination 123.0.0.1
 tunnel mode ipsec ipv4
 ip mtu 1400
 tunnel protection ipsec profile IPSEC

router eigrp 123
 network 10.0.123.0 0.0.0.255
 network 192.168.0.2 0.0.0.0
```

Static VTI

Sunday, December 13, 2015
11:25 AM

Static Virtual Tunnel Interfaces (VTI)

- Default tunnel mode is GRE, optionally change to IPsec IPv4 / IPv6.
- The VPN will be initiated even if no traffic is generated.

```
crypto isakmp policy 10
 encr aes 256
 hash sha256
 authentication pre-share
 group 14
 lifetime 3600

crypto isakmp key cisco address 12.0.0.2

crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode tunnel

crypto ipsec profile IPSEC
 set transform-set TS

int tun0
 ip add 10.0.12.1 255.255.255.0
 tunnel source 12.0.0.1
 tunnel destination 12.0.0.2
 ip mtu 1400
 tunnel protection ipsec profile IPSEC
 qos pre-classify

router eigrp 12
  network 10.0.12.0 0.0.0.255
  network 192.168.0.1 0.0.0.0
```

DMVPN

Monday, December 14, 2015
1:26 PM

Dynamic Multipoint VPN (DMVPN) Phases

Phase	GRE-Mode	Dynamic Tunnels	Summarization / Default-Route
1	mGRE on hub, GRE on spokes	no	Allowed
2	mGRE on all	yes	Allowed, but lose dynamic tunnel functionality
3	mGRE on all	yes	Allowed

DMVPN Phase 1

- Basically traditional Hub-and-Spoke topology without dynamic tunnels.
- Configure spokes with tunnel destination <hub nbma address> and tunnel mode gre on the tunnel interface.

DMVPN Phase 3 Additions

- The ip nhrp redirect is configured on the hub and works similar to IP redirect. Informs spokes of the location of others.
- When a hub receives and forwards packet out of same interface it will sent a NHRP redirect message back to the source.
- The original packet from the source is not dropped but forwarded down to other spoke via RIB.
- The ip nhrp shortcut is configured on spokes and rewrites the CEF entry after getting the redirect message.

Next Hop Resolution Protocol (NHRP)

- ARP-like protocol that dynamically maps a NBMA network.
- NHRP works similar to Frame-Relay DLCI but instead maps NMBA addresses to tunnel addresses.
- The hub can only communicate with spokes after spokes have registered to the hub.
- The hub is the Next Hop Server (NHS) and the spokes are the Next Hop Clients (NHCs).
- NHRP allows one NHC to dynamically discover another NHC within the network and build tunnels dynamically (resolution).

NHRP Dynamic Flags

Authoritative	Obtained from NHS.
Implicit	Obtained from forwarded NHRP packet .
Negative	Could not be obtained.
Unique	Request packets are unique, disable if spoke has a dynamic outside IP address.
Registered	Obtained from NHRP registration request (Seen on hub). The spoke has instructed the hub not to take a registration from another identical NBMA address.
Used	Set when data packets are process switched and mapping entry is in use, 60s timer.
Router	NHRP mapping entries for the remote router for access to network.
Local	Local network mapping.

BGP

Monday, December 14, 2015
1:26 PM

iBGP DMVPN Recommendations

- Configure Hub as Route-Reflector.
- Configure Spokes as Route-Reflector Clients.
- Use Peer-Groups for scalable config.

P3 hub configuration:

```
int fa0/0
 ip add 123.0.0.1 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip address 10.0.123.1 255.255.255.0
 ip nhrp redirect

router bgp 123
 bgp listen range 10.0.123.0/24 peer-group DMVPN
 bgp listen limit 2
 neighbor DMVPN peer-group
 neighbor DMVPN remote-as 123
 address-family ipv4
  neighbor DMVPN route-reflector-client
  neighbor DMVPN activate
```

P3 spoke configuration:

```
int fa0/0
 ip add 123.0.0.2 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoin
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast 123.0.0.1
 ip nhrp nhs 10.0.123.1 nbma 123.0.0.1
 ip address 10.0.123.2 255.255.255.0
 ip nhrp shortcut

router bgp 123
 neighbor 10.0.123.1 remote-as 123
 address-family ipv4
  neighbor 10.0.123.1 activate
```

EIGRP

Monday, December 14, 2015
1:26 PM

EIGRP DMVPN Recommendations

Phase	Next-Hop-Self	Split-Horizon
1	Enabled	Disable when using specific routes Enable when using default summary
2	Disable	Disable
3	Enabled	Disable when using specific routes Enable when using default summary

EIGRP add-path support for DMVPN

- Enables hubs to advertise up to four best paths to connected spokes.
- Disable next-hop-self for add-paths to operate.
- Can only be enabled named configuration.
- Should not be configured alongside variance.

P3 hub configuration:

```
int fa0/0
 ip add 123.0.0.1 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip address 10.0.123.1 255.255.255.0
 ip nhrp redirect

router eigrp DMVPN
 add ipv4 au 123
  network 10.0.123.0 0.0.0.255
 af-interface tun0
  summary-address 0.0.0.0/0
  no next-hop-self no-ecmp-mode
  no split-horizon
  add-paths 4
```

P3 spoke configuration:

```
int fa0/0
 ip add 123.0.0.2 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast 123.0.0.1
 ip nhrp nhs 10.0.123.1 nbma 123.0.0.1
 ip address 10.0.123.2 255.255.255.0
 ip nhrp shortcut
```

```
router eigrp DMVPN
 add ipv4 au 123
  network 10.0.123.0 0.0.0.255
```

Misc

Tuesday, December 15, 2015
9:21 AM

DMVPN Encryption

- Mode tunnel adds 20 bytes overhead and is only used in a multi-tier DMVPN hub.
- One set of routers running cryptography and another set performing NHRP services.
- Mode transport is preferred for single-hub configurations.

```
crypto ipsec transform-set TS esp-aes 256 esp-sha256-hmac
 mode transport
```

DMVPN QoS

- Apply a QoS policy on a DMVPN hub on a tunnel instance in the egress direction.
- Shape traffic for individual spokes (parent policy). Policy data flows going though the tunnel (child policy).
- Defined by NHRP group, each spoke can be managed individually using Per-Tunnel QoS.
- The spoke can belong to only one NHRP group per GRE tunnel interface.

```
policy-map R2_PM
 class class-default
  shape average 50 m
policy-map R3_POLICY
 class class-default
  shape average 25 m

int tun0
 ip nhrp map group R2 service-policy output R2_PM
 ip nhrp map group R3 service-policy output R3_PM

show policy-map multipoint
```

Spoke configuration:

```
int tun0
 ip nhrp group R2
```

DMVPN PIMv2

- PIM does not allow multicast traffic to leave the same interface it was received on.
- Override this behavior with ip pim nbma-mode configured on the tunnel interface.

ODR

Monday, December 14, 2015
1:26 PM

DMVPN On-Demand Routing (ODR)

- Disable CDP on outside (physical) interfaces. Enable CDP on tunnel interfaces.
- ODR timers and CDP timers should be the same.
- Configure on hub only, spokes will receive a default route via ODR and will advertise their connected subnets.
- In P2 all traffic will still flow through the hub, ODR needs P3 in order to create dynamic tunnels.

P3 hub configuration:

```
int fa0/0
 ip add 123.0.0.1 255.255.255.0

cdp timer 20
cdp holdtime 60

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip address 10.0.123.1 255.255.255.0
 ip nhrp redirect
 cdp enable

router odr
 timers 20 60 60 90
```

P3 spoke configuration:

```
int fa0/0
 ip add 123.0.0.2 255.255.255.0

cdp timer 20
cdp holdtime 60

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast 123.0.0.1
 ip nhrp nhs 10.0.123.1 nbma 123.0.0.1
 ip address 10.0.123.2 255.255.255.0
 ip nhrp shortcut
 cdp enable
```

OSPF

Monday, December 14, 2015
1:26 PM

OSPF DMVPN Recommendations

- It is vital to the workings of OSPF that the hub is the DR in broadcast and non-broadcast networks.
- OSPF is not recommended for DMVPN because it is not possible to summarize within the same area.
- The DMVPN topology is most likely going to be part of the backbone area.
- PTMP will not create dynamic tunnels in P2. Broadcast is recommended for P2.
- PTMP is recommended for P1 and P3.

P3 hub configuration:

```
int fa0/0
 ip add 123.0.0.1 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip address 10.0.123.1 255.255.255.0
 ip ospf network point-to-multipoint
 ip ospf 123 area 0
 ip nhrp redirect
```

P3 spoke configuration:

```
int fa0/0
 ip add 123.0.0.2 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast 123.0.0.1
 ip address 10.0.123.2 255.255.255.0
 ip nhrp nhs 10.0.123.1 nbma 123.0.0.1
 ip ospf network point-to-multipoint
 ip ospf 123 area 0
 ip nhrp shortcut
```

OSPF DMVPN Filter

- You can only filter routes from being added to the RIB, not LSAs from being received.
- Basically filter every route on the spokes except for the default route.

Hub configuration:

```
router ospf 123
 default-information originate always
```

Spoke configuration:

```
access-list 1 permit host 0.0.0.0
router ospf 123
 distribute-list 1 in
```

RIP

Monday, December 14, 2015
1:26 PM

RIP DMVPN Recommendations

Phase	Split-Horizon
1	Disable when using specific routes Enable when using default summary
2	Disable
3	Disable when using specific routes Enable when using default summary

P3 hub configuration:

```
int fa0/0
 ip add 123.0.0.1 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip address 10.0.123.1 255.255.255.0
 ip nhrp redirect

router rip
 version 2
 network 10.0.0.0
 no auto-summary
```

P3 spoke configuration:

```
int fa0/0
 ip add 123.0.0.2 255.255.255.0

int tun0
 tunnel source fa0/0
 tunnel mode gre multipoint
 ip nhrp network-id 123
 ip mtu 1400
 ip nhrp map multicast 123.0.0.1
 ip nhrp nhs 10.0.123.1 nbma 123.0.0.1
 ip address 10.0.123.2 255.255.255.0
 ip nhrp shortcut

router rip
 version 2
 network 10.0.0.0
 no auto-summary
```

IPv6

Monday, December 14, 2015
1:27 PM

IPv6 DMVPN over IPv4 NBMA

- IPv4 for the physical interface.
- IPv6 for the tunnel interface.
- Uses IPv4 NBMA address on nhs command.
- Uses IPv6 NHRP commands.

P3 hub configuration:

```
int fa0/0
 ip add 10.0.123.1 255.255.255.0

interface tun0
 tunnel source 10.0.123.1
 tunnel mode gre multipoint ipv6
 ipv6 nhrp map multicast dynamic
 ipv6 nhrp network-id 123
 ipv6 mtu 1400
 ipv6 address FE80::1 link-local
 ipv6 address 2001:10:0:123::1/64
 ipv6 nhrp redirect
```

P3 spoke configuration:

```
int fa0/0
 ip add 10.0.123.2 255.255.255.0

interface tun0
 tunnel source 10.0.123.2
 tunnel mode gre multipoint
 ipv6 nhrp map multicast 123.0.0.1
 ipv6 nhrp network-id 123
 ipv6 nhrp nhs 2001:10:0:123::1 nbma 123.0.0.1
 ipv6 mtu 1400
 ipv6 address FE80::2 link-local
 ipv6 address 2001:10:0:123::2/64
 ipv6 nhrp shortcut
```

IPv6 DMVPN over IPv6 NBMA

- IPv6 for the physical interface.
- IPv6 for the tunnel interface.
- Uses IPv6 NBMA address on nhs command.
- Uses IPv6 NHRP commands.

P3 hub configuration:

```
int fa0/0
 ipv6 add 2001:123::1/64
 ipv6 address FE80::1 link-local

int tun0
 ipv6 mtu 1400
 tunnel source 2001:123::1/64
 tunnel mode gre multipoint ipv6
 ipv6 address FE80::1 link-local
 ipv6 address 2001:10:0:123::1/64
 ipv6 mtu 1400
 ipv6 nhrp map multicast dynamic
```

```
 ipv6 nhrp network-id 123
 ipv6 nhrp redirect
```

P3 spoke configuration:

```
int fa0/0
 ipv6 add 2001:123::2/64
 ipv6 address FE80::2 link-local

int tun0
 ipv6 mtu 1400
 tunnel source 2001:123::2/64
 tunnel mode gre multipoint ipv6
 ipv6 address FE80::2 link-local
 ipv6 address 2001:10:0:123::2/64
 ipv6 nhrp map multicast 2001:123:0:0::1
 ipv6 nhrp network-id 123
 ipv6 nhrp nhs 2001:10:0:123::1 nbma 2001:123::1
 ipv6 mtu 1400
 ipv6 nhrp shortcut
```

IPv4 DMVPN over IPv6 NBMA

- IPv6 for the physical interface.
- IPv4 for the tunnel interface.
- Uses IPv6 NBMA address on nhs command.
- Uses IPv4 NHRP commands.

P3 hub configuration:

```
int fa0/0
 ipv6 add 2001:123::1/64
 ipv6 address FE80::1 link-local

interface Tun 0
 ip mtu 1400
 tunnel source 2001:123::1
 tunnel mode gre multipoint ipv6
 ip nhrp map multicast dynamic
 ip nhrp network-id 123
 ip address 10.0.123.1 255.255.255.0
 ip nhrp redirect
```

P3 spoke configuration:

```
int fa0/0
 ipv6 add 2001:123::2/64
 ipv6 address FE80::2 link-local

interface Tun 0
 ip mtu 1400
 tunnel source 2001:123::2
 tunnel mode gre multipoint ipv6
 ip nhrp map multicast dynamic
 ip nhrp network-id 123
 ip address 10.0.123.2 255.255.255.0
 ip nhrp nhs 10.0.123.1 nbma 2001:123::1
 ip nhrp shortcut
```

EIGRP

Monday, December 14, 2015
1:27 PM

EIGRP Vector Metrics

- Changing the bandwidth or delay on the interface will cause the router to re-advertise the route.
- Changing the load or reliability will not cause to router to re-advertise the route.

EIGRP Best Path Selection

- EIGRP uses Reliable Transport Protocol (RTP) for guaranteed, ordered delivery of EIGRP packets to all neighbors.
- Supports multicast and unicast, and uses IP protocol 88. Multicast address is 224.0.0.10 and FF02::A.
- Only some EIGRP packets are sent reliably. For efficiency, reliability is provided only when necessary.
- EIGRP updates are sent to the multicast address and acknowledgements are replied via unicast.

EIGRP Timers

- The configured hold-time is communicated to the neighbor on the segment. This hold-time is included in the hello message.
- The neighbor receives this and will expect a new hello from the router within this time. These timers do not have to match.

Filtering

Monday, December 14, 2015
1:27 PM

Extended Access-List

Deny even from R2 and odd from R3:

```
ip access-list extended 100
 deny ip host 10.0.12.2 4.0.0.0 0.0.0.254
 deny ip host 10.0.13.3 4.0.0.1 0.0.0.254
 permit ip any any

router eigrp 1
 distribute-list 100 in
```

Prefix-List

Deny all prefixes from R2:

```
ip prefix-list DENY_R2 deny 10.0.12.2/32
ip prefix-list DENY_R2 permit 0.0.0.0/0 le 32
ip prefix-list PREFIXES permit 0.0.0.0/0 le 32

router eigrp 1
  distribute-list prefix PREFIXES gateway DENY_R2 in
```

Only accept prefixes from R3:

```
ip prefix-list R3 permit 10.0.13.3/32

router eigrp 1
 distribute-list gateway R3 in
```

Deny specific prefixes from R2:

```
ip prefix-list R2 permit 10.0.12.2/32
ip prefix-list NETWORKS deny 4.0.0.1/32
ip prefix-list NETWORKS deny 4.0.0.2/32
ip prefix-list NETWORKS permit 0.0.0.0/0 le 32

router eigrp 1
  distribute-list prefix NETWORKS gateway R2 in fa0/0
```

Metric

Monday, December 14, 2015
1:27 PM

EIGRP Composite Metric (Weight Calculation)

- EIGRP uses metric weights along with a set of vector metrics to compute the composite metric for local RIB and route selections.
- Type of service (first K value) must always be zero.
- The formula is [K1*bandwidth + (K2*bandwidth)/(256 - load) + K3*delay] * [K5/(reliability + K4)].

256*[(10^7/slowest bandwidth in kbps) + all link delays in tens microseconds]

```
router eigrp
 metric weights 0 1 0 1 0 0
```

EIGRP Wide Metrics

- The EIGRP Wide Metric feature supports 64-bit metric calculations and Routing Information Base (RIB) scaling.
- The lowest delay that can be configured for an interface is 10 microseconds with 32-bit metrics (normal).
- Use on interfaces that are faster than 1Gbps.
- The new metrics no longer fit in the output of the RIB, because this is limited to 32bit numbers.
- The topology table will show the correct metric numbers which is divided by the value specified in the rib-scaling.
- The formula is [K1*bandwidth + (K2*bandwidth)/(256 - load) + K3*delay +K6*Ext Attr] * [K5/(reliability + K4)]
- K6 is an additional K value for future use. Other K values remain the same with K1 and K3 set to 1, and K2,K4,K5 set to 0.

```
router eigrp EIGRP
 address-family ipv4 unicast autonomous-system 1
  metric rib-scale 128
```

The 64-bit metric calculations work only in EIGRP named mode configurations. EIGRP classic mode uses 32-bit metric calculations.

- EIGRP named mode automatically uses wide metrics when speaking to another EIGRP named mode process.

EIGRP Offset Lists

- The offset list increases the existing metric by a specified amount. It is not possible to decrease a metric.
- The offset list modifies the delay value, not the bandwidth value. Meaning that it is included in the cumulative delay.
- Using offset-list 0 will apply the offset to all networks. Using an empty access list has the same effect.

An offset list will only influence the calculated metric, and thus the composite metric. The composite metric is only used for local calculation and is not communicated to neighbors. However, routing paths through the router between neighbors will still add the offset value to the metric. This is because the offset list changes the delay value which is cumulative in the total metric for the route.

EIGRP Unequal Cost Multi-Path Load-Sharing

- Only feasible successors are candidate for load balancing.
- The traffic-share min accross-interfaces only uses the primary path in case of a UCMP configuration using variance.
- This will stop the route from using UCMP. The point of this configuration is to speed up convergence.
- When calculating the eigrp traffic share count, divide numbers to figure out metric.
- A preffered traffic share count of two routes, one 10 the other 3 -> 10/3 = 3.33 * metric value.

```
router eigrp 1
 variance 5
 traffic-share min accross-interfaces
```

Default Metric & Redistribution

- Only connected and static routes can be redistributed without a default metric. This metric is set to 0.
- IPv6 networks add the include-connected keyword to include connected networks as well.

- In IPv4 connected networks are included in the redistribution by default.

Metric Maximum-Hops

- The maximum hops is a greater than statement. Default is 100 hops.
- If 10 is entered for example, the prefixes 10 hops away are still valid. The prefixes 11 hops away are denied.

```
router eigrp 1
 metric maximum-hops 10
```

Misc

Monday, December 14, 2015
1:27 PM

EIGRP Neighbor Statement

- Limit neighbor ship with specific neighbors by configuring neighbor statements (on the same segment).
- Alternative is to use ACL to block specific link-local addresses (IPv6).

```
ipv6 router eigrp 1
 neighbor FE80::2 fa0/0
 neighbor FE80::3 fa0/0
```

Or:

```
ipv6 access-list EIGRP_BLOCK
 deny 88 fe80::2/128  any
 deny 88 fe80::3/128  any
 permit ipv6 any any

int fa0/0
 ipv6 traffic-filter EIGRP_BLOCK in
```

EIGRP Router-ID

- Duplicate router-IDs (RID) do not show up in the logging or debug. Use the show ip eigrp events | i Ignored command.
- EIGRP uses the concept of the RID as a loop-prevention mechanism to filter out a routers own routes.
- In the event of a duplicate RID the neighbors routes will not be installed.

EIGRP Authentication

- It is possible to configure multiple keys. Only one authentication packet is sent, regardless of how many keys exist.
- The software examines the key numbers in the order from lowest to highest, and uses the first valid key that it encounters.
- HMAC-SHA-256 authentication is only available in named mode. Does not support key-chains.
- Named mode ignores all authentication (and other) commands configured on the local interface using the old method.

```
key chain EIGRP_KEY
 key 1
  key-string cisco

int fa0/0
 ip authentication key-chain eigrp 1 EIGRP_KEY
 ip authentication mode eigrp 1 md5
```

HMAC-SHA-256 using named configuration:

```
router eigrp EIGRP
 address-family ipv4 autonomous-system 1
  af-interface fa0/0
   authentication mode hmac-sha-256 cisco
```

Dampening Change & Interval

- Dampening controls the update of metric changes of routes advertised by neighbors.
- Dampening compares the old metric for the route with the new metric.
- Using dampening-change, if this new metric is within the percentage threshold, the update will be ignored.
- Using dampening-interval, if this new metric is within the configured time interval, the update will be ignored.
- Dampening is disabled by default.

```
int fa0/0
```

```
ip dampening-change eigrp 1 percent 75
ip dampening-interval eigrp 1 seconds 60
```

Next-Hop-Self

- Disable next-hop-self on the redistributing router when using 3rd party next hop.
- This is used when OSPF and EIGRP coexist on the same segment, and one router is used for redistributing both protocols.
- Normally all traffic would go through the redistributing router, with 3rd party next hop the neighbors can communicate directly.
- Requirement is to disable next-hop-self on the interface of the redistributing router towards the shared segment.
- Other situations where next-hop-self might be disabled is in DMVPN solutions.

```
int fa0/0
 description SHARED_OSPF_EIGRP_SEGMENT
 ip add 10.0.123.1 255.255.255.0
 no ip next-hop-self eigrp 1

router eigrp 1
 network 10.0.123.0 0.0.0.255
 redistribute ospf 1

router ospf 1
 network 10.0.123.0 0.0.0.255 area 0
 redistribute eigrp 1 subnets
```

Bandwidth Percentage

- By default, EIGRP packets are allowed to consume a maximum of 50 percent of the link bandwidth.
- The bandwidth is the configured bandwidth, not the original interface bandwidth.
- Note that values greater than 100 percent may be configured.
- This configuration option may be useful if the bandwidth is set artificially low for other reasons.

```
int fa0/0
 ip bandwidth-percent eigrp 75
```

Administrative Distance

- Changes to the administrative distance of all internal and external routes is limited to the router itself.
- Can be limited to a specific neighbor, only one specific distance command can be entered per neighbor.

```
ip access-list standard NETWORK
 permit host 172.16.0.0

router eigrp 1
 distance eigrp 90 170
 distance 85 10.0.12.2 0.0.0.0
 distance 75 10.0.13.3 0.0.0.0 NETWORK
```

EIGRP Loop-Free Alternate Fast Reroute (FRR)

- Uses repair paths or backup routes and installs these paths in the RIB.
- FRR picks the best feasible successor and places it in the FIB as a backup route.

```
router eigrp EIGRP
 address-family ipv4 autonomous-system 1
  topology base
   fast-reroute per-prefix all
```

Route-Tags

Monday, December 14, 2015
1:27 PM

Route Tag Enhancements

- The route tag enhancements allow the route tag to be formatted as a dotted decimal tag.
- These can me matched either directly (in the traditional route tag method in route-map) or via a route-tag list.
- EIGRP named mode provides the option of assigning a default-route-tag to all routes sourced by the router.

```
route-tag notation dotted-decimal
ip access-list standard Lo1
 permit host 4.0.0.1
ip access-list standard Lo2
 permit host 4.0.0.2

ip prefix-list Lo3 permit 4.0.0.3/32

route-map LOOPBACKS permit 10
 match ip address Lo1
 set tag 44.44.1.1
route-map LOOPBACKS permit 20
 match ip address Lo2
 set tag 44.44.2.1
route-map LOOPBACKS permit 30
 match ip address prefix-list Lo3
 set tag 44.44.3.1
route-map LOOPBACKS permit 40
 match interface loopback 4
 set tag 44.44.4.1

router eigrp 1
 redistribute connected route-map LOOPBACKS
```

Route-tag Even Filtering:

```
route-tag notation dotted-decimal
route-tag list LOOPBACKS permit 44.44.0.0 0.0.254.255

route-map LOOPBACKS permit 10
 match tag list LOOPBACKS
 set metric 50000 581 255 1 1500
route-map LOOPBACKS permit 20

router eigrp 1
 distribute-list route-map LOOPBACKS in fa0/0
```

Default Route-Tags:

```
router eigrp EIGRP
 address-family ipv4 unicast autonomous-system 1
  eigrp default-route-tag 192.168.0.1
```

Summarization

Monday, December 14, 2015
1:27 PM

EIGRP Route Summarization

- Summary routes are always internal, even if external routes are summarized.
- The summary route has an AD of 5 and is called the discard route which points to Null0 on the summarizing router.
- The neighbor will receive an internal route with an AD of 90.
- Disable the discard route by specifying a summary-metric distance of 255.
- Configure a leak-map to allow components of the summary to advertised alongside the summary.
- Stop IP ICMP unreachables based on discard route with no ip unreachables configured on null0 interface.

```
int fa0/0
 ip summary-address eigrp 1 0.0.0.0/0

router eigrp 1
 summary-metric 0.0.0.0/0 distance 255
```

Leak-Map

Allow the 10.0.12.0/24 network to be advertised alongside the summary:

```
ip prefix-list NETWORK permit 10.0.12.0/24
route-map LEAK_MAP permit 10
 match ip address prefix-list NETWORK

int fa0/0
 ip summary-address eigrp 1 10.0.0.0/16 leak-map LEAK_MAP
```

Stub

Monday, December 14, 2015
1:27 PM

EIGRP Stub

- EIGRP query messages live for 3 minutes by default. Can be modified with the timers active-time command.
- Queries are not sent if a feasible successor is present in the topology for the route.
- A router will wait for a reply for all neighbors before failing over to a different path. This is called Stuck in Active (SIA).
- Fix SIA with summary routes or stub routers.
- A leak map can allow routes that would normally have been suppressed.

Allow prefixes from 172.16.1.0/24 alongside the connected and summary routes:

```
int fa0/0
 ip summary-address eigrp 1 172.16.0.0/16

ip prefix-list STUB_PREFIX permit 172.16.1.0/24
route-map STUB_LEAK_MAP permit 10
 match ip address prefix STUB_PREFIX

router eigrp 1
 eigrp stub connected summary leak-map STUB_LEAK_MAP
```

Frame-Relay (v4)

Monday, December 14, 2015
1:28 PM

L1 and Frame-Relay Terminology

- L1 DTE = Data Termination Equipment
- L1 DCE = Data Communication Equipment
- Frame-Relay switch is called the DTE, endpoints are called DCE. This is unrelated to L1 DCE/DTE.
- FR DCE responds to LMI inquiries by sending LMI status, never sends LMI inquiries.
- FR DTE sends LMI inquiries, never sends LMI status.
- DLCIs are only locally significant.

Inverse ARP

- Because serial interfaces and FR are NBMA networks, there is no ARP flooding to discover neighbor addresses.
- Instead FR relies on inverse ARP to allow clients to notify neighbors of their presence and reply to requests.
- Disabling inverse ARP will disable the notification, these are called dynamic mappings.
- Has to be disabled on all routers to disable automatic learning of mappings.
- Inverse-ARP does not allow self-ping.

```
int s1/0
 encapsulation frame-relay
 no frame-relay inverse-arp

clear frame inarp
```

Frame-Relay Mappings

- If dynamic mappings have been disabled, static ones need to be created on the clients (R1,R3).
- The broadcast keyword allows broadcast packets over NBMA.
- Only configure broadcast on one statement, preferably the statement that configures the router to ping itself.

```
int se1/0
 ip address 10.0.13.1 255.255.255.0
 frame-relay map ip 10.0.13.1 103 broadcast
 frame-relay map ip 10.0.13.3 103

int se1/0
 ip address 10.0.13.3 255.255.255.0
 frame-relay map ip 10.0.13.1 301
 frame-relay map ip 10.0.13.3 301 broadcast

show frame pvc
show frame map
```

Frame-Relay Switch

- Set encapsulation to frame-relay.
- Configure DCE on interfaces
- Create FR PVCs using either frame routes (FR switching framework) or connections (L2VPN framework).

Create FR PVCs using frame routes on FR Switch:

```
int se1/0
 encapsulation frame-relay
 frame-relay intf-type dce
 frame route 103 interface se1/1 301

int se1/1
 encapsulation frame-relay
 frame-relay intf-type dce
```

```
 frame route 301 interface se1/0 103

show frame route
show frame pvc
```

Create FR PVCs using connections on FR Switch:

```
connect R1-R3 se1/0 103 se1/1 301
show connection all
```

Frame-Relay with Sub-interfaces

- If sub-interfaces are being used, the DLCI needs to be linked to the sub-interface.
- Inverse ARP configuration is not inherited by sub-interfaces.
- With PTP mode sub-interfaces are configured on the hub, each to single location.
- With PTMP mode a single interface is configured on the hub to multiple locations. Disable split-horizon if using EIGRP.
- The no inverse-arp configuration on the physical interface is not inherit by the sub-interface.

```
int se1/0.123 point-to-multipoint
 frame-relay interface-dlci 123

int se1/0.102 point-to-point
 frame-relay interface-dlci 102
int se1/0.103 point-to-point
 frame-relay interface-dlci 103
```

Frame-Relay Authentication

- Use point-to-point sub-interfaces configured with PPP encapsulation.
- Use ip unnumbered in order to enable self-ping.

R1 configuration:

```
username R3 password cisco
int se1/0.103 point-to-point
 frame-relay interface-dlci 103 ppp virtual-template 1

interface virtual-template 1
 ip address 10.0.13.1 255.255.255.0
 encapsulation ppp
 ppp authentication chap
 ppp pap sent-username R1 password cisco
```

R3 configuration:

```
username R1 password cisco
int s1/1.301 point-to-point
 frame-relay interface-dlci 301 ppp virtual-template 1

interface virtual-template 1
 ip address 10.0.13.3 255.255.255.0
 encapsulation ppp
 ppp chap hostname R2
 ppp chap password cisco
 ppp authentication pap
```

IP Routing

Tuesday, December 8, 2015
8:59 PM

Administrative Distance and Route Selection

- Hardcoded original administrative distance will win if different routing protocols are configured to use the same AD.
- The metrics between different routing protocols or different routing processes are not compared in the route selection.
- When receiving the same route from different OSPF processes with the same AD, the route learned first wins.
- OSPF does not differentiate between internal and external routes to the same destination. Only the AD matters.
- When receiving the same route from different EIGRP ASs with the same AD, the route from the lower AS wins.

IP Source-Routing

- Allows the originator of a packet to decide which routers the packet will flow through.
- Basically a custom path of all hops specified at the source and set in the actual IP header by the source.
- Enabled by default but is a security risk. Disable with no ip source-route command.

IP Accounting

- Counts the number of IP packets and logs source/destination.
- Only works for transit egress traffic, not local traffic.

```
int fa0/0
 ip accounting output-packets

show ip accounting
```

IP Redirects

- ICMP redirect messages are send by default when routers have to forward a packet on the same interface it was received.
- Routers will notify hosts of better next-hop through redirects.

```
int fa0/0
 no ip redirects
 no ipv6 redirects
```

IP Unreachables

- By default the router will respond with an IP unreachable ICMP message in case the neighbor router pings an unknown address.
- Disable to block UDP port scans. These have a destination of an unused or unreachable UDP port.
- IP packets for unknown destinations are sent to null0. Disable redirects for unknown traffic on the null0 interface.
- Other interfaces will still respond with unreachable if traffic is destined to the local interface address.

```
int fa0/0
 no ip unreachables
 no ipv6 unreachables

int null0
 no ip unreachables
 no ipv6 unreachables
```

IP Local Proxy-ARP

- Enable proxy of ARP request on the same subnet with the ip local proxy-arp interface command.
- The ip proxy-arp feature is for ARP requests to different subnets. Enabled by default.
- Use in Private-VLANs to allow communication between isolated hosts. Configure on the promiscuous port.
- Instead of configuring local proxy-arp, you can also statically configured the IP to MAC mappings for individual hosts.

```
arp 10.0.123.2 abcd.1234.abcd arpa
```

```
arp 10.0.123.3 1234.abcd.1234 arpa
```

IP Directed-Broadcast (SMURF)

- Disabled by default. Exploited in SMURF attacks.
- Enable with `ip directed-broadcasts` command on interfaces.

Gratuitous ARP

- Update ARP tables after a MAC address for an IP changes, or a MAC address is now on a different port.
- Sends special ARP packet when interface goes up to notify other hosts in advance so that ARP requests are not needed.
- Does not expect a reply. When a reply is received there is an IP address conflict in the network.
- Used by FHRP to update MAC tables on L2 devices with the virtual MAC address.

IP Event Dampening

- Suppress flapping interface effects on routing protocols and routing tables.
- Can only be configured on physical interfaces, not on sub-interfaces or virtual-templates.

```
int fa0/0
 dampening 5 1000 2000 20

show interfaces dampening
```

Redistribution

Tuesday, December 8, 2015
8:53 PM

Redistribution

- Redistribution only redistributes routes that are present in the RIB.
- There is no direct redistribution between protocols.
- Routing protocol redistribution also redistributes the connected networks that the protocol is enabled for.
- Include IGP interfaces when filtering redistributed connected routes (loopbacks).
- Another way to include the connected interfaces is to advertise them into the protocols and optionally configure as passive.
- OSPF default static route cannot be redistributed with the redistribute static command, even if a route-map is specified.
- Always redistribute the default route into OSPF using the default-information originate command.

BGP Redistribution

- Only internal OSPF routes will be redistributed into BGP by default.
- A default route learned by an IGP is not redistributed into BGP by default.
- Advertise the redistributed default route with network 0.0.0.0 mask 0.0.0.0.

```
ip route profile
show ip route profile
debug ip routing
```

Redistribution using Direct Tags

```
route-map EIGRP_ROUTES deny 10
 match tag 90
route-map EIGRP_ROUTES permit 99

router ospf 1
 redistribute eigrp 1 subnets tag 90
  distribute-list route-map EIGRP_ROUTES in
```

Mutual MultiPoint Redistribution using Prefix-Lists

```
ip prefix-list OSPF_ROUTES permit 3.0.0.1/32
ip prefix-list OSPF_ROUTES permit 3.0.0.2/32
ip prefix-list OSPF_ROUTES permit 3.0.0.3/32

ip prefix-list EIGRP_ROUTES permit 4.0.0.1/32
ip prefix-list EIGRP_ROUTES permit 4.0.0.2/32
ip prefix-list EIGRP_ROUTES permit 4.0.0.3/32

route-map EIGRP_TO_OSPF deny 10
 match ip address prefix-list OSPF_ROUTES
route-map EIGRP_TO_OSPF permit 20
 match ip address prefix-list EIGRP_ROUTES

route-map OSPF_TO_EIGRP deny 10
  match ip address prefix-list EIGRP_ROUTES
route-map OSPF_TO_EIGRP permit 20
 match ip address prefix-list OSPF_ROUTES

router eigrp 1
 redistribute ospf 1 metric 1000000 10 255 1 1500 route-map OSPF_TO_EIGRP

router ospf 1
 redistribute eigrp 1 metric-type 1 subnets route-map EIGRP_TO_OSPF
```

Three-Way Redistribution using Tags

```
route-map EIGRP_TO_RIP deny 10
 match tag 120
route-map EIGRP_TO_RIP permit 20
 match tag 110
 set tag 110
route-map EIGRP_TO_RIP permit 30
 set tag 90

route-map RIP_TO_EIGRP deny 10
 match tag 90
route-map RIP_TO_EIGRP permit 20
 match tag 110
 set tag 110
route-map RIP_TO_EIGRP permit 30
 set tag 120

route-map OSPF_TO_RIP deny 10
 match tag 120
route-map OSPF_TO_RIP permit 20
 match tag 90
 set tag 90
route-map OSPF_TO_RIP permit 30
 set tag 110

route-map RIP_TO_OSPF deny 10
 match tag 110
route-map RIP_TO_OSPF permit 20
 match tag 90
 set tag 90
route-map RIP_TO_OSPF permit 30
 set tag 120

route-map OSPF_TO_EIGRP deny 10
 match tag 90
route-map OSPF_TO_EIGRP permit 20
 match tag 120
 set tag 120
route-map OSPF_TO_EIGRP permit 30
 set tag 110

route-map EIGRP_TO_OSPF deny 10
 match tag 110
route-map EIGRP_TO_OSPF permit 20
 match tag 120
 set tag 120
route-map EIGRP_TO_OSPF permit 30
 set tag 90

router eigrp 1
 redistribute ospf 1 metric 1000000 10 255 1 1500 route-map OSPF_TO_EIGRP
 redistribute rip metric 1000000 10 255 1 1500 route-map RIP_TO_EIGRP

router ospf 1
 redistribute eigrp 1 metric-type 1 subnets route-map EIGRP_TO_OSPF
 redistribute rip metric 1 subnets route-map RIP_TO_OSPF
```

```
router rip
 redistribute eigrp 1 metric 3 route-map EIGRP_TO_RIP
 redistribute ospf 1 metric 3 route-map OSPF_TO_RIP
```

VRF-Lite

Monday, December 7, 2015
8:45 AM

VRF-Lite

- Divide interfaces into VRFs and create separate routing tables.
- Do not implement L3VPN afterwards, this is why its called VRF-Lite.

```
vrf definition 10
 add ipv4
vrf definition 172
 add ipv4
int fa0/0
 vrf forwarding 10
 ip add 10.0.12.1 255.255.255.0
int fa0/1
 vrf forwarding 10
 ip add 10.0.13.1 255.255.255.0
int se1/0
 vrf forwarding 172
 ip add 172.0.12.1 255.255.255.0
int se1/1
 vrf forwarding 172
 ip add 172.0.13.1 255.255.255.0

router eigrp EIGRP
 address-family ipv4 unicast vrf 10 autonomous-system 10
  no auto-summary
  network 10.0.12.0 0.0.0.255
  network 10.0.13.0 0.0.0.255
address-family ipv4 unicast vrf 172 autonomous-system 172
  no auto-summary
  network 172.0.12.0 0.0.0.255
  network 172.0.13.0 0.0.0.255
```

Non-VRF neighbors:

```
router eigrp 10
  no auto-summary
  network 10.0.12.0 0.0.0.255
router eigrp 172
  no auto-summary
  network 172.0.12.0 0.0.0.255
```

IP services

Tuesday, December 8, 2015
9:13 PM

Secure Copy Protocol (SCP)

- Requires SSH and AAA Authorization.
- Has to be enabled on both routers to allow mutual copying.

```
aaa new-model
aaa authentication login default local
aaa authorization exec default local
username bpin privilege 15 password cisco
ip scp server enable

copy scp://bpin@10.0.12.1/nvram:startup-config null:
```

RCMD Remote-Copy (RCP) and Remote-Shell (RSH)

- Allow remote users to copy files to router with RCP.
- Allow remote users to execute commands with RSH.
- Server side has two names in the rcmd command.
    - First one must match /user on client.
    - Second one must match client hostname or client remote-username command.
    - The enable keyword allows execution of exec commands.

Configure the server:

```
ip rcmd rcp-enable
ip rcmd rsh-enable
ip rcmd remote-host remoteadmin 10.0.12.2 R2 enable
```

Configure the client:

```
ip rcmd remote-username R2

rsh 10.0.12.1 /user remoteadmin show ip interface brief
copy rcp://remoteadmin@10.0.12.1/nvram:startup-config null:
```

The boot/service config enables auto-loading of configuration files from a network server:

```
service config
ip rcmd remote-username R2
boot network rcp://10.0.12.1/BOOT
boot network tftp:BOOT
```

Local TFTP-Server

- Specify all files that are eligible for TFTP transfer separately.
- Optionally create an alias for the file and limit access.

```
access-list 1 permit host 192.168.0.2
tftp-server nvram:startup-config alias STARTUP 1
```

Configure client:

```
ip tftp source-interface loopback0
copy tftp://10.0.12.1/STARTUP null:
```

DNS Services

- Create individual host entries on the DNS server.

Server:

```
ip domain-lookup
```

ip domain-name lab.local
ip dns server
ip host Server1 2.2.2.2
ipv6 host Server2 2::2

Client:
ip domain-lookup
ip name-server 1.1.1.1
ip name-server 1::1

Configuration Generation Performance Enhancement (Parser)

- Caches interface configuration in memory, thus allowing faster execution of show run, write memory, etc.
- Enable with the parser config cache interface command.

TCP small servers

- Echo (7): Echoes back whatever you type.
- Chargen (19): Generates a stream of ASCII data.
- Discard (9): Throws away whatever you type.
- Daytime (13): Returns system date and time.

telnet 10.0.12.1 19

Terminate on server with:
show tcp brief
clear tcp tcb <tcb-value>

UDP small servers:

- Echo (7): Echoes the payload of the datagram you send.
- Discard (9): Silently pitches the datagram you send.
- Chargen (19): Pitches the datagram you send, and generates a stream of ASCII data.

Misc. Services
The X28 editor is enabled by default:
no service pad

Ensure that abnormally terminated TCP sessions are removed:
service tcp-keepalives-in
service tcp-keepalives-out

The finger service (TCP port 79) gives line information and is disabled by default:
ip finger
service finger

BFD

Bidirectional Forwarding Detection (BFD)

- Requires CEF, sent unicast to UDP 3784.
- Only supports asynchronous mode, must be enabled on both sides.
- Only works for directly connected neighbors, BFD itself has no neighbor detection.
- Is not tied to any routing protocol, and can be used as a generic and consistent failure detection mechanism.
- Parts of BFD can be distributed to the data plane (echo), better than reduced IGP timers that exist only at the control plane.

BFD Echo Mode

- BFD echo packets are are sourced from UDP 3785 and sent to 3785.
- Enabled by default and can be enabled on either side. Does not work alongside ip redirects or uRPF (or IPv6 on CSR).
- Echo mode is supported on single-hop only. The packets are sent on the negotiated BFD timer interval.
- BFD packets are processed in fast switching instead of the control plane.
- Control plane packets are still sent but they are transmitted at the slow timers speed (1000 ms by default).

BFD Timers

- The time at which 'hello' messages are sent is configured with the interval timer.
- The min_rx timer is the receive timer, if no message is received within this time the neighbor is considered timed-out
- The multiplier specifies how many BFD messages can be missed before neighbor interface is considered down.
- BFD timers work like EIGRP. The send and receive timer do not have to match on both sides
- The slower receive timer of the neighbor will decide the value of the local send timer.

```
bfd slow-timers 1000
int gi0/0
 bfd echo
 bfd interval 500 min_rx 500 multiplier 3
```

BFD Authentication

```
key chain BFD_KEY
 key 1
  key-string cisco

bfd-template single-hop BFD
 echo
 interval both 500 multiplier 3
 authentication md5 keychain BFD_KEY

int gi0/0
 bfd template BFD
```

BFD Static

- Static routes that support BFD must specify an egress interface in single-hop mode.
- The neighbor must point back with a static route, or an unassociated route.
- Static routes can be dependent on a group. If one location becomes inaccessible the depending (passive) routes are also removed from the routing table.

Configure R1:

```
ip route static bfd gi0/0 10.0.12.2
ip route 0.0.0.0 0.0.0.0 gi0/0 10.0.12.2
```

Configure R2:

```
ip route static bfd gi0/0 10.0.12.1 unassociate
```

BFD Static Groups

```
ip route static bfd gi0/0 10.0.12.2 group BFD
```

```
ip route 33.33.33.0 255.255.255.0 gi0/1 10.0.13.3
ip route static bfd gi0/1 10.0.13.3 group BFD passive
```

CPPr

Wednesday, December 2, 2015
7:09 PM

IOS Control Plane

- Handles packets that are not CEF switched, meaning the CPU takes time to handle these packets.
- Maintains keep-alives for routing adjacencies.
- Handles traffic directed at the device itself (management traffic).

Control Plane Protection (CPPr)

- Framework that consists of traffic classifiers, protection and policing.
- Improvement over Control Plane Policing (CoPP) by allowing finer policing granularity.
- Management Plane Protection (MPP) is a part of CPPr and is basically just specifying a management-interface.
- Depends on CEF. When disabled, CPPr is disabled on sub-interfaces but not on the aggregate interface.

Control Plane Interfaces

- Host. Handles traffic destined for the router or one of its own interfaces (MGMT, EIGRP, iBGP)
- Transit. Handles software switched IP traffic.
- CEF-Exception. Handles non-IP related packets such as OSPF, eBGP, ARP, LDP and CDP (or packets with TTL <=1) .
- Aggregate interface <cr>. Configuration applied here applies to all the sub-interfaces.

- It is not possible to apply a L3 policy-map to the aggregate and any of the other interfaces at the same time.
- A L3 policy-map applied to the control plane can only use police or drop, not shape...etc.
- The port-filter keyword polices packets going to closed/non-listening TCP/UDP ports.
- The queue-threshold keyword limits the number of protocol packets that are allowed in the input queue.
    - Rate limit OSPF and eBGP on the cef-exception sub-interface, iBGP on the host and EIGRP on the aggregate.

Police all ICMP traffic:

```
ip access-list extended ICMP_ACL
 permit icmp any any

class-map match-all ICMP_CM
 match access-group name ICMP_ACL
policy-map ICMP_PM
 class ICMP_CM
  police 10000 conform-action transmit exceed-action drop

control-plane host
 service-policy input ICMP_PM
```

Drop connections to closed ports:

```
class-map type port-filter match-all CLOSED_PORTS_CM
 match closed-ports

policy-map type port-filter CLOSED_PORTS_PM
 class CLOSED_PORTS_CM
  drop

control-plane host
 service-policy type port-filter input CLOSED_PORTS_PM
```

Queue SNMP traffic to 75 and any other open UDP/TCP ports to 100:

```
class-map type queue-threshold SNMP_CM
 match protocol snmp
class-map type queue-threshold HOST_CM
 match host-protocols
```

```
policy-map type queue-threshold QUEUE_PM
 class SNMP_CM
  queue-limit 75
 class HOST_CM
  queue-limit 100

control-plane host
 service-policy type queue-threshold input QUEUE_PM
```

Rate limit EIGRP traffic (requires egress direction):

```
ip access-list extended EIGRP
permit eigrp any any

class-map match-all EIGRP_CM
 match access-group name EIGRP

policy-map EIGRP_PM
 class EIGRP_CM
  police 10000 conform-action transmit exceed-action drop

control-plane
 service-policy output EIGRP_PM
```

Management Plane Protection (MPP)

Multiple interfaces can be specified for different protocols:

```
control-plane host
 management-interface fa0/0 allow ssh
 management-interface fa0/1 allow snmp
```

DHCPv4

Saturday, December 5, 2015
11:16 AM

DHCP Messages

DHCPDiscover	Sent by client to 0.0.0.0 to find a DHCP server (broadcast).
DHCPOffer	Response to client with DHCP server information and IP address assignment (unicast or broadcast).
DHCPRequest	Sent by client in response to DHCPOffer, client accepts IP address assignment (broadcast).
DHCPAcknowledge	Acknowledgement by the DHCP server (unicast or broadcast).

The DHCPOFFER and DHCPACK are sent broadcast by default. Disable with the no ip dhcp-client broadcast-flag command.

Disable DHCP and do not reply to Bootstrap Protocol request packets received:

```
no service dhcp
ip dhcp bootp ignore
```

DHCP Conflict Logging

Similar function to excluded-addresses. Logs conflicts with a syslog message and stores the address on an exclusion list.

- Conflicted addresses are stored and need clearing or restart to become usable again.
- Enabled by default. Disable with the no ip dhcp conflict logging command.

```
service dhcp
no ip dhcp conflict logging
ip dhcp excluded-address 10.0.123.1 10.0.123.99
ip dhcp excluded-address 10.0.123.200 10.0.123.254
ip dhcp pool DHCP
 default-router 10.0.123.1
 network 10.0.123.0 /24
 dns-server 192.168.0.1

int fa0/0
 no ip dhcp-client broadcast-flag
 ip address dhcp
```

DHCP Reservation

Cisco routers use the Client-Identifier to identify themselves.

- This is a combination of the hardware address, interface name and cisco.
- This client-identifier is then turned into a HEX string and presented to the server.
- Add 00 to the beginning of a client-identifiers hex string.
- The easiest way to acquire the client-identifier is by first giving out a regular DHCP address and debug ip dhcp server packets.

```
ip dhcp pool R2
 host 10.0.123.2 /24
 client-identifier 0063.6973.636f.2d63.6130.332e.3065.3434.2e30.3030.382d.4661.302f.30

ip dhcp pool R3
 host 10.0.123.3 /24
 client-identifier 0063.6973.636f.2d63.6130.342e.3138.6638.2e30.3030.382d.4661.302f.30
```

DHCP Relay Agent and Information Option 82

- Relay agents receive DHCP messages and then generate a new DHCP message to send out on another interface.

- DHCP option 82 (relay information option) identifies hosts by both the MAC-Address and the switchport. Disabled by default on routers, enabled by default on switches. It allows DHCP relays to inform the DHCP server where the original request came from.
- The reply from the server is forwarded back to the client after removing option 82.
- Enable the option with either the global or interface level command, ip dhcp relay information option. (interface takes preference)

```
service dhcp
int fa0/0
 description CLIENTS
 ip helper address 10.0.12.1
```

A DHCP relay agent may receive a message from another DHCP relay agent that already contains relay information (Relayed twice).

- By default, the relay information from the previous agent is replaced. Customize with ip dhcp relay information policy command.
- If the information policy is changed, also disable the information check with the no ip dhcp relay information check command.

EEM

Tuesday, December 8, 2015
9:13 PM

Embedded Event Manager (EEM)

- The skip keyword prevents the command from being executed. Default is skip no.
- The sync keyword runs the script before the command. Default is sync yes.
- _exit_status 1 means that the command is run.
- _exit_status 0 means that the command is skipped.
- $_cli_result is the outcome of a cli command that was executed, can be pasted into the console with the puts keyword.
- $_cli_msg is the pattern matched with the event keyword. Can be pasted into the console with the syslog msg keyword.
- The cli command is not executed until the EEM policy exits.

```
show event manager policy registered
debug event manager action cli
debug event manager action mail
```

EEM Examples

Disable show running-config command:

```
event manager applet DIS_SH_RUN
 event cli pattern "show run" skip yes sync no
 action 1.0 cli command "enable"
 action 1.1 syslog msg "$_cli_msg not executed, function disabled"
 action 1.2 mail server ....
```

Hide interfaces from the running configuration:

```
event manager applet SH_RUN_NO_INT
 event cli pattern "show run" sync yes
 action 1.0 syslog msg "$_cli_msg executed"
 action 1.1 cli command "enable"
 action 1.2 cli command "show run | section exclude interface"
 action 1.3 puts "$_cli_result"
```

Re-enable manually shut down interfaces:

```
event manager applet NO_SHUT_INT
 event syslog pattern "Interface FastEthernet0/0, changed state to administratively down"
 action 1.0 cli command "enable"
 action 1.1 cli command "configure terminal"
 action 1.2 cli command "interface Fa0/0"
 action 1.3 cli command "no shut"
```

Print confirmation to the terminal:

```
event manager applet WRITE_MEMORY
 event cli pattern "write memory" sync yes
 action 1.0 syslog msg "$_cli_msg Command Executed"
 set 2.0 _exit_status 1
```

Disable OSPF and EIGRP:

```
event manager applet DIS_OSPF_EIGRP
 event cli pattern "router [eEoO].*" sync no skip yes
 action 1.0 syslog msg "Routing protocols OSPF and EIGRP have been disabled"
```

Send ICMP requests based on tracking object:

```
event manager applet TRACK_1_DOWN
 event syslog pattern "1 ip sla 1 state Up->Down"
```

```
 action 1.0 syslog msg "IP SLA 1 Transferred to Down State, Testing ICMP"
 action 1.1 cli command "enable"
 action 1.2 cli command "ping 10.0.12.2 repeat 5 time 1"
 action 1.3 syslog msg "ping 10.0.12.2 repeat 5 time 1"
 action 1.4 puts "$_cli_result"

event manager applet TRACK_1_UP
 event syslog pattern "1 ip sla 1 state Down->Up"
 action 1.0 syslog msg "IP SLA 1 Returned to UP State, Testing ICMP"
 action 1.1 cli command "enable"
 action 1.2 cli command "ping 10.0.12.2 repeat 5 time 1"
 action 1.3 syslog msg "ping 10.0.12.2 repeat 5 time 1"
 action 1.4 puts "$_cli_result"
```

Or:

```
event manager applet TRACK_1_DOWN
 event track 1 state down
 ...
```

Periodically send output to the console:

```
ip sla 1
 udp-jitter 192.168.0.2 16384 codec g729a
  frequency 5

event manager applet IP_SLA_1
 event timer watchdog time 3600
 action 1.0 cli command "show ip sla statistics 1"
 action 1.2 puts "$_cli_result"
 action 1.3 mail ...
```

Create a log message based on added routes (does not disable function):

```
event manager applet STATIC_ROUTES
 event routing type add protocol static network 0.0.0.0/0  le 32
 action 1.0 syslog msg "Static routes are not allowed, notifying admin"
 action 1.1 mail server ....
```

EPC

Wednesday, December 2, 2015
7:09 PM

Old Method

Association (and disassociation) are actions that can be performed in order to bind a capture point to a capture buffer.

- A capture point can only be associated with one capture buffer (an ACL filter can also be applied).
- A capture buffer can be associated with many capture points.
- A buffer can collect data from many points but a point can send data to only one buffer.
- Capture local traffic with the monitor capture point ip process-switched LOCAL from-us command.

monitor capture buffer BUFFER
monitor capture point ip cef PCAP fa0/0 both
monitor capture point associate PCAP BUFFER

show monitor capture buffer BUFFER dump
show monitor capture buffer BUFFER parameters
show monitor capture point PCAP

monitor capture buffer BUFFER export

New Method

monitor capture PCAP match any int gi0/0 both
monitor capture PCAP start

show monitor capture PCAP

monitor capture PCAP export

IPCP

Internet Protocol Control Protocol (IPCP)

- IPCP relies on PPP.
- IOS ignores mask requests and offers. This is a problem when running RIP.
- Use PPPoE with import IPCP into DHCP to acquire the correct mask.
- Disable validate-update source in RIP when using IPCP.

```
ip local pool IPCP 10.0.12.2
int se1/0
 ip add 10.0.12.1 255.255.255.0
 encapsulation ppp
 peer default ip address pool IPCP
 peer default ip address 10.0.12.2
 ppp ipcp mask 255.255.255.0
 peer neighbor-route

int se1/0
 encapsulation ppp
 ip address negotiated
 ppp ipcp mask request
 peer neighbor-route
 no shut

router rip
 no auto
 version 2
 network 10.0.12.0
 network 192.168.0.0
 no validate-update source
```

Import IPCP subnet settings to local DHCP

- This will allow the import of the correct subnet mask.
- The imported IPCP pool will always start at the first address (.1) even if only a single address is specified in the pool.
- This will ignore additional IPCP settings, such as the default route installed through IPCP.

```
ip dhcp pool LOCAL
 import all
 origin ipcp

int se1/0
 encapsulation ppp
 ppp ipcp mask request
 no ppp ipcp route default
 no ip add negotiated
 ip add pool LOCAL
```

IRDP

Tuesday, December 8, 2015
9:00 PM

ICMP Router Discovery Protocol (IRDP)

- Allows routers to discover gateways.
- IP routing has to be disabled on client.

```
int fa0/0
 ip irdp
 ip irdp maxadvertinterval 30
 ip irdp minadvertinterval 10
 ip irdp holdtime 90
 ip irdp preference 200

show ip irdp fa0/0
```

Client:

```
no ip routing
ip gdp irdp
```

FHRP

Tuesday, December 8, 2015
9:13 PM

Hot Standby Router Protocol (HSRP)

- Preemption is disabled by default.
- Highest priority router is active, if priority is equal the highest IP-address wins.
- Hello timer is 3 seconds by default, and hold timer is 10 seconds by default.

HSRPv1	0000.0c07.acXX	UDP 1985	224.0.0.2
HSRPv2	0000.0C9F.F000 through 0000.0C9F.FFFF	UDP 1985	224.0.0.102
HSRPv6	0005.73A0.0000 through 0005.73A0.0FFF	UDP 2029	FF02::66

XX = Group number in HEX

Virtual Router Redundancy Protocol (VRRP)

- Preemption is enabled by default.
- Highest priority router is active, if priority is equal the highest IP-address wins.
- Advertisement interval (hello timer) is 1 second by default, automatically sets the hold-timer to 3x hello.
- IPv6 VRRP requires version3.

VRRPv2	0000.5e00.01XX	UDP 112	224.0.0.18
VRRPv3	0000.5e00.01XX	UDP 112	FF02::12

XX = Group number in HEX

```
fhrp version vrrp v3
int fa0/0
 vrrp 1 address-family ipv6
  address 2001:10:0:12::254/64
  address fe80::254 primary
```

Gateway Load Balancing Protocol (GLBP)

- AVG preemption is disabled by default.
- AVF preemption is enabled by default, with a delay of 30 seconds. Useful when preempting lower weighted routers.
- Highest priority router is AVG, if priority is equal the highest IP-address wins.
- Hello timer is 3 seconds by default, and hold timer is 10 seconds by default.
- Up to 4 AVF per group, this includes the AVG.

GLBP	0007.b400.XXYY	UDP 3222	224.0.0.102
GLBPv6	0007.b400.XXYY	UDP 3222	FF02::66

XX = Group number in HEX
YY = AVF number

GLBP Load-Balancing and Weighting

- Host-dependent. Same AVF is used for the same host (based on mac address).
- Round-robin. Default. Each AVF is used in turn.
- Weighted. Dependent on the weighting value.
    - The maximum weighting value is the default 'normal' value (100 by default).
    - If devices go below the lower weighting value, they lose their AVF status (1 by default)
    - If devices go above the upper weighting value, they regain their AVF status after a loss (same as maximum by default)

Device will lose AVF status if only 1 tracking object is up. Will regain status if 2/3 tracking objects are up:

```
int fa0/0
 glbp 1 load-balancing weighted
 glbp 1 weighting 100 51 74
 glbp 1 weighting track 1 decrement 25
 glbp 1 weighting track 2 decrement 25
 glbp 1 weighting track 3 decrement 25
```

GLBP Redirection

- Redirect timer is 10 minutes, timeout timer is 4 hours by default.

In case of an unreachable AVF the AVG redirects traffic:

- During redirecting time, the AVG points a new AVF for any new request with old virtual MAC address.
- After the redirect timer expires, the AVG stops pointing a new AVF for any new request with old virtual MAC address.
- Hosts that using old mac-address can get responses and are able to use old mac address until the timeout timer expires.
- If the AVF doesn't return until the timeout timer expires, all GLBP peers flush the record of the old MAC address and old AVF.

KRON

Command Scheduler (KRON)

- Only works for exec mode commands, not global or interface configuration commands.
- Choose either the oneshot or recurring keyword to schedule KRON occurrence once or repeatedly.
- The system-startup keyword will set the occurrence to be at system startup.

Show routes every 5 minutes:

```
kron policy-list KRON_POLICY
 cli show ip route

kron occurrence KRON_OCC in 5 recurring
 policy-list KRON_POLICY

sh kron schedule
debug kron all
```

Logging

Wednesday, December 2, 2015
7:09 PM

Archive Log
Configure archiving and optionally log commands to syslog:

```
archive
 log config
  logging enable
  ;notify syslog
  exit
alias exec sal show archive log config all provisioning
```

Archive Config
Configure archiving of configs to TFTP server:

```
archive
 path tftp://192.168.10.1/archive

archive config

show archive
show archive config differences
```

Logging

- Logging to the console and or buffer might be disabled in the lab. Verify with the show run all | i logging command.
- Show timestamps (on by default) on debug messages with the service timestamps debug datetime command.
- Show timestamps (on by default) on log messages with the service timestamps log datetime command.
- Show sequence numbers (off by default) on log messages with the service sequence-numbers command.

Enable logging:

```
logging console guaranteed
logging buffered 8192 debugging
logging console debugging
logging monitor debugging
logging on
```

History

```
show run all | i history
show history
```

Reset history:

```
term history 0
term history 10
```

NetFlow

Tuesday, December 8, 2015
9:14 PM

NetFlow Versions

5	NetFlow v5 is fixed format, cannot be extended or added to. IPv4 only. Added BGP AS information and sequence numbers. Exports data from main cache only.
8	Added support for data export from aggregation caches.
9	NetFlow v9 can add additional information to flows, template based. Added support for MPLS, BGP next-hop and IPv6 headers. Exports data from main and aggregation cache.

NetFlow IP Flows

- NetFlow Requires CEF in order to function.
- In original NetFlow if all of characteristics match, they're considered the same flow.
- An IP Flow can be characterized by a set of 5 and up to 7 packet attributes:
    - Source / destination IP address
    - Source / destination port
    - L3 protocol type
    - Class of Service
    - Router or switch interface

NetFlow Collector

- NetFlow Collector = NetFlow server.
- The Collection Engine (local) sends NetFlow data to the collector with 1.5% export data overhead.
- The NetFlow Cache creates cache entries (flow records) for every active flow.
- Flow records store IP flow information.
- NetFlow export, unlike SNMP polling, pushes information periodically to the collector.
- Flows that have terminated or expired (Based on cache) are exported as well.

NetFlow v5

- NetFlow v5 does not have a concept of 'ingress' and 'egress' flows.
- The collector engine reverses the information behind the scenes without any additional configuration.
- The ip route-cache flow command is the old way of configuring NetFlow. This is called the Flow fast-switching cache.
- The old command will also enable NetFlow on all sub-interfaces, the newer command does not.

```
int fa0/0
 description Inside
 ip flow ingress
int fa0/0
 description Outside
 ip flow ingress
```

```
show ip flow interface
show ip cache flow
```

NetFlow v9

- NetFlow v9 introduces the 'egress' flows concept.
- With egress, it is possible to configure ingress and egress on the same interface and capture both traffic directions.
- Multicast traffic can't be effectively matched on ingress.
- Ingress calculates before compression, this is a problem if WAN links are using compression of packets.
- Egress calculates flows after compression.
- Configuring NetFlow v9 with the egress keyword uses a default template behind the scenes.

- Only when templates are specified Flexible NetFlow is being used.

```
int fa0/0
 description Inside
 ip flow ingress
int fa0/0
 description Outside
 ip flow egress
```

Export flows to collector:

```
ip flow-export destination 1.1.1.1 9995
ip flow-export version 9
ip flow-export source loopback 0
```

```
show ip flow export
```

NetFlow Aggregation Cache (v8 and v9)

- Enables specification of which type of flows will be exported to the collector.
- All flows are still captured on the device (using v5 or v9) but only when flows are exported they are filtered.

Only export flow entries that have a /32 mask:

```
ip flow-aggregation cache destination-prefix
 cache entries 1024
 export version 9
 export destination 2.2.2.2 9995
 mask destination minimum 32
 enabled
```

```
show ip cache flow aggregation destination-prefix
```

NetFlow Top Talkers

- Useful if no collector server is present to analyze data flows.
- Shows top talkers based on bytes or packets.

```
ip flow-top-talkers
 top 10
 sort by packets
```

```
show ip flow top-talkers
```

NetFlow Sampler

- Sampled mode lets you collect only for a subset of traffic.
- Can be linked directly to the interface, or be part of a policy-map.
- Can not be used alongside the ingress command. Either capture all flows or a subset of flows.

```
flow-sampler-map RANDOM
 mode random one-out-of 10
```

```
int fa0/0
 flow-sampler RANDOM
```

```
show flow-sampler
```

Add a sampler to a policy-map and match one out of 10 ICMP packets:

```
flow-sampler-map RANDOM
 mode random one-out-of 10
flow-sampler-map ONE_ONE
 mode random one-out-of 1
```

```
class-map match-all ICMP
 match protocol icmp
policy-map SAMPLER
 class ICMP
  netflow-sampler RANDOM
 class class-default
  netflow-sampler ONE_ONE

int fa0/0
 service-policy input SAMPLER
```

Flexible NetFlow (FNF)

Consists of three parts:

- Flow Records, which set key and non-key fields.
- Flow Exporter, which details where and how to send the exports.
- Flow Monitors, which match the flow records and exporters, and are then applied to an interface.

```
flow exporter FNF_EXPORT
 destination 1.1.1.1
 transport udp 9995
 export-protocol netflow-v9

flow monitor FNF
 record netflow ipv4 original-input
 exporter FNF_EXPORT

int fa0/0
 ip flow monitor FNF input

show flow exporter
show flow monitor FNF cache format table
```

FNF Sampler

```
sampler FNF_SAMPLER
 mode random 1 out-of 10

int fa0/0
 ip flow monitor FNF sampler FNF_SAMPLER input

show sampler
```

NTP

Tuesday, December 8, 2015
9:14 PM

Network Time Protocol (NTP)

- NTP Peers can both act as either a client or a server at the same time and offer bidirectional synchronization.
- When the connection to the NTP server fails, the peer will be regarded as the new server.
- Masters on older IOS versions (12.4) use the 127.127.7.1 local address to peer with itself.
- Newer IOS versions use the 127.127.1.1 local address.

- The source local address is always one stratum lower than the configured value. Default configured stratum is 8.
- Stratum is the tie-breaker. If two servers offer the same stratum, the prefer keyword can be added to prefer one over the other.

```
ntp master 8
ntp server 192.168.0.1 prefer
ntp peer 192.168.0.2
```

```
show ntp status
show ntp association detail
debug ntp packet
debug ntp events
```

The offset value is the time difference in milliseconds between the local clock and the NTP server's reference clock.

- The offset must be < 1000 msec (1 second) off in order for the server to be considered sane.
- NTP does not shift the clock instantaneously, instead the router slowly drifts towards the time.
- If the offset value between the client and the server is large, this process can take a long time.
- After the offset value is < 1 second off, the router will adjust its stratum from 16 (infinite) to the appropriate stratum.

Time Zones

- NTP updates are always sent in UTC/GMT.
- EU and US summer time dates are different. Default is US, configure with clock summer-time US recurring.
- US summer time begins second Sunday in March, ends first Sunday in November.
- EU summer time begins last Sunday in March, ends last Sunday in October.

```
clock timezone CaPc -8
clock summer-time US recurring 2 Sun Mar 02:00 First Sun Nov 02:00
clock summer-time EU recurring Last Sun Mar 02:00 Last Sun Oct 02:00
```

NTP Authentication

- The client authenticates the server, it is more important to receive time from the correct source over giving time to devices.
- Other NTP clients will still be able to request time without authentication.

```
ntp trusted-keys 1
ntp authenthication-key 1 md5 cisco
```

Client:

```
ntp authenthication-key 1 md5 cisco
ntp trusted-keys 1
ntp authenticate
ntp server 192.168.0.1 key 1
```

NTP Access Control

- Control messages (queries) are for reading and writing internal NTP variables and status information. Not synchronization.
- NTP request/update messages are used for actual time synchronization.

- The serve-only keyword allows only time requests from NTP clients.
- The peer keyword allows time requests and NTP control queries from clients. But also allows bidirectional synchronization.
- Masters on older IOS versions (12.4) need to specifically allow peering with the own loopback address (127.127.7.1).
- Access-groups associated with access types are scanned in the order most permissive to most restrictive. Peer -> Serve-Only.
    - This means that denying a client in serve-only but allowing with peer, the client will still be able to peer.

```
ntp master
access-list 1 deny host 192.168.0.2
access-list 2 permit host 192.168.0.2
ntp access-group serve-only 1
ntp access-group peer 2

access-list 127 permit host 127.127.7.1
ntp access-group peer 127
```

NTP Broadcast and Multicast

- Default multicast address is 224.0.1.1.

```
int fa0/0
 ntp broadcast
 ntp broadcast destination 10.0.12.2
 ntp multicast 224.0.1.1
```

Client:

```
int fa0/0
 ntp broadcast client
 ntp multicast client
```

PPP

Tuesday, December 8, 2015
9:14 PM

PPP Authentication

- The router that enables PPP authentication requests credentials from the remote router.
- The credentials supplied by the remote router has to match the local user database.
- Local authentication is based on usernames.
- EAP requires the addition of the local keyword to authenticate using the local database.
- PPP usernames can still be used for line management. Use the username PPP-USER autocommand logout to prevent this.

R1 requests CHAP from R2:

```
username R2 password cisco
username R3 password cisco
int se1/0
 encapsulation ppp
 ppp authentication chap
 ppp pap sent-username R1 password cisco
```

R2 requests PAP from R1:

```
username R1 password cisco
int se1/0
 encapsulation ppp
 ppp chap hostname R2
 ppp chap password cisco
 ppp authentication pap

show users
who
debug ppp authentication
```

AAA Authentication for PPP

- When using AAA the autocommand will only function if aaa authorization is also configured.
- Preferably use a private Radius or Tacacs+ server in combination with PPP authentication.

R1 requests EAP from R2 and authenticates using Radius:

```
aaa new-model
aaa group server radius MYRADIUS
 server-private 1.1.1.1 timeout 5 retransmit 0 key cisco

aaa authentication ppp PPP_R1_R2 group MYRADIUS local
aaa authorization exec default group MYRADIUS local
username R2 password cisco
username R2 autocommand logout
int se1/0
 encapsulation ppp
 ppp authentication eap PPP_R1_R2
 ppp eap local
 ppp chap hostname R1
 ppp chap password cisco

show aaa servers
show radius server-group all
debug radius
```

R2 requests MS-CHAP-V2 from R2 and authenticates using Tacacs+:

```
aaa new-model
aaa group server tacacs+ MYTACACS
 server-private 2.2.2.2 single-connection key cisco

aaa authentication ppp PPP_R1_R2 group MYTACACS local-case
aaa authorization exec default group MYTACACS local
username R1 password cisco
username R1 autocommand logout
int se1/0
 encapsulation ppp
 ppp eap identity R2
 ppp eap password cisco
 ppp authentication ms-chap-v2 PPP_R1_R2

show aaa servers
show tacacs private
debug tacacs
debug tacacs events
debug tacacs packets
```

Multilink PPP (MLPPP)

- Uses a fragmentation scheme where large packets are sliced in pieces and sequence numbers are added using headers.
- Fragments are sent over multiple links and reassembled at the opposite end.
- Small voice packets are interleaved with fragments of large packets using a special priority queue.

The interleave keyword enables real-time packet interleaving.

- Allows large packets to be MLPPP encapsulated and fragmented into a small enough size to satisfy delay requirements.

```
int multilink 1
 ppp multilink interleave
 ppp multilink

int se1/0
 ppp multilink
 ppp multilink group 1
```

PPPoE

Tuesday, December 8, 2015
9:15 PM

PPP over Ethernet (PPPoE)

- PPPoE provides a standard method of employing the authentication methods of PPP over an Ethernet network.
- Allows authenticated assignment of IP addresses.
- The MTU size is automatically set to 1492 bytes.

PPPoE IPCP

- IOS ignores mask requests and offers. This is a problem when running RIP.
- Use PPPoE with import IPCP into DHCP to acquire the correct mask.

Server:

```
bba-group pppoe R2
 virtual-template 12
int fa0/0
 pppoe enable group R2

interface Virtual-Template 12
 description R2
 ip address 10.0.12.1
 ip mtu 1492
 encapsulation ppp
 ppp authentication chap
 peer default ip address pool IPCP
 ppp ipcp mask 255.255.255.0

username R2 password cisco
ip local pool IPCP 10.0.12.2
```

Client:

```
int fa0/0
  pppoe-client dial-pool-number 12

interface Dialer 1
 ip address negotiated
 ip mtu 1492
 encapsulation ppp
 ppp chap username R2
 ppp chap password cisco
 dialer pool 12
 ppp ipcp mask request
 ppp ipcp route default
```

PPPoE IPCP with local DHCP

- Import IPCP subnet settings to local DHCP. Will allow the import of the correct subnet mask.
- The imported IPCP pool will always start at the default first address (.1) even if only a single address is specified in a pool.
- This will ignore additional IPCP settings, such as the default route installed through IPCP.
- Requires bouncing of the interfaces if overwriting an existing IPCP configuration.

Client:

```
ip dhcp pool IMPORT_IPCP
 import all
 origin ipcp
```

```
int dialer1
 dialer pool 12
 ip mtu 1492
 encapsulation ppp
 ppp chap username R2
 ppp chap password cisco
 ppp ipcp mask request
 ip address pool IMPORT_IPCP

int fa0/0
 pppoe-client dial-pool-number 12
```

PPPoE DHCP

Server:

```
bba-group pppoe R2
 virtual-template 12
int fa0/1
 pppoe enable group R2

interface Virtual-Template12
 description R2
 ip address 10.0.12.1 255.255.255.0
 ip mtu 1492
 peer default ip address dhcp-pool DHCP
 ppp authentication pap

username R2 password cisco
ip dhcp excluded-address 10.0.12.1
ip dhcp excluded-address 10.0.12.3 10.0.12.254
ip dhcp pool DHCP
 network 10.0.12.0 /24
 default-router 10.0.12.1
```

Client:

```
int fa0/1
  pppoe-client dial-pool-number 12

interface Dialer 1
 dialer pool 12
 ip mtu 1492
 encapsulation ppp
 ppp pap sent-username R2 password cisco
 ip address dhcp
```

PBR

Wednesday, December 2, 2015
4:03 PM

Policy-Based Routing (PBR)

- The set ip next-hop and set interface are used unconditionally, meaning that the RIB is not used in case of failure.
- The set ip default next-hop and set default interface apply to the default route and are used before the regular default route.
- The idea is to specify an alternate default route for hosts matched in the access-list.

There's an order of operations to PBR set statements. ip next-hop -> interface -> ip default next-hop -> default interface.

- If the first statement fails the next will be evaluated. Remember that addresses are preferred over interfaces.
- The recursive keyword can be used to specify a next-hop that is not directly connected.
- Only statement 10 is matched in route-maps. Meaning that if there is no match in 10, traffic is routed normally.
- For this reason PBR route-maps allow more than one set statement in route-map sequences.

```
ip access-list standard PBR_ACL
 permit 172.16.0.0 0.0.0.255

route-map PBR permit 10
 match ip address PBR_ACL
 set ip default next-hop 10.0.12.2
 set default interface se1/0

interface fa0/0
 ip policy route-map PBR
```

Local PBR

- Applying the PBR on an interface does not affect traffic locally generated by the router (even when sourcing the interface)
- Create an ip local policy using loopbacks to policy-route local traffic.
- Use extended access-lists to have more granular control over which local traffic is policy routed.
- Using standard access-lists will policy route all local traffic.

```
ip access-list standard PBR_LOCAL_ACL
 permit host 192.168.0.1

route-map PBR_LOCAL permit 10
 match ip address PBR_LOCAL_ACL
 set ip next-hop 10.0.12.2
 set interface se1/0
ip local policy route-map PBR_LOCAL

show ip local policy
```

SLA

TCP Connect

- Control messages communicates the port that will be used from the sender to the receiver (enabled by default).
- Disable control packets when using a well-known TCP port (telnet for example). This also does not need a responder.
- SLA messages to an unknown port requires a SLA responder at the destination.

```
ip sla 1
 tcp-connect 192.168.0.2 23 control disable
 threshold 500
 timeout 1000
 frequency 5
ip sla schedule 1 life forever start-time now

show ip sla statistics
show ip sla configuration
show tcp brief
```

IP SLA Responder

- Does not calculate processing time, allowing for more accurate measurements on the speed of the link.
- Enable globally with the ip sla responder command. General IP SLA responder uses port 1967 for control messages.
- Can also be configured to listen on a specific port for UDP or TCP. However this specific port must be configured on both sides. It is not possible to use control messages at sender and specific port at the receiver.

UPP Echo

- UDP echo always requires a responder at the destination.
- If control messages are disabled, the responder must be configured to listen on the specific port (55555 in this case).

Source:

```
ip sla 1
 udp-echo 192.168.0.2 55555 control disable
 threshold 500
 timeout 1000
 frequency 5
ip sla schedule 1 life forever start-time now
```

Destination:

```
ip sla responder udp-echo ipaddress 192.168.0.2 port 55555

show ip sla responder
```

UDP Jitter

- Per-direction jitter (source to destination and destination to source).
- Per-direction packet loss.
- Per-direction delay (one-way delay).
- Round-trip delay (average round-trip time).
- Same as UDP echo, also requires a responder configured at the destination.
- Success/failures will only be updated when all packets are analyzed (10 packets are sent by default).

IP SLA Authentication

- The authentication hash is MD5.
- Enabled for all SLAs present on the device.
- Only applied on SLAs where both sides need to participate (using a responder).

```
key chain SLA
 key 1
  key-string cisco
```

ip sla key-chain SLA

show ip sla authentication

SNMP

Wednesday, December 2, 2015
8:52 AM

Simple Network Management Protocol (SNMP)

- SNMP GET. Allows a NMS to request information from device. (UDP 161)
- SNMP SET. Allows a NMS to make changes to the device. (UDP 161)
- SNMP TRAP / INFORM. Directly send urgent information to the NMS. (UDP 162)
    - TRAP is unacknowledged packet.
    - INFORM is acknowledged packet.
- The downside of SNMP set is that the NMS can only poll on intervals, so events may be lost in between polls.
- INFORMS makes sure that all messages arrive at the NMS (more memory intensive). The NMS responds with a PDU.

SNMP Communities

- Both v1 and v2 groups are created when configuring a SNMP community.
- Disable the v1 group with the no snmp-server group public v1 command.
- Also disable the Interim Local Management Interface (ILMI) SNMP groups. The ILMI community itself cannot be deleted.

```
snmp-server community public ro
no snmp-server group public v1
no snmp-server group ILMI v1
no snmp-server group ILMI v2c

show snmp community
show snmp group
```

SNMP Host

- Only SNMP Traps will be sent to the host, unless you specify the inform keyword.
- SNMP v1 is the default when not specifying a version.

```
snmp-server enable traps
snmp-server host 192.168.0.1 traps version 2c public udp-port 162
snmp-server host 192.168.0.1 inform version 2c public udp-port 162
```

SNMPv3

- The SNMP group security level is a minimum allowed security level.
- The actual security level for the user is defined in the snmp-server user command. This is the minimum level for that user.
- Other users may still connect using the minimum allowed group security level.

CLI command	Authentication Method	Encryption Support
noAuthNoPriv	Username	No encryption
authNoPriv	MD5 or SHA	No encryption
authPriv	MD5 or SHA	DES, 3DES or AES

```
snmp-server group GROUP v3 priv
snmp-server user USER GROUP remote 192.168.0.1 v3 auth sha cisco priv aes 256 cisco cisco
snmp-server host 192.168.0.1 informs version 3 priv USER

show snmp user
show snmp group
```

SNMP FIltering

```
ip access-list standard SNMP
 permit host 192.168.0.1
 deny any log

snmp-server community public ro SNMP
```

SNMPv3 FIltering

```
ip access-list standard SNMP_USER
 permit host 192.168.0.1
ip access-list standard SNMP_GROUP
 permit 192.168.0.0 0.0.0.255

snmp-server user USER GROUP v3 auth sha cisco priv aes 256 cisco access SNMP_USER
snmp-server group GROUP v3 priv access SNMP_GORUP
```

SNMP Engine-ID

- SNMPv3 user passwords are hashed based on the value of the local Engine-ID.
- If the Engine-ID changes, the security digests of SNMPv3 users will be invalid, and the users will have to be reconfigured.
- Trailing zeroes will be added automatically to create 24 characters when changing the Engine-ID.

```
snmp-server engine-id local 123412341234
snmp-server engine-id remote 192.168.0.1 123412341234
```

System

Tuesday, December 8, 2015
9:18 PM

Memory Reservations
memory free low-watermark processor
memory reserve critical
memory reserve console

show memory console reserved

CPU Threshold

- The rising and falling commands trigger a syslog message when CPU is above/below threshold.

snmp-server enable traps cpu threshold
process cpu threshold type total rising 50 interval 5 falling 10 interval 5

TCL

Tuesday, December 8, 2015
9:14 PM

TCL Scripting

```
tclsh
foreach X {
192.168.0.1
192.168.0.2
192.168.0.3
192.168.0.4
192.168.0.5
192.168.0.6
192.168.0.7
192.168.0.8
192.168.0.9
192.168.0.10
2001:192:168::1
2001:192:168::2
2001:192:168::3
2001:192:168::4
2001:192:168::5
2001:192:168::6
2001:192:168::7
2001:192:168::8
2001:192:168::9
2001:192:168::10
} { ping $X repeat 100 time 1 source loopback 0}
```

Tracking

Monday, November 30, 2015
4:53 PM

Object Tracking

- Interface IP or line-protocol. Track either the line-protocol (up/down) or the presence of an IP address.
- IP route. Track a route present in the routing table.
- IP SLA. Track an IP SLA object.
- List. Combine multiple track objects based on percentage/weight or boolean AND/OR.

Tracking Options

- State. Up or down (default setting).
- Reachability. UP or down + within configured threshold (timeout settings).

```
track 1 int fa0/0 ip routing
track 2 int fa0/1 line-protocol
track 3 ip sla 1 state
track 4 ip route 0.0.0.0/0 reachability
```

IP route metric threshold:

- Whatever the metric is will be given a value from 1-255 based on the track resolution.
- If route metric is under the up threshold, tracking is up.
- Optionally change resolution of specific routing protocol.

```
track 5 ip route 0.0.0.0/0 metric threshold
 metric threshold up 50

track resolution ip route eigrp 2560
show track resolution
```

Combine Track Objects

Boolean tracking:

- If both objects up, track 12 is up.
- If either object is down, track 12 is down.

```
track 12 list boolean and
 object 1
 object 2
```

Percentage tracking:

- If two out of three objects are up(66%), track 123 is up.
- If one out of three objects is up (33%), track 123 is down.
- Track 123 will stay down until two out of tree objects are back up (66%).

```
track 123 list threshold percentage
 object 1
 object 2
 object 3
 threshold percentage up 66 down 33
```

Weighted tracking:

- If object 2 is down, track 234 is down.
- If either object 3 or 4 is down, track 234 is up.

```
track 234 list threshold weight
 object 2 weight 100
 object 3 weight 50
```

object 4 weight 50
threshold weight up 150 down 100

Track Timers

- Not every object type is tracked at the same rate.
- Manually change track timers to allow for faster detection.

track timer ip route 5
show track timers

uRPF

Tuesday, December 8, 2015
8:53 PM

Unicast Reverse Path Forwarding (uRPF)

- Verifies reachability of source address in packets being forwarded. Requires CEF.
- Use optional ACL to log dropped packets by uRPF, or use it to allow specific subnets that fail the check.

- With strict mode (rx) packets must be received on the interface that the router uses to forward the return packet.
- With loose mode (any) it is only required that the source address appears in the routing table.
- Use loose mode when asymmetric routing paths are present in the network.
- The allow-default keyword also includes the default route in valid route list.
- The allow-self-ping keyword allows a router to ping itself on that particular interface.

```
int fa0/0
 ip verify unicast source reachable-via any allow-default allow-self-ping
```

IPv6

Monday, December 21, 2015
7:46 PM

Neighbor Discovery

- Uses ICMP and solicited-node multicast addresses to discover neighbors on the local link segment and verify reachability.
- A solicited-node multicast address starts with FF02:0:0:0:0:1:FF::/104.
- Is formed by taking the low-order 24 bits of an address and appending those bits to the solicited-node prefix.

- Afterwards, Duplicate Address Detection (DAD) sends a ping to the solicited node multicast address.
- If another node responds, then the router will not use the address.
- Is performed first on a new, link-local IPv6 address before the address is assigned to an interface.
- The new address remains in a tentative state [TEN] while DAD is performed.

Global Address	Link-Local Address	Solicited-Node Multicast Address
2001:10:0:12:a212:7aff:fecb:6b40/64	fe80::a212:7aff:fecb:6b40	ff02::1:ffcb:6b40/104
2001:10:0:12:b414:9bff:feaa:12fe/64	fe80::b414:9bff:feaa:12fe	ff02::1:ffaa:12fe/104

Neighbor Solicitation and Router Advertisements

Neighbor Solicitation (NS)	Neighbor solicitations are used by nodes to determine the link layer address of a neighbor. Or to verify that a neighbor is still reachable via a cached link layer address. Also used in SLAAC (by hosts) to verify uniqueness of a local address before it is assigned.
Neighbor Advertisement (NA)	Used by nodes to respond to a Neighbor Solicitation message.
Router Solicitation (RS)	Requests neighbor address and advertises own, Also used in SLAAC (by routers) to verify uniqueness of a local address before it is assigned.
Router Advertisement (RA)	Periodically sent out between neighboring routers (200s default). Used by routers to notify hosts of presence of a router and a default-gateway address. Can also be a reply to a RS message. Will only be advertised if ipv6 unicast-routing is enabled.

IPv6 Address Summarization

Addresses	Differences in Binary	Summary Range Start	Summary
2001:db8:24:131a:: 2001:db8:24:131b::	131a = 0001 0011 0001 1010 131b = 0001 0011 0001 1011	0001 0011 0001 101x = 131a Subnets differ at 63th bit = /63	2001:10:0:12:24:131a::/63
2001:cfb:14:: 2001:cfb:15:: 2001:cfb:16:: 2001:cfb:17::	0014 = 0000 0000 0001 0100 0015 = 0000 0000 0001 0101 0016 = 0000 0000 0001 0110 0017 = 0000 0000 0001 0111	0000 0000 0001 01xx = 0014 Subnets differ at 46th bit = /46	2001:cfb:14::/46

Addresses

Wednesday, December 23, 2015
1:02 PM

IPv6 Address Ranges

2000::/3	Global
fe80::/10	Link-Local
ff00::/8	Multicast
fc00::/7	Unique-Local

Multicast Addresses

FF02::1	All Nodes
FF02::2	All Routers
FF02::5	OSPF
FF02::6	OSPF DR
FF02::9	RIPng
FF02::A	EIGRP
FF02::1:FF	Solicited-Node

Anycast Address

- An anycast address is an address shared by multiple systems, with the closest system being the receiver of the packet.
- An address with :: at the end specifies an anycast interface.
- The anycast keyword is optional and only required if there are multiple hosts on the same subnet (connected to the same interface on the router) using the same address.

```
int fa0/0
 description TO_SERVERS
 ipv6 address 2001:10:0:12::/64 anycast
```

EUI-64 address

- An EUI-64 address is an auto configured address using the MAC address of the interface.
- A MAC address is split in the middle and a 16-bit hex value FFFE is inserted between these two to form a 64-bit address.
- Afterwards the 7th bit is flipped in the OUI part of the MAC address.

Mac-Address	Mac-Address + FFFE	Flip 7th bit in OUI	EUI-64 Address
a0:12:7a:cb:6b:40	a0:12:7a:ff:fe:cb:6b:40	1010 0000 = a0 1010 0010 = a2	::a212:7aff:fecb:6b40
b6:14:9b:aa:12:fe	b6:14:9b:ff:fe:aa:12:fe	1011 0110 = b6 1011 0100 = b4	::b414:9bff:feaa:12fe

```
int fa0/0
 mac-address a012.7acb.6b40
 ipv6 address 2001:10:0:12::/64 eui-64
```

Same as:

```
int fa0/0
 mac-address a012.7acb.6b40
 ipv6 address 2001:10:0:12:a212:7aff:fecb:6b40/64
```

DHCPv6

Monday, December 21, 2015
7:46 PM

DHCPv6 Messages

Message Type	Description	IPv4 Equivalent
Solicit	Sent by a client to locate servers.	DHCPDiscover
Advertise	Sent by a server in response to a Solicit message to indicate availability.	DHCPOffer
Request	Sent by a client to request addresses or configuration settings from a specific server.	DHCPRequest
Reply	Sent by a server to a specific client in response to a Solicit, Request, Renew, Rebind Information-Request, Confirm, Release, or Decline message.	DHCPAck

DHCPv6 rapid configuration only uses the Solicit and Reply message. Enable with ipv6 dhcp server DHCP_POOL rapid-commit command.

DHCPv6 multicast groups

FF02::1:2	Link-local DHCP using UDP 546, 547
FF02::1:3	Link-local multicast DNS using UDP 5355
FF02::FB	Multicast DNS using UDP 5353
FF05::1:3	Site wide DHCP using UDP 546, 547
FEC0::	Deprecated

DHCPv6 Modes

- Stateful. DHCPv6 is used for address and other information (DNS). Uses managed-config-flag (M-bit) and other-config-flag (O-bit).
- Stateless. SLAAC is used for address, other information (DNS) is obtained via DHCPv6. Uses other-config-flag (O-bit) only.

Default-Gateway

- The default gateway is a link-local address that is discovered through RAs.
- Routers will only send out RAs if ipv6 unicast-routing is enabled.
- Routers will not receive a default-gateway if ipv6 unicast-routing is enabled.
- Hosts normally pick the first router they discover through ND as the default-gateway.
- This decision can be influenced with the router-preference command (default preference is medium). Routers will preempt
- Optionally advertise the DNS server address through RAs (not supported on all platforms).

```
int fa0/0
 ipv6 nd router-preference high
 ipv6 nd ra dns server 1::1
```

Stop a router from becoming the default gateway with the ipv6 nd ra lifetime 0 interface command.

- This does not disable SLAAC or RAs, hosts will just not install a default route towards the router.

Stop a router from sending RA messages and becoming the default gateway:

```
int fa0/0
 ipv6 nd ra suppress all
```

The suppress keyword indicates not to send periodic RAs.
The all keyword will instruct the router not to respond to RS.

SLAAC
Configure hosts to obtain an address from the routers subnet and add a default gateway:

```
int fa0/0
 ipv6 address autoconfig default
```

The default keyword will also add the advertising (remote) router as the default-gateway even if ipv6-unicast routing is enabled.

Stateless DHCP

```
ipv6 dhcp pool STATELESS
 dns-server 1::1
 domain-name lab.local

int fa0/0
 ipv6 dhcp server STATELESS
 ipv6 nd other-config-flag
```

Clients:

```
int fa0/0
 ipv6 address autoconfig
```

Stateful DHCP

```
ipv6 dhcp pool STATEFUL
 address prefix 2001:10:0:12::/64
 dns-server 1::1
 domain-name lab.local

int fa0/0
 ipv6 dhcp server STATEFUL
 ipv6 nd managed-config-flag
```

Clients:

```
int fa0/0
 ipv6 enable
 ipv6 address dhcp
```

DHCPv6 Relay

- The interface on the DHCP router that the relay address points to has to be configured for a DHCP pool.
- If relay destination is link-local, specify the outgoing interface.

Relay server:

```
int fa0/1
 description TO_CLIENTS
 ipv6 address 2001:10:0:12::254/64
 ipv6 nd managed-config-flag
 ipv6 dhcp relay destination fe80::1 fa0/0
```

DHCPv6 server:

```
ipv6 dhcp pool STATEFUL
 address prefix 2001:10:0:12::/64
 dns-server 1::1
 domain-name lab.local

int fa0/0
```

```
 description TO_DHCP_RELAY
 ipv6 address fe80::1 link-local
 ipv6 dhcp server STATEFUL
```

Prefix-Delegation

Monday, December 21, 2015
7:47 PM

IPv6 Prefix-Delegation (PD)

```
ipv6 local pool LOCAL 2001:db8:abcd::/56 60
ipv6 dhcp pool PD_POOL
 prefix-delegation pool LOCAL lifetime infinite infinite

int fa0/0
 description TO_CLIENT
 ipv6 address 2001:db8:abcd::1/64
 ipv6 dhcp server PD_POOL
```

Client:

```
int fa0/0
 description OUTSIDE
 ipv6 address autoconfig default
 ipv6 dhcp client pd PD_PREFIXES
int fa0/1
 description INSIDE
 ipv6 address PD_PREFIXES ::0:0:0:0:254/64
```

IPv6 Tunneling

Monday, December 21, 2015
7:47 PM

Automatic

Wednesday, December 23, 2015
1:05 PM

Automatic Tunneling Methods

Automatic 6to4	Treats the underlying IPv4 network as an NBMA cloud. Point-to-Multipoint. Uses 2002::/16 address space. Encapsulates IPv4 address into IPv6 address (converted to HEX). Does not support dynamic routing protocols.
Automatic IPv4 Compatible	Uses IPv4-compatible IPv6 addresses for the tunnel interfaces. Point-to-Multipoint. Uses ::/96 address space (::IPv4-Address/96). Deprecated.
ISATAP	Treats the underlying IPv4 network as an NBMA cloud. Point-to-Multipoint. 0000:5EFE is the ISATAP address identifier. Designed for tunneling within a site, not between sites. Supports routing protocols using NBMA (neighbor statements).

Automatic 6to4

- The first 32-bits after the 2002::/16 address space as used for the converted IPv4 address.
- All addresses that need to be reachable over the tunnel need to be configured with the same 2002::/16 prefix.
- Alternatively, a route can be created for non 2002::/16 prefixes pointing to the tunnel address of the neighbor.

IPv4 Address	Converted to HEX	6to4 Address
192.168.0.1	c0.a8.00.01	2002:c0a8:0001::/64
172.16.30.254	ac.10.1e.fe	2002:ac10:1efe::/64
10.0.12.2	0a.00.0c.02	2002:0a00:0c02::/64

6RD

The 6RD feature is an extension of the 6to4 feature that uses encapsulation.

- Does not require 2002::/16 prefix or all 32bits of the IPv4 address.
- Embeds IPv6 into IPv4 using protocol type 41.

```
ipv6 unicast-routing
int se1/0
 ip address 10.0.123.1 255.255.255.0
int lo0
 ip address 192.168.0.1 255.255.255.255

int tun0
 ipv6 address 2002:c0a8:1::1/64
 tunnel source lo0
 tunnel mode ipv6ip 6to4

int lo1
 description PREFIXES_OVER_TUNNEL
 ipv6 address 2002:c0a8:1:1::1/64
```

```
 ipv6 address 2002:c0a8:1:2::1/64
 ipv6 address 2002:c0a8:1:3::1/64

ipv6 route 2002::/16 tunnel0
;ipv6 route 2::2/128 20002:c0a8:2::1
```

ISATAP

- Automatically converts IPv4 addresses and inserts these in the IPv6 address.
- Preferably use eui-64 addressing. Free to choose network portion of the address.
- The last 32-bits after 0000:5EFE are used for the converted IPv4 address (in the host portion of the address)
- By default, tunnel interfaces disable periodic router advertisements (RA).
- RAs must be enabled on ISATAP tunnels to support client auto configuration. Enable with no ipv6 nd ra suppress.

IPv4 Address	Converted to HEX	ISATAP Address with custom Network Range
192.168.0.1	c0.a8.00.01	2001:10:0:123:0000:5efe:c0a8:0001/64
172.16.30.254	ac.10.1e.fe	2001:10:0:123:0000:5efe:ac10:1efe/64
10.0.12.2	0a.00.0c.02	2001:10:0:123:0000:5efe:0a00:0c02/64

```
ipv6 unicast-routing
int fa0/0
 ip address 10.0.123.1 255.255.255.0
int lo0
 ip address 192.168.0.1 255.255.255.255

int tun0
 ipv6 address 2001:10:0:123::/64 eui-64
 tunnel source lo0
 tunnel mode ipv6ip isatap
 no ipv6 nd ra suppress
```

Same as:

```
int tun0
 ipv6 address fe80::5efe:c0a8:1 link-local
 ipv6 address 2001:10:0:123:0:5EFE:C0A8:1/64
 tunnel source lo0
 tunnel mode ipv6ip isatap
 no ipv6 nd ra suppress

router eigrp ISATAP
 address-family ipv4 autonomous-system 1
  network 192.168.0.1 0.0.0.0
  network 10.0.123.0 0.0.0.255
 address-family ipv6 autonomous-system 1
  neighbor fe80::5efe:c0a8:2 tun0
  neighbor fe80::5efe:c0a8:3 tun0

int lo1
 description PREFIXES_OVER_TUNNEL
 ipv6 address 1::1/128
 ipv6 address 11::11/128
```

Static

Wednesday, December 23, 2015
1:05 PM

Static Tunneling Methods

Static IPv6IP	Carries only IPv6 packets over IPv4. Point-to-Point. Uses protocol 41.
Static GRE	Carries IPv6, CLNS + other traffic over IPv4. Point-to-Point. Uses protocol 47. (default method)

The only difference between GRE and IPv6IP configuration is the tunnel mode.

```
ipv6 unicast-routing
int se1/0
 ip add 10.0.12.1 255.255.255.0

int tun0
 tunnel source se1/0
 tunnel destination 10.0.12.2
 ipv6 address fe80::1 link-local
 ipv6 address 2001:10:0:12::1/64
 tunnel mode ipv6ip
```

MPLS

Tuesday, January 12, 2016
12:16 PM

MPLS Label Binding

- MPLS assigns a local label and receives a remote label for routes.
- The local label is distributed to neighbors and stored in their forwarding table.
- Traffic arriving with a label matching a local label can then immediately be forwarded to the remote label that is assigned for the same route, not requiring any form of lookup.

Forwarding Equivalence Class (FEC)

- A label represents a FEC which is a group of IP packets which are forwarded in the same manner, over the same path.
- A FEC might correspond to a destination IP subnet or an IP precedence value.

MPLS Infrastructure

- MPLS is globally enabled and requires CEF to operate.
- MPLS traffic will follow the same path as IP traffic by default.
- All IGPs (including connected and static) routes will have a label assigned by default.

MPLS Label Operations

- PUSH. Label is installed on the packet, the label is pushed onto the stack.
- SWAP. The top-most label is swapped.
- PULL. The top-most label is removed from the stack.
- DELETE. The entire stack is deleted.

MPLS Label Switching Routers (LSR) / Label Edge Router (LER)

- Traffic flows upstream to downstream in order to reach a network prefix.
- A MPLS downstream router is closest to the subnet that is being reached.
- Routers that reside in the core of the network are called LSRs.

Routers that reside connecting to CE devices are called Edge-LSRs (LER).

- Where traffic originating in the MPLS domain is called an Ingress Edge-LSR.
- The router that forwards the traffic to the CE is called an Egress Edge-LSR.

Penultimate Hop Popping (PHP)

- The top-most label is removed by the LSR adjacent to the Edge-LSR.
- The label is popped one hop earlier than the Edge Egress LSR.

L3VPN

Tuesday, January 12, 2016
12:16 PM

Route Distinguisher (RD) and Route Target (RT)

- The only purpose of the RD is to make routes unique in the mBGP.
- Doesn't have to match on neighboring routers that are part of the same VPNv4 neighborship.
- There is no relationship between RT and RD, they do not have to match on the same router or on neighboring router.
- RT helps sort routes in the appropriate routing table.
- The router imports the RT that the other router exports, this does not have to match the RD.

Labels

- The top-most label (MPLS) is the transport label. This gets swapped between LSRs and popped at the egress PE.
- One transport label per VRF, not per route.
- The label between the top-most and the prefix is the mBGP label (VPN).
- One mBGP label per route in the VRF.

L3VPN Configuration Steps

- Create VRFs and associate interfaces.
- MPLS and routing infrastructure is operational.
- Create VPNv4 infrastructure with mBGP peerings.
- Configure PE-CE routing.
- Configure mBGP -> PE-CE redistribution, this is not needed if PE-CE connection is using eBGP.

```
vrf definition 1
 rd 192.168.0.1:1
 address-family ipv4
  route-target both 1:1

router bgp 1
 no bgp default ipv4-unicast
 neighbor 192.168.0.2 remote-as 1
 neighbor 192.168.0.2 update-source Loopback0
 add vpnv4
  neighbor 192.168.0.2 activate
  neighbor 192.168.0.2 send-community both
```

VPNv4 mBGP Peering Rules

- BGP needs a peering with a loopback address.
- If peered with physical address the PHP pops the label to soon because its a directly connected network.
- Fix non-loopback peering MPLS mBGP with route-map.
- A P router can also be configured as a RR. This will automatically disable the route-target filter.
- When using eBGP between VPNv4 peers the RT filter has to de explicitly disabled on the P router with the `no bgp default route-target filter` command.

```
route-map NEXT_HOP
 set ip next-hop 192.168.0.1
router bgp 1
 add vpnv4
  neighbor 10.0.12.2 route-map NEXT_HOP out
  neighbor 192.168.0.2 route-reflector-client
```

VRF Route-Leaking between Sites:

```
ip prefix-list Lo1 permit 11.11.11.11/32

route-map EXPORT_MAP permit 10
 match ip address prefix Lo1
```

```
 set extcommunity rt 11:11 additive

vrf definition 1
 rd 192.168.0.1:1
 address-family ipv4
  route-target both 1:1
  export map EXPORT_MAP

vrf definition 2
 rd 192.168.0.1:2
 address-family ipv4
  route-target both 2:2
  route-target import 11:11
```

BGP

Tuesday, January 12, 2016
12:16 PM

BGP PE-CE

Either associate each CE with a different AS (65000+) or give each CE the same AS.
If same AS on CE:

- Configure CEs to allow their own AS inbound.
- Override CEs AS when forwarding BGP prefixes.
- Prevent loops between CE backdoors by setting the Site of Origin (SoO) on the PE

```
ip bgp-community new-format
router bgp 1
 no bgp default ipv4-unicast
 neighbor 192.168.0.3 remote-as 1
 neighbor 192.168.0.3 update-source Loopback0
 address-family vpnv4
  neighbor 192.168.0.3 activate
  neighbor 192.168.0.3 send-community both
 address-family ipv4 vrf 2
  neighbor 172.0.58.8 remote-as 65001
  neighbor 172.0.58.8 activate
  neighbor 172.0.58.8 send-community
  neighbor 172.0.58.8 as-override
  neighbor 172.0.58.8 soo 1:1

router bgp 65001
 address-family ipv4
  neighbor 172.0.58.5 remote-as 1
  neighbor 172.0.58.5 allowas-in
```

BGP Multipathing

- Same rules apply for normal multipathing.
- Routes from different PEs must have the same values in relation to cost, med, local-preference etc.
- When using full-mesh peering, the RD can be the same on all PEs.
- When using RR peering, the RD must be unique between the PEs that advertise the same routes.
- The route-import and export values can be the same between PEs.
- Specify the eibgp keyword otherwise multipathing is only applied for iBGP routes.

```
 address-family ipv4 vrf 2
  maximum-path eibgp 32
```

BGP GRE

Tuesday, January 12, 2016
12:16 PM

BGP GRE Tunnels (Old way)

- Alternative to using a MPLS / L3VPN configuration with VRFs:

```
int tun0
 ip address 13.0.0.1 255.255.255.0
 tunnel source lo0
 tunnel destination 192.168.0.3

route-map NEXT_HOP_TUNNEL
 set ip next-hop 13.0.0.1
router bgp 3
 neighbor 192.168.0.3 remote-as 3
 neighbor 192.168.0.3 update-source lo0
 address-family ipv4
 neighbor 192.168.0.3 route-map NEXT_HOP_TUNNEL out
```

EIGRP / RIP

Tuesday, January 12, 2016
12:16 PM

EIGRP

- EIGRP routes that are redistributed into BGP receive a Cost Community ID of 128 by default.
- This will ensure that routes over the L3VPN are considered internal.
- This value is compared before all other BGP path selection attributes (including weight).
- The value/cost of the pre-bestpath community is the composite metric of the redistributed EIGRP route.
- Routes without this cost community are evaluated as if they had a cost value of 2147483647, which is half of the maximum possible value.
- Possible to modify the Cost Community ID. Lower values are better. Apply on EIGRP redistribution point or between VPNv4 neighbors.
- Ignore the cost community with the bgp bestpath cost-community ignore command.

```
router eigrp VPNV4
 address-family ipv4 unicast vrf 2 autonomous-system 1
 network 172.0.17.0 0.0.0.255
  topology base
   redistribute bgp 1 metric 100000 10 255 1 1500

router bgp 1
 add ipv4 vrf 2
  redistribute eigrp 1

ip prefix-list CE_Lo0 permit 192.168.0.6/32

route-map SET_EXT_COMM permit 10
 match ip address prefix CE_Lo0
 set extcommunity cost pre-bestpath 127 2662400
route-map SET_EXT_COMM permit 99

router bgp 1
address-family ipv4 vrf 2
 redistribute eigrp 1 route-map SET_EXT_COMM
```

RIP

- Backdoor CE-CE using Offset-Lists
- An offset list of 7 in both directions will ensure that routes are not looped around the MPLS backbone.

```
router bgp 1
address-family ipv4 vrf 2
 redistribute rip route-map RIP

router rip
 address-family ipv4 vrf 2
  no auto
  version 2
  redistribute bgp 1 metric 1

router rip
 no auto
 version 2
 offset-list 0 in 7 FastEthernet0/0
 offset-list 0 out 7 FastEthernet0/0
```

LDP / TDP

Tuesday, January 12, 2016
12:16 PM

Tag Distribution Protocol (TDP)

- Uses UDP 711 to create adjacency.
- Uses destination port TCP 711 to create session, random TCP source port.
- Does not support authentication.

TDP uses same commands as LDP. The only requirement is that it is either enabled globally or separately per interface.

- If not specified, LDP will be the default label protocol.
- TDP and LDP can coexist on the same router. However they must match on interfaces between neighbors.
- Interface configuration takes precedence over global configuration.
- TDP and LDP use a default hello timer of 5 seconds and a hold timer of 15 seconds.

Label Distribution Protocol (LDP)

- Uses UDP 646 to create adjacency.
- Uses destination port TCP 646 to create session, random TCP source port.
- Supports authentication.

The LDP Router-ID is highest loopback by default followed by highest interface.

- The LDP RID is an actual IP address, unlike the BGP or OSPF RID.
- The LDP RID will be used as the transport address by default, meaning that it must be reachable.
- Between neighbors the LSR with the highest RID will initiate the TCP session.

Uses the all router multicast address 224.0.0.2 (UDP 646 or 711) to form the neighborship.

- Label Switched Paths (LSPs) are unidirectional.

```
ip cef
mpls ldp router-id loopback0 force
no mpls ip
mpls label range 100 199
mpls ip
int fa0/0
 mpls ip

show mpls ldp discovery detail
show mpls ldp neighbor detail
show tcp brief
```

MPLS Transport Address

- The transport address can circumvent the RID for MPLS LDP neighbor advertisement.

```
int fa0/0
 mpls ldp discovery transport-address 10.0.12.1
```

MPLS LDP Authentication

- If the session is already active, the password will have no effect until the session is cleared.
- Specifying the required keyword will require the local router to specify a password before neighborship can form.
- The LDP password applies to the RID, not the transport address. This RID must be reachable by the neighboring router.

```
mpls ldp neighbor 192.168.0.2 password cisco
mpls ldp password required
mpls ldp password rollover duration
```

A password rollover takes effect after the duration when passwords are configured without the use of a key chain.

- This feature is used when statically configured neighbor passwords (not with the option) need to be changed on the router.

```
ip access-list standard LDP_AUTH
 permit 192.168.0.0 0.0.0.255

key chain LDP
 key 1
  key-string cisco

mpls ldp password required
mpls ldp password option 1 for LDP_AUTH key-chain LDP
```

Globally configure password for all peers.

- Only used if neighbor password and password option are not configured

```
mpls ldp password fallback cisco
```

Targeted Session

Can speed up label convergence time when the connection is restored after a failure.

- With a targeted session the session state between neighbors is kept after a failure,
- This is done by setting the holdtime to infinite.
- The targeted session can also be enabled globally for all peers or per prefix. Default duration is 24h.

```
mpls ldp neighbor 192.168.0.1 targeted ldp | tdp
mpls ldp session protection
```

MPLS Filtering

Only add labels for /32 prefixes (preferred):

```
mpls ldp label
 allocate global host-routes
```

Other methods (not preferred):

```
ip prefix-list LOOPBACKS permit 192.168.0.1/32
ip prefix-list LOOPBACKS permit 192.168.0.2/32
ip prefix-list LOOPBACKS permit 192.168.0.3/32
ip prefix-list LOOPBACKS permit 192.168.0.4/32

mpls ldp label
 allocate global prefix-list LOOPBACKS
```

Or:

```
ip access-list standard LDP_ADV
 permit host 192.168.0.1
 permit host 192.168.0.2
 permit host 192.168.0.3
 permit host 192.168.0.4
no mpls ldp advertise-labels
mpls ldp advertise-labels for LDP_ADV

ip access-list standard LDP_REC
 permit host 192.168.0.1
 permit host 192.168.0.2
 permit host 192.168.0.3
 permit host 192.168.0.4
mpls ldp neighbor 192.168.0.2 labels accept LDP_REC
```

Or, in the case of OSPF IGP:

```
router ospf 1
 prefix-suppression

int fa0/0
 ip ospf prefix-suppression
```

MPLS TTL

Hide the MPLS backbone by setting the TTL of traceroute traffic to 255.

- Forwarded applies to transit traffic.
- Local applies to locally generated traffic.
- Default is both.

```
no mpls ip propagate-ttl
```

Handling of TTL expiring on packets:

```
mpls ip ttl-expiration pop 1-6 (default is 0)
```

With 0 a packet with an expired TTL is forwarded by the global routing table.
With 1-6 a packet is forwarded by the underlying label, if more than 1 label is present.

Allow the default route to be associated with a label:

```
mpls ip default-route
```

OSPF

Tuesday, January 12, 2016
1:04 PM

mBGP OSPF Super Backbone

The super backbone exists as the BGP cloud itself, as an area logically above Area 0.
If both CE routers connect to the PE are using OSPF area 0, the process ID that is configured on the CE-PE connection matters.

- If the process ID is different on the CEs the routes through the super backbone will be seen as external.
- If the process ID matches, the routes through the super backbone will be seen as inter-area routes.
- The OSPF domain ID is based on the process ID.
- If the CE's are using VRF-lite then it is required to disable the downward bit (D-bit) loop-prevention check when using the same domain-id.

```
router ospf 17 vrf 2
 network 172.0.17.0 0.0.0.255 area 0
 redistribute bgp 1 subnets
 domain-id 12.12.12.12
 capability vrf-lite

router bgp 1
 add ipv4 vrf 2
  redistribute ospf 26 vrf 2
```

OSPF Backdoor CE-CE using Sham-Links

- The OSPF link through the MPLS cloud would be an inter-area link despite both CEs being in area 0.
- Internal routes are always preferred over inter-area and external routes, so all traffic will flow over the backdoor link.

Configuration steps:

- Create new loopback interfaces on PE routers.
- Associate loopback interfaces with VRF instance and assign unique IP.
- Advertise loopbacks in mBGP, but not in OSPF (Route-Map Filter).
- PE routers must be an ASBR, redistribute mBGP -> OSPF.
- Create sham-links on PE routers between new loopbacks.
- Modify cost on CE routers preferred internal interfaces.

Sham-links are Type-5 external LSAs. Networks sent over the sham-link are Type-1 LSA.

```
int lo1
 description SHAM LOOPBACK
 vrf forwarding 2
 ip add 1.1.1.1 255.255.255.255

router bgp 1
 address-family ipv4 vrf 2
  network 1.1.1.1 mask 255.255.255.255

ip prefix-list SHAM_LOOPBACK permit 1.1.1.1/32
ip prefix-list SHAM_LOOPBACK permit 2.2.2.2/32

route-map BLOCK_SHAM deny 10
 match ip address prefix-list SHAM_LOOPBACK
route-map BLOCK_SHAM permit 99

router ospf 17 vrf 2
 area 0 sham-link 1.1.1.1 2.2.2.2
 redistribute bgp 1 subnets route-map BLOCK_SHAM
```

QoS

Tuesday, January 12, 2016
1:03 PM

MPLS Experimental Bits (EXP bits)

The DSCP value set on the IP packet is not changed by MPLS by default. Uses two modes:

- Pipe Mode. Uses egress queues based on EXP bits.
- Short Pipe Mode. Uses egress queues based on the 'original' ToS (DSCP) bits.

The DSCP value in the IP packet can also be replaced by the EXP bits, this is called 'Uniform Mode'.

QoS Matching on PE using Groups

- On ingress interface, it is not possible to match on a IPP or DSCP value, because the MPLS header is still on the frame.
- On egress interface, it is not possible to match on the EXP bits to set IPP / DSCP bits, because the label is already popped.

The solution is to use QoS groups, which are local to the device itself.

- A packet is marked with a QoS group value only while it is being processed within the device.
- The QoS group value is not included in the packet's header when the packet is transmitted over the output interface.

```
class-map match-all EXP5
 match mpls experimental topmost 5
policy-map MPLS_INGRESS
 class EXP5
  set qos-group 5

int fa0/0
 description MPLS_CORE
 service-policy input MPLS_INGRESS

class-map match-all QOS_GROUP5
 match qos-group 5

policy-map MPLS_EGGRESS
 class QOS_GROUP5
  set ip dscp af41

int s1/0
 description VRF_CE
 service-policy output MPLS_EGGRESS
```

Table-Maps

- Rewrite all ingress traffic using a Table Map.
- Table-Maps map a QoS group to a specific ToS value using DSCP values.

```
class-map match-all EXP5
 match mpls experimental topmost 5
policy-map MPLS_INGRESS
 class EXP5
  set qos-group 5

int fa0/0
 description MPLS_CORE
 service-policy input MPLS_INGRESS

table-map TABLE_MAP
 map from 1 to 8
```

```
 map from 2 to 16
 map from 3 to 24
 map from 4 to 32
 map from 5 to 40
 map from 6 to 48
 map from 7 to 56

policy-map MPLS_EGGRESS
 class class-default
  set dscp qos-group table TABLE_MAP

int s1/0
 description VRF_CE
 service-policy output MPLS_EGGRESS
```

MPLS Implicit-null / Explicit-null

An implicit-null label is set to instruct upstream routers that they should perform PHP.

- The implicit-null label is set for directly connected prefixes on each LSR.

An explicit-null label is used in QoS combined with MPLS.

- When a packets gets encapsulated in MPLS, there is the option of copying the IP precedence to the MPLS header (EXP bits).
- If a POP is performed (implicit-null) at the penultimate LSR, the EXP bits in the MPLS header are removed as well.
- With explicit-null the MPLS header is left intact until it reaches the Egress LSR.
- Explicit Null is advertised in place of Implicit Null for directly connected prefixes.
- Configure with the mpls ldp explicit-null global command. Default is to enable explicit-null for all local prefixes.

Limit explicit-null for route 10.10.10.0 only:

```
ip access-list standard EXP_NULL
 permit host 10.10.10.0
mpls ldp explicit-null for EXP_NULL
```

Limit explicit-null to peer 192.168.0.2 only:

```
ip access-list standard EXP_NULL
 permit host 192.168.0.2
mpls ldp explicit-null to EXP_NULL
```

Multicast

Saturday, December 19, 2015
12:34 PM

Multicast Addressing

Link-Local	224.0.0.0/24	Used by network protocols on a local network segment. Non-Routable traffic.
Globally Scoped	224.0.1.0 - 238.255.255.255	Normal range. Can send between organizations and across the Internet.
SSM	232.0.0.0/24	Source-Specific Multicast (SSM)
Private Multicast	239.0.0.0/8	Administratively scoped addresses. Equivalent of RFC1918 address space.
GLOP	233.0.0.0/8	Maps 16bit AS to multicast groups.

GLOP Addresses

- Convert AS to Hex, then take the separate parts of the 4-part hex and convert them back into decimal groups.
- Or convert straight to binary and separate the 16bit value into two groups of 8 and convert back into decimal.
- 0 and 255 are valid addresses in these ranges.

AS	AS in binary	AS in hex	AS in decimal	GLOP address
65053	11111110.00011101	FE.1D	254.29	233.254.29.0/24
64512	11111100.00000000	FC.00	252.00	233.252.0.0/24

Protocol Independent Multicast (PIM)

- Forms adjacencies with neighboring PIM routers.
- Default hello timer is 30 seconds, hold-time is 3x hello.
- PIMv2 hello uses IP protocol 103 and 224.0.0.13 (ALL-PIM-Routers address).

An interface can be configured in three different modes (RPF must succeed):

- Dense-Mode. Traffic is flooded on all enabled PIM interfaces.
- Sparse-Mode. Traffic is only forwarded only on interfaces with downstream clients. Uses RPs.
- Sparse-Dense-Mode. If RP is not known, operate in dense mode. If RP is known, operate in Sparse-Mode.

Internet Group Messaging Protocol (IGMP)

- IGMP messages are sent in IP datagrams with IP protocol number 2, and a TTL of 1.
- IGMP packets pass only over a LAN and are not forwarded by routers, because of their TTL field values.
- IGMPv1. Clients join a multicast group.
- IGMPv2. Clients can also leave (all) groups. Backwards compatible with v1.
- IGMPv3. Clients can join and leave specific groups. Support for SSM.

Anycast RP

Saturday, December 19, 2015
12:34 PM

Anycast RP

- PIM Register and Join messages go to the closest RP in the topology.
- When PIM Register is received, MSDP Source Active (SA) is sent to MSDP peers which synchronize (S,G) information.
- The originator-id must point to a unique address on the router (do not use the same loopback used for the MSDP peering).
- The anycast loopback (not the peering loopback) is specified with the rp-candidate (BSR) or send-rp-announce (Auto-RP) command.

```
int lo0
 description MSDP_PEERING
 ip address 192.168.0.1 255.255.255.0255
 ip pim sparse-mode
int lo1
 description ANYCAST_RP
 ip address 12.12.12.12 255.255.255.255
 ip pim sparse-mode

ip msdp originator-id lo0
ip msdp peer 192.168.0.2 connect-source lo0
ip msdp mesh-group MSDP 192.168.0.2

ip pim bsr-candidate lo0
ip pim rp-candidate lo1

show ip msdp peer
show ip msdp sa-cache
debug ip msdp detail
debug ip msdp peer
```

Exchange Multicast Information Without Anycast RP

- You do not necessarily need the same loopback IP when you want to exchange multicast information between multiple RPs.
- These RP's do not have the same IP address configured, instead it is just two separate RP's that will exchange multicast information.
- Can be used to link multicast areas together and have multiple RPs coexist with each other.

```
int lo0
 ip address 192.168.0.1 255.255.255.0255
 ip pim sparse-mode

ip msdp peer 192.168.0.2 connect-source lo0
ip pim bsr-candidate lo0
ip pim rp-candidate lo0
```

Auto-RP

Saturday, December 19, 2015
12:34 PM

Auto-RP Discovery and Announcements

- Auto-RP needs a RP to form multicast trees and allow traffic to flow.
- However the location of the RP has to be discovered through multicast as well.
- This creates a chicken and the egg situation. In order to find the RP, some kind of dense-mode solution is needed.
    - Statically assign the mapping agent and the RP for the 224.0.1.39-40 addresses. Kind of defeats the purpose of Auto-RP.
    - Configure interfaces with ip pim parse-dense mode. Uses dense mode for all groups without an RP, sparse for all others.
    - Configure the ip pim autorp listener. Allows usage of sparse-mode only interfaces and basically configures an ACL for the 224.0.1.39-40 addresses to be allowed to run in dense mode (preferred method).

Mapping Agent

- Receive candidate messages (announcements) and decide which one will be the RP (Highest IP address wins).
- The mapping agents listen on 224.0.1.39 and propagate the decision to all other routers via 224.0.1.40.
- All routers join the 224.0.1.40 group by default, but only the mapping agents join 224.0.1.39.
- Configure with ip pim send-rp-discovery, the scope has be large enough to reach the DR for the PIM segments.

Rendezvous Point (RP)

- Send multicast announcements (Dense Mode) to announce their RP candidacy to 224.0.1.39.
- Configure with ip pim send-rp-announce, the scope has be large enough to reach the mapping agent.
- If the mapping agent and the RP are the same router, a scope of 1 is enough.

```
ip pim autorp listener
ip pim send-rp-announce Loopback0 scope 255
ip pim send-rp-discovery Loopback0 scope 255
```

Static Auto-RP groups without listener (configure on all mrouters, including RP):

```
ip access-list standard AUTO_RP
 permit host 224.0.1.40
 permit host 224.0.1.39

ip pim rp-address 192.168.0.1 AUTO_RP
```

Auto-RP Filtering

- Filter RP announcement messages on the mapping agent to only allow specific RPs, or bind RPs to specific groups.
- Configure on mapping agent only.

```
ip access-list standard RP1
 permit host 192.168.0.1
ip access-list standard RP2
 permit host 192.168.0.2

ip access-list standard GROUP_224_231
 permit 224.0.0.0 7.255.255.255

ip access-list standard GROUP_232_239
 permit 232.0.0.0 7.255.255.255

ip pim rp-announce-filter rp-list RP1 group-list GROUP_224_231
ip pim rp-announce-filter rp-list RP2 group-list GROUP_232_239
```

Deny all other RPs:

```
ip access-list standard OTHER_RP
```

```
 deny host 192.168.0.3
 deny host 192.168.0.4
 permit any

ip access-list standard GROUP_224_239
 permit 224.0.0.0 15.255.255.255

ip pim rp-announce-filter rp-list OTHER_RP group-list GROUP_224_239
```

Auto-RP Cache Filtering

- Accept only (*, G) join messages destined for the specified Auto-RP cached address.
- Accept join and prune messages only for RPs in Auto-RP cache.
- Configure with the `ip pim accept-rp auto-rp` command on all mrouters.

BSR

Saturday, December 19, 2015
12:34 PM

Bootstrap Router (BSR)

- The rp-candidate is the actual RP. The bsr-candidate is the mapping agent
- Messages are flooded hop-by-hop by all multicast routers, this is because 224.0.0.13 is a link-local address.

Designed for Sparse-Mode, there is no need for Dense-Mode. Flooding is a control-plane feature and can be debugged.

- Auto-RP uses routable addresses that are outside the 224.0.0.0/24 range (224.0.1.39-40).
- BSR on the other hand uses link-local addresses, so its easier to control where traffic is flooded.
- Even though these messages are flooded, they are still subject to the RPF check.
- The edge of the BSR network can be specified with the ip pim bsr-border interface command.

```
ip access-list standard GROUP_224_231
 permit 224.0.0.0 7.255.255.255

ip pim rp-candidate Loopback 0 group-list GROUP_224_231
ip pim bsr-candidate Loopback 0

show ip pim bsr-router
debug ip pim bsr
```

BIDIR-PIM

Saturday, December 19, 2015
12:34 PM

Bidirectional PIM (BIDIR-PIM)

- No source-based (S,G) trees.
- RP builds a shared tree through which source routers forward traffic downstream toward the RP.
- The RP in BIDIR-PIM is always in the data plane, so placement is important.
- It's possible to limit which groups will be enabled for BIDIR, and use another or the same RP for regular ASM.

```
ip access-list standard BIDIR_RANGE
 permit 228.0.0.0 0.255.255.255
 permit 229.0.0.0 0.255.255.255

ip pim bidir-enable
ip pim bsr-candidate Loopback0
ip pim rp-candidate Loopback0 bidir BIDIR_RANGE
```

PIM-DM

Saturday, December 19, 2015
12:34 PM

PIM Dense Mode (PIM-DM)

- Forwards multicast traffic out of all interfaces, except the one received.
- Does not forward if no active downstream router and no hosts joined group.
- If both are true, router informs upstream router to stop sending via a prune message.
- If host joins the network after prune, then routers will use graft message to override prune.
- Only uses SPT.

State refresh messages can be used to 'refresh' the state before the 3min prune timer.

- This will stop all routers from pruning and un-pruning on the specified interval.
- Only has to be enabled on interface pointing to source, mrouter closest to source will relay state-refresh messages.
- Disable state-refresh with the ip pim state-refresh disable command.

```
int fa0/0
 ip pim state-refresh origination-interval 60
```

PIM-DM Assert

- Prevents multiple senders from replicating the same multicast stream on to the wire.
- Used in dense-mode and enabled automatically.

In order to trigger a PIM Assert the (S,G) has to match exactly.

- IE both transferring routers need to connected to the same segment.
- Specifying a (loopback) source on the sender of the multicast traffic has no effect.

The winner is decided by:

- Lowest AD back to the source.
- In a tie, best metric value.
- In a tie, highest IP address.

PIM-SM

Saturday, December 19, 2015
12:34 PM

PIM Sparse Mode (PIM-SM)

- Assumes that no clients want to receive multicast packets until they specifically ask to receive them.
- Downstream routers request multicast traffic using PIM Join messages.
- Routers keep sending Joins, otherwise they are pruned.

Rendezvous Point (RP)

- A common, agreed place in the network where clients can meet multicast sources.
- Not necessarily the center of the network.
- If there are many clients, it is better to make the router closest to the clients the RP.
- If there is only one source, it is better to make the router closes to the source the RP.
- All mrouters need to be configured with the location of the RP.

RP Messages

- Multicast receivers (clients) inform the router that they want to receive multicast traffic, this is the (*,G) state.
- Multicast sources also inform the RP that they are sending multicast traffic, this is the (S,G) state (register message).
- This information is propagated through the network, by all routers that know the location of the RP.

- 3min state, after 3min the client will re-register with the RP. The RP informs routers to stop sending with register-stop message.
- As long as the source is transmitting this register-stop-register-stop state will continue
- This is similar to Dense-Mode, except that it is only between the RP and the source router (instead of all routers).

RPF Failures/Fixes

Monday, December 21, 2015
11:46 AM

RPF Failures/Fixes

- Fix with static mroutes, multicast-BGP or tweaking unicast routing.

```
traceroute
mrinfo
show ip route multicast
show ip mfib
show ip mroute
show ip mroute count
show ip rpf
mtrace
debug ip pim
debug ip mfib pak
```

Tunneling to fix RPF Failures

- Usage of loopback source/destination is preferred.
- Don't forget to enable PIM on the tunnel interface as well.

```
interface Tunnel 12
 ip address 12.0.0.1 255.255.255.0
 ip pim sparse-mode
 tunnel source Loopback 0
 tunnel destination 192.168.0.2

ip mroute 0.0.0.0 0.0.0.0 Tunnel 12
```

Multicast-BGP to fix RPF Failures

- Static mroute is preferred over dynamic MBGP routes.
- Administrative distance of eBGP will make sure that MBGP routes are preferred over unicast routes (EIGRP or OSPF).
- Administrative distance of the IGP needs to be lowered on the router closest to the receiver in order for iBGP to be preferred.
- Advertise the source of the multicast traffic, and the location of the RP into MBGP on the router closest to the source.
- Works similar in concept to a static mroute, only the information is propagated by BGP.
- Advertise the network that needs to go over a different path instead of the unicast routing path.

Change the next_hop to the next BGP destination that the neighboring router must take.

- Remember that multicast will only try to go over PIM enabled interfaces.

Instruct R3 to choose R4 as the next-hop for mtraffic destined towards 172.16.0.0/24 (R1 is the RP):

```
router bgp 234
 neighbor 10.0.234.3 remote-as 234
 add ipv4 multicast
  network 172.16.0.0 m 255.255.255.0
  network 1.1.1.1 m 255.255.255.255
  neighbor 10.0.234.3 activate
  neighbor 10.0.234.3 route-map NEXT_HOP_MC out
  distance bgp 20 20 200

route-map NEXT_HOP_MC permit 10
   set ip next-hop 10.0.234.4
```

SSM

Monday, December 21, 2015
11:46 AM

Source Specific Multicast (SSM)

- Does not require RP, BSR or Auto-RP.
- Receiver specifies the source address, RPF is still applied.
- Specify a custom range with the range statement (must fall within the 232.0.0.0/8 range).

Configure router closest to receiver on same link:

```
ip access-list standard SSM_RANGE
 permit host 232.0.0.1
 permit host 232.0.0.2

ip pim ssm range SSM_RANGE
int fa0/0
 description MULTICAST_SOURCE
 ip igmp version 3
 ip pim sparse-mode
```

Configure receiver of multicast traffic:

```
ip pim ssm default
int fa0/0
 description MULTICAST_RECEIVER
 ip pim-sparse mode
 ip igmp version 3

int lo0
 ip pim sparse-mode
 ip igmp join-group 232.0.0.1 source 192.168.0.1
 ip igmp join-group 232.0.0.2 source 192.168.0.1
 ip igmp join-group etc..
```

SSM IGMP Filtering

Configure router closest to receiver on same link to filter specific groups:

```
ip access-list extended SSM_GROUPS
 permit igmp any host 232.0.0.1

int fa0/0
 ip igmp access-group SSM_GROUPS
```

Static RP

Monday, December 21, 2015
11:46 AM

Static RP Configuration

- Statically configure each router with the location of the RP.
- This also has to be configured on the RP to point to itself.
- By default dynamically learned RP (BSR, Auto-RP) is preferred over static. Override this behavior with the override keyword.
- Between BSR and Auto-RP there is no preferred mapping order, the last mapping learned is preferred.

```
ip access-list standard GROUP_224_231
 permit 224.0.0.0 7.255.255.255

ip access-list standard GROUP_232_239
 permit 232.0.0.0 7.255.255.255

ip pim rp-address 192.168.0.1 GROUP_224_231 override
ip pim rp-address 192.168.0.2 GROUP_232_239 override
```

RP Register Filtering

- Prevent unauthorized sources from registering with the RP (S,G). Configure on RP.
- If an unauthorized source sends a register message to the RP, the RP will immediately send back a register-stop message.

```
ip access-list standard GROUP_224_231
 permit 224.0.0.0 7.255.255.255

ip pim accept-register list GROUP_224_231
```

RP Join Filtering

- Accept only (*, G) join messages destined for the specified RP address.
- Configure on mrouters, and optionally on RP.
- The group address must be in the range specified by the access list.
- If the RP points to itself, the RP will only accept registers from that particular multicast range.
- This is basically the same as the ip pim accept-register command.

```
ip access-list standard GROUP_224_231
 permit 224.0.0.0 7.255.255.255

ip access-list standard GROUP_232_239
 permit 232.0.0.0 7.255.255.255

ip pim accept-rp 192.168.0.1 GROUP_224_231
ip pim accept-rp 192.168.0.2 GROUP_232_239
```

Dense-Mode Fallback

- Dense mode fallback allows the usage of dense mode if the RP becomes unreachable.
- Requires sparse-dense-mode configured on interfaces.

```
ip pim dm-fallback
int fa0/0
 ip pim sparse-dense-mode
```

NAT

Tuesday, December 8, 2015
9:14 PM

Application Level Gateway (ALG)

- Some protocols embed IP address information in the Application Level payload.
- Regular NAT does not check the application level for protocols such as FTP, HTTP, DNS, SIP.
- ALG allows the use of dynamic ports by clients.
- ALG is on by default. Disable ALG by specifying the no-payload command.

NVI

Monday, December 21, 2015
5:23 PM

NAT Virtual Interface (NVI)

- No more concept of nat inside and nat outside interfaces.
- The add-route keyword also adds the NAT_POOL route to the RIB, this can then be redistributed into BGP.
- Using this method, the outside interface address does not necessarily have to match the NAT_POOL ip range.

```
int fa0/0
 description PRIVATE_TO_R2
 ip address 10.0.12.1 255.255.255.0
 ip nat enable
int se1/0
 description PUBLIC_TO_R4
 ip address 14.0.0.1 255.255.255.0
 ip nat enable

ip access-list standard NAT_ACL
 permit 10.0.12.0 0.0.0.255

ip nat pool NAT_POOL 12.0.0.2 12.0.0.10 prefix-length 24 add-route
ip nat source list NAT_ACL pool NAT_POOL

router bgp 1
 neighbor 14.0.0.4 remote-as 4
 address-family ipv4
  network 12.0.0.0 mask 255.255.255.0
```

Dynamic

Monday, December 21, 2015
5:22 PM

Dynamic NAT

- Common NAT usage to map multiple Inside Local addresses to a single Inside Global address.
- Only 65536 ports are available, extend the range by specifying more address in a pool.
- The main differentiator between dynamic and static NAT is the overload keyword.

```
int fa0/0
 description PRIVATE_TO_R2
 ip address 10.0.12.1 255.255.255.0
 ip nat inside
int se1/0
 ip address 14.0.0.1 255.255.255.0
```

```
 description PUBLIC_TO_R4
 ip nat outside

ip access-list standard NAT_ACL
 permit 10.0.12.0 0.0.0.255

ip nat inside source list NAT_ACL int se1/0 overload
```

Dynamic NAT Pool

```
ip access-list standard NAT_ACL
 permit 10.0.12.0 0.0.0.255

ip nat pool NAT_POOL 14.0.0.2 14.0.0.10 prefix-length 24
ip nat inside source list NAT_ACL pool NAT_POOL overload
```

Dynamic NAT Pool using Route-Maps

- Use route-maps alongside dynamic NAT pools to provide more granular control.
- Can translate different traffic types to different outside addresses.

```
ip access-list extended NAT_ICMP
 permit icmp 10.0.12.0 0.0.0.255 any
ip access-list extended NAT_TCP
 permit tcp 10.0.12.0 0.0.0.255 any
ip access-list extended NAT_UDP
 permit udp 10.0.12.0 0.0.0.255 any

route-map NAT_ICMP_RM permit 10
 match ip address NAT_ICMP
route-map NAT_TCP_RM permit 10
 match ip address NAT_TCP
route-map NAT_UDP_RM permit 10
 match ip address NAT_UDP

ip nat pool NAT_POOL_ICMP 14.0.0.2 14.0.0.10 prefix-length 24
ip nat inside source route-map NAT_ICMP_RM pool NAT_POOL_ICMP overload

ip nat pool NAT_POOL_TCP 14.0.0.11 14.0.0.20 prefix-length 24
ip nat inside source route-map NAT_TCP_RM pool NAT_POOL_TCP overload

ip nat pool NAT_POOL_UDP 14.0.0.21 14.0.0.30 prefix-length 24
ip nat inside source route-map NAT_UDP_RM pool NAT_POOL_UDP overload
```

Policy

Monday, December 21, 2015
5:22 PM

Policy NAT

- Uses tracking alongside route-maps in order to provide continuous NAT services on multiple outside interfaces.
- If main neighbor goes offline, remote the static route and switch over to the other neighbor.
- Poor man's NAT redundancy.
- Does not work with NAT pools.

```
int fa0/0
 description PRIVATE_TO_R2
 ip address 10.0.12.1 255.255.255.0
 ip nat inside
int se1/0
 description PUBLIC_TO_R3
 ip address 13.0.0.1 255.255.255.0
 ip nat outside
int se1/1
 description PUBLIC_TO_R4
 ip address 14.0.0.1 255.255.255.0
 ip nat outside

ip sla 1
 icmp-echo 13.0.0.3 source-interface se1/0
  frequency 5
track 1 ip sla 1

ip route 0.0.0.0 0.0.0.0 13.0.0.3 track 1
ip route 0.0.0.0 0.0.0.0 14.0.0.4 5

ip access-list standard NAT_ACL
 permit 10.0.12.0 0.0.0.255

route-map NAT_13 permit 10
 match ip address NAT_ACL
 match int se1/0
route-map NAT_14 permit 10
 match ip address NAT_ACL
 match int se1/1

ip nat inside source route-map NAT_13 interface se1/0 overload
ip nat inside source route-map NAT_14 interface se1/1 overload
```

Rotary

Monday, December 21, 2015
5:25 PM

Rotary NAT (Round-Robin Load-Balancing)

- Can load-balance for servers located in the Private subnet range.
- This will only work for TCP traffic and the connections are forwarded in a round-robin (rotary) style to the inside network.
- If the outside connection is using ethernet interfaces, an ip alias needs to be created for the specific TCP traffic.
- Uses a destination list instead of a source list.

Send Telnet traffic to R2 and R3 in a round-robin fashion:

```
int fa0/0
 description PRIVATE_TO_R2_R3
 ip address 10.0.123.1 255.255.255.0
 ip nat inside
int fa0/1
 description PUBLIC_TO_R4
 ip address 14.0.0.1 255.255.255.0
 ip nat outside

ip alias 14.0.0.100 23
ip access-list standard ROTARY_NAT
 permit host 14.0.0.100

ip nat pool NAT_POOL 10.0.123.2 10.0.123.3 prefix-length 24 type rotary
ip nat inside destination list ROTARY_NAT pool NAT_POOL
```

Static

Monday, December 21, 2015
5:27 PM

Static NAT

- Map a single Inside Local (IL) address to a single Inside Global (IG) address.
- The no-alias keyword will stop the creation of an alias for the global address space.
- The extendable keyword is added automatically in IOS, and is used when the same IL address is mapped to multiple IG addresses.

```
int fa0/0
 description PRIVATE_TO_R2
 ip address 10.0.12.1 255.255.255.0
 ip nat inside
int se1/0
 ip address 14.0.0.1 255.255.255.0
 description PUBLIC_TO_R4
 ip nat outside

ip nat inside source static 10.0.12.2 14.0.0.100 extendable
ip nat inside source static 10.0.12.2 14.0.0.200 extendable
```

Static PAT

- Map specific ports to IG addresses (or the outside interface address).

```
ip nat inside source static tcp 10.0.12.2 23 int se1/0 2323
```

Static NAT Network Range

- Map an entire IL network range to an IG range, preserving host portion addressing.
- This will create a static 1:1 range, meaning that 10.0.12.2 will be mapped to 14.0.0.2 for example.

```
ip nat inside source static network 10.0.12.0 14.0.0.0 /24
```

Static NAT Pool

Configure a pool that contains IG addresses and map these to an access-list that contains the IL addresses.

```
ip access-list standard NAT_ACL
 permit 10.0.12.0 0.0.0.255

ip nat pool NAT_POOL 14.0.0.2 14.0.0.10 prefix-length 24
ip nat inside source list NAT_ACL pool NAT_POOL
```

Reversible Static NAT Pool using Route-Maps

- Use route-maps alongside static NAT pools to provide more granular control.
- The reversible keyword enables outside-to-inside initiated sessions to use route_maps for destination-based NAT.

```
ip access-list extended NAT_ICMP
 permit icmp 10.0.12.0 0.0.0.255 any
ip access-list extended NAT_TCP
 permit tcp 10.0.12.0 0.0.0.255 any
ip access-list extended NAT_UDP
 permit udp 10.0.12.0 0.0.0.255 any

route-map NAT_ICMP_RM permit 10
 match ip address NAT_ICMP
route-map NAT_TCP_RM permit 10
 match ip address NAT_TCP
```

```
route-map NAT_UDP_RM permit 10
 match ip address NAT_UDP

ip nat pool NAT_POOL_ICMP 14.0.0.2 14.0.0.10 prefix-length 24
ip nat inside source route-map NAT_ICMP_RM pool NAT_POOL_ICMP reversible

ip nat pool NAT_POOL_TCP 14.0.0.11 14.0.0.20 prefix-length 24
ip nat inside source route-map NAT_TCP_RM pool NAT_POOL_TCP reversible

ip nat pool NAT_POOL_UDP 14.0.0.21 14.0.0.30 prefix-length 24
ip nat inside source route-map NAT_UDP_RM pool NAT_POOL_UDP reversible
```

SNAT

Monday, December 21, 2015
5:23 PM

Stateful NAT (SNAT)

- The SNAT feature allows multiple routers to share NAT tables.
- When used alongside HSRP, the standby router can take over the NAT translations.
- The standby router can share state with the active router, keeping the NAT sessions alive.
- The mapping-id must be the same between peers. The redundancy string must match the standby name.
- The configuration of the standby router is identical, with the exception of the ip nat stateful id.

```
int fa0/0
 standby 1 name SNAT
 standby 1 ip 10.0.123.254
 ip nat inside
int se1/0
 ip nat outside

ip nat stateful id 1
  redundancy SNAT
   mapping-id 12

access-list standard NAT
 permit 10.0.123.0 0.0.0.255
ip nat pool NAT_POOL 12.0.0.100 12.0.0.100 prefix-length 24 add-route
ip nat inside source list NAT pool NAT_POOL mapping-id 12 overload

router bgp 12
 neighbor 14.0.0.4 remote-as 4
add ipv4
 network 12.0.0.0 mask 255.255.255.0

show ip snat peer 10.0.123.2
show ip snat distributed
```

OSPF

Saturday, January 9, 2016
1:10 PM

OSPF Networking

- Intra-area are preferred over inter-area routes.
- Intra-area and inter-area routes are preferred over external routes.
- E1 routes are preferred over E2, even though they might have higher cost.
- Type-3 LSAs received from the backbone area are never re-advertised into the backbone area.

- Internal routes (network statement) are advertised as Type-1 LSAs and are translated to Type-3 by ABRs.
- The next-hop of redistributed routes are advertised as Type-1 LSAs with the E-bit set by the ASBR and are translated to Type-4 LSAs by the next ABR. The Type-4 LSA is updated by each ABR which changes the next-hop value.
- Redistributed routes are also advertised as Type-5 LSAs, this is the actual route, not the way to reach it.
- The Type-5 LSA next-hop is unchanged, Type-4 is used to reach the Type-5 next-hop.
- The Type-5 external LSA advertising router is not updated, only the Type-4 LSA indicating how to reach this prefix is updated.

OSPF DR / BDR Election

- The DR/BDR election is not preemptive, the first router that boots will be selected.
- If all routers boot at the same time, the DR is chosen based on the highest priority, or the highest RID.
- The DR originates the network LSA on behalf of the network and forms adjacencies with all DROTHERS to synchronize the LSDB.
- BDR and DR perform the same function, except the BDR only sends out updates if the DR goes down.
- In NBMA networks the BDR is elected before the DR.

Primary DR Role

- All routing updates will be forwarded from the DROTHER routers to the DR/BDR using 224.0.0.6.
- The DR will update the LSAs and propagate the changes to the rest of the DROTHER routers using 224.0.0.5.

Secondary DR Role

- Informs all routers in the area of the shared segment using Type-2 LSA.
- This is only performed by the DR, not the BDR. The BDR only receives LSAs and is promoted if the DR fails.

OSPF Network Types

Name	Timers	Hello	Updates	Neighbor	DR / BDR	Mask	Default on type
Point-to-Point	10 / 40	Multicast	Multicast	No	No	-	Serial / Tunnel / FR P2P
Broadcast	10 / 40	Multicast	Multicast	No	Yes	-	Ethernet
Non-Broadcast	30 / 120	Unicast	Unicast	Yes	Yes	-	FR Physical / FR P2MP
Point-to-Multipoint	30 / 120	Multicast	Unicast	No	No	/32	-
Point-to-Multipoint Non-Broadcast	30 / 120	Unicast	Unicast	Yes	No	/32	-
Loopback	-	-	-	-	-	/32	Loopbacks

OSPF Area Types Summary

Type	LSA Allowed	Default route inserted by default	Default route type	Default route generated by

Normal	1,2,3,4,5	No	Type-5	default-information-originate [always]
Stub	1,2,3	Yes	Type-3	area 1 stub
T-Stub	1,2,[3]	Yes	Type-3	area 1 stub no-summary
NSSA	1,2,3,7	No	Type-7	area 1 nssa default-information-originate
T-NSSA	1,2,[3],7	Yes	Type-3 Type-7	area 1 nssa no-summary area 1 nssa no-summary default-information-originate

OSPF ABR / ASBR

- ABR routers must be connected to the backbone areas.
- ABRs are filtering and summarization point for inter-area routes.
- ABR summarizes Type-1 and Type-2 and Type-3 from other ABRs into Type-3 LSAs.
- Summarizes other received Type 3 information.
- ABR announce their ABR status using the Type-1 LSA other flag, the B-bit (Border-bit).
- ASBR set the E-bit to 1 (Edge-bit).

OSPF Non-NSSA Forward Address

- The forward address (FA) is the highest address local enabled for OSPF. This is an actual ip-address not a RID.

If two routers are border routers, and only one is redistributing a route, then this router will be chosen as the next hop regardless of cost.
The above is true unless the following conditions are met:

- Exit interface pointing towards external destination must be enabled for OSPF and must not be passive.
- To be considered valid the external destination forward address must be known as OSPF route.
- The network attached to the external destination must be either broadcast or NBMA.

Authentication

Saturday, January 9, 2016
1:10 PM

OSPF IPv4 Authentication

- Authentication can be configured for an entire area, or per interface.
- The key-strings themselves have to be configured per interface.
- Interface authentication modes overrides area authentication.
- When using MD5 the Key-ID must match between neighbors.

OSPF Supports the following authentication types:

- Null (Type-0). Default.
- Plain-Text (Type-1). Simple authentication.
- MD5 (Type-2). Message-Digest authentication.

Simple Authentication (Type-1)

```
router ospf 1
 area 0 authentication

int fa0/0
 ip ospf authentication-key cisco
 ip ospf authentication
```

MD5 Authentication (Type 2)

```
router ospf 1
 area 0 authentication message-digest

int fa0/0
 ip ospf message-digest-key 1 md5 cisco
 ip ospf authentication message-digest
```

Virtual Link Authentication

- Virtual-Links are always considered to be area 0.
- Virtual-Links use area 0 authentication.

```
router ospf 1
 area 0 authentication message-digest
 area 1 virtual-link 192.168.0.2 authentication message-digest message-digest-key 1 md5 0 cisco
```

OSPF IPv6 Authentication

- Requires the use of IPsec to enable authentication. Only supports full hexadecimal keys.
- To use the IPsec AH header, you use only the ipv6 ospf authentication command.
- When MD5 authentication is used, the key must be 32 hex digits long.
- When SHA-1 authentication is used, the key must be 40 hex digits long.

```
ipv6 router ospf 1
 area 0 authentication ipsec spi 256 md5 1234567890abcdef1234567890abcdef

int fa0/0
 ipv6 ospf authentication ipsec spi 256 md5 1234567890abcdef1234567890abcdef
```

OSPF IPv6 Encryption

- Requires the use of IPsec to enable encryption. Only supports full hexadecimal keys.
- To use the IPsec ESP header, you use the ipv6 ospf encryption command.
- When ESP is set to a non-null value, both encryption and authentication are provided.
- It is not possible to configure encryption and authentication using different commands.

ESP Null:

```
ipv6 router ospf 1
 area 0 encryption ipsec spi 256 esp null md5 1234567890abcdef1234567890abcdef

int fa0/0
 ipv6 ospf encryption ipsec spi 256 esp null md5 1234567890abcdef1234567890abcdef
```

ESP AES-CBC 128:

```
ipv6 router ospf 1
 area 0 encryption ipsec spi 256 esp aes-cbc 128 1234567890abcdef1234567890abcdef md5
1234567890abcdef1234567890abcdef

int fa0/0
 ipv6 ospf encryption ipsec spi 256 esp aes-cbc 128 1234567890abcdef1234567890abcdef md5
1234567890abcdef1234567890abcdef
```

Adjacencies

Saturday, January 9, 2016
1:10 PM

OSPFv2 Adjacency Requirements

- Matching Hello/Dead Timers.
- Matching Area ID.
- Matching Subnet.
- Matching Area type.
- Matching Authentication.
- Matching Network type.
- Matching MTU.

OSPFv3 Adjacency Requirements

- Matching Instance ID.
- Above list except matching subnet, because of usage of link-local address.

OSPF Adjacency Troubleshooting / Stuck States

Stuck in WAIT-State

- Reason: Unreasonably long dead-interval on broadcast and non-broadcast network type. This is because routers spend the dead-interval time (40 seconds by default) in the wait state before becoming FULL neighbors.
- Behavior: Routers will appear as DROTHERS even with priority set to non-zero value.
  show ip ospf interface: State will show as WAIT.
  show ip ospf neighbor: State will show as TWO-WAY.

Stuck in INIT-State

- Reason: One-way communication. OSPF not enabled on one side, L2 problem, mismatched Hello or Dead interval on ethernet interfaces. Dead interval can differ on P2P interfaces.
- Behavior: Routers will not show up as neighbors (blank), even though ospf has been enabled on both sides.
  show ip ospf interface: State will show DR for both routers, and correctly P2P on P2P interfaces.
  show ip ospf neighbor: No state.
- Troubleshoot local sent Hellos with show ip ospf interface, will show own router in INIT phase.
- Troubleshoot remote received Hellos with show ip ospf neighbor, if the neighbor is also shown in INIT phase then the problem is local.

Stuck in TWO-WAY-State

- Reason: Not always a problem (stable state for DROTHERS), problem when both routers on the same segment are set to priority 0 for a broadcast network. No DR is elected so no FULL neighborship can form.
  show ip ospf neighbor: State will show DROTHER for both routers.

Stuck in EXSTART-State

- Reason: MTU mismatch on both sides, this relates to unicast reachability. This indicates a unicast problem where multicast hellos are received properly, the master communicates to the slave that they should move on to exchange using unicast.
- Behavior: One side will show stuck in EXSTART (master) the other side will have moved on to EXCHANGE state (slave). The master will then tear down the adjacency after a number of failed retransmissions. After this the adjacency will restore and the process starts from the start.
  show ip ospf neighbor: State will show EXSTART for both master and EXCHANGE on slave.

Stuck in EXCHANGE-State

- Reason: Same as EXSTART.

Stuck in LOADING-State

- Reason: MTU mismatch but not between neighbors, but when intermediate device (switch) is configured with wrong MTU settings.
- Behavior: One router will show FULL state, the other will show LOADING state, or both will show LOADING state.
  show ip ospf interface: State will show DR for both routers, and correctly P2P on P2P interfaces.
  show ip ospf neighbor: No state.

Stuck in FULL-State

- Reason: No problem unless no routes are received. This indicates an OSPF database mismatch or some sort of route filtering configured on neighbors.

Filtering

Saturday, January 9, 2016
1:10 PM

Filter-Lists

- Filter-lists only affect Type-3 LSAs and can only be used with prefix-lists.

```
ip prefix-list Lo3 deny 192.168.0.3/32
ip prefix-list Lo3 permit 0.0.0.0/0 le 32

router ospf 1
 area 0 filter-list prefix Lo3 in
```

Filtering Type-3 LSAs using Summarization

- The area range command can also be used to filter all routes that match the criteria.
- This command only works on inter-area routes, not redistributed routes.
- The not-advertise keyword will filter all the more specific routes and the summary, basically blocking the Type-3 LSAs.
- Use the summary-address keyword for external routes.

```
router ospf 1
 area 0 range 172.16.0.0 255.255.0.0 not-advertise
```

Distribute-Lists

- Filtering routes using distribute-lists is only possible in the inbound direction.
- Only one distribute list can be applied per process.
- It is not possible to filter LSAs from being received within the same area.
- Distribute-Lists are a local filter of the LSA database and the RIB.
- Only prevents routes from being installed in the RIB. They do not actually block the LSA from being received.

```
ip prefix-list Lo3 deny 192.168.0.3/32
ip prefix-list Lo3 permit 0.0.0.0/0 le 32

router ospf 1
 distribute-list prefix Lo3 in
```

Alongside gateway:

```
ip prefix-list R2 deny 10.0.12.2/32
ip prefix-list R2 permit 0.0.0.0/0 le 32

ip prefix-list PREFIXES permit 0.0.0.0/0 le 32

router ospf 1
 distribute-list prefix PREFIXES gateway R2 in
```

Distribute-List Filtering from NSSA area

- Filtering from NSSA into normal areas is only needed on the router that performs the Type-7 translation (Highest RID).
- NSSA ABRs translate Type-7 LSAs into Type-5. The requirement is that the routes they translate exist in the routing table.
- Because distribute-lists filter routes from reaching the routing table, the function can be used to filter routes from both the NSSA ABR and all routers that exist in the backbone and other areas.

```
ip prefix-list Lo3 deny 192.168.0.3/32
ip prefix-list Lo3 permit 0.0.0.0/0 le 32

router ospf 1
 area 13 nssa
 distribute-list prefix Lo3 in
```

Outgoing Database Filter

- Filter all outgoing LSAs on the specified interface.
- Filter all outgoing LSAs to the specified neighbor using the neighbor statement.

```
int fa0/0
 ip ospf database-filter all out

router ospf 1
 neighbor 10.0.12.2 database-filter all out
```

Prefix-Suppression

- Only advertises prefixes associated with secondary IP addresses, and passive interfaces.
- This can be configured on a per interface basis or for the entire process.
- Basically, all primary addresses will be suppressed.
- Secondary IP addresses are only advertised by enabling OSPF on the interface, not the network statement.
- If prefix suppression was enabled for the entire process, only secondary addresses and loopbacks would be advertised.
- By specifying the secondaries none keyword, the secondary address is not advertised into OSPF.

```
router ospf 1
 prefix-suppression
 ip ospf 1 area 0 secondaries none

int lo0
 ip ospf prefix-suppression disable
```

LFA / FRR

Saturday, January 9, 2016
1:10 PM

Fast Reroute (FRR) Direct LSA

- IOS only supports per-link LFA.
- The high priority enables FRR for /32 prefixes only, the low priority enables FRR for all prefixes.
- The fast-reroute keep-all-paths option keeps all information in the table, including paths that were not chosen.
- When an area is specified, external routes are not a candidate for FRR. This is because they do not belong to an area.

```
router ospf 1
 fast-reroute per-prefix enable area 0 prefix-priority high
 fast-reroute per-prefix enable prefix-priority high
 fast-reroute keep-all-paths
```

Configure a custom high prefix priority:

```
ip prefix-list FRR permit 0.0.0.0/0 ge 30

route-map FRR permit 10
 match ip address prefix FRR

router ospf 1
 prefix-priority high route-map FRR
```

Exclude interface in calculation:

```
int fa0/0
 ip ospf fast-reroute per-prefix candidate disable
```

FRR Tie Breakers

On by default:

- SRLG 10 - Shared Risk Link Group. Connected to same switch for example. Configure with srlg gid interface command.
- Primary Path 20 - Prefer backup path that is ECMP.
- Interface Disjoint 30 - Prefer backup path that exits through a different (sub)interface.
- Lowest-Metric 40 - Prefer backup path with the lowest metric.
- Linecard-disjoint 50 - Prefer backup path that exits through a different line-card
- Node protecting 60 - Prefer backup path that doesn't lead to the same router.
- Broadcast interface disjoint 70 - Prefer backup path that is in the same broadcast range.
- Load Sharing 256 - If no tie breakers, share backup paths in ECMP.

Off by default

- Downstream - Similar to feasibility condition.
- Secondary-Path - Prefer backup path that is not ECMP.

Manually specify tie breakers and index number (lower is more preferred).

- The required keyword forces matching. If no match, do not go to next-tie breaker and don't use the path.
- When manually configuring tie-breakers, others not included will not be used.

```
router ospf 1
 fast-reroute per-prefix tie-break lowest-metric required index 10
 fast-reroute per-prefix tie-break node-protecting required index 20
 fast-reroute per-prefix tie-break srlg required index 30

show ip ospf fast-reroute prefix
show ip route repair-paths
show ip ospf rib
```

LSAs / Packets

Saturday, January 9, 2016
1:10 PM

OSPF Link-State Advertisement (LSA)

- LSA maximum age is 60min, refresh time is 30min.
- Check summing is performed on all LSAs every 10 minutes.
- The router keeps track of LSAs that it generates and LSAs that it receives from other routers.
- The router refreshes LSAs that it generated and it ages the LSAs that it received from other routers.
- Prior to the LSA group pacing feature, all LSAs would be refreshed and check summed at the same intervals.
- This process wasted CPU resources because only a small portion of the database needed to be refreshed.

OSPF LSA Types

Type 1: Router (All Routers)

- Contents: Router ID, router interfaces & neighbors (not always!). Prefix information is not included.
- Originator: All OSPF routers.
- Triggers an SPF recalculation.
- Flooding scope: Own area, information remains unchanged, not altered by others.

Type 2: Network (DR)

- Contents: Router IDs of all connected routers, netmask of subnets. (This is not included in Type-1 LSA in multi-access segments. Prefix information is not included.
- Originator: DR on the shared (broadcast, non-broadcast) segment. BDR in case of DR failure.
- Triggers an SPF recalculation.
- Flooding scope: Own Area, information remains unchanged, not altered by others.

Type 3: Summary (ABR)

- Contents: Calculated routing information for area routes,
- Based on the information from Type-1 and Type-2 LSAs. The must be in the routing table.
- Summarizes Type-1 and Type-2 LSAs, not the actual route prefixes.
- Summarizes other received Type 3 information.
- Originator: ABR sends this LSA into the area, summarizing Type-1 and Type-2 LSAs from other areas.
- Flooding scope: Own Area, information remains unchanged, not altered by others.

Type 4: ASBR Summary (ABR)

- Contents: Router ID of ASBR from another area, contains the route to the ASBR. The next hop of this route is updated to the RID of each ABR that forwards it.
- Originator: ABR of the area with an ASBR present.
- Flooding scope: Area, information is updated by each ABR.
- Renamed to Inter Area Router Link State in OSPFv3.

Type 5: External (ASBR) / NSSA-ABR

- Contents: Redistributed external routes. Includes the link cost to an external destination (E1 or E2)(E2 is default).
- Originator: ASBR, the next hop of the external route, remains the ASBR as is not updated by ABRs. Only the route towards the ASBR is updated using a Type-4 LSA.
- The NSSA ABR that translates Type-7 external LSAs into Type-5 also functions as an ASBR.
- Flooding scope: Entire domain (Except stub and NSSA areas).

Type 7: NSSA External (ASBR)

- Contents: Redistributed external routes from within an NSSA. These are translated into Type-5 by the NSSA ABR.
- Originator: ASBR in an NSSA area.
- Flooding scope: Own area (NSSA in which it was injected).

IPv6 OSPF LSA Types

Type 8: Link LSA (All Routers)

- Contents: Link-local address and IPv6 prefix for the each link connected to the router.
- Originator: All OSPF routers.
- Flooding scope: Link-local.

Type 9: Inter-Area Prefix LSA (All Routers)

- Contents: IPv6 prefix or link-state changes.
- Originator: All OSPF routers.
- Flooding scope: Own area. The next hop of this prefix is updated to the RID of each ABR that forwards it.

- Basically replaces Type-3 LSA for each individual prefix.
- Does not trigger an SPF recalculation.

Type 11: Grace LSA (All Routers)

- Contents: Informing others that the router is undergoing a graceful restart.
- Originator: Restarting OSPF routers.
- Flooding scope: Link-local.

OSPF Packet Types

Type 1	Hello Message	Discovers and monitors neighbors. Sent periodically to 224.0.0.5 on all interfaces (link-local in scope). Virtual-Links use unicast Hello packets. On broadcast and NBMA networks, Hello packets are used to elect DR and BDR.
Type 2	Database Descriptor (DD/DBD)	Synchronizes the link-state databases for all routers. The routers only exchange the list of all LSAs they posses and update the ones that are missing from the database. No actual LSAs are exchanged.
Type 3	Request (LSR)	Requests for individual neighbors LSA details. After DBD packets exchange process, the router may find it does not have an up-to-date database. The LSR packet is used to request pieces of neighbor database that is more up-to-date or missing.
Type 4	Update (LSU)	Response to LSR with LSA details. Implement the flooding of LSAs. The local router advertises LSA with an LSU packet to its neighboring routers. The local router also advertises the LSU packet with information in response to an LSR.
Type 5	Acknowledgement (LSAck)	Confirmation of the reception of an LSU in response to an LSR.

Misc

Saturday, January 9, 2016
1:10 PM

OSPF TTL Security

- Normally OSPF packets are sent with a TTL of 1 or 2 for directly connected neighbors.
- With TTL Security the TTL for sent packets is set to 255 and received packets must match the configured value.
- Must be configured and on both sides.
- If TTL Security is configured with a hop count of 1, the router will only accept packets with a TTL of 254.

```
int fa0/0
 ip ospf ttl-security hops 1

router ospf 1
 ttl-security all-interfaces hops 1
int fa0/0
 ip ospf ttl-security disable
```

OSPF Ignore MTU

- If two neighbors use different MTU settings on the link, the neighborship will not form.
- Override with the ip ospf mtu-ignore interface command.

```
int fa0/0
 ip ospf mtu-ignore
```

Incremental SPF (iSPF)

- Faster convergence means that the routing protocol is more sensitive to oscillating processes, which in turn makes it less stable.
- iSPF keeps the SPT structure after the first SPF calculation and using it for further computation optimizations.

```
router ospf 1
 ispf
```

OSPFv3

Saturday, January 9, 2016
1:18 PM

OSPFv3

- Multiple instances of OSPFv3 can be run on a link.
- An OSPFv3 process can be configured to be either IPv4 or IPv6.
- Not compatible with OSPFv2.

```
ipv6 unicast-routing
int fa0/0
 ip add 10.0.12.1 255.255.255.0
 ipv6 add fe80::1 link-local
 ospfv3 1 ipv4 area 0
 ospfv3 1 ipv6 area 0
 ospfv3 network broadcast

int lo0
 ip add 192.168.0.1 255.255.255.255
 ipv6 add 1::1/128
 ospfv3 network point-to-point
 ospfv3 1 ipv4 area 0
 ospfv3 1 ipv6 area 0
```

Stub / NSSA

Saturday, January 9, 2016
1:18 PM

OSPF Stub Areas

- Ignores received Type 5 LSA (External) by ABR, does not sent type Type-4 and Type-5 LSA.
- Has no knowledge of external prefixes besides default route that is converted to Type-3.
- Totally stubby area that also stops LSA Type-3 (Summary) except default route.
- Default route is installed automatically into STUB and TS areas.
- Only the ABRs need to be configured with no-summary keyword.

```
router ospf 1
 area 123 stub no-summary
```

OSPF NSSA Areas

- NSSA is a stubby area that allows creation of Type-7 LSA that injects external routes into other areas.
- NSSA Totally stubby (NSSA-TS) is a totally stubby area that allows creation of Type-7 LSA and default route.
- Default route is not installed automatically NSSA areas
- Default route is installed automatically into NSSA-TS areas.
- Only the ABRs need to be configured with the no-summary keyword.
- Specify no-redistribution on the NSSA ABR to have local redistributed routes only be redistributed into normal areas, not the NSSA.
- N1 and N2 type routes will show inside the NSSA area (every router in the area) but they will be converted to regular E1 and E2 when the ABR sends those routes into other areas.
- When using multiple NSSA ABRs, the router with the highest RID is responsible for translating Type-7 LSAs into the normal area.

```
router ospf 1
 area 123 nssa no-summary no-redistribution
 area 123 nssa default-info-originate metric 1 metric-value metric-type 2 nssa-only
 area 123 default-cost 1
```

The default-cost keyword will affect the cost of the redistributed default-route. Not other routes.
NSSA default-information-originate does not require the presence of a default route when configured on the NSSA ABR.

- Originating the default on routers within the NSSA the presence of a default route is required.
- The nssa-only keyword will limit the default route to the NSSA area only by setting the propagate (P) bit in the type-7 LSA to zero.
- The F-bit indicates that a forwarding address is included in the LSA when set. (Used to forward Type 7 LSA)

OSPF NSSA Forward Address

- Always inserted into Type-7 LSA (interface IP-address) loopback has preference.
- Translated by default into Type-5 LSA with forward address intact.
- Based on this forward address, routers outside the NSSA will choose the best path towards the NSSA ASBR.

```
router opsf 1
 area 123 nssa translate type7 suppress-fa
```

The suppress-fa keyword will stop the forwarding address of the Type-7 LSAs from being placed in the Type-5 LSAs.

- This keyword takes effect only on an NSSA ABR or an NSSA ASBR.
- The P-bit is used in order to tell the NSSA ABR whether to translate type 7 into type 5. P=1 means translate.
- When using multiple NSSA ABRs, if suppression is enabled on the translating NSSA ABR (highest RID) ECMP will stop functioning.
- In this case the translating router (highest RID) will become the next-hop, because the forward address is not known (0.0.0.0).

Summarization

Saturday, January 9, 2016
1:19 PM

OSPF Summarization

- Routes can only be summarized manually at the area boundaries (ABR) or at the ASBR.
- The area range command summarizes inter-area routes.
- The area that is specified is the area where the routes are located, not the area that is being summarized into.
- The discard route is installed pointing to Null0 for internal (inter-area, 110) and for external (redistributed, 254) routes.
- Disable the discard route with the no discard-route command.

```
router ospf 1
 area 1 range 172.16.0.0 255.255.0.0
```

```
router ospf 1
 no discard-route internal
 no discard-route external
```

The summary-address command only summarizes external routes.

- This can only be performed on the ASBR (usually the router that redistributes the routes).
- In the case of NSSA, the ABR to the NSSA will also be an ASBR and will be able to further summarize external routes.
- The nssa-only keyword will keep the summary (and the more specific routes) inside the NSSA.
- Configure the nssa-only on the redistributing ASBR, not the ABR that connects tot the NSSA area.

```
router ospf 1
 summary-address 172.16.0.0 255.255.0.0 nssa-only
```

VL / GRE

Saturday, January 9, 2016
1:19 PM

OSPF Virtual-Links

- Only work over normal (transit) areas and do not operate over stub areas and NSSA areas.
- Always in area 0, even if they are configured on other areas. Use area 0 authentication by default.
- Endpoints are the RIDs, not actual ip-addresses.
- Can only cross one area.
- P2P in nature and unnumbered, and they carry only OSPF communication such as hellos and LSAs.
- Existing VLs can easily be spotted by the DNA bit set in the OSPF database, VLs also set the V-bit to 1.
- Default hello-timer is 10 seconds and dead timer is 40 seconds.
- The ttl-security hops keyword specifies over how many hops the Virtual-Link is allowed to travel.

```
router ospf 1
 area 23 virtual-link 192.168.0.3 ttl-security hops 2
```

OSPF GRE

- Beware of recursive routing when choosing tunnel endpoints.
- Tunnel between ABRs, not the routers that should be linked in the same area.

```
int fa0/0
 ip ospf 1 area 1234
int fa0/1
 ip ospf 1 area 0
int tun0
 ip unnumbered fa0/1
 tunnel source fa0/1
 tunnel destination 10.0.23.2
 tunnel mode gre ip
 ip ospf 1 area 1234
```

Quality of Service

Wednesday, December 9, 2015
8:44 AM

VoIP

Different applications require different treatment, the most important parameters are:

- Delay: The time it takes from the sending endpoint to reach the receiving endpoint.
- Jitter: The variation in end to end delay between sequential packets.
- Packet loss: The number of packets sent compared to the number of received as a percentage.

One-way requirements for voice:

- Latency ≤ 150 ms (Delay)
- Jitter ≤ 30 ms
- Loss ≤ 1%
- Bandwidth (30-128Kbps)

Hardware Queue

- Hardware transmit (Tx) queue is FIFO by default for ethernet interfaces. Tx queue is 256.
- Hardware queue is WFQ by default for serial interfaces. Tx queue is 64.

```
int fa0/0
 tx-ring-limit 256

show controllers fa0/0 | i tx
```

Class of Service (CoS)

- 802.1p is L2 information.
- ToS / DSCP is L3 information, it will remain constant between endpoints.

802.1p

111	7	Reserved
110	6	Reserved
101	5	Voice
100	4	Video
011	3	Voice Signal
010	2	High Data
001	1	Low Data
000	0	Best Effort

L3 Type of Service (ToS)

- ToS / DSCP uses left 3 bits with IPP, uses left 6 bits with DSCP and 8 bits with ToS.
- Drop Probability (DP) right-most bit is always 0.
- Flow Control (FC) is always 0 unless ECN is used.
- IP Precedence (IPP) part of DSCP is called the Per-Hop Behavior (PHB).

Class	Binary	DSCP	ToS
cs1	001|000|00	8	32
af11	001|010|00	10	40

af12	001|100|00	12	48
af13	001|110|00	14	56
cs2	010|000|00	16	64
af21	010|010|00	18	72
af22	010|100|00	20	80
af23	010|110|00	22	88
cs3	011|000|00	24	96
af31	011|010|00	26	104
af32	011|100|00	28	112
af33	011|110|00	30	120
cs4	100|000|00	32	128
af41	100|010|00	34	136
af42	100|100|00	36	144
af43	100|110|00	38	152
cs5	101|000|00	40	160
ef	101|110|00	46	184
cs6	110|000|00	48	192
cs7	111|000|00	56	224

Maps

Wednesday, December 9, 2015
8:44 AM

Class-Maps

- With match-all, all criteria must be met in order to have a match. Default.
- With match-any, only one of the criteria has to be met in order to have a match.
- The ip precedence and ip dscp keyword only match on IPv4 traffic.

Match telnet traffic that is not marked with the default value of cs6:

```
class-map match-all TELNET
 match protocol telnet
 match not dscp cs6

policy-map TELNET
 class TELNET
  drop

int fa0/0
 service-policy input TELNET
```

Hierarchical Policy-Map

- Police traffic matching the ACL to 64000 bps.
- Police traffic matching a subset of this ACL with IPP0, IPP1, IPP2 to 32000 bps.
- Police traffic matching a subset of this ACL with IPP2 to 16000 bps.
- Always nest the most specific subset into the upper level policy-map.
- In this case IPP2 with 8000 bps is the most specific, so this will be the lowest level policy-map.

```
ip access-list standard QOS_TRAFFIC
 permit 10.10.10.0 0.0.0.255

class-map match-all LEVEL_1_CM
 match access-group name QOS_TRAFFIC
class-map match-all LEVEL_2_CM
 match precedence 0 1 2
 match access-group name QOS_TRAFFIC
class-map match-all LEVEL_3_CM
 match precedence 2
 match access-group name QOS_TRAFFIC

policy-map LEVEL_3_PM
 class LEVEL_3_CM
 police 16000
policy-map LEVEL_2_PM
 class LEVEL_2_CM
 police 32000
 service-policy LEVEL_3_PM
policy-map LEVEL_1_PM
 class LEVEL_1_CM
 police 64000
 service-policy LEVEL_2_PM

int fa0/0
 service-policy output LEVEL_1_PM
```

NBAR

Tuesday, December 8, 2015
9:09 PM

NBAR

- NBAR uses deep packet inspection instead of just matching on the specified port.
- This is more CPU-intensive than matching with an ACL.
- Managing with an ACL should be used if a previous devices has already performed deep-packet inspection with NBAR.
- List all known ports with show ip nbar port-map.
- Map a well-known port of a protocol to a new port with ip nbar port-map http 80 8080.

NBAR Protocol-Discovery

- Monitor traffic protocols known to NBAR on a specific interface. CPU intensive.

```
int fa0/0
 ip nbar protocol-discovery

show ip nbar protocol-discovery
```

Policing / Shaping

Wednesday, December 9, 2015
8:44 AM

Policy-Maps

- Policy-maps can be applied in both the ingress and egress direction.
- CBWFQ and LLQ policy-maps can only be applied in the egress direction.
- Shaping should be applied in the egress direction. But can be applied ingress.
- Policing should be applied in the ingress direction. But can be applied egress.

Terminology

- Access rate (AR). This is the actual speed of the physical port.
- Committed Information Rate (CIR). Average rate the shaper is targeting in bps.
- Time Committed (Tc). Time interval in ms to emit traffic bursts.
- Burst Committed (Bc). Amount of bits that should be sent every Tc.
- Burst Excessive (Be). Amount of bits exceeding Bc that could be sent during Tc. Accumulated by idle periods.

Policing

- Can be used to drop incoming packets that do not conform to the policy.
- Cisco routers have a Tc default value of 125 ms = 8 times a second.
- If the goal is 50Mbit/sec transmit speed on a 100Mbit/sec interface. The CIR would be specified as 50Mbit.
- The IOS will automatically calculate the Bc based on the configured CIR (recommended).
- In order to calculate the Bc the CIR needs to be divided by 8 or 4 (depending on the platform).
- Conform. Traffic is under Bc.
- Exceed. Burst size exceeds Bc, but under Bc+Be.
- Violate. Burst size exceeds Bc+Be.

Tc = CIR/Bc
Bc = CIR/Tc

CIR = 50 Mbit/sec = 50000000 bits/sec
Bc = 50000000 / 8 = 6250000 bits / 8 = 781250 bytes

```
policy-map POLICER
class class-default
 police cir 50000000 bc 781250
  conform-action transmit
  exceed-action drop

int fa0/0
 service-policy input POLICER
```

Shaping

- Shaping is applied by altering the time in which traffic is allowed to send, not the speed of the port.
- This means that the average traffic sent will be less over time.
- The IOS will automatically calculate the Bc and Be based on the configured CIR (recommended).
- With a shape average (CIR) of 50 Mbit, the calculated Bc and Be will be 200000.
- The Tc will be 250, meaning that every 4ms (1 second / 250) the interface will forward at the line speed.
- The shaper can also be based on a percentage of the link speed, however this is dependent on the manually configured bandwidth, not the hardcoded line speed.

Tc = CIR/Bc
Bc = CIR/Tc

CIR = 50 Mbit/sec = 50000000 bits/sec
Bc = 50000000 / 200000 = 250

```
policy-map SHAPER
 class class-default
  shape average 50000000 200000 200000

int fa0/0
 service-policy output SHAPER
```

Pre-Classify

Tuesday, December 8, 2015
9:09 PM

QoS Pre-Classification

- By default tunneling and VPN operations are applied before the QoS policy. QoS pre-classify (PQ) reverses the order.
- PQ on crypto map affects all tunnels on that physical interface.
- PQ on tunnel affects only that specific tunnel interface.
- PQ is needed when classification is based on IP address, ports, etc. Or a crypto map is used.
- PQ is not needed when classification is based on ToS. Or a tunnel interface is used.
- Can be enabled regardless, very little impact on performance.

```
interface tun0
 qos pre-classify

crypto map CMAP 10 ipsec-isakmp
  qos pre-classify
```

Rate-Limiting

Rate-Limiting

- Requires CEF. Can be configured on physical or sub-interface.
- Works similar to policer without MQC configuration.
- Like a policer, the CIR is configured in bits. The Bc and Be are configured in bytes.
- Match on all traffic or specific DHCP, QoS group, ACL (no named ACL support).

```
int fa0/0
 rate-limit input 50000000 781250 781250 conform-action transmit exceed-action drop

show interfaces rate-limit
```

WRED

Wednesday, December 9, 2015
8:45 AM

Weighted Random Early Detection (WRED)

- Only works for TCP traffic. Enable in the egress direction.
- Drop preference values (AF) are used by WRED.
- WRED turns any queue on the interface into FIFO. The minimum threshold is the FIFO's queue depth before WRED is activated.
- The overall size of the queue depth is specified by the hold-queue command.
- The hold-queue must be higher than the minimum threshold configured in WRED.

```
int fa0/0
 hold-queue 40 out

show interface | i queue
```

The default WRED values differ per precedence and DSCP values.

- Minimum-threshold. Above this threshold WRED engages and starts randomly dropping packets.
- Maximum-threshold. Above this threshold TAIL-DROP engages and starts dropping packets (WRED is basically disabled).
- Mark-probability. Amount of packets dropped up until maximum threshold is reached (1 out of 10, 1 out of 5, etc).
- The Exponential Weighting Constant (EWC) alters how quick WRED reacts.
- Higher EWC value makes WRED react more slowly, lower EWC value makes WRED react more quickly (default is 9).

Change IPP 0 to 15 min, 30 max and 1/5:

```
policy-map WRED_IPP
 class class-default
  random-detect precedence-based
  random-detect precedence 0 15 30 5
 random-detect exponential-weighting-constant 9

int fa0/0
 service-policy WRED_IPP out
```

Change DSCP 46 (EF) to 35 min, 40 max, and 1/2:

```
policy-map WRED_DSCP
 class class-default
  random-detect dscp-based
  random-detect dscp 46 35 40 2
  random-detect exponential-weighting-constant 9

int fa0/0
 service-policy WRED_DSCP out
```

AF Values

- AF11 is IP Precedence 1 (cs1) with low drop preference.
- AF23 is IP Precedence 2 (cs2) with high drop preference.
- AF32 is IP Precedence 3 (cs3) with medium drop preference.

Drop Chance	Class #1	Class #2	Class #3	Class #4
Low	(AF11) 001010	(AF21) 010010	(AF31) 011010	(AF41) 100010

Medium	(AF12) 001100	(AF22) 010100	(AF32) 011100	(AF42) 100100
High	(AF13) 001110	(AF23) 010110	(AF33) 011110	(AF43) 100110

Explicit Congestion Notification (ECN)

- Mode of WRED that can be enabled to suggests a traffic flow to slow down, instead of actually dropping packets.
- Uses the last two bits of ToS (Flow Control value). Instead of dropped a packet WRED sets both these bits (ECT and CE) to 1.
- When the destination receives a TCP packet with both ECT and CE set to 1, it sets the ECE (Explicit Congestion Experienced) flag on its next TCP packet back to the sender. When the sender receives the packet, it is instructed to slow down.

```
policy-map WRED_ECN
 class class-default
  random-detect ecn

int fa0/0
 service-policy WRED_ECN out
```

ECN Values

- The 7th bit is the ECT bit and the 8th bit is the CE bit.
- If a TCP host supports ECN, it sets either (but not both) of the low-order bits in the DSCP byte - ECT or CE - to 1.
- If a TCP host doesn't support ECN, these will both be set to 0.

00	Not ECN capable
01	Endpoints are ECN capable
10	Endpoints are ECN capable
11	Congestion experienced

RIP

Wednesday, December 9, 2015
8:45 AM

RIP Passive Interfaces

- Passive interfaces transmit no protocol-related data, but still receive data and networks. This includes RIP updates.
- RIP passive interfaces still receive routing updates because RIP does not use a hello mechanism.
- Neighbor relations can still be setup over passive-interfaces by using the neighbor command.

```
router rip
 passive-interface default
 neighbor 10.0.12.1
```

RIP Redistribution

- Routing protocols or static routes redistributed into RIP require a set metric.
- With a transparent the redistributed route will inherit the metric that is seen in the routing table of the redistributing router.
- Use the transparent keyword only when you know that the metric is lower than 16.
- Redistributed OSPF E2 routes (cost 20) will always lead to an infinite metric when using the transparent keyword.

```
router rip
 redistribute ospf 1
 default-metric 5
```

RIP Summarization

- RIP does not install a discard route by default. Instead it relies on route poisoning when a routers own summary is received.
- Manually add a discard route to null0 with ip route 0.0.0.0 0.0.0.0 null0.
- RIP will generate a default route with the default-route originate command whether it exists or not in the routing table.

Originate a default-route out of a specific interface (may lead to routing loops):

```
route-map RIP permit 10
 set interface se0/0

router rip
 default-information originate route-map RIP
```

RIP Filtering

- Filter routes using either the distribute-list or the offset-list command.

Filter all routes from se1/0 (R2):

```
ip prefix-list PREFIX permit 0.0.0.0/0 le 32
ip prefix-list NOT_R2 deny 10.0.12.2/32
ip prefix-list NOT_R2 permit 0.0.0.0/0 le 32

router rip
 distribute-list prefix PREFIX gateway NOT_R2 in se1/0
 offset-list 0 in 16 se1/0
```

Validate Update Source

- Ensures that the source IP address of incoming routing updates is on the same IP network. Enabled by default.
- Disabling split horizon on the incoming interface will also cause the system to perform this validation check.
- Disable when using PPP IPCP addressing.
- For unnumbered IP interfaces no checking is performed.

```
router rip
```

validate-update-source

RIP Split Horizon

- Does not advertise the same networks out of interfaces from which they are learned. On by default (except FR and ATM).
- Poison Reverse is a stronger variant of this. The routes are advertised out of the interfaces, but with an unreachable metric.
- Poisoned Reverse overrides Split Horizon, however it is not implemented in Cisco RIPv2. Not enabled by default in RIPng.

RIP Authentication

- The default authentication mode for RIP is plain-text, even if not specifically configured.
- RIPv1 does not support authentication. RIPng does not support authentication or encryption by IPsec.

RIPng

Wednesday, December 9, 2015
9:18 AM

RIPng

- Passive interfaces are not supported.
- Static (manual) neighbors cannot be configured (no neighbor command).
- Split Horizon and Poison Reverse can be activated only on a per-process basis, not on individual interfaces.
- RIPng Updates are send to FF02::9 using UDP port 521 by default.

```
ipv6 router rip RIPng
 timers 30 180 0 120
 maximum-paths 16
 distance 120
 split-horizon
 poison-reverse
 port 521 multicast-group FF02::9

show ipv6 rip
show ipv6 rip database
show ipv6 rip next-hops
```

RIPng Summarization

- The default-information originate command will originate a default route in addition to all other specific routes.
- The default-information only command will originate a default route and suppress all other specific routes.
- Like RIPv2, the default-route does not have to exist in order to be advertised.
- When the default-route is advertised, the router will ignore reception of all other default routes received on any interface.
- The default metric of the default-route is 1.
- The summary address copies the metric of the more specific routes it is summarizing.

```
int fa0/0
 ipv6 rip RIPng enable
 ipv6 rip RIPng default-information originate metric 1
 ipv6 rip RIPng summary-address 2001::/16
```

RIPng Metric

- The metric is incremented by the receiver instead of the sender.
- Offset lists are configured only on interfaces and are for all received subnets, use the metric-offset command.
- The metric-offset specified replaces the original increment of the metric (1). Meaning that a metric-offset with value 4 will increase the metric by adding 3 to the existing value (4-1).

```
int fa0/0
 ipv6 rip RIPng enable
 ipv6 rip RIPng metric-offset 4
```

Updates

Wednesday, December 9, 2015
9:11 AM

RIP Updates

RIP sends routing update based on two conditions:

- Regular update is sent at a 30 second interval by default. (full update)
- Request message (flash update) is sent immediately when there is a change to the topology. (full update)
- Full update is sent when RIP is started, a RIP interface comes up, or when routes are cleared from the RIB.
- Full updates contains learned and connected RIP routes.
- RIP Route Poisoning advertises a truly unreachable route to quickly flush it from routing tables.
- RIPv2 updates are sent to 224.0.0.9 using UDP port 520 by default.
- RIPv1 updates are sent to 255.255.255.255 using UDP port 520 by default.
- Send v2 updates to the broadcast address using the ip rip v2-broadcast interface command.

By default IOS routers will send v1 and receive v2. Set both to v2 with the version 2 command in the router sub-configuration.

- Alternatively specify the send and receive version on specific interfaces. Interface configuration takes preference over router.

```
int fa0/0
 ip rip send version 1 2
 ip rip receive version 1 2
 ip rip v2-broadcast
```

RIP Flash-Update-Threshold

If there is a change in the network topology at 27 seconds, RIP will send its full routing table in a Flash Update, and then again a Regular Full Update at the 30 second interval. So in 3 seconds 2 of the exact same messages are send flooding the network.

- This can be an unwanted situation on slow links, the flash-update-threshold can be configured to stop this behavior.

For example if the threshold is set to 10 seconds, a change occurring at 19 seconds (after the last Regular Update) will send a Flash Update, and then again at 30 seconds a Regular Full Update. A change occurring at 20 seconds (after the last Regular Update) is within the 10 second threshold and will be suppressed.

RIP Timers

invalid	180 seconds by default. Reset when update is received. Declares route invalid after timer expires.
hold-down	180 seconds by default. Starts after invalid after timer has expired. Marks route as unreachable and does not accept any updates about the route.
flush	240 seconds by default. Reset when update is received. Removes route from routing table after timer expires.
sleep	Disabled by default. The interval in milliseconds for postponing routing updates in the event of a Flash update.

Because the flushed after timer expires after 240 seconds, the effective hold-down period is only 60 seconds.

Triggered Updates

When triggered extensions to RIP are enabled, updates are sent on the WAN only if one of the following events occurs:

- The router receives a specific request for a routing update. (Full database is sent.)
- Information from another interface modifies the routing database. (Only latest changes are sent.)
- The interface comes up or goes down. (Partial database is sent.)
- The router is first powered on, to ensure that at least one update is sent. (Full database is sent.)

```
int se1/0
 ip rip triggered
```

Spanning Tree

Monday, November 30, 2015
8:22 PM

PVST+ / Rapid PVST+ (802.1W)

- The + means that the STP instance is backwards compatible with IEEE standard STP.
- PVST creates a separate logical STP instance for each VLAN. This appears as one instance to other non-cisco switches.
- RSTP is not faster because of increased timers, but because of synchronization.
- The synchronization process only works on ports that are P2P and full-duplex.
- If a port negotiates half-duplex with its connected neighbor, the switch assumes that the neighbor is a hub (shared link).
- If a port negotiates full-duplex operation, the switch will assume that the neighbor is a switch (P2P link).

RSTP port roles:

- Root Port (maintains its usual meaning)
- Designated Port. Continues to process received and send BPDUs. And receive DTP, VTP, CDP, etc.
- Alternate Port (a prospective replacement for the Root Port). Basically blocked 802.1D port + uplinkfast.
- Backup Port (a prospective replacement for the Designated Port on a shared segment)

STP Cost

4 Mbit/s	250	5000000
10 Mbit/s	100	2000000
16 Mbit/s	62	1250000
100 Mbit/s	19	200000
1 Gbit/s	4	20000
2 Gbit/s	3	10000
10 Gbit/s	2	2000

STP Timers

MessageAge	Estimation of BPDU age since it was originated by root bridge (0 at root). Every bridge increments this by 1 before forwarding the BPDU.
MaxAge-MessageAge	Remaining lifetime of a BPDU after being received by a bridge. If port stores BPDU it must be received again within the MaxAge-MessageAge interval.
Hello	Root switch creates and sends a hello every 2 seconds (default). Contains the RB-ID and SB-ID set to the ID of the root, RPC set to 0 and SPID set to egress port. Other switches add root path cost to RPC (increment). Hellos are always received on root port, and forwarded out designated ports with RPC, SBID, SPID, and MessageAge fields updated.
Max-Age	20 seconds by default, set by root. Maximum time a BPDU is stored. The time a bridge waits after it stops receiving BPDUs (from the root) on the specified port.
Forward-Delay	15 seconds by default, set by root.

STP BPDUs

- Only root bridge sends BPDUs in a converged network. Other bridges forward BPDUs.
- Configuration BPDUs are sent out only from Designated Ports.
- Designated Ports store the BPDU they send.

- Root and Blocking ports store the best BPDU they receive.
- Received superior stored BPDUs will expire in MaxAge-MessageAge seconds if not received within this time period.

- Access ports send only BPDUs relevant to their access VLAN.
- Trunk ports always send a set of BPDUs E-formatted BPDUs for VLAN1, always untagged.
- PVST+ BPDUs tags for VLAN1, but not for native VLAN.
- Original 802.1D (STP) frames are sent in the native VLAN. This is important when questions are asked to 'Send 802.1D frames in VLAN 50'. This means that the native VLAN should be set to 50.

Root Bridge ID (RBID)	Root switch creates and sends a hello every 2 seconds (default). Contains the RB-ID and SB-ID set to the ID of the root, RPC set to 0 and SPID set to egress port.
Root Path Cost (RPC)	Path cost to Root Bridge. Set to 0 on root. Other switches add root path cost to RPC (increment), this is a value of 19 (200000) on FastEthernet and 4 (20000) on GigabitEthernet.
Sender Bridge ID (SBID)	Bridge that originated the BPDU. Set by each bridge to own ID. Hellos are always received on root port, and forwarded out designated ports with RPC, SBID, SPID, and MessageAge fields updated.
Sender Port ID (SPID)	Bridge port that originated the BPDU. Set by each bridge to own ID.
Receiver Port ID (RPID)	Not included in the BPDU message, evaluated locally. Port where BPDU was received, not forwarded.

<u>STP Root Election (Bridge-ID)</u>
- Bridge Priority.
- Mac-Address.
    - The actual mac-address is the lowest mac-address (Base ethernet Mac-Address) in use on the system.
    - This address is one address below the lowest interface address (usually GigabitEthernet 0/1).

<u>STP Path Selection</u>
- Lowest root bridge ID
- Lowest path cost to root bridge
- Lowest bridge ID
- Lowest port ID (on the root bridge)
    - Configure priority settings on the root bridge. Lower is better, in increments of 16. (128 is default)
    - Configure cost settings on the non-root switches. Lower is better.
    - Cost is used over priority settings.

MSTP

Tuesday, December 1, 2015
1:15 PM

Multiple Spanning Tree Protocol (802.1S MSTP)

- MSTP uses the long path cost notation by default.
- Common Spanning Tree (CST) is the entire spanning-tree domain including other versions of STP.
- Inside the MST region the IST (Internal STP) is present, this is MST0. The IST is the only STP instance that sends and receives BPDUs.
- Hello, ForwardTime and MaxAge timers can only be tuned for the IST. All other MIST inherit the timers from the IST.
- Configure the maximum number of switches between any two end stations with the diameter command (MST0 only).
- The hello-time is the timer at which configuration messages are sent by the root switch.
- The MST extended system-id is made up of the instance number instead of the VLAN number.

```
spanning-tree mst 0 root primary diameter 7 hello-time 2
```

MST Configuration with VTPv3:

```
vtp version 3
vtp mode server mst
do vtp primary mst force

spanning-tree mst configuration
 name cisco
 revision 1
 instance 1 vlan 100-200
 show pending

show spanning-tree mst configuration
```

MST Caveats

If one VLAN is active on the link in the instance, the entire instance is active.

- This can be a problem if other VLANs are pruned (administratively disallowed) that are part of the instance.

Make sure that all VLANs that are part of the instance are allowed on the trunk.

- If not, either add the VLAN to the allowed list (if not restricted).
- Or remove other VLANs from the allowed list (prune them) so the MST instance switches to another link.
- Or put the disallowed VLAN in another instance by itself and make it go over a separate link.

MST Interoperability

The idea is to hide the internal MST from other versions of STP. But maintain backwards compatibility and fast convergence.

- Other regions will view the MST region as a single switch.
- Every MST region runs a special instance of spanning-tree known as IST or Internal Spanning Tree (MST0).
- IST has a root bridge, elected based on the lowest Bridge ID.
- With multiple MST regions in the network, a switch that receives BPDUs from another region is a boundary switch.
- Another region can be RSTP or PVST+, the ports to these regions are marked as MST boundary ports.
- When multiple regions connect together, every region constructs its own IST and all regions build a common CIST.
- A CIST Root is elected among all regions and CIST Regional Root (IST root) is elected in every region.

CIST Root	The bridge that has the lowest Bridge ID among ALL regions. This could be a bridge inside a region or a boundary switch in a region.
CIST Regional Root	A boundary switch elected for every region based on the shortest external path cost to reach the CIST Root. Path cost is calculated based on costs of the links connecting the regions, excluding the internal regional paths. CIST Regional Root becomes the root of the IST for the given region as well.

PortFast

Monday, November 30, 2015
8:23 PM

PortFast
Interface is moved directly to forwarding state, bypassing learning without forward-time delay (15sec).

- This only happens when the port transitions from a disabled state, not from a blocking / alternate state.
- PortFast ports will still transmit BPDUs, but will lose their PortFast state if a BPDU is received.
- When configured globally PortFast will be enabled on all access ports.
- Just because PortFast is configured on the port or in the global configuration, doesn't mean that it is operational.
- Add the trunk keyword to enable PortFast on trunk ports.

```
spanning-tree portfast default

int fa0/0
 spanning-tree portfast
```

BPDUGuard

- BPDUGuard supersedes PortFast and can be configured per port or globally. Errdisables port if BPDU is received.
- Global BPDUGuard is part of PortFast and will only be active if PortFast is also enabled (meaning it is operational).
- When configured on the port, BPDUGuard is enabled unconditionally and does not need PortFast to function.

```
spanning-tree portfast bpduguard default

int fa0/0
 spanning-tree bpduguard enable
```

BPDUFilter

- Filters BPDUs on the port in egress direction (global) or both directions (per interface).
- Global BPDUFilter stops sending outgoing BPDUs on interfaces that have an operational PortFast status.
- Global BPDUFilter will still sent 11 BPDUs when the interface comes online, this is to prevent misconfigurations.
- PortFast enabled ports that receive BPDUs will lose their PortFast status, and thus global BPDUFilter will also be disabled.

BPDUFilter configured on the port itself will filter all incoming and outgoing BPDUs.

- This is the equivalent of turning off STP on the port.
- PortFast does not have to be enabled on the port for BPDUFilter to be active.

```
spanning-tree portfast bpdufilter default

int fa0/0
 spanning-tree bpdufilter enable
```

Interoperability PortFast, BPDUFilter and BPDUGuard
Global BPDUFilter + BPDUGuard:

- Guard is triggered first.
- PortFast triggered second.
- Filter is triggered third.

This is a valid configuration that err-disables ports that receive BPDUs, and filters outgoing BPDUs.

Per port BPDUFilter + BPDUGuard:

- Filter is applied first.
- Guard is triggered second.
- PortFast is applied third.

BPDUFilter configured on the port itself will filter BPDUs in both directions, and will supersede PortFast and BPDUGuard on the port.

This is not a valid configuration because incoming and outgoing BPDUs will be filtered by BPDUFilter. The guard never receives BPDUs and will never be triggered.

BBFast / ULFast

Monday, November 30, 2015
8:21 PM

UplinkFast (802.1D STP)

- Deals with direct failures, basically RSTP alternate port on IEEE STP.
- Blocked ports immediately transition to the forwarding state by skipping listening and learning state.
- When you enable UplinkFast the priority for all VLANs is set to 49152.
- The spanning tree port cost is increased by 3000 (short) or 10^7 (long).
- UplinkFast is enabled / disabled for all VLANs. Only enabled if switch has a blocked port.
- Use on access layer switches.

spanning-tree uplinkfast

BackboneFast (802.1D STP)

- Deals with indirect failures.
- Max age is the timer a bridge waits after it stops receiving BPDUs (from the root) on the specified port.
- Max age unblocks links after 20 seconds, or when 10 BPDUs are missed.
- Under normal condition this failure process takes 20 seconds + listening + learning = 50 seconds.

spanning-tree backbonefast

BackboneFast removes the max-age timer if inferior BPDUs are received on a (separate) blocked port.

- The switch will generate a RLQ (Root Link Query) to the root bridge to verify that the other switch has lost connectivity to the root, and thus declares itself root. The root will respond to the RLQ and the switch will remove max-age timer from the blocked port.
- Must be enabled on all switches in the network. Not supported on Token Ring networks.

Loop/RootGuard

Monday, November 30, 2015
8:24 PM

Loopguard

Prevents alternate or root ports from becoming designated in response to a unidirectional link.

- Puts ports into a loop-inconsistent state if BPDUs are no longer received on the interface.
- Prevents loops by detecting the sudden loss of BPDUs.
- Only superior BPDUs are considered.
- Incompatible with RootGuard.
- Loopguard-enabled ports may send BPDUs when inferior BPDUs are received.
- UDLD is a better option for etherchannels, because it can block individual ports.

Protects from:

- Unidirectional links.
- Switches/links that block BPDUs.

```
int fa0/0
 spanning-tree guard loop

spanning-tree loopguard default

show spanning-tree inconsistentports
```

Etherchannel Guard

- Detect etherchannel misconfiguration between the switch and a connected device.
- Places interfaces in errdisable state in the event of a misconfiguration.

```
spanning-tree etherchannel guard misconfig
```

Rootguard

Ignores superior BPDUs on specified port. Apply on interfaces that connect to switches that should never become the root.

- Allows interface to participate in STP, but ignores superior BPDUs, state is cleared when superior BPDUs stop.
- In MST the port does not become root-inconsistent, but is forced to become a designated port.
- Not VLAN-aware, enabled for all VLANs present on the port. Cannot be enabled globally.
- Do not enable the root guard on interfaces to be used by the UplinkFast feature.
    - Backup (blocking) ports will go to root-inconsistent state instead of forwarding state.

```
int fa0/0
 spanning tree guard root
```

UDLD

Monday, November 30, 2015
8:24 PM

Unidirectional Link Detection (UDLD)

- Exchanges protocol packets that contains the device + port ID and neighbor device + port ID.
- If device does not see its own ID echoed back it considers the link unidirectional.
- Message interval is 15 sec by default on FastEthernet, 7 seconds on GigabitEthernet.
- Detection interval (Expiration time) is 3x message interval = 45 sec.
- Must be enabled on both sides. When the feature is enabled globally, it only effects fiber ports.
- Use aggressive mode on Copper links.
- Enabling UDLD globally will only activate it on fiber ports
- Aggressive mode detects ports that are stuck in up state (neither transmit/receive but are up). And ports that are up on only one side (fiber).
- UDLD is a better option for Etherchannels, because it can block individual ports.

```
udld enable
udld aggresive
```

Normal mode	Port is marked undetermined, behaves according to STP state.
Aggressive mode	Tries to re-establish port state, if not successful port is put in errdisable state.

```
int fa0/0
 udld port
 udld port aggressive

show udld neighbors
```

Switching

Dynamic Trunking Protocol (DTP)

- Point-to-Point (P2P) protocol.
- ISL is the preferred encapsulation when using DTP.
- DTP carries VTP domain information.
- If VTP domain does not matches, DTP will not negotiate a trunk.
- In this case DTP desirable on both sides will still lead to an operational mode of static access.
- Desirable has priority over auto. If one side is set to auto, and one side is desirable the port will become a trunk.
- ISL encapsulates the entire frame.
- Dot1Q adds a 4-byte header (tag) to the frame, and has a native VLAN concept.
- Dynamic desirable prefers trunk, dynamic auto prefers access.

Turn off DTP with:

```
switchport nonegotiate
switchport mode access
no switchport
switchport mode private-vlan
switchport mode dot1q-tunnel
```

802.1X

Monday, November 30, 2015
4:36 PM

802.1X Authentication Using EAP

- User authentication requires the user to supply a username and password, verified by a RADIUS server.
- Supplicant identifies using EAP over LAN (EAPoL).
- Radius server only supports radius type messages, therefore the switch translates between EAPoL and RADIUS.

802.1X Terminology

- Supplicant. Supplies a username/password prompt to the user and sends/receives the EAPoL messages. (client)
- Authenticator. Translates between EAPoL and RADIUS messages in both directions, and enables/disables ports based on the success/failure of authentication. (switch)
- Authentication Server. Stores usernames/passwords and verifies that the correct values were submitted before authenticating the user. (radius server)

```
dot1x system-auth-control
aaa authentication dot1x default group radius

int fa0/0
 dot1x portcontrol auto | force-authorized | force-unauthorized
```

Auto	Using 802.1X. This is the default.
Force-Authorized	Not using 802.1X, but the interface is automatically authorized.
Force-Unauthorized	Not using 802.1X, but the interface is automatically unauthorized.

Bridge Groups

Saturday, December 5, 2015
7:13 PM

Bridge Groups (Fallback Bridging)

- Allows non-IP traffic to be communicated between hosts that reside in different VLANs or are connected on routed ports.
- SVIs allow hosts in different VLANs to communicate using IP addressing. Bridge groups extends this functionality to L2 as well.
- Can only be applied to SVI interfaces or L3 routed ports (no switchport).
- An actual IP address is not required to be configured on these SVIs or routed ports for fallback bridging to work.

```
bridge 1 protocol vlan-bridge
int fa0/0
 no switchport
 bridge group 1

int vlan 10
 bridge-group 1

show bridge 1 group
```

DHCP Snooping

Saturday, December 5, 2015
2:02 PM

DHCP Snooping

- The gateway address (giaddr) is automatically set to 0.0.0.0 by the snooping switch.
- This all-zeroes address needs to be trusted by the DHCP server.
- Alternative disable the information option 82 on the switch. This will stop the giaddr being set to 0.0.0.0.
- DHCP option 82 (relay information option) identifies hosts by both the MAC-Address and the switchport. Disabled by default on routers, enabled by default on switches. It allows DHCP relays to inform the DHCP server where the original request came from.

```
ip dhcp snooping
ip dhcp snooping vlan 10

int fa0/0
 description DHCP_SERVER
 ip dhcp snooping trust

int range fa0/1 - 10
 description CLIENTS
 ip dhcp snooping limit rate 10
```

DHCP server configuration:

```
int fa0/0
 description DHCP_SERVER
 ip dhcp relay information trusted
```

Or:

```
ip dhcp relay information trust-all
```

DHCP Snooping + Relay

When using a relay router / switch the port to that device also needs to be trusted:

```
int fa0/0
 description DHCP_RELAY
 ip dhcp relay information trusted
```

When using multiple DHCP Snooping devices untrusted information must be allowed on the switch closest to the server.

- Alternatively Option82 can be disabled on the switch not connected to the server.

```
ip dhcp snooping information option allow-untrusted
no ip dhcp information option
```

DAI

Saturday, December 5, 2015
4:00 PM

Dynamic Arp Inspection (DAI)

- Validates ARP packets in a network, looks at the IP-to-MAC-address bindings.
- Uses the DHCP snooping database by default, alternatively use static entries with an ARP access-list.
- Only untrusted interfaces (all by default) will be validated. Use the ip arp inspection trust command to trust an interface.
- Untrusted interfaces are also rate-limited to 15 pps, customize with the ip arp inspection limit command on an interface.
- Only works at the ingress level. Optionally err-disable violating ports with errdisable detect cause arp-inspection command

```
ip dhcp snooping vlan 10
ip arp inspection vlan 10

int fa0/0
 description DHCP_SERVER
 ip arp inspection trust

int range fa0/1 -10
 description CLIENTS
 ip arp inspection limit 10

show ip arp inspection
```

DAI with ARP Access-List

- ARP access-lists are checked before the snooping database.
- Can also be configured without DHCP snooping enabled.
- Configure an explicit deny in the ACL to stop consulting the snooping database if no match is found.
- Alternatively add the static keyword behind the ip arp inspection filter command to not check the snooping database.

```
arp access-list HOST1
 permit ip host 10.0.10.1 mac host 1234.abcd.1234 log
 deny ip any mac any

ip arp inspection filter HOST1 vlan 10 static
```

Etherchannel

Link Aggregation Control Protocol (LACP)

- Requirements. Full-duplex, same STP cost and same speed.
- LACP Suspended ports (not meeting above criteria) can receive, but not send BPDUs.
- Only 8 LACP links can be active at a time.
- The master port is the port that is responsible for the STP traffic.

Hot-Standby

- Places additional links (up to 16) in hot-standby mode to replace links that have become inactive.
- Only the LACP master switch will replace the link.
- The system-ID is a combination of the LACP system priority and the switch MAC address. (same as STP, 32678).
- The LACP master switch is based on the system-ID. Meaning that the default STP root will also be the LACP master.
- Manually change the priority with the lacp system-priority command. Lower priority is better.

Links are activated based on the LACP port priority followed by the port number. Lower priority is better.

- The priority is equal by default, meaning that the highest port will become hot-standby (fa0/9).

```
int range fa0/1 - 9
 shutdown
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 99
 switchport mode trunk
 channel-group 1 mode active

int port-channel 1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 99
 switchport mode trunk

int range fa0/1 - 9
 no shutdown

show lacp sys-id
show lacp internal
show lacp neighbor
show etherchannel summary
```

Port Aggregation Protocol (PAgP)

- Only 8 PAgP links can be active at a time.
- Uses desirable, auto instead of active, passive.
- Does not support Hot-Standby ports, port-priority settings have nothing to do with standby interfaces.
- Port-priority and learn methods are used for interoperability between different platforms, not actual traffic distribution.

```
define interface-range PAGP fa0/1 - 8

int range macro PAGP
 shutdown
 switchport trunk encapsulation dot1q
 switchport mode trunk
 channel-group 1 mode desirable

int port-channel 1
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 99
 switchport mode trunk
```

```
int range macro PAGP
 no shutdown

show pagp internal
show pagp neighbor
show etherchannel summary
```

Load-Balancing (Load-Sharing)

- With simple methods the same traffic type will always choose the same link.
- With complex methods the same traffic type will always choose the same link.
- Etherchannels actually use load-sharing because one interface is chosen. Traffic is not sent over multiple paths.
- Change the load-sharing scheme with the port-channel load-balance command.

Simple Methods:

- scr-mac
- scr-ip
- dst-mac
- dst-ip

Complex Methods:

- scr-dst-mac
- scr-dst-ip
- scr-dst-port (N/A)

Link-State Tracking

- Ports connected to servers are configured as downstream ports,
- Ports connected to other switches are configured as upstream ports.
- If the upstream trunk ports (Etherchannel) fails, the downstream ports are put in an err-disable state.
- This is useful when a server connects to two separate switches with two NICs.
- If one switch loses its connection upstream, the NIC port to that switch is shutdown and the server will move over to the other switch as the primary NIC port.

```
link state track 1
int range fa0/23 - 24
 switchport trunk encapsulation dot1q
 switchport mode trunk
 channel-group 1 mode active
 link state group 1 upstream

int po1
 switchport trunk encapsulation dot1q
 switchport mode trunk
 link state group 1 upstream

int fa0/0
 switchport mode access
 switchport access vlan 10
 link state group 1 downstream

show link state group detail
```

IGMP Snooping

Monday, December 21, 2015
10:38 AM

L2 Multicast Address Conversion

- 1st octet is irrelevant, replace with 01-00-5E.
- Convert the 2nd octet to binary and set the first bit to 0.
- Convert the 2nd, 3rd and 4th octet to hex.

- In this case the 230.255.124.2 and 227.127.124.2 lead to exactly the same L2 address.
- A 2nd octet of 255 or 127 will always result in the same value.

IPv4 address	1st octet	2nd octet	3rd octet	4th octet	L2 address
226.144.154.4	01-00-5E	00100000 = 10	9A	04	01-00-5E-10-9A-04
230.255.124.2	01-00-5E	01111111 = 7F	7C	02	01-00-5E-7F-7C-02
227.127.124.2	01-00-5E	01111111 = 7F	7C	02	01-00-5E-7F-7C-02

IGMP Snooping / Querier

- The switch examines IGMP messages and learns the location of mrouters and hosts.
- Listens for IGMP Reports/Leaves and limits multicast traffic to specific ports only.
- Enable IGMP snooping globally or per VLAN (enabled by default on all VLANs).
- Can also bind static groups (using L2 address conversion) to interfaces.

```
ip igmp snooping
ip igmp snooping vlan 10
ip igmp snooping vlan 1 static 01-00-5E-7F-7C-02 interface fa0/0

show mac address-table multicast
```

IGMP Profiles

- IGMP Profile allows IGMP access-control at Layer 2.
- Can only be applied to L2 interfaces.
- Permit mode. Allows specified groups and blocks all others.
- Deny mode. Blocks specified groups and allows all other.

Only allow the specific multicast range:

```
ip igmp profile 1
 permit
 range 224.0.O.O 229.255.255.255

int fa0/0
 ip igmp filter 1

show ip igmp profile
```

PIM Snooping

- IGMP Snooping only limits traffic to and from hosts, PIM Snooping also limits traffic to mrouters.
- Limits multicast traffic to interfaces that have downstream receivers joined to the same multicast group.
- Listens to PIM hello, join, forward-election and prune messages.
- Requires IGMP snooping and is applied on VLAN SVIs.

```
interface vlan 10
 ip pim snooping
```

IP Source Guard

Saturday, December 5, 2015
4:01 PM

IP Source Guard

- Checks the source IP address of received packets against the DHCP snooping binding database.
- IP source guard is a port-based feature that automatically creates an implicit port access control list (PACL).
- With prefer port-mode IP source guard is preferred over VACL configurations.
- With merge mode IP source guard and VACL configurations are both implemented (default mode).

```
int fa0/0
 ip verify source

ip source binding 1234.abcd.1234 vlan 10 10.0.12.1 int fa0/0
```

IPv6 Security

Monday, December 21, 2015
7:47 PM

- Router Advertisements (RA) Guard
- Apply on ports that should not receive router advertisements (RAs).
- In other words, another router should not be present on the link.
- Configure on the switch to allow only RA from a single router on the specified port.

```
ipv6 prefix-list RA_PREFIX permit 2001:10:0:12::/64
ipv6 access-list RA_SOURCE
 permit ipv6 host FE80::1 any

ipv6 nd raguard policy ROUTER
 device-role router
 match ra prefix-list RA_PREFIX (optional)
 match ipv6 access-list RA_SOURCE (optional)

int fa0/0
 description TRUSTED_ROUTER_PORT
 ipv6 nd raguard attach-policy ROUTER
```

- Block all other RAs from other sources on the specified VLAN:

```
ipv6 nd raguard policy HOSTS
 device-role host

ipv6 snooping logging packet drop
vlan configuration 1
 ipv6 nd raguard attach-policy HOSTS
```

- Or:

```
interface vlan 1
 ipv6 nd raguard attach-policy HOSTS
```

- There is no need to specify a policy when applying RA Guard to untrusted hosts. The ipv6 nd raguard command will suffice.
- The ipv6 snooping logging packet drop command is needed for logging untrusted RA messages.

- RA Guard PACL
- It is also possible to configure the concept of RA Guard using a PACL to block unwanted RAs.
- The undetermintedtransport keyword must be included to capture all unwanted traffic.

```
ipv6 access-list RA_PACL
 deny icmp any any routeradvertisement
 deny ipv6 any any undeterminedtransport
 permit ipv6 any any

int fa0/0
 description UNTRUSTED_PORT
 ipv6 traffic-filter RA_PACL in
```

- DHCPv6 Guard
- Trust all ports but require matching on link-local source and address range reply:

```
ipv6 prefix-list TRUSTED_PREFIX permit 2001:10:0:12::/64 le 128
ipv6 access-list TRUSTED_SERVER
 permit ipv6 host FE80::1 any
```

```
ipv6 dhcp guard policy DHCP
 device-role server
 match server access-list TRUSTED_SERVER
 match reply prefix-list TRUSTED_PREFIX

vlan configuration 1
 ipv6 dhcp guard attach-policy DHCP
```

- Or:

```
interface vlan 1
 ipv6 dhcp guard attach-policy DHCP
```

- ND Inspection
- Control plane feature only, it doesn't inspect actual data traffic and only looks at ND ICMP packets.
- Builds a table based on NS/NA messages. It then enforces the table.
- If there is a link local address on the network, the switch will send a NS from the IPv6 address.
- If there isn't an IPv6 address on the VLAN (L2 switching only), it will send an NS from the IPv6 unspecified address.
- ND Tracking Policy is optional when enabling ND Inspection.
- Tracking is basically IP SLA echo only with ND packets. And is useful for two reasons:
  - Since the table is first-come first-serve, this frees up address space if it's actually not in use.
  - It allows for a host to move ports by aging out information.

```
vlan configuration 1
 ipv6 nd inspection
show ipv6 neighbor binding

Create a static binding to allow a certain host:
ipv6 neighbor binding vlan 1 FE80::1 interface gi1 abcd.abcd.abcd

show ipv6 snooping capture-policy interface
```

- Enable ND Inspection alongside a tracking policy:

```
ipv6 nd inspection policy ND_POLICY
 tracking enable [reachable-lifetime] 300

vlan configuration 1
 ipv6 nd inspection attach-policy ND_POLICY
```

- Destination Guard
- Destination Guard is a "last hop" security feature, the last hop router is the only one that is heavily impacted.
- Interim routers don't have to NS for the final destination, they just CEF-switch the packet.
- Needed because of the size of /64 IPv6 subnets and the possible amount of destinations.

```
ipv6 destination-guard policy DESTINATION_POLICY
 enforcement always | stressed

vlan configuration 1
 ipv6 destination-guard attach-policy DESTINATION_POLICY

show ipv6 destination-guard policy
```

- The stressed option will only enable Destination Guard during high usage. Default is always.

IPv6 Snooping

Monday, December 21, 2015
7:47 PM

- IPv6 Snooping
- Builds the neighbor database, similar to IPv6 ND inspection.
- Glean ports are trusted ports.
- The difference is that it uses and enforces more methods all at once. It can use:
    - Information from DHCP (Default).
    - Information from ND (Default).
    - Static bindings.

```
ipv6 snooping policy UNTRUSTED_HOSTS
 security-level guard | inspect

vlan configuration 1
 ipv6 snooping attach-policy UNTRUSTED_HOSTS
```

- The guard keyword enables DHCP Guard, RA Guard, and ND Inspection.
- The inspect keyword only enforces ND Inspection.

- The policy is optional when enabling IPv6 Snooping on untrusted ports.
- By default, IPv6 snooping enables its version of RA Guard, DHCPv6 Guard and ND Inspection.
- Optionally disable ND Inspection with the no protocol ndp command.
- Optionally disable DHCPv6 Guard with the no protocol dhcp policy command.

```
ipv6 snooping policy TRUSTED_ROUTER
 security-level glean

int gi1
 ipv6 snooping attach-policy TRUST_ROUTER

show ipv6 neighbor binding
```

- IPv6 Source Guard
- Any traffic other than IPv6 ND/RA and DHCP that doesn't match the source address present in the prebuilt binding table, will get dropped.

```
vlan configuration 1
 ipv6 source-guard
```

Link-State Track

Saturday, December 5, 2015
7:12 PM

Link-State Tracking

- Downstream ports connect to servers (with more than two NICs)
- Upstream ports are part of an etherchannel.
- If the etherchannel fails, the downstream ports will be err-disabled triggering a switchover to another NIC on the server.

```
link state track 1

int po1
 link state group 1 upstream

int fa0/0
 description SERVER
 link state group 1 downstream

do show link state group detail
```

MAC

Monday, November 30, 2015
10:08 AM

MAC Aging

- Default is 300 seconds for all VLANs.
- It is possible to apply different aging timers for different VLANs.

```
mac address-table aging-time 500 vlan 10
show mac address-table aging-time
```

MAC Learning

- All interfaces and VLANs can learn MAC-addresses.
- It is possible to disable learning for specific VLANs.

```
no mac address-table learning vlan 10
show mac address-table learning
```

Static MAC

- Does not expire and can be associated with multiple ports.
- The specified output interface cannot be an SVI.
- When configuring static MAC-address in a primary private-VLAN, also configure the same address in the secondary VLAN.

```
mac address-table static 1234.abcd.1234 vlan 10 int fa0/0 fa0/1
show mac address-table static
```

MAC Filtering

- Disabled by default and only supports unicast static addresses.
- Multicast MAC addresses, broadcast MAC-addresses, and router MAC addresses are not supported (forwarded to CPU).
- Only works on access-ports, L2 ACL takes precedence over L3 ACL.
- You cannot apply named MAC extended ACLs to Layer 3 interfaces.
- When filtering the same MAC-address that is also configured statically, the command configured last is applied.

```
mac access-list extended MAC_ACL
 deny   any any appletalk
 permit any any

int fa0/0
 mac access-group MAC_ACL in

show access-lists
show mac access-group
```

Be careful when only permitting only certain addresses. MAC-addresses will time out and the MAC ACL will block the ARP.

- In this case create static entries that match the ones permitted in the MAC ACL.

```
mac access-list extended MAC_ACL
 permit host 1234.abcd.1234 host abcd.1234.abcd
 deny any any

int fa0/1
 mac access-group MAC_ACL in

mac address-table static 1234.abcd.1234 vlan 10 int fa0/0
mac address-table static abcd.1234.abcd vlan 10 int fa0/0
```

Macros

Tuesday, December 8, 2015
9:03 PM

Switch Macros

- End custom macros with the @ sign.
- Show the individual commands in the cli with the trace keyword.

Global macro:

```
macro name ACCESS_PORT
 default int $int
 int $int
 switchport mode access
 switchport access vlan $vlan
 switchport voice vlan $voice
 no shut
end
@

macro global trace ACCESS_PORT $int fa0/0 $vlan 10 $voice 60
```

Interface macro:

```
macro name ACCESS_PORT
 switchport mode access
 switchport access vlan $vlan
 switchport voice vlan $voice

int fa0/0
 macro trace ACCESS_PORT $vlan 10 $voice 60

show parser macro
```

Native VLAN

Native VLAN

- All traffic that is untagged is placed in the native VLAN.
- Primarily used for management traffic nowadays.
- LLDP and CDP use the native VLAN.
- Used for the switch behind an IP-Phone. Voice traffic will use the voice VLAN and data traffic from PC will be untagged.
- 802.1D (STP) frames are sent in the native VLAN.

Native VLAN mismatch errors apply to the CDP protocol. Traffic can actually flow between mismatched switches.

- PVST will give errors with this configuration. Other STP configurations will work.
- Native VLANs are useful when the same VLANs exist on opposite ends of the switches, but are not used for the same purpose. For example: Cat2--->Native VLAN 2------Trunk Connection------Native VLAN 3<---Cat3

Native VLAN Mismatch Trunk

Cat2:

```
int fa0/1
 description CON_TO_R2
 switchport access vlan 2

no spanning-tree vlan 2

int fa0/24
 description CON_TO_CAT3
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 2
 switchport mode trunk
 no cdp enable
```

Cat3:

```
int fa0/1
 description CON_TO_R3
 switchport access vlan 3

no spanning-tree vlan 3

int fa0/24
 description CON_TO_CAT2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 3
 switchport mode trunk
 no cdp enable
```

Routers:

```
int fa0/0
 description CON_TO_CAT2
 ip address 10.0.23.2 255.255.255.0
 no shut

int fa0/0
 description CON_TO_CAT3
 ip address 10.0.23.3 255.255.255.0
 no shut
```

PACL / VACL

Ingress Direction	PACL is applied first, then VACL, then ACL on same VLAN.
Egress Direction	ACL on VLAN is applied first, then VACL, no support for PACL.

Port-Based Access Control Lists (PACL)

- Can only be applied in the ingress direction.
- Can be configured on a trunk port only with Prefer Port Mode.

Two modes:

- Prefer Port Mode. Overwrites the effect of other ACL or VACL (not supported on 3750).
- Merge Mode. PACL, VACL, and ACLs are merged in the ingress direction (default mode).

```
ip access-list extended PACL
 deny icmp any any
 permit ip any any

int fa0/0
 ip access-group mode merge
 ip access-group PACL in
```

VLAN Access Control Lists (VACL)

- Unlike ACLs that are applied on routed packets only, VACLs apply to all packets (L2 and L3).
- Can be applied to any VLAN and can filter traffic between devices in the same VLAN or between VLANs.
- Uses traffic matched in ACL with permit statements.
- Important. When matching on MAC ACL the explicit deny is for L2 traffic only, all unmatched L3 traffic is allowed.
- The same is true for the reverse (matching on L3 ACL will only explicitly deny L3 traffic).

VACL for Private-VLANs

- Host port to promiscuous port traffic is matched on secondary VLAN access-list.
- Promiscuous port to host port traffic is matched on primary VLAN access-list.
- To filter out specific IP traffic apply the VACL to both the primary and secondary VLANs.

L3 Filtering:

```
ip access-list extended ICMP
 permit icmp any any

vlan access-map VACL 10
 match ip address ICMP
 action drop
vlan access-map VACL 20
 action forward

vlan filter VACL vlan-list 10,20
```

L2 Filtering:

```
mac access-list extended MAC
 permit host 1234.abcd.1234 any

vlan access-map VACL 10
 match mac address MAC
 action drop
vlan access-map VACL 20
 action forward

vlan filter VACL vlan-list 10, 20
```

VLAN Access Logging

- The content of the log table can be deleted by setting the maxflow to 0.
- When the log table is full, the software drops logged packets from new flows (500 is default).
- The default threshold is 0, which means that a syslog message is generated every 5 minutes.

```
vlan access-log maxflow 500
vlan access-log threshold 0
```

Port Security

Monday, November 30, 2015
12:08 PM

Protected Ports

- Does not forward any traffic to any other port that is also a protected port.
- Only control traffic (PIM for example) is forwarded.
- All data traffic passing between protected ports must be forwarded through a Layer 3 device.

```
int fa0/0
 switchport protected
```

Port Blocking

- Block flooding of unknown multicast and unicast traffic out of a port.
- The default is to flood unknown destinations out of all ports.

```
int fa0/0
 switchport block multicast
 switchport block unicast

show int fa0/0 switchport
```

Port-Security

- Can only be configured on static access ports or trunk ports, not dynamic (DTP) ports.
- Not supported on Private-VLAN ports.
- When using the voice VLAN, set the maximum number of MAC-addresses for both the access and voice VLAN.
- The port does not forward packets with source addresses outside the group of defined addresses.

Violation modes:

- Protect. Stops forwarding traffic of new MAC-addresses after limit is reached.
- Restrict. Works the same as protect, but also sends a trap and syslog message.
- Shutdown. Err-disables the port, or the entire VLAN when the vlan keyword is added.

Secure MAC address types:

- Static. Manually defined and stored in running-config.
- Dynamic. Learned automatically until limit is reached. Removed when switch is rebooted.
- Sticky. Dynamically learned and stored in running config (requires write).
- The switchport port-security mac-address sticky command converts all the dynamic secure MAC addresses to sticky.

Port-Security Aging

- Absolute. The secure addresses on the port are deleted after the specified aging time (default).
- Inactivity. The secure addresses on the port are deleted only if they are inactive.
- Static address aging is disabled by default. Enable with the switchport port-security aging static command.
- Sticky addresses do not age, dynamic addresses that are converted to sticky will lose aging limitations.

Dynamic port-security with aging:

```
int fa0/0
 switchport mode access
 switchport access vlan 10
 switchport voice vlan 60
 switchport port-security
 switchport port-security maximum 3
 switchport port-security maximum 2 vlan access
 switchport port-security maximum 1 vlan voice
 switchport port-security violation restrict
 switchport port-security aging type inactivity
 switchport port-security aging time 60
```

Static port-security:

```
int fa0/1
 switchport mode access
 switchport access vlan 20
 switchport port-security maximum 2
 switchport port-security mac-address ca02.1598.0008
 switchport port-security mac-address 000e.04b7.801a
 switchport port-security violation shutdown
 switchport port-security

show port-security address
show port-security interface fa0/0
```

Private VLANs

Private VLANs

- All hosts in the private VLAN belong to the same IP subnet as the primary VLAN.
- Private VLANs require VTPv3 or VTP mode transparent on older versions of VTP.
- Isolated VLAN forwards frames from isolated to promiscuous ports. (only 1 allowed)
- Community VLAN forwards frames from community to promiscuous and other community ports. (multiple allowed)

Promiscuous PVLAN Trunk rewrites secondary VLAN ID to primary VLAN ID.
Isolated PVLAN Trunk translates primary VLAN ID to isolated VLAN ID.

- Used for connecting to switches without PVLAN support (protected ports).

Private VLANs Communication

It is possible for two isolated PVLAN ports to communicate through the promiscuous port if ip local-proxy-arp is configured.

- The ip proxy-arp is used by routers to answer to ARP queries that are outside the network (on by default).
- The ip local-proxy-arp is used by routers to answer to ARP queries that are inside the local subnet, used by PVLAN.
- IP Local-Proxy-ARP needs IP Proxy-ARP active in order to function.
- Another workaround is to use OSPF PTMP network type.
- Enabling DHCP Snooping, ARP Inspection and Source Guard on primary VLAN will also enable it on secondary VLANS.

```
vtp mode transparent
vlan 100
 private-vlan primary
vlan 110
 private-vlan community
vlan 120
 private-vlan isolated
vlan 100
 private-vlan association 110,120

int fa0/1
 description PROMISCUOUS
 switchport mode private-vlan promiscuous
 switchport private-vlan mapping 100 110,120

int range fa0/2 - 3
 description COMMUNITY
 switchport mode private-vlan host
 switchport private-vlan association host 100 110

int fa0/4
 description ISOLATED
 switchport mode private-vlan host
 switchport private-vlan association host 100 120

show vlan private-vlan
```

SVIs with Private-VLANs

- Only work if the SVI is configured the same way as the promiscuous port.
- The primary VLAN does not have to be specified in the SVI configuration.
- Using an SVI is required to perform DHCP services on the same switch that has Private-VLANs configured.

```
interface vlan 100
 private-vlan mapping 110,120
```

QinQ

QinQ Tunnel

- The native VLAN is not QinQ tagged by default, enforce using the vlan dot1q tag native command.
- QinQ adds a 4 byte overhead to ethernet frames (1504).
- Make sure switches increase the system MTU by at least 4 bytes (reload required).
- It is possible to send over CDP, STP, and other L2 protocols using QinQ.

Configure Cat1:

```
vlan dot1q tag native
int fa0/0
 switchport access vlan 30
 switchport mode dot1q-tunnel
 l2protocol-tunnel cdp
 l2protocol-tunnel dtp

show dot1q-tunnel
```

SPAN

Switch-Port Analyzer (SPAN)

- SPAN does not affect the switching of traffic on sources. You must dedicate the destination port for SPAN use.
- Source can be multiple ports or a VLAN, but not both. Multiple ports can be combined into a single session.
- Only a single session per destination port is allowed.
- Does not capture traffic that is forwarded within the switch (Inter-VLAN).
- Local SPAN does not copy locally sourced RSPAN and ERSPAN traffic.
- Allows up to 64 span destination ports per switch.

Capture options:

- Rx. Received traffic without modification.
- Tx. Sent traffic after modification by ACL, QoS, etc.
- Both. Default setting, both directions are captured.

To capture L2 frames such as CDP, BPDU, VTP, DTP add the encapsulation replicate to the destination command.

- The filter keyword can filter on specific VLANs or subnets (using ACL) when the source is a trunk port.
- Filtering using ACL is called FSPAN and is not supported on the 3750.
- Whatever you specify in the filter will be the only thing that matched on.
- The ingress keyword allows traffic from the monitoring station to enter the switch on the SPAN destination port.

```
mon session 1 source interface fa0/0 both
mon session 1 destination interface fa0/3 encapsulation replicate
mon session 1 filter vlan 10
```

Remote Switch-Port Analyzer (RSPAN)

- Sacrifice a VLAN to use for RSPAN with the remote-span keyword under the VLAN configuration.
- VTPv3 can propagate RPSAN VLAN information.
- Sacrifice a port to use for RPSAN with the reflector-port keyword.
    - This is the port that copies packets onto a RSPAN VLAN.
    - The reflector port cannot be used for any other purpose, however it must be online and connected.

Configuring Remote SPAN Source:

```
vlan 100
 name RSPAN
 remote-span
monitor session 1 source interface fa0/0 both
monitor session 1 destination remote vlan 100 reflector-port fa0/24
```

Configuring Remote SPAN Destination:

```
vlan 100
 name RSPAN
 remote-span
monitor session 1 source remote vlan 100
monitor session 1 destination interface fa0/3
```

```
show vlan remote-span
```

Encapsulated Remote Switch-Port Analyzer (ERSPAN)

- Changes only take effect when exiting the sub-configuration mode. Sessions also have to be enabled with no shutdown.
- The origin ip address keyword specifies the source IP on the ERSPAN source.
- The ip address keyword specifies the destination IP on the ERSPAN source.
- The ip address keyword specifies the source IP on the ERSPAN destination.
- The values for ip address have to match on both sides.

Configuring ERSPAN Source:

```
monitor session 1 type erspan-source
 source interface gi1
 destination
  erspan-id 12
  ip address 192.168.0.1
  origin ip address 192.168.0.2
  no shutdown
```

Configuring ERSPAN Destination:

```
monitor session 1 type erspan-destination
 destination interface gi2
 source
  erspan-id 12
  ip address 192.168.0.1
  no shutdown
```

Storm Control

Monday, November 30, 2015
4:34 PM

Storm Control

- Traffic storm control monitors traffic in 1-second traffic storm control intervals.
- The normal transmission restarts when traffic drops below the falling threshold.
- If you don't specify a falling value, the rising value will be copied.
- 100 percent means no traffic storm control, 0 percent suppresses all traffic.
- The default action is drop traffic (no action configured).
- Optionally send a trap or err-disable the port with the storm-control action command.

```
int fa0/0
 storm-control broadcast level bps 500 100
 storm-control action shutdown

show storm-control fa0/0 broadcast
```

Small-Frame Storm Control

- Incoming VLAN-tagged packets smaller than 67 bytes are considered small frames.
- Do not cause the switch storm-control counters to increment.

```
int fa0/0
 small violation-rate 1000

show int fa0/0
```

Voice VLAN

Connecting a phone to a switch, two options:

- Trunk interface. Phone will use VLAN, PC will use Native VLAN.
- Access interface. Phone will use Voice VLAN, PC will use Access VLAN.

Voice VLAN

- The voice VLAN feature enables access ports to carry IP voice traffic from an IP phone.
- Voice traffic is sent with Layer 3 IPP (5) and Layer 2 CoS (5).
- The internal IP Phone switch will tag VoIP traffic with the Voice VLAN tag. Data from PC is sent untagged to switch.
- The switchport voice detect cisco-phone is a security measure to make sure a cisco phone is connected.

```
int fa0/0
 switchport access vlan 10
 switchport voice vlan 60
 spanning-tree portfast
```

Voice VLAN Dot1p (802.1p)

- Use a single VLAN for Data and Voice but the 802.1p CoS tag is added to the voice traffic.
- The 802.1p CoS field is set to 5 for VoIP and the access VLAN on the switch is used for data frames.
- Call control CoS field is set to 3. PC traffic should be the default CoS of 0.
- The benefit of this mode is that you get QoS abilities without needing a separate voice VLAN.

```
int fa0/0
 switchport access vlan 10
 switchport voice vlan dot1p
 spanning-tree portfast
```

Voice VLAN None

- Allow the phone to use its own configuration to send untagged voice traffic.
- This provides the same result as Dot1p, however the decisions are left to the phone.

```
int fa0/0
 switchport access vlan 10
 switchport voice vlan none
 spanning-tree portfast
```

Voice VLAN Untagged

- Forces the phone to use untagged voice traffic (native VLAN)

```
int fa0/0
 switchport mode trunk
 switchport trunk native vlan 60
 switchport voice vlan untagged
```

Voice VLAN CoS Extension

- Untrusted mode is default and will remark all packets from the PC to CoS 0.
- It is also possible to remark the PC traffic to a specific value using the extend cos command.
- Trusted mode will accept whatever CoS value the PC is sending packets with.

```
int fa0/0
 switchport priority extend cos 5
 switchport priority extend trust
```

LLDP Voice Vlan

- Default DSCP value is EF (46) when configuring the Voice VLAN under the network-policy profile.

```
network-policy profile 1
 voice vlan 60 cos 5
 voice vlan 60 dscp 40
 voice vlan dot1p cos 5
 voice vlan none
 voice vlan untagged

int fa0/0
 switchport mode access
 switchport access vlan 10
 network-policy profile 1
 lldp med-tlv-select network-policy

show network-policy profile 1
```

VLANs

Default VLANs

VLAN 1002	fddi-default
VLAN 1003	token-ring-default
VLAN 1004	fddi-net-default
VLAN 1005	token-ring-net-default

VLAN Database

- When the switch is in VTP server or transparent mode, you can configure VLANs in the VLAN database mode.
- When you configure VLANs in VLAN database mode, the VLAN configuration is saved in the vlan.dat file.
- Supports creation of VLAN 1 to 1001, not extended range.

Routed Interface VLAN Allocation

- A routed interface on the switch (SVI or interface-based) is assigned a VLAN from 1006 and up by default.
- These VLANs will become unusable for their normal purpose.
- Can be altered to descend from VLAN 4094 downwards with the vlan internal allocation policy command.

VSS

Virtual Switching System (VSS)

- Virtual Switch Link (VSL) is an Etherchannel connection between Active and Standby chassis with up to 8 links.
- Carries control and data traffic between Active and Passive chassis.
- Control traffic has a higher priority than data, data traffic is load-balanced.
- Does not preempt by default.

Active chassis	Controls the VSS, provides management functions. Online insertion and Removal (OIR) + console Switches L2 and L3 Packet forwarding on local interfaces
Standby chassis	Packet forwarding on local interfaces Sends management traffic to Active chassis

Virtual Switch Link Protocol (VSLP)

- Contains the Link Management Protocol (LMP) and Role Resolution Protocol (RRP).
- The LMP runs on all VSL links and exchanges information required to establish communication between the two chassis.
- After a LMP connection is established, the peer chassis use RRP to negotiate the role (active or passive) for each chassis.
- With Stateful Switchover (SSO) redundancy, the VSS standby supervisor engine is always ready to assume control following a fault on the VSS active supervisor engine.

VSS Initiation process:

- Preparse config.
- Bring up VSL links.
- Run VSLP (LMP).
- Run RRP.
- Interchassis SSO.
- Continue system bootup.

```
switch convert mode virtual
int port-channel 10
 switch virtual link 1
```

VSS Dual Active Recovery Detection

- Provides means of communication between both VSS chassis outside the VSL link.
- If the VSL connection goes down, the active switch will be informed and will go into recovery mode.
- In this mode, all ports except the VSL ports are shut down.
- Upon seeing the VSL ports come active again, the switch will reload and come back as the standby chassis with all its ports up.
- An enhanced version of PAgP is used on the EtherChannel and provides the Dual-Active Detection.

```
switch virtual domain 1
 dual-active detection pagp
 dual-active trust channel-group port channel 10

show switch virtual dual-active pagp
show pagp dual-active
```

VTP

VLAN Trunking Protocol (VTP)

- When the domain is NULL no VTP messages will be sent.
- The switch is waiting for the first reception of a VTP message with the domain name set.
- VTP only operates on trunk ports.
- VTP does not merge VLAN databases, it only compares the MD5 hash and if it is not equal the highest revision number will win and overwrite the existing database.
- Devices with same revision number will not update each other when first coming online, because the hash is different.
- VTP authentication alters the MD5 hash.
- VTPv1 and v2 are designed to work with ISL, this is why extended range VLANs are not supported.
- VTP off mode is the same as VTP transparent mode except that VTP advertisements are not forwarded.
- Extended VLANs can only be configured with VTPv3 or VTPv2 transparent mode. If you configure extended range VLANs with any other mode, the changes are lost when the switch is rebooted.
- MST and PVST databases are separate, this is why you can configure switches with vtp mode server mst or mode server vlan.
- You can configure the switch as a VTP server for the VLAN database but with VTP off for the MST database.
- The MST database is the instance to VLAN mapping, not the actual VLAN list.

VTP Pruning

- Switches communicate which VLANs it doesn't need frames for. (no trunks or active access ports in a specific VLAN)
- Other switches will stop forwarding these frames until the switch communicates that it has an active port in the VLAN.
- Only servers can enable VTP pruning.
- Specify VLANs that are eligible for pruning with the switchport trunk pruning vlan command on the interfaces.
- VLANs specified with this command will be pruned, anything not specified will not be pruned. Default is prune all VLANs.

(VTPv3)

- VTPv3 provides enhanced authentication, the hidden option hides the password in the configuration.
- VTPv3 supports extended VLAN range (1006-4094) and private VLANs (and also propagates RPSAN VLAN information).
- VTPv3 is backwards compatible with v2 and will send VTPv3 and v2 packets to devices over a trunk port.
- Devices using v1 will automatically update to v2 when communicating with v3 neighbors.
- If transparent or off mode is selected, VLANs are also present in the running-config.
- VTPv3 removes possibility to reset the revision to 0 by changing to transparent mode and back.
- The revision number will be reset to 0 only by modifying the VTP domain name or by configuring a VTP password.
- Before making any changes to the VTP configuration, change switches to transparent mode.

All configuration changes should be made on the v3 switches with the v2 devices configured as clients.

- If changes are made on the v2 devices, only v2 neighbors will update their revision numbers.
- Updates from v3 neighbors will be rejected later on, because of the revision number.

VTPv3 Server modes

- Primary. Can modify vtp, only 1 per domain. Configure with vtp primary command.
- Secondary. Cannot modify vtp, can be promoted.

A primary server is the only switch in a VTPv3 domain whose VLAN database can be propagated throughout the domain.

- Demote a Primary Server to a Secondary (normal) server by changing modes or by setting a VTP password.
- MST and VLAN are separate, meaning that different switches can be the primary server for MST or PVST.
- If the VTP password is configured, you need to enter it when promoting the server to primary.

VTP client can update database if:

- The new link connecting the new switch is trunking.
- The new switch has the same VTP domain name as the other switches.
- The new switch's revision number is higher than that of the existing switches.
- The new switch must have the same password, if configured on the existing switches.

Even in VTPv3, a secondary server or a client switch with a higher revision number can overwrite a VLAN database.

- This can only occur if the above is true, plus the switches agree on the identity of the primary server.

```
vtp file flash:VTP.dat
vtp version 3
vtp mode server
vtp domain cisco
vtp password cisco hidden
do vtp primary vlan force
cisco
```

```
show vtp password
show vtp device
show vtp status
debug sw-vlans vtp events
```

With the force keyword, the switch does not check for conflicting devices.
The hidden keyword hides the password in the configuration.

- After configuring a password it is required in order to promote a server to primary.
- The hidden password cannot be used with v2 neighbors.

Disable VTP

```
vtp mode off
int fa0/0
 no vtp
```

VTPv3 Feature Unknown

The unknown feature allows you to configure the behavior of the switch databases that it cannot interpret.

- These databases will be features handled by future extensions of VTP version 3.
- Default is off for unknown instances, can set to transparent.

```
vtp mode off unknown
vtp mode transparent unknown
```