Part No. 013001-01, Rev. A
March 2011

OmniSwitch Feature Guidelines
AOS Release 6.4.3.R01

Alcatel-Lucent EDS
26801 West Agoura Road, Calabasas, CA 91301
(818) 880-3500
www.alcatel-lucent.com

Table of Contents

1. About This Guide

The *OmniSwitch Feature Guidelines* guide provides recommendations and guidelines for release 6.4.3.R01 of the OmniSwitch 6400, 6850, 6855, 9000, and 9000E.

Who Should Read this Guide?

The intended audience for this document is anyone wanting to become familiar with the features of the OmniSwitch products. Reading the 6.4.3.R01 Release Notes prior to reading this guide is highly recommended.

What is in this Guide?

This Section …	Contains …
New Feature Guidelines	Information about new features and enhancements introduced in the 6.4.3.R01 release. This includes a brief description, platforms supported, guidelines, sample configurations, and references to other related documentation.
Existing Feature Guidelines	Information about existing features (introduced in a previous release). These features are included in this guide to provide additional information about general switch guidelines and functionality.
Interoperability Test Results	Interoperability test information and results for tests conducted between the OmniSwitch and other vendor platforms.

2. New Feature Guidelines

This section contains guidelines for the following new and enhanced features that were introduced in the OmniSwitch AOS Release 6.4.3.R01:

Ethernet Interfaces

- **Traffic Storm Control**
- **BPDU Shutdown Auto-Recovery Timer**

Bridging

- **Dual Home Link Aggregation**
- **Load Balancing Non-Unicast on a Link Aggregate**
- **MVRP**
- **LLDP – VLAN Assignment**

Routing

- **Recursive Static Route**
- **Extended Ping and Traceroute Functionality**

QoS

- **Egress Policy Support**
- **Tri-Color Marking (srTCM/trTCM) Support**
- **Ingress Policy Rule Statistics Enhancements**
- **802.1ad DEI/CFI Bit Support**
- **Egress Port/Queue QoS and Statistics Enhancements**
- **Policy Condition Enhancements**
- **Map Several Inner DSCP/ToS Values to Same Outer 802.1p**

Security

- **DHCP Option-82 ASCII Enhancement**
- **MAC-Forced Forwarding (Dynamic Proxy ARP)**

Ethernet Access (Metro)

- **VLAN Stacking –Tunneling L2 Protocols**
- **SVLAN Routing**
- **Wire-Speed Ethernet Loopback**
- **Ethernet OAM 802.1ag Version 8 and ITU Y1731**
- **SAA for ETHOAM**
- **VPLS-MPLS**

Management

- **Internal DHCP Server**
- **DHCP Client**
- **Out-of-the-Box Auto-Configuration**
- **RADIUS Service-Type Attribute**
- **IP Managed Interface**

Hardware

- **XNI-U12E**

2.1. Traffic Storm Control

A traffic storm occurs when packets flood the LAN, creating excessive traffic and degrading network performance. The traffic storm control feature prevents LAN ports from being disrupted by a broadcast, multicast or unicast traffic storm on physical interfaces.

Traffic storm control monitors the level of incoming traffic, compares the level of traffic with the configured thresholds, and drops traffic when the traffic exceeds the specified thresholds. In addition, this feature allows the configuring of different threshold values (bits-per-second, percentage of the port speed, or packets-per-second) for different traffic types (broadcast, multicast, and unknown unicast).

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- Three traffic types supported:
  - Broadcast (destination MAC address is FF:FF:FF:FF:FF:FF)
  - Multicast (destination MAC address is multicast address)
  - Unknown unicast (destination MAC address is unknown)
- Three different threshold types supported:
  - Mega bits per second
  - Packets per second
  - Percentage of the port speed
- Default flood limit (broadcast, multicast and unknown-unicast):
  - 4 mbps if the port speed is 10 mbps
  - 49 mbps if the port speed is 100 mbps
  - 496 mbps if the port speed is 1G
  - 997 mbps if the port speed is 10G
- Traffic storm control threshold cannot be accurate because of hardware limitation.
- No counter for dropped packet.
- Only rate limiting supported, no traffic suppression support.
- No indication that flood rate limit has been reached.
- All user inputs can be changed slightly because of unit conversion and hardware limitation.

CLI Commands and Examples

Two new CLI commands were added to support this feature:

interfaces [*slot/port* **|** *slot/port1-port2***] flood [broadcast | multicast | unknown-unicast] [enable | disable]**

interfaces [*slot/port* **|** *slot/port1-port2***] flood [broadcast | multicast | unknown-unicast] rate [mbps *num* | pps *num* | percentage *num*]**

The **show interfaces flood rate** command displays the flood rate settings. For example:

```
→show interfaces flood rate

Slot/    Bcast      Bcast  Bcast      Ucast    Ucast  Ucast      Mcast    Mcast  Mcast
Port     Value      Type   Status     Value    Type   Status     Value    Type   Status
--------+----------+-----+---------+----------+-----+---------+----------+-----+---------+
 1/1      14880952   pps    enable       1000   pps    enable       49      %    disable
 1/2           100   pps   disable      10000   pps    enable       49      %    disable
```

The following CLI command was deprecated, but remains supported during the boot up process in the boot.cfg file:

interfaces [*slot/port* **|** *slot/port1-port2***] flood rate *num***

References

- Chapter 1, “Ethernet Port Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 1, “Configuring Ethernet Ports”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.2. BPDU Shutdown Auto-Recovery Timer

This enhancement is mainly for the BPDU Guard feature, also known as BPDU Shutdown, which is configured through the **qos user-ports shutdown** and **policy port group UserPorts** commands. Once **qos user-ports shutdown** is configured for BPDU, all ports that are configured in the UserPorts port group will be shutdown (disabled) when these ports receive a spanning tree BPDU.

In previous releases, the disabled ports had to be manually re-enabled. This feature provides a configurable timer that will automatically re-enable disabled UserPorts ports when the specified time has expired.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- The auto-recovery timer is global for all ports and all protocols.
- This feature is not supported for link aggregate ports.

Two new CLI commands to support auto-recovery:

interfaces violation-recovery-time *num*
interface violation-recovery-trap {**enable** | **disable**}

Configuration Examples

```
→vlan 1001
→vlan 1001 port default 1/2
→ip interface 1001 address 1.1.1.1 mask 255.255.255.0 vlan 1001
→qos user-port shutdown bpdu
→policy port group UserPorts 1/2
→qos apply
→interfaces violation-recovery-time 60

→show interfaces violation-recovery
UserPorts Shutdown Recovery Time: 60,
UserPorts Shutdown Recovery Trap: Disable
```

The port goes down when it receives a BPDU packet.

```
→show interfaces 1/2 port
Slot/    Admin     Link    Violations                 Alias
Port     Status   Status
-----+----------+---------+----------+-----------------------------------------
 1/2    enable      down      STP     ""
```

The timer is set to 60 seconds, after 60 seconds the switch will re-enable the port.

```
→show interfaces 1/2 port
Slot/    Admin     Link    Violations                 Alias
Port     Status   Status
-----+----------+---------+----------+-----------------------------------------
 1/2    enable      up        none    ""
```

References

- Chapter 1, "Ethernet Port Commands", *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).

- Chapter 1, “Configuring Ethernet Ports”, and Chapter 30, “Configuring ACLs”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.3. Dual Home Link Aggregation

The 6.4.3.R01 Release introduces Dual Home Link (DHL) Aggregation to provide fast failover for edge devices in a network. DHL enhances link aggregation by providing fast failover between core and edge switches without implementing Spanning Tree. An edge switch is connected to two core switches via LACP; one port is active and the other is on standby. The standby port provides failover to the active link.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

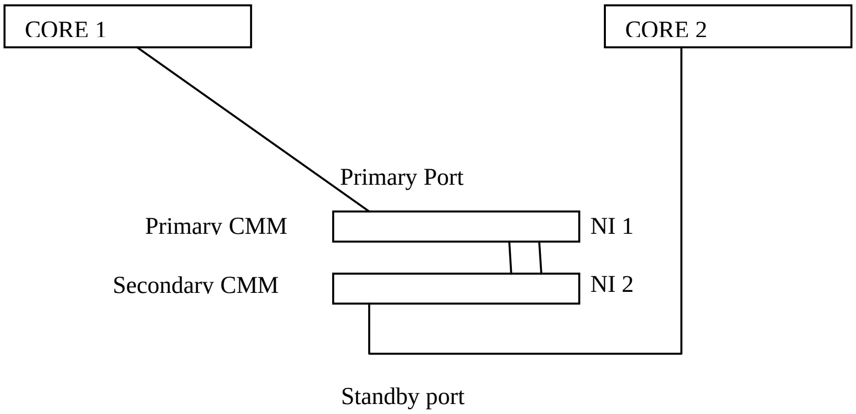
Guidelines

Analysis of DHL Convergence Across NIs with Link Down and Power Down

Convergence with Link Down

When the Primary port and the Standby port are configured across the stacking units, the convergence time increases to a value around 400-500ms.

The following is an example setup of this scenario:



When the Primary port goes down in this setup, the Standby port (port connected to a non-primary CMM) comes up and continues to send the traffic that was previously flowing through the port connected to the Primary CMM. However, before the Standby port comes up, there are a few messages communicated across the stacking link. The following events occur after the switch detects that the Primary port (in NI 1) is down:

1. NI 1 sends a message to CMM and NI 2 indicating that the Primary port is down.
2. NI 2 on receiving the message from NI 1, sends a Join Message to CMM (NI 1 currently)
3. Primary CMM on receiving the Join from NI 2 sends the message to 802.1Q CMM.
4. 802.1Q CMM sends the Join message to 802.1Q NI, which is NI 2, to program the new port in hardware.
5. 802.1Q NI in NI 2 on receiving the message from Primary CMM programs the port in Hardware.

Communication between the NI 1-NI 2, and NI 2 – Primary CMM (NI 1) occurs via the Stacking links. Stack links are the medium for communication between NI's for all protocols/features. So the message to make the

Standby port become active depends on the time taken for the message to reach from one NI to another. This has a definite increase in the amount of time taken than the message communication happening within the same NI.

One change that can be attempted is to pass the Join message for the standby port directly to 802.1Q on the NI. But this is a change from the current design of Link Agg and 802.1Q communication. The benefit could be a reduction of another ~100ms.

Convergence with NI Power Down

When the unit containing the active DHL port is powered down, the convergence time will be high. In the current Link Aggregation NI architecture, tokens are being used to decide the master and slave NIs. This ownership is required to synchronize events such as port join, port leave, aggregation up and aggregation down when the link aggregation exists across NIs.

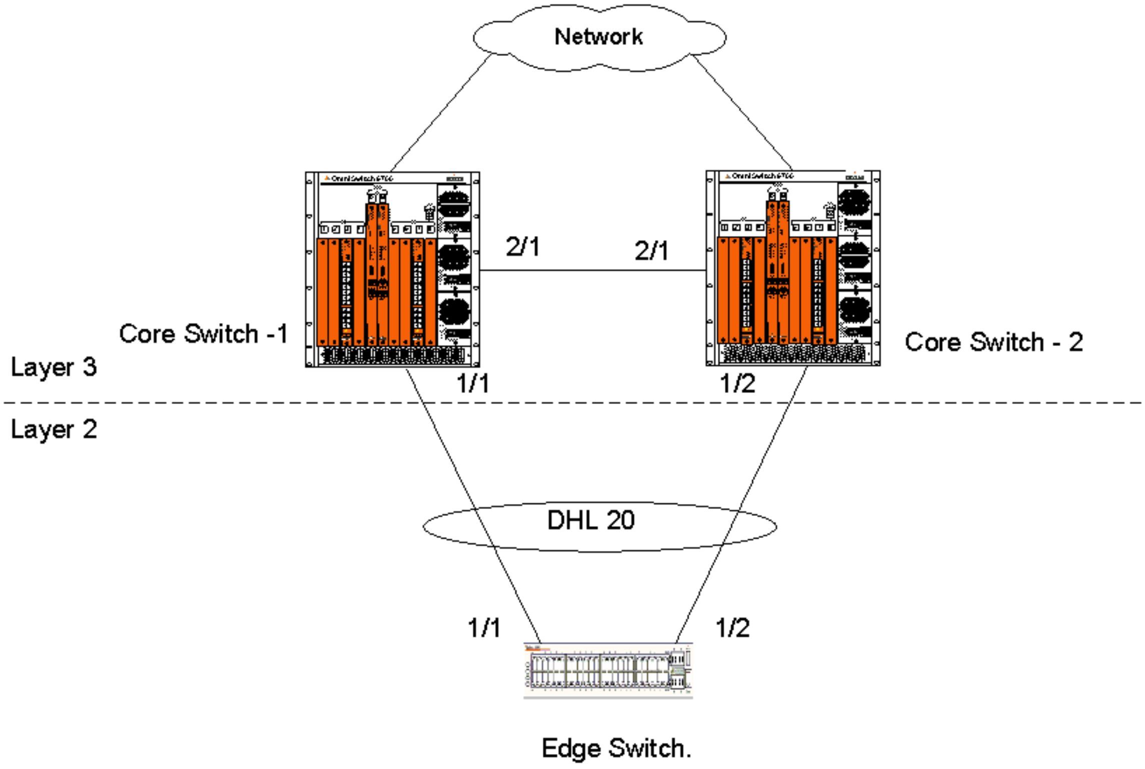
Any NI which is interested in declaring the events would fetch the token ownership from the master NI. Once the master NI relinquishes the token to the slave NIs, aggregation events are triggered.

If a port leave or a port join is needed due to one of the NIs going down, link aggregation NI sends a token request message to the failing NI. This is to make sure that the slave NI can now own the token and declare the port leave for the outgoing NI ports. This specific token request would not have any response, since the NI is already down. So only after the token expiry, would a port leave be done and a response to the NI down event. This timer expiry is 10 seconds. Note that these protocol timers hold these values since inception of 6.x release, or earlier. So the convergence time in this case is 12-14 seconds.

Configuration Examples

The following sample configurations were used to test Dual Home Link-Aggregation.

Example 1



Example 1 CLI Configuration

The following CLI commands provide an example of how to configure Example 1 for Dual Home Link-Aggregation.

Edge Switch

6. Create VLAN 2 with STP disabled.

```
→vlan 2 stp disable
```

1. Create an IP interface for VLAN 2.

```
→ip interface vlan-2 address 2.2.2.1/24 vlan 2
```

2. Create the link aggregate.

```
→lacp linkagg 1 size 2 actor admin key 1
→lacp agg 1/1 standby enable
→lacp agg 1/1 actor admin key 1
→lacp agg 1/2 actor admin key 1
```

3. Tag the DHL with VLAN 2.

```
→vlan 2 802.1q 1
```

Core Switch-1

1. Create VLAN 2 with STP disabled.

```
→ vlan 2 stp disabled
```

2. Create an IP interface for VLAN 2.

```
→ ip interface vlan-2 address 2.2.2.2/24 vlan 2
```

3. Create the link aggregate at the core.

```
→ lacp linkagg 21 size 1 actor admin key 1
→ lacp linkagg 21 actor system id xx:xx:xx:xx:xx:xx (where xx:xx:xx:xx:xx:xx is a
  mac address)
→ lacp lagg 1/1 actor admin key 1
→ lacp agg 1/1 actor system id xx:xx:xx:xx:xx:xx
```

4. Tag the DHL with VLAN 2.

```
→ vlan 2 802.1q 1
```

5. Tag the link between Core Switch-1 and Core Switch-2 with VLAN 2.

```
→ vlan 2 802.1q 2/1
```

Core Switch-2

1. Create VLAN 2 with STP disabled.

```
→ vlan 2 stp disabled
```

2. Create an IP interface for VLAN 2.

```
→ ip interface vlan-2 address 2.2.2.3/24 vlan 2
```

3. Create the link aggregate at the core.

```
→ lacp linkagg 21 size 1 actor admin key 1
→ lacp linkagg 21 actor system id xx:xx:xx:xx:xx:xx (where xx:xx:xx:xx:xx:xx is a
  mac address)
→ lacp lagg 1/2 actor admin key 1
→ lacp agg 1/2 actor system id xx:xx:xx:xx:xx:xx
```

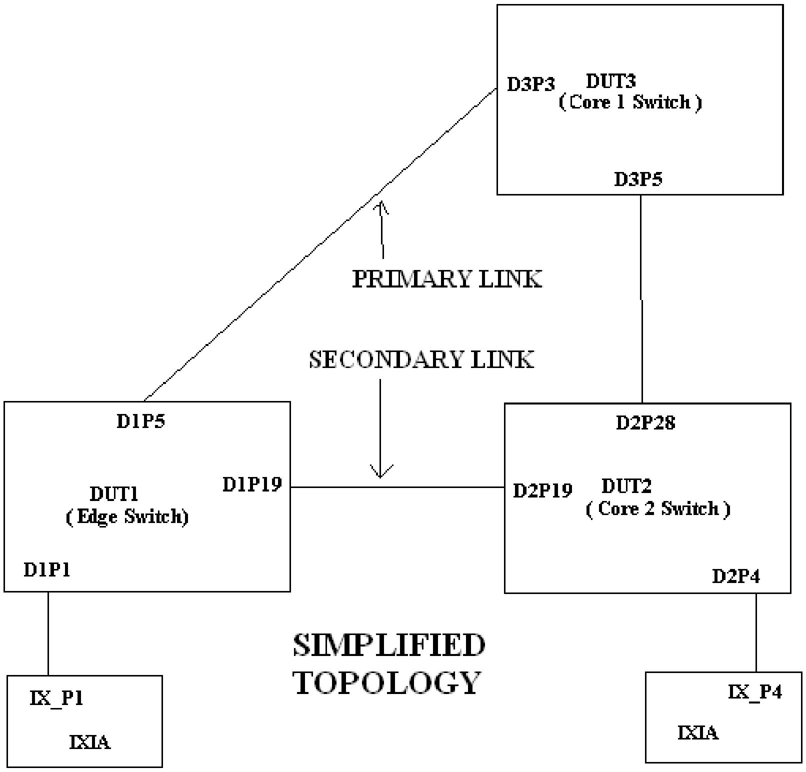
4. Tag the DHL with VLAN 2.

```
→ vlan 2 802.1q 1
```

5. Tag the link between Core Switch-1 and Core Switch-2 with VLAN 2.

```
→ vlan 2 802.1q 2/1
```

Example 2



Example 2 Configuration Guidelines

1. Use the same system ID for both Core 1 and Core 2 DUT. The MAC address used for the system ID can be any MAC address (even a dummy MAC) as long as the same address is used on both DUTs.
2. Multicast settings are enabled only when multicast traffic is required to flow across the linkagg. If only unicast and broadcast traffic is sent, then multicast settings can be skipped.

Example 2 CLI Configuration

The following CLI commands were used to configure the Dual Home Link Aggregation topology shown in the above Example 2 illustration:

Edge Switch (DUT 1)

```
→vlan <vlan_id> 1x1 stp disable flat stp disable name "<vlan_name>"
→vlan <vlan_id> port default <D1P1>
→ip interface "<interface_name>" address <interface_ip> mask <mask> vlan <vlan_id>
→ip multicast status enable
→ip multicast querying enable
→lacp linkagg <linkagg_id> size 2 admin state enable
→lacp linkagg <linkagg_id>  actor admin key <linkagg_key>
→lacp agg <D1P5> actor admin key <linkagg_key>
→lacp agg <D1P19> standby enable
→lacp agg <D1P19> actor admin key <linkagg_key>
→vlan <vlan_id> port default <linkagg_id>
```

Core 1 (DUT 3)

```
→vlan <vlan_id> 1x1 stp disable flat stp disable name "<vlan_name>"
→vlan <vlan_id>  port default <D3P5>
→ip interface "<interface_name>" address <interface_ip>  mask <mask> vlan <vlan_id>
→ip multicast status enable
→ip multicast querying enable
→lacp linkagg <linkagg_id>  size 1 admin state enable
→lacp linkagg <linkagg_id>  actor admin key <linkagg_key>
→lacp agg <D3P3> actor admin key <linkagg_key>
→lacp agg <D3P3> actor system id <DUMMY_MAC>
→vlan <vlan_id> port default <linkagg_id>
```

Core 2: (DUT 2)

```
→vlan <vlan_id> 1x1 stp disable flat stp disable name "<vlan_name>"
→vlan <vlan_id> port default <D2P4>
→vlan <vlan_id> port default <D2P28>
→ip interface "<interface_name>" address <interface_ip>  mask <mask> vlan <vlan_id>
→ip multicast status enable
→ip multicast querying enable
→lacp linkagg <linkagg_id> size 1 admin state enable
→lacp linkagg <linkagg_id> actor admin key <linkagg_key>
→lacp agg <D2P19> actor admin key <linkagg_key>
→lacp agg <D3P3> actor system id <DUMMY_MAC>
→vlan <vlan_id> port default <linkagg_id>
```

References

- Chapter 8, “Link Aggregation Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 10, “Configuring Dynamic Link Aggregation”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.4. Load Balancing Non-Unicast on a Link Aggregate

Hashing is an algorithm widely used by link aggregation to load balance traffic on link aggregate ports. The OmniSwitch hardware has a set of hash register settings with which this algorithm is calculated and applied to decide the destination egress on which packets are sent.

By default, non-unicast traffic is not load balanced but sent through the primary port for the link aggregate. The 6.4.3.R01 release introduces the ability to enable load balancing of non-unicast traffic (multicast, broadcast, and destination lookup failure packets) over a link aggregate.

Load balancing of non-unicast traffic is supported only with the normal hashing mode. As a result, the hashing computation of source IP and destination IP is used to determine the destination egress port for non-unicast traffic.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

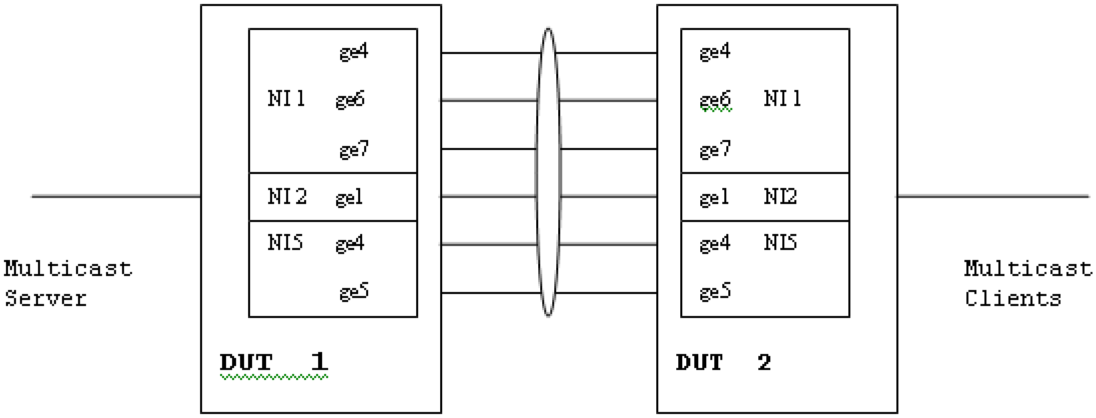
- The **hash-control** CLI command was extended to include a **loadbalance non-unicast** parameter to **enable** or **disable** load balancing of non-unicast traffic over a link aggregate. For example:

```
→ hash-control loadbalance non-unicast enable
→ hash-control loadbalance non-unicast disable
```

- Enabling or disabling (disabled by default) this feature is done on a global basis for the switch. As a result, load balancing of non-unicast traffic is applied to all link aggregates when the feature is enabled.
- This feature applies to all non-unicast traffic (for example, broadcast, L2 multicast, L3 multicast and unknown unicast).

Configuration Examples

The following sample configuration was used to test load balancing non-unicast traffic on a link aggregate:



The following **show linkagg port** command displays the Link Aggregate configuration in this example:

```
→ show linkagg 10 port

Slot/Port Aggregate SNMP Id   Status   Agg  Oper Link Prim
---------+---------+-------+----------+----+----+----+----
   1/5    Static      1005  ATTACHED    10  UP   UP   NO
   1/7    Static      1007  ATTACHED    10  UP   UP   NO
   1/8    Static      1008  ATTACHED    10  UP   UP   YES
   2/2    Static      2002  ATTACHED    10  UP   UP   NO
   5/5    Static      5005  ATTACHED    10  UP   UP   NO
   5/6    Static      5006  ATTACHED    10  UP   UP   NO
```

References

- Chapter 54, “Chassis Management and Monitoring Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).

2.5. MVRP

The Multiple VLAN Registration Protocol (MVRP) is defined in the IEEE standard 802.1ak-2007, an amendment to IEEE 802.1Q 2005. Intended as a replacement for GVRP, MVRP offers several enhancements and changes to address the scale and requirements of large bridged networks.

MVRP fills the role of GVRP and operates in a manner similar to GVRP, in that MVRP controls and signals dynamic VLAN registration entries across the bridged network. This implementation of MVRP provides the following improvements over GVRP:

- Improved PDU format to fit all 4094 VLANs in a single PDU, instead of 11 PDU for GVRP. See sections 10.8 and 11.2.3 in the standard for more detail.
- Reduced unnecessary flushing of dynamic VLANs triggered by STP topology changes that do not impact the Dynamic VLAN topology.

The 802.1ak standard defines the MRP, MMRP, and MVRP protocols. This implementation uses the MRP and MVRP protocols. MRP replaces GARP, and MVRP is an application of MRP, the same as GVRP is an application of GARP.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

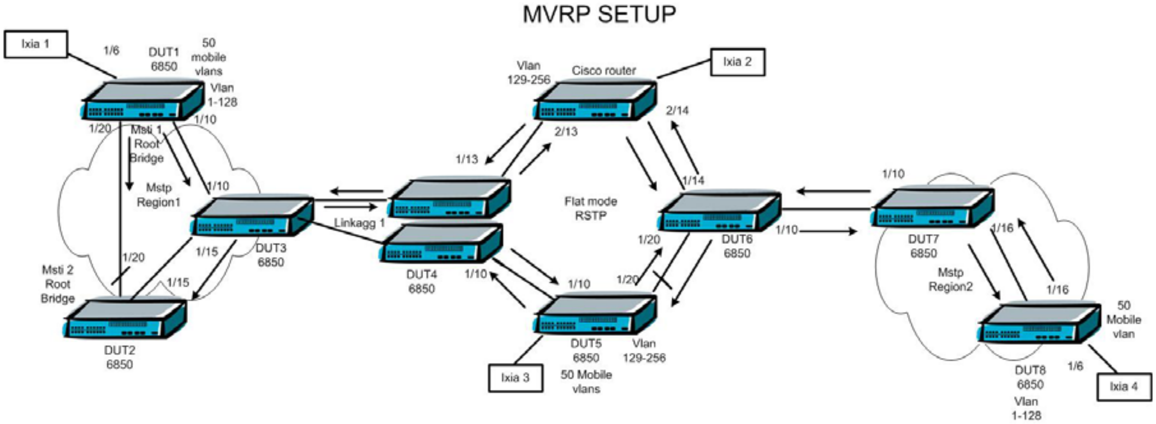
Guidelines

- MVRP is supported only in Flat mode RSTP and MSTP; 1x1 mode is not supported.
- By Default, MVRP will work in Applicant active mode in AOS System
- MVRP and GVRP are mutually exclusive; if the operational mode is changed from one protocol to another, the existing protocol configuration is automatically removed from the switch.
- When MVRP is enabled, GVRP frames are ignored; when GVRP is enabled, MVRP frames are ignored.
- MVRP is not supported on mirroring ports, mobile ports, aggregate ports, VLPS access ports and VLAN stacking user ports. Attempting to configure any of these features on an MVRP-enabled port is not allowed.
- MVRP is not supported on IPMVLANs
- MVRP processing is limited to static or dynamic VLAN(s) only on conventional ports
- Flushing for MVRP dynamic VLANs is controlled by MVRP and not by STP.
- PVID for a port is advertised in the MVRP topology. This behavior is different from GVRP in that PVID advertisement was restricted in the system.
- By default, all the customer MVRP frames received on VLAN Stacking ports will be flooded in hardware on the SVLAN.
- Management VLANs should be restricted on MVRP ports, otherwise there is a potential for an STP loop.
- When Group Mobility and MVRP are enabled on an Access switch, the uplink ports should also be tagged with static VLANs in addition to enabling MVRP. If the uplink ports are not tagged, unnecessary flushing and relearning of VLANs may occur.
- 256 Dynamic VLANs are supported in MVRP.
- STP Convergence with MVRP dynamic VLANs is not as fast as RRSTP.

Configuration Examples

The following sample configurations were used to test MVRP functionality and the tunneling of MVRP PDUs using VLAN Stacking Ethernet Services.

Example 1: MVRP Test Configuration



Example 1 Configuration

The test configuration in Example 1 was used to verify MVRP functionality between two endpoints. In this configuration:

- There are four Ixia end points: Ixia 1, Ixia 2 , Ixia 3 and Ixia 4.
- DUT 1, DUT 2, and DUT 3 are configured with MSTP in region 1.
- DUT 7 and DUT 8 are configured with MSTP in region 2.
- DUT4, DUT5, and DUT6 are running Flat mode RSTP.
- DUT1 will propagate 256 VLANs (1-256 including 50 mobile VLANs).
- Other switches, such as the Cisco router, will propagate 256 VLANs (257-512 including 50 mobile VLANs).
- DUT5 will propagate 256 VLANs (257-512 including 50 mobile VLANs).
- DUT 8 will also propagate 256 VLANs (1-256 Including 50 mobile VLANs).

Example 1 CLI Configuration

All switches are configured to run in STP flat mode and have the MVRP configuration as per the Example 1 topology.

Configure MVRP:

```
→mvrp port 1/10 restrict-vlan-registration vlan 172
→mvrp port 1/20 restrict-vlan-registration vlan 172
→mvrp port 1/6 enable
→mvrp port 1/10 enable
→mvrp port 1/20 enable
```

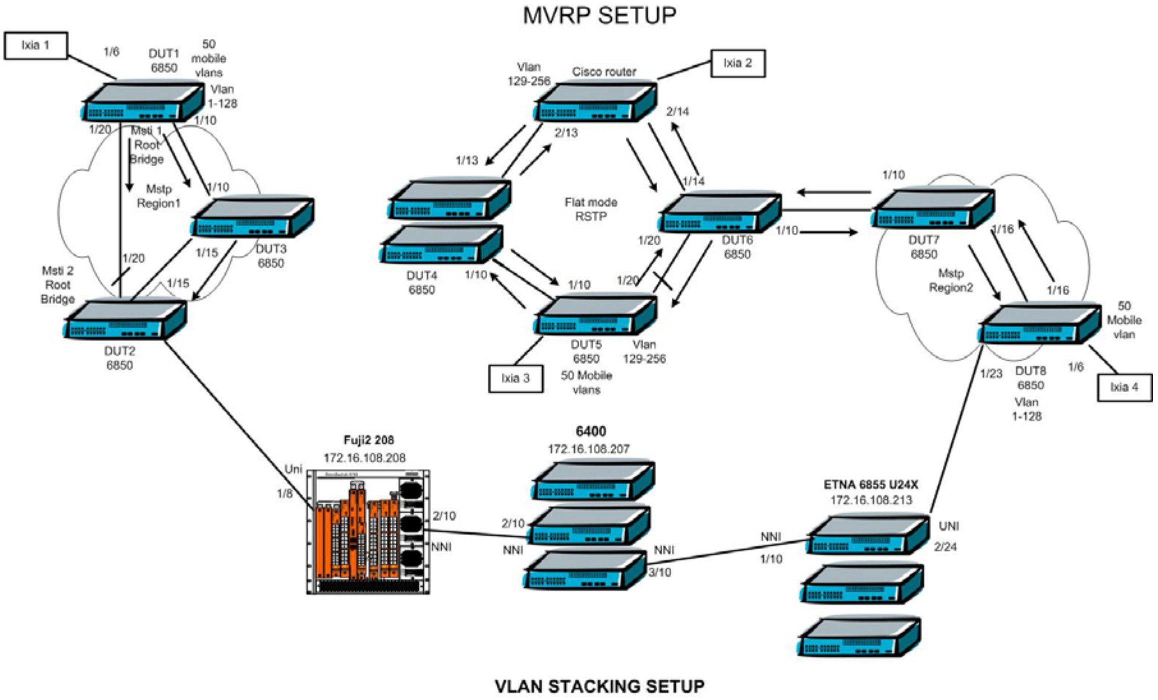
Configure MSTP Region 1 (DUT1, DUT2, DUT3)

```
→Bridge mode flat
→Bridge mst region name region_1
→Bridge mst region revision_level_1
→Bridge protocol mstp.
→Bridge msti 1
→Bridge msti 2
→Bridge msti 1 vlan 1-128
→Bridge msti 2 vlan 129-256
```

Configure MSTP Region 2 (DUT7, DUT8)

```
→Bridge mode flat
→Bridge mst region name region_2
→Bridge mst region revision_level_2
→Bridge protocol mstp.
→Bridge msti 1
→Bridge msti 2
→Bridge msti 1 vlan 1-128
→Bridge msti 2 vlan 129-256
```

Example 2: MVRP/VLAN Stacking Test Configuration



Example 2 Configuration

- The Example 2 test configuration includes a VLAN Stacking setup that was used to validate the tunneling of MVRP PDUs through Ethernet Services.
- By default, MVRP PDUs are tunneled by VLAN Stacking Ethernet Services. As a result, there is no need to specifically configure the service to tunnel MVRP PDU.

References

- Chapter11, “MVRP Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 13, “Configuring MVRP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.6. LLDP – VLAN Assignment

In the previous AOS implementation, AVA (Automatic VLAN Assignment) was used for assigning VLANs to IP phones (which is an Endpoint device). AVA implementation is often complicated as it requires two DHCP servers: one for VLAN assignment and one for IP address assignment. The "LLDP for IP Phone & OmniSwitch" feature attempts to provide a network-friendly solution to replace AVA by providing a LLDP MED Network Policy that allows the OmniSwitch to advertise the VLAN to the connected IP Phones.

In release 6.4.3.R01, only VLAN assignment through explicit LLDP MED Network Policy is supported. VLAN advertisements based on classification in voice UNP on 802.1x ports will be implemented in a future release.

When the IP phone (or any MED endpoint deice) connects to a port with an Explicit MED Network Policy configuration, the OmniSwitch advertises the policy in the LLDP PDU with MED Network TLVs encapsulated. The endpoint device will then configure itself according to the advertised policy. This advertisement will take place only if the transmission of the Network Policy TLV is enabled by the user.

The Fast restart (as described in IEEE 802.1ab rev) was implemented to transmit the LLDP-MED Network Policy TLV as soon as AOS detects a new MED endpoint.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

The LLDP feature was tested with the following software and IP phones:

- OXE version 9.1.
- IP phone models 4028 and 4038 EE.
- NOE version 4.25.51 on the IP phones.
- AOS Release 6.4.3.R01

Consider the following guidelines when configuring LLDP-MED Network Policies:

- If the user modifies the MED-Network policy on a specific port at runtime, the port needs to be disabled and enabled again to classify the port in the proper network policy VLAN.
- LLDP-MED Network Policy TLV advertisement is disabled by default. It must be enabled on the port, slot or chassis in order for LLDPDU TLVs to be transmitted via either the fast start mechanism or periodic transmission interval.
- Fast-start LLDPDU TLV transmission only occurs the first time a MED device is detected on a port. Fast-start will not occur again on that port unless the device is disconnected and connected again (either physically or with an admin down/up) or the device sends a TTL=0.

- The configuration for Network Policy is done explicitly through LLDP MED CLI. However, there is no corresponding read-write object in the standard MIB. As a result, a new proprietary MIB was defined for LLDP MED with the same objects as defined by the lldpXMedLocMediaPolicyTable.
- A maximum of 32 network policies (0 – 31) is allowed.
- For the feature to work on fixed ports, the port should be assigned to the advertised VLAN.
- The Group Mobility configuration for the VLAN and port should be in sync with the LLDP configuration for the feature to work on mobile ports.
- For LLDP phones to be operational on 802.1x ports, configure the following:
  - A voice-policy, using LLDP CLI commands, with voice VLAN (e.g., x) and priority (e.g., p) and bind the policy to the given 8021x port.
  - A User Network Profile, using AAA CLI commands, with the same VLAN (VLAN x) and configure the Radius server to return that UNP name.
  - Create a VLAN port association between the 802.1x port and VLAN x (for example, **vlan x 802.1q 1/10** or **vlan x mobile tag enable**). For 802.1x ports, a VPA for “voice” VLAN will be created by AAA when the IP Phone is classified in UNP.
- Traffic from IP phones is tagged with the voice VLAN ID.
- Traffic sent to IP Phone is tagged with the voice VLAN ID.
- Only topologies with one IP Phone per port are supported.

References

- Chapter12, “802.1AB Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 14, “Configuring 802.1AB”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.7. Recursive Static Route

Recursive static routes are similar to static routes. However, with a recursive static route the gateway does not have to be a directly connected to the router. If the OmniSwitch is unable to find a route in the routing table for a packet it can use the recursive static route. The OmniSwitch will use the routing table to lookup a route for the gateway instead of having to use a directly connected router. This feature can be used in large networks to configure a uniform static route for all routers on a network. Each router will use the same gateway but the path to reach the gateway may differ for each router.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- With recursive static routes, the nexthop (or gateway) address no longer must be tied to a particular interface. Instead, you can tie the destination route to the best route used to reach a particular host. This may be an interface or it may be a dynamically learned route (i.e. BGP, OSPF, RIP, etc) and it may change over time, with the recursive route switching to pick up the “best route's” gateway to use as its gateway.
- As long as there is a route to the address that was used in configuring the recursive route, the recursive destination route will be installed and remain active in the switch.
- Recursive static routes are differentiated from regular static routes by substituting the keyword “*follows*” for the keyword “*gateway*” in the “*ip static-route*” command. Example: *ip static-route 2.0.0.0/8 follows 129.2.2.3*
- Only IPv4 static routes can be installed as recursive routes.
- A BFD-enabled route cannot be installed as a recursive route.
- A maximum of 32 recursive static routes can be configured per VRF. (This means inactives and backups are included in the count.). So OS6855-U24X can have a max of 256 routes (8VRF’s * 32) and OS9000E will support a max of 2048 routes (64 VRF’s * 32).
- If the IP address used in configuring a recursive route becomes unreachable, the recursive static route also becomes unreachable. In this case, it transitions to the inactive list. (Later, if a route to reach the IP address is activated, the recursive route will be re-activated, possibly using a new gateway/nexthop address.)
- Recursive routes can be configured as ECMP routes. A maximum of 16 ECMP recursive static routes are supported. For example, a user types the following two commands:

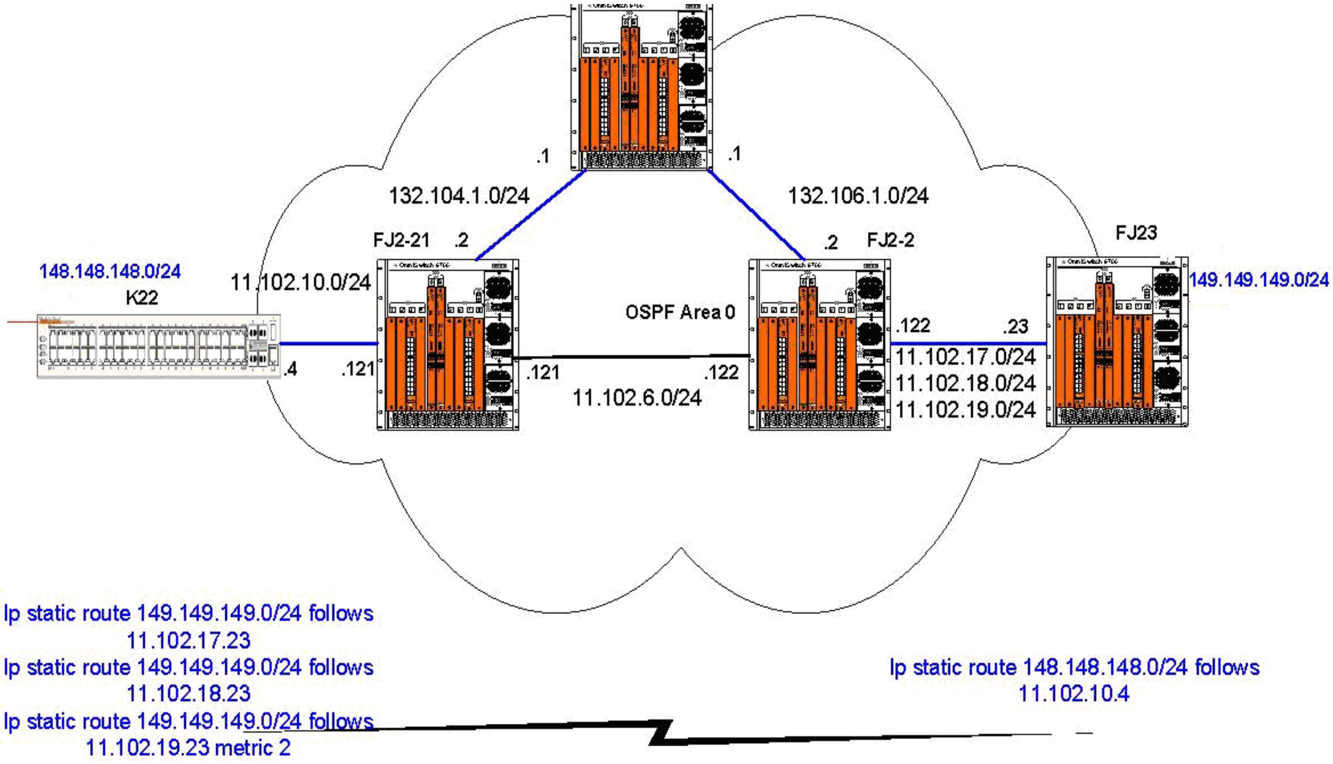
```
→ ip static-route 10.0.0.0/8 follows 2.0.0.22
→ ip static-route 10.0.0.0/8 follows 3.0.0.33
```

  In this example, host 2.0.0.22 is reached via RIP route 2.0.0.0/8 using gateway 22.22.22.1, and host 3.0.0.33 is reached via RIP route 3.0.0.0/8 using gateway 33.33.33.1. Thus, the recursive route 10.0.0.0/8 will be installed as two ECMP routes using the gateways 22.22.22.1 and 33.33.33.1, respectively.

- Duplicate routes are not allowed in the system. To continue with the previous example (item 12), if both host 2.0.0.22 and host 3.0.0.33 are reached using the same gateway address 4.0.0.1, only one recursive route will be installed. In this case the route 10.0.0.0/8 will be installed using the gateway address 4.0.0.1, with one of the host following addresses listed. The other recursive route, using the other following host address, will end up on the inactive list.
- If a recursive route maps to an existing ECMP route chain, the recursive route will install one copy of itself using the first route in that ECMP chain as its gateway/nexthop. If the ECMP route it uses becomes

inactive, it will transition to the next one in the ECMP chain. This transition happens until the last gateway in the ECMP chain is active. Recursive routes transitions to inactive state when none of the ECMP gateways are active.

Configuration Example



In this configuration:

- Static route to network 149.149.149.0 is required but the administrator does not know which router to assign as the next hop, or wants the assignment to be able to change from one router to another.
- Destination Network 149.149.149.0 is always reachable through another network 11.102.17.0.
- Recursive static route for 149.149.149.0 with 11.102.17.23 as a next hop can be configured.
- When the dynamic routing protocols such as RIP, OSPF, and BGP add a route to 11.102.17.0 and assign the next hop as 132.104.1.1, the routing stack knows that it can add a route for 149.149.149.0 with the next hop as 132.104.1.1

The following CLI **show** commands display the switch configuration as described in this recursive static route example:

```
→ show configuration snapshot ip-routing
ip static-route 149.149.149.0/24 follows 11.102.17.23 metric 1

→ show ip route

  + = Equal cost multipath routes
  * = BFD Enabled static route
  Total 5 routes
  Dest Address    Subnet Mask      Gateway Addr    Age      Protocol
---------------+-----------------+--------------+---------+-----------
  11.102.17.0    255.255.255.0      132.104.1.1   00:00:29   OSPF
  11.102.18.0    255.255.255.0      132.104.1.1   00:00:29   OSPF
  11.102.19.0    255.255.255.0      132.104.1.1   00:00:29   OSPF
  132.104.1.0    255.255.255.0      132.104.1.2   00:00:12   LOCAL
  149.149.149.   255.255.255.0      132.104.1.1   00:00:08   NETMGMT

→  show ip router database
Legend: + indicates routes in-use
        * indicates BFD-enabled static route
        r indicates recursive static route, with following address in brackets

Total IPRM IPv4 routes: 5

Destination          Gateway       Interface    Protocol  Metric Tag   Misc-Info
------------------+--------------+------------+---------+-------+----+-----------
+  11.102.17.0/24    132.104.1.1   vlan-20      OSPF      2       0
+  11.102.18.0/24    132.104.1.1   vlan-20      OSPF      2       0
+  11.102.19.0/24    132.104.1.1   vlan-20      OSPF      2       0
+  132.104.1.0/24    132.104.1.2   vlan-20      LOCAL     1       0
+r 149.149.149.0/24  132.104.1.1   vlan-20      STATIC    1       0   [11.102.17.23]


Inactive Static Routes
   Destination       Gateway           Metric
--------------------+-----------------+---------
```

References

- Chapter14, “IP Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 16, “Configuring IP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.8. Extended Ping and Traceroute Functionality

The extended **ping** and **traceroute** applications are used to perform a more advanced check of host reachability and network connectivity. The extended **ping** command is used to determine the type of connectivity problem, then the extended **traceroute** command is used to narrow down where the problem is occurring.

The extended **ping** has the following options:

- Source IP - Source interface whose IP address is to be used as source IP for the ping packets.
- TOS value - Specifies the Type of Service field in the IP header.
- DF bit in IP header - Sets the fragment bit to 1 in the IP packet
- Data pattern - Specifies the data pattern to be used the data field of the ping packets
- Sweep range - This contains 3 parameters: *start_size* specifies the size in bytes of the first packet to be sent, *diff_size* mentions the increment factor of size for the next packet and *end_size* specifies the maximum size of the packet.

The extended **traceroute** has the following options:

- Source IP - Specifies the source IP interface to be used in the traceroute packets
- Timeout value - The time in seconds to wait for the response of each probe packet
- Probe count - The number of packets (retry) that will be sent for each hop-count
- Minimum and Maximum Hops - The minimum number of hops to be set for the first packet and the maximum number of hops for the destination address
- Port number - The destination port number to be used in the probing packets

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- The source address can also be the loopback0 address.
- The source address must be reachable from the destination.
- The source IP address and destination IP address must be in the same VRF instance.
- The port number in traceroute must be greater than 1024.

References

- Chapter14, “IP Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 16, “Configuring IP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.9. Egress Policy Support

OmniSwitch egress policy rules allow administrators to enforce traffic controls on the egress queues as a "last resort" action. By default, QoS policy rules are applied to traffic ingressing the port. The QoS Policy List feature includes an "egress" policy list option to create a list of rules that are applied to traffic egressing a destination port(s). If a policy rule is not associated with an egress policy list, the rule will only apply to ingress traffic.

Platforms Supported

OmniSwitch 6400, 6855-U24X, 9000E

Guidelines

- Refer to the 6.4.3R01 Network Configuration Guide for supported policy conditions and actions for Egress rules.
- Maximum TCAM entries:
- OS9000E has 1022 TCAM available for Egress policy rules.
- OS6400 and OS6855-U24X have 510 TCAM available for Egress policy rules.
- Total number of Policy rules per DUT is 2048, which includes both ingress policies and egress policies.

Egress policy condition guidelines:

- Source port group is not supported for egress policy condition.
- IPv6 conditions are not supported for egress policy.
- Egress policy with TCP/UDP port range will consume multiple TCAM entries. However software tries to convert TCP/UDP range into maskable range to reduce TCAM usage.

Egress policy action guidelines:

- EFP doesn't have action to assign internal priority/CoS hence it would not be supported.
- Policy based routing, redirect to port/LinkAgg and Policy based mirroring are not supported using Egress policy rule.
- srTCM/trTCM metering are not supported for egress policy condition.

Other software guidelines:

- There is no support for destination/egress Linkagg for Egress policy. This behavior is similar to Ingress policy. The workaround is to create a port group of all the ports that are part of a Linkagg and apply a single rule.
- Logging of policy rule is not supported.
- XNI-U12E based NI sometimes doesn't increment the policy match counter.

Configuration Examples

```
→policy condition c1 destination port 4/1 inner source vlan 10
→policy action a1 maximum bandwidth 512k
→policy rule r1 condition c1 action a1 no default-list
→policy port group g2 7/5 7/6
→policy condition c2 destination port group g2 inner source vlan 20
→policy action a2 maximum bandwidth 50.00M
→policy rule r2 condition c2 action a2 no default-list
→policy list egress-list  type egress rules r1 r2
→qos apply
```

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, and Chapter 30, “Configuring ACLs”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.10. Tri-Color Marking (srTCM/trTCM) Support

The srTCM/trTCM rate-limiting provides the ability to have bandwidth guarantee through Committed Information Rate (CIR) configurations. TCM polices network traffic by limiting the rate at which traffic is sent or received on a switch interface. The TCM policer meters traffic based on user-configured packet rates and burst sizes and then marks the metered packets as green, yellow, or red based on the metering results.

This implementation complies with RFC 2697 and RFC 2698.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

The TCM Meter operates in Color-Blind mode (the Color-Aware mode is not supported). In the Color- Blind mode, the Meter assumes that the incoming packet stream is uncolored. However incoming packets with the CFI/DEI bit set are automatically given an internal lower priority.

There are two types of TCM marking supported:

- **Single-Rate TCM (srTCM) according to RFC 2697**—Packets are marked based on a Committed Information Rate (CIR) and two associated burst size values: Committed Burst Size (CBS) and Peak Burst Size (PBS).
- **Two-Rate TCM (trTCM) according to RFC 2698**—Packets are marked based on a CIR value *and* a Peak Information Rate (PIR) value and two associated burst size values: CBS and PBS.

Both srTCM and trTCM handle the burst in the same manner. The main difference between the two types is that srTCM uses one rate limiting value (CIR) and trTCM uses two rate limiting values (CIR and PIR) to determine packet marking.

QoS manager treats the different colored packets as follows:

- Packets violating the SLA (Red color packets) are dropped.
- Packets exceeding the SLA (Yellow color packets) are given the high drop precedence and dropped during the congestion at the egress side.
- Packets conforming to the SLA (Green color packets) are allowed as is.

Following are the maximum Tri Color Marker (TCM) rules allowed on different platforms. If the user tries to configure more than the maximum allowed, *ERROR: Out of TCAM processors on X/0(0)* will appear on screen.

Chipset	Slices	Available Slices	Num entries per slice	Counter per slice	Meter Per slice	TCM Rules (available Slice * meter)/ 2
OS6850 OS6855	16	12	128	128	128	768
OS6400	8	4	256	256	256	640
OS6855-U24X OS9000E	16	12	512	512	512	3072
XNI-U12E	12	8	128	128	128	512

Other Guidelines

- Color aware mode is not supported.
- Exceed-action is not configurable. Red packets are always dropped.
- srTCM and trTCM is not supported for egress policy rule list.
- This release does not support srTCM and trTCM in the Ethernet Service SAP Profile.
- In srTCM, Yellow packet marking only happens in the first burst. For continuous burst, packets will not be marked yellow in srTCM.

Configuration Examples

```
→ policy network group g1 10.10.10.2 10.10.10.3 10.10.10.60
→ policy condition c1 source ip 10.10.10.1 dscp 24
→ policy condition c2 source network group g1 tos 3
→ policy action a1 CIR 10.0M CBS 20.0M PIR 15.0M
→ policy action a2 CIR 15.0M CBS 30.0M PIR 20.0M
→ policy rule r1 condition c1 action a1
→ policy rule r2 condition c2 action a2
→ qos port 3/9 dei egress
→ qos apply
```

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.11. Ingress Policy Rule Statistics Enhancements

This QoS enhancement provides green, yellow and red counters. These counters will provide information about the traffic which is conformant to the SLA, not conformant to SLA and violating the SLA.

The Tri-Color Marking (TCM) policy action now includes a counter color mode option. This option determines which metered packets are counted based on the color the packet was marked by the TCM policy. Enabling this option also allows the display of the counter color statistics using existing QoS **show** commands.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E.

Guidelines

The **policy action cir** command includes a **counter-color** parameter that is used to configure color-based statistics for packets marked by TCM. By default, the counter color mode for a TCM action is set to count red and yellow packets (green packets are not counted).

To change this mode, use the **counter-color** parameter. For example:

```
 → policy action A4 cir 10m cbs 4k pir 20m counter-color green-nongreen
```

This command sets the counter color mode to count all green and non-green (red and yellow combined) packets. The following color combination options are supported with the **counter-color** parameter:

- **green-red**—counts the number of packets marked green (low drop precedence) and the number of packets marked red (packet is dropped). Packets marked yellow (high drop precedence) are not counted.
- **green-yellow—**counts the number of green and yellow packets. Red packets are not counted.
- **red-yellow**—counts the number red and yellow packets. Green packets are not counted.
- **red-nonred—**counts the number of red and non-red (yellow and green) packets.
- **green-nongreen**—counts the number of green and non-green (yellow and red) packets.

The **show active policy rule meter-statistics** command is used to view color-based packet counts generated by a TCM policy rule. All of the color combinations are displayed with this command, however, statistics only show for the combination that was selected for the TCM policy action.

Configuration Examples

```
 → policy condition c1 destination ip 10.10.10.10 source port 3/24 destination tcp
   port 112
 → policy condition c2 destination ip 10.10.10.10 source port 1/9 destination tcp
   port 120-150
 → policy action a1 CIR 10.0M CBS 20.0M PIR 15.0M counter-color RED-NONRED
 → policy action a2 CIR 15.0M CBS 30.0M PIR 20.0M counter-color RED-NONRED
 → policy rule r1 condition c1 action a1
 → policy rule r2 condition c2 action a2
 → qos port 2/19 dei egress
 → qos apply

 → show active policy rule meter-statistics

Policy:r1,  Counter Color Mode:RED_NONRED
  Green    :                       -,   Non-Green:                         -,
  Red        :             844372,   Non-Red   :             155627,
  Yellow   :                  -
  Policy:r2,  Counter Color Mode:RED_NONRED
    Green     :                         -,   Non-Green:                        -,
    Red        :              792500,   Non-Red   :              207499,
    Yellow   :                   -
```

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.12. 802.1ad DEI/CFI Bit Support

When the sr/trTCM ingress rate limiter is used, frames that are non-conforming to the SLA (yellow) might still be delivered to the egress port when the port is not congested. By enabling CFI/DEI bit stamping on these frames, a color-aware upstream switch would be able to treat these frames differently and drop them first when the network is congested.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E.

Guidelines

By default, DEI bit marking (egress) and mapping (ingress) is disabled on all switch ports. The DEI bit setting operation may be configured globally on the switch, or on a per-port basis.

To configure the global DEI bit setting operation to mark traffic egressing on QoS destination ports, use the **"qos dei"** command with the **egress** parameter option. For example:

```
→qos dei egress
```

To configure the switch to map ingress traffic marked with the DEI bit, use the **qos dei** command with the **ingress** parameter option. For example:

```
→qos dei ingress
```

To configure the DEI bit operation for an individual port, use the **qos port dei** command with the **ingress** or **egress** parameter option. For example:

```
→qos port 1/10 dei egress
→qos port 1/11 dei ingress
```

Note that the CFI/DEI marking is applicable only for the outer VLAN tag.

References

- Chapter 32, "QoS Commands", and Chapter 33, "QoS Policy Commands", *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, "Configuring QoS", *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.13. Egress Port/Queue QoS and Statistics Enhancements

QoS statistics monitoring allows the gathering of egress CoS drop and transmit packet statistics for individual ports. Enabling this type of monitoring also allows the user to display egress CoS queue statistics on a per port basis using existing QoS show commands.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- Enabling QoS port monitoring is required to capture statistics on a per port basis, except on OS9000E ports.
- The OS9000E automatically gathers egress CoS statistics on a per port basis, so enabling QoS port monitoring is not required or supported.
- Only one port per slot can be configured to monitor CoS statistics. This restriction does not apply on the OS9000E because all ports on this switch gather egress CoS statistics.

Configuration Examples

```
→ qos port 2/19 monitor

→ show qos queue 2/19
Slot/      Q         Bandwidth        Packets
Port  VPN  No Pri Wt Min   Max        Xmit Drop      Type
-----+----+--+---+--+-----+-----+---------+---------+----
2/19   82   0   -  1     -     -         0         0  WRR
2/19   82   1   -  1     -     -      8925     91074  WRR
2/19   82   2   -  1     -     -         0         0  WRR
2/19   82   3   -  1     -     -      8929     91071  WRR
2/19   82   4   -  1     -     -         1         0  WRR
2/19   82   5   - 10     -     -     87116    112882  WRR
2/19   82   6   -  1     -     -         0         0  WRR
2/19   82   7   -  1     -     -         0         0  WRR
Total Xmit Packets:    104971,
Total Drop Packets:    295027
```

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, and Chapter 30, “Configuring ACLs”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.14. Policy Condition Enhancements

The following policy condition enhancements are now available in the 6.4.3.R01 release:

- VLAN IDs can be grouped together into a single VLAN group. Similar to other QoS group types, such as MAC and port groups, creating a VLAN group avoids having to configure a separate policy condition for multiple VLAN IDs.
- Specifying a range of 802.1p values for a policy condition is now supported. A range of values is supported when configuring both inner and outer 802.1p policy conditions. A condition must use either a single 802.1p value or a range of 802.1p values; both are not supported at the same time.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- Maximum number of VLAN groups supported is 1024, the same limit that applies to other policy groups.
- The VLAN range and 802.1p range command may program into multiple TCAM entries if it cannot fit into the calculated maskable range.

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, and Chapter 30, “Configuring ACLs”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.15. Map Several Inner DSCP/ToS Values to Same Outer 802.1p

The ability to specify a range of 802.1p values is particularly useful when classifying Ethernet Services SAP traffic. A new option in a SAP profile suspends the use of SAP bandwidth and priority actions. This allows the use of QoS rules for advanced classification of SAP traffic, such as mapping several DSCP/ToS values to the same outer 802.1p value.

When the priority is configured in the Ethernet-service sap-profile, user can not override the priority assignment with a policy rule. This is required when the user wants specific DSCP or inner 802.1p to outer 802.1p/priority mapping. By adding “not-assigned” priority option in the sap-profile configuration, the user can configure its own set of policy rules for priority mapping.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E.

Guidelines

A new parameter option, **priority not-assigned**, was added to the **ethernet-service sap-profile** command to prevent the profile from triggering QoS allocation of switch resources. When a profile is created using this parameter, QoS policy rules/ACLs are then available to define more custom priority settings for profile traffic. For example, mapping several inner DSCP/ToS values to the same outer 802.1p value.

Configuration Examples

```
→ethernet-service svlan 1001 nni 1/1
→ethernet-service svlan 1001 nni 4/5
→ethernet-service svlan 1002 nni 1/1
→ethernet-service svlan 1002 nni 4/5
→ethernet-service sap-profile "sp1" priority not-assigned
→ethernet-service service-name "VLAN1001" svlan 1001
→ethernet-service sap 1 service-name "VLAN1001"
→ethernet-service sap 1 uni 1/6
→ethernet-service sap 1 cvlan 2000
→ethernet-service service-name "VLAN1002" svlan 1002
→ethernet-service sap 2 service-name "VLAN1002"
→ethernet-service sap 2 sap-profile "sp1"
→ethernet-service sap 2 uni 1/6
→ethernet-service sap 2 cvlan 2001
```

References

- Chapter 32, “QoS Commands”, and Chapter 33, “QoS Policy Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 29, “Configuring QoS”, and Chapter 30, “Configuring ACLs”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.16. DHCP Option-82 ASCII Enhancement

The DHCP Snooping Option-82 feature has been enhanced to allow the configuration of a flexible ASCII string for the Circuit-ID value.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- The circuit level properties include the following:

  Base MAC
  System name
  User string
  VLAN
  Slot/port
  Port alias

  The VLAN, slot/port, and port-alias properties refer to the port where the DHCP Discover/Request packet was received.

- In format type ASCII, a maximum of five fields from a predefined set of six different circuit level properties can be copied to the circuit ID of the option82 field in the DHCP packet (Bootp request). At least one field should be configured. When there is more than one field, configuring a delimiter character is required. The delimiter can be any of the following characters:

  “|”, “-“, “_”, “/”, “\” and **“ “(space).**

- Once the DHCP Snooping Option-82 format is configured as ASCII, an ASCII string will be populated based on the configured fields with the delimiter. In addition, the remote ID field is no longer configurable and is automatically populated with the DHCP host name (Option-12 field).

- Configuring the Circuit ID suboption in ASCII format allows up to five fields (types) of information within the ASCII string. However, if the contents of all the fields combined exceeds 127 characters, then the ASCII string is truncated.CLI Commands and Examples

CLI Command and Examples

The following command is an extension of the existing DHCP Option-82 command that now includes an ASCII formatting option:

ip helper dhcp-snooping option-82 format ascii {base-mac | system-name | vlan | user-string *string* **| interface-alias | auto-interface-alias} {delimiter** *character***}**

The following command example shows how this command is used to configure the ASCII string:

```
→ ip helper dhcp-snooping option-82 format ascii user-string “Bldg. B Server” base-
  mac system name vlan interface-alias auto-interface-alias delimiter |
```

The **show ip helper** command displays the Option-82 ASCII configuration:

```
→Show ip helper
Ip helper:
Forward Delay(seconds) = 3,
Max number of hops = 4,
Relay Agent Information = Disabled,
DHCP Snooping Status = Switch-Level Enabled,
Option 82 Data Insertion Per Switch = Enabled,
MAC Address Verification Per Switch = Enabled,
DHCP Snooping Bypass Opt82-Check = Disabled,
DHCP Snooping Opt82 Format = ASCII,
DHCP Snooping Opt82 String = 00:d0:95:ae:3b:f6 system 4 1/2 hostname,
DHCP Snooping Binding DB Status = Enabled,
Database Sync Timeout = 300,
Database Last Sync Time = Mar 19 2007 14:32,
PXE Support = Disabled,
Forward option = standard
Vlan Number NA
Bootup Option Disable
Forwarding Address :
1.1.1.1
21.2.2.10
172.19.4.1
```

References

- Chapter 20, “DHCP Relay Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).

2.17. MAC-Forced Forwarding (Dynamic Proxy ARP)

MAC-Forced Forwarding (Dynamic Proxy ARP) is a mechanism to ensure the L2 separation of stations in the same VLAN beyond the local switch. The current port mapping functionality is limited to isolate user ports in the same switch. With MAC-FF the capability is extended to shared topologies such as rings or daisy chains to prevent users from communicating directly and ensuring that all communication happens via their default gateway. In order to accomplish this, the OmniSwitch supports Dynamic Proxy ARP which combines the functionality of port mapping and dhcp-snooping to dynamically learn a router's addresses and act as a local arp proxy for the VLAN's router. Dynamic Proxy ARP - MAC Forced Forwarding uses the following features:

- **Port Mapping -** Port Mapping forwards traffic from user-ports only to network-ports, preventing communication between L2 clients in the same VLAN in the same switch. This prevents direct communication between clients in the same VLAN forcing all traffic to be forwarded to the head end router.
- **Dynamic Proxy ARP -** All ARP requests received on port mapping user-ports are answered with the MAC address of the head end router. Dynamic Proxy ARP dynamically learns the IP and MAC address of a head end router and responds with that router's MAC address instead of flooding the ARP request.
- **DHCP Snooping -** Snoops the DHCP packets between the server and clients. DHCP snooping is used to dynamically learn the IP address of the head end router.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- When Dynamic Proxy ARP (DPA) and DHCP Snooping are enabled, the access router IP for the different VLANs will be collected from the DHCPACK packets sent to the clients on those VLANs.
- When any of the L2 clients (identified using the port mapping configurations) in that VLAN sends out an ARP request for the router IP, the same would be flooded on all the network ports with the MAC and IP of the L2 client.
- When the ARP reply from the access router is received, the MAC information of the access router would be stored. For subsequent ARP requests from the L2 clients to the access router or to any other L2 client in the same VLAN/subnet, an ARP reply with the MAC address of the access router will be sent. This would ensure that all traffic from the L2 clients are forwarded through the access router.
- When the access router MAC is yet to be resolved and one of the L2 clients sends out an ARP request for the other clients in the VLAN/subnet, the ARP packet would get dropped.
- DHCP snooping is required and must be enabled on the VLAN level for the Dynamic Proxy ARP feature to work.
- The switch does not support static configuration of router IP and related ARP information.
- Works in pure L2 environment, no IP interface is configured in any of the ports in DPA enabled port mapping session.
- Dynamic Proxy ARP works only in the default VRF instance.
- When DHCP Snooping is enabled on the fly, the feature would work based upon the next DHCP ACK received for a client. This is because the router IP will not be obtained unless the client requests for the IP through DHCP. When a DHCP ACK comes for this request, the router IP would be learned. Only after this the feature will start working.

- When DHCP binding database is used in sync with persistency, DHCP bind table will not get deleted and the client will not lose the IP info. So, there won't be any new DHCP ACK in the network. This is possible on the network where L2 device and hosts are connected with a HUB.
- ARP handling on User Ports is not in sync with RFC 4562. The RFC specifies differentiation of ARP request generated from one premise to another. This will not be done as there is no tracking per user level. Therefore, all ARP requests from User Ports would be trapped and responded by an L2 Dynamic Proxy ARP enabled switch.
- When router IP is not resolved, all the ARP requests from User Ports for the router IP/any other client will be dropped.
- When a takeover happens, UDP relay will not have any router IP information, the same has to be learned later through DHCKACK packets. However, the router IP information present in IPNI would be passed on to IPEDR and dynamic proxy ARP functionality for the existing clients would work.
- When a router IP is modified for a VLAN, the packets from the clients will be dropped until the ARP is resolved for the new router IP.

References

- Chapter 14, "IP Commands", *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 16, "Configuring IP", *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.18. VLAN Stacking –Tunneling L2 Protocols

The Ethernet Services (VLAN Stacking) UNI profile feature allows the transparent tunneling of L2 control frames. However, this feature cannot be deployed in a network where:

- The provider switches process such frames.
- The remote access switch is not an OmniSwitch and does not support transparent tunneling.

To overcome this limitation, the 6.4.3.R01 release provides enhancements to the UNI profile to support a MAC tunneling action that will change the destination MAC address of the L2 control frames to a unique tunnel MAC address.

In addition to MAC tunneling from UNI to NNI, this feature also includes a de-tunnel operation that is performed when the MAC-tunneled L2 control frame is received on a NNI port; the destination MAC address is changed back to the functional MAC address of the L2 control frame.

By default, a global tunnel MAC address of 01:00:0C:CD:CD:D0 is used for the MAC-tunnel actions. The default tunnel MAC address can be changed on a per-UNI profile basis.

The UNI profile enhancements in this release also add support for the following protocols:

- MVRP
- 802.3ad (LACP)
- 802.3ah (OAM)
- LACP Marker
- Cisco PAPG, CDP, DTP, VTP, PVST, VLAN BRIDGE and FAST UPLINK

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000

Guidelines

- The feature is only supported on VLAN Stacking UNI ports.
- The existing **ethernet-service uni-profile** command was updated to include parameters for the additional protocols now supported and parameters to configure the tunnel MAC operation. See Chapter 44, “VLAN Stacking Commands”, in the *OmniSwitch CLI Reference Guide* for more information.
- MAC tunneling is only supported in software, whereas the discard and tunnel actions (i.e. transparent tunneling) are supported in hardware. However, when L2 protocols that have the same functional MAC address (802.3ad/802.3ah, Cisco protocols) are set with different actions, the processing of such protocols will always be in software. For example, when 802.3ad is set to tunnel and 802.3ah is set to discard, the tunnel and discard is done in software.
- The MAC tunnel action (same for tunnel action) allows flooding of the L2 control frame from UNI to UNI and NNI to NNI.
- Overall software processing is limited to 512 packets per second for tunneling and de-tunneling L2 protocol packets.
- Rate limiting is done only for packets ingressing on a UNI, so if more than 1000 control packets are sent (512 on UNI and another 512 on NNI) the CPU would spike.
- No port-shutdown when the receiving rate of a L2 protocol on a port exceeds the rate limit.

- The Cisco protocols can only be tunneled or discarded; no peer action. The exception is UDLD.
- Configuring a tunnel MAC that is within a multicast MAC range is not allowed.
- The software tunneling does not provide any statistics. This can be achieved by policy rules.

Protocol Identification Table

Protocol	MAC address	Ether type	LLC DSAP/SSAP	Additional Info Protocol Identifier	Comment
STP	0180c2000000	n/a	0x42	STP: 0 RSTP: 2 MSTP: 3 GVRP: 1	
GVRP	0180c2000021				
AMAP	0020da007004				
802.3ad (LACP) Include 802.3ah	0180c2000002	0x8809	n/a		802.3AD: Subtype = 1 802.3AH: Subtype =3
802.1x	0180c2000003	0x888e	n/a		
802.1ab (LLDP)	0180c200000e	0x88cc	n/a		

New Protocols Supported

Protocol	MAC address	Ether type	LLC DSAP/SSAP	Additional Info (PID)	Comment
Cisco PAPG	01000ccccccc	n/a	0xAA	0x0104	Port link Aggregation Protocol
Cisco UDLD				0x0111	
Cisco CDP				0x2000	Discovery Protocol
Cisco DTP				0x2004	Dynamic Trunk Protocol
Cisco VTP				0x2003	Vlan Trunk Protocol
Cisco PVST	01000ccccccd	n/a	0xAA	0x010b	
Cisco Vlan Bridge	01000ccdcdce	n/a	0xAA	0x010c	Not supported
Cisco Uplink Fast	01000ccdcdcd	n/a	0xAA	0x200a	Not supported
802.3ad (LACP)	0180c2000002	0x8809	n/a		Subtype = 1
LACP Marker	0180c2000002	0x8809	n/a		Subtype = 2 Not supported (included with LACP)
802.3ah(OAM)	0180c2000002	0x8809	n/a		Subtype = 3

Configuration Examples

The following example configures the “U2” profile with a tunnel MAC address and specifies the L2 protocols to tunnel using this MAC address:

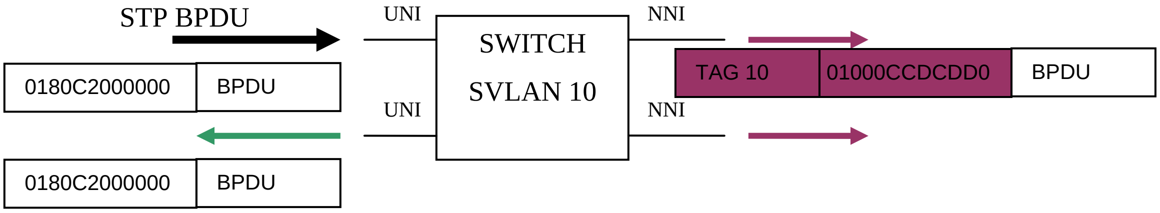
```
→ethernet-service uni-profile "U2" tunnel-mac 00:00:00:11:11:11
→ethernet-service uni-profile "U2" l2-protocol pagp mac-tunnel udld mac-tunnel vtp
 mac-tunnel dtp mac-tunnel cdp mac-tunnel pvst mac-tunnel
```

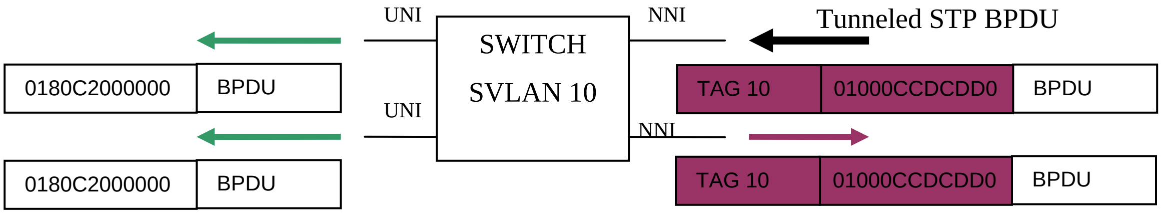
In this example, the 00:00:00:11:11:11 address is configured as the tunnel MAC address for the “U2” profile. By default, the CISCO DA MAC, 01:00:0c:cd:cd:d0, is used if a tunnel MAC is not specified.

The **show ethernet-service uni-profile** command displays the UNI profile configuration. The following example shows the default UNI profile:

```
→Show ethernet-service uni-profile
Profile Name: default-uni-profile
 Tunnel MAC : 01:00:0c:cd:cd:d0,
 STP : tunnel,      802.1x : drop,       802.3ad : peer,       802.1ab    : drop,
 GVRP: tunnel,      AMAP   : drop,       OAM     : peer,       LACPMARKER : peer,
 UDLD: drop,        PAGP   : drop,       CDP     : drop,       VTP        : drop,
 DTP : drop,        PVST   : drop,       VLAN    : drop,       UPLINK     : drop,
 MVRP: tunnel
```

The following example shows the tunneling and de-tunneling of an STP BPDU using the global default MAC address (CISCO DA 01:00:0c:cd:cd:d0):





References

- Chapter 44, “VLAN Stacking Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 42, “Configuring VLAN Stacking”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.19. SVLAN Routing

VLAN Stacking provides a mechanism to tunnel multiple customer VLANs (CVLAN) through a service provider network using one or more service provider VLANs (SVLAN) by way of 802.1Q double-tagging or VLAN Translation. This feature enables service providers to offer their customers Transparent LAN Services (TLS). This service is multipoint in nature so as to support multiple customer sites or networks distributed over the edges of a service provider network.

Although the SVLAN is used to bridge CVLANs through the provider network, communicating with other network VLANs may be required. To allow SVLAN routing, the ability to configure an IP interface for an SVLAN is now supported.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- Although the hardware supports routing of double-tagged frames, the CPU is only able to send single tagged IP frames to the SVLAN. This means that the Ethernet service configured with a “routed” SVLAN must guarantee that only single tagged frame would be routed. As a result, the SAP must be configured in one of the following two ways:
    - Untagged SAP. Customer has access to the routed service with untagged frames only
    - Tagged SAP with a translation SAP profile. Customer has access to the routed service with tagged frames with a specific VLAN ID only. The customer tag VLAN ID is translated to the SVLAN ID.
- Configuring an IP interface on a SVLAN can make the switch vulnerable to DoS attacks from the customer premises that are connected to a SAP associated with this SVLAN. To minimize the security impact, it is highly recommended that the provider configure ACLs to restrict switch access only to “trusted” networks or hosts. For example (the following policy uses the built-in “Switch” network group):

```
→policy service ftp destination tcp port 21
→policy service ssh destination tcp port 22
→policy service telnet destination tcp port 23
→…
→policy service http destination tcp port 80
→policy service group All_Services ftp ssh telnet http
→policy condition Switch_Services service group All_Services destination network
 group Switch
→policy action deny disposition deny
→policy rule Deny_Switch_Services condition Switch_Services action deny
→policy network group Trusted 10.0.0.1 11.0.0.1 12.0.0.1
→policy action accept disposition accept
→policy condition Trusted_Services service group All_Services source network group
 Trusted
→policy rule Allow_Switch_Services condition Trusted_Services action accept
 precedence 10000
→qos apply
```

- Only IPv4 interfaces are allowed on an SVLAN; IPv6 interfaces are not supported.

Configuration Example

The following configuration snapshot displays an Ethernet Service configuration for translating CVLAN 2001 to 1001. In this example, SVLAN 1001 is created with an IP interface; RIP is enabled and the SVLAN interface is configured as a RIP interface.

```
→show configuration snapshot vlan ip ip-routing linkagg rip

! VLAN :
vlan 1 enable name "VLAN 1"
vlan 172 enable name "VLAN 172"
vlan 172 port default 2/10
ethernet-service svlan 1000 name "VLAN 1000"
ethernet-service svlan 1001 name "VLAN 1001"
! VLAN SL:
! IP :
debug ip set ipedrArpPoisonLrn 1
ip service all
ip interface "172" address 172.35.10.20 mask 255.255.255.0 vlan 172 ifindex 1
ip interface "SVLAN-1000" address 201.10.10.3 mask 255.255.255.0 vlan 1000
ifindex 2
ip interface "SVLAN-1001" address 201.10.20.1 mask 255.255.255.0 vlan 1001
ifindex 3
! RIP :
ip load rip
ip rip interface "SVLAN-1001"
ip rip interface "SVLAN-1001" status enable
ip rip status enable
ip static-route 10.145.59.0/24 gateway 172.35.10.1 metric 1
ip static-route 172.25.200.0/24 gateway 172.35.10.1 metric 1
ip static-route 172.35.200.0/24 gateway 172.35.10.1 metric 1
ip static-route 172.65.200.0/24 gateway 172.35.10.1 metric 1
! Link Aggregate :
lacp linkagg 1 size 4 admin state enable
lacp linkagg 1 actor admin key 1
lacp agg 1/11 actor admin key 1
lacp agg 1/20 actor admin key 1
lacp agg 2/9 actor admin key 1
lacp agg 2/11 actor admin key 1
! VLAN AGG:
! VLAN STACKING:
ethernet-service svlan 1000 nni linkagg 1
ethernet-service svlan 1001 nni 1/19
ethernet-service sap-profile "SAP2" cvlan-tag translate
ethernet-service service-name "SVLAN-1000" svlan 1000
ethernet-service service-name "SVLAN-1001" svlan 1001
ethernet-service sap 1001 service-name "SVLAN-1001"
ethernet-service sap 1001 sap-profile "SAP2"
ethernet-service sap 1001 uni 1/10
ethernet-service sap 1001 cvlan 2001
```

References

- Chapter 44, “VLAN Stacking Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 42, “Configuring VLAN Stacking”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.20. Wire-Speed Ethernet Loopback

A Wire-Speed Ethernet loopback test function is available to perform In-Service and Out-of-Service throughput testing during initial turn-up or on-the-fly in an active network. The loopback tests can be used to validate the configured Service Level Agreements (SLAs) and QoS parameters that are associated with a service or a flow.

The loopback test capability provided allows the use of an external test head to send traffic at wire-rate speed to a specific switch port which then loops the traffic back to the test head. The test head measures and collects statistics on frame loss, delay, and latency of the loopback traffic.

There are two types of loopback tests supported with this implementation: inward loopback and outward loopback. The inward test loops back test head frames ingressing on a given port. The outward test loops back test head frames egressing on a given port.

Platforms Supported

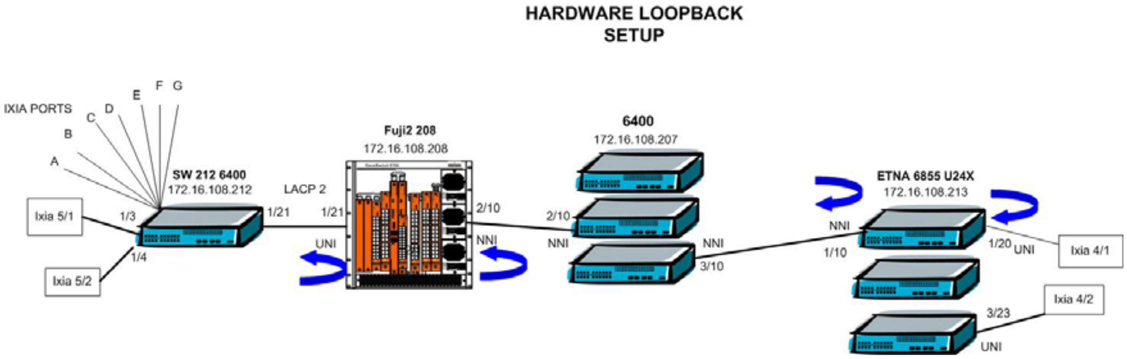
OmniSwitch 6400, 6850, 6855, 9000

Guidelines

- The loopback test profile specifies the switch attributes that are required to conduct an inward or outward loopback operation on a switch port. Up to eight profiles are supported.
- The loopback operation is not active until the loopback profile is enabled. A separate CLI command, **loopback-test** *profile_name* **{enable | disable}**, is used to start and stop the loopback operation.
- Once a UNI or NNI port is designated as a loopback port, the port is no longer eligible to participate in other switch functions.
- An outward loopback port goes “out-of-service” and will no longer carry customer traffic but remains active for test frame traffic. However, an inward loopback port remains “in-service” and will continue to carry customer traffic as well as test frame traffic.
- Only pure L2 IP traffic (Ethertype 0x800 frames) is supported. Non-IP packets, including Layer2 MAC only flows, are not supported.
- Configuring an inward and an outward profile with the same port is not allowed. In addition, configuring a profile for a port that is a member of a link aggregate is not supported.
- The switch creates a static MAC address entry for the egress port when the outward loopback profile is applied on that port. The static address created is the destination MAC address specified in the profile. If the switch receives a non-test frame that contains the same MAC address, both the test and non-test frames are filtered even if they were received on different ports.
- If the MAC addresses specified in the loopback test profile are actual network address (for example, 02:da:95:e1:22:10, not aa:aa:aa:aa:aa:aa), flush the MAC address table for the switch when the loopback test is finished.
- Only the profile configuration is stored in the boot.cfg file, not the commands to enable or disable the profile status. As a result, the profile status must be manually enabled after a system reload or CMM takeover.

Configuration Examples

The following sample network configuration was used to test the Wire-Speed Ethernet Loopback feature:



This configuration contains a VLAN Stacking setup that is used to perform the loopback operation on UNI and NNI ports. Test traffic is sent through the VLAN Stacking Ethernet Service where the Service Access Profile (SAP) applies bandwidth and priority values. Traffic is then looped back to verify the traffic and that the SAP values were applied.

Example CLI Configuration

OmniSwitch 6850 VLAN Stacking Configuration

```
→ethernet-service svlan 100 nni 2/10
→ethernet-service sap-profile “Customer A” bandwidth 10
→ethernet-service sap-profile “Customer A” priority map-inner-to-outer-p
→ethernet-service service-name “Customer A” svlan 100
→ethernet-service sap 10 service-name “Customer A”
→ethernet-service sap 10 uni 1/21
→ethernet-service sap 10 cvlan 10
```

OmniSwitch 6850-U24X VLAN Stacking Configuration

```
→ethernet-service svlan 100 nni 1/10
→ethernet-service sap-profile “Customer A” bandwidth 10
→ethernet-service sap-profile “Customer A” priority map-inner-to-outer-p
→ethernet-service service-name “Customer A” svlan 100
→ethernet-service sap 10 service-name “Customer A”
→ethernet-service sap 10 uni 1/20
→ethernet-service sap 10 cvlan 10
```

Loopback Configuration

The following commands configure and enable the loopback test profile:

```
→ loopback-test test1 source-mac 00:00:00:00:01:01 destination-mac 00:00:00:00:01:02
  vlan 100 loopback-port 1/21 type inward
→ loopback-test test1 enable
```

The following command disables the profile:

```
→ loopback-test test1 disable
```

References

- Chapter 44, "VLAN Stacking Commands", *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 42, "Configuring VLAN Stacking", *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.21. Ethernet OAM 802.1ag Version 8 and ITU Y1731

IEEE 802.1ag Connectivity Fault Management (CFM) defines protocols and practices for OAM (Operations, Administration and Maintenance) for paths through 802.1 bridges and local area networks (LANs). It is an amendment to IEEE 802.1Q-2005 and was approved in 2007. Note that previous releases of AOS support an earlier version of IEEE 802.1ag (Draft Revision 7).

Also known as Service OAM, the IEEE 802.1ag CFM is used to monitor and troubleshoot end-to-end Ethernet services. This implementation of Ethernet Service OAM supports both IEEE 802.1ag and ITU-T Y.1731 for Connectivity Fault Management (plus performance monitoring provided by ITU-T Y.1731).

In compliance with the ITU-T Y.1731 performance monitoring definition, the OmniSwitch supports Ethernet frame delay measurement (one-way and two-way). However, the OmniSwitch implementation is agnostic to either IEEE 802.1ag or ITU-T Y.1731 in that delay measurement can be performed for maintenance points that comply with either standard.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 6855-U24X, 9000, 9000E

Guidelines

This section contains general information and configuration guidelines for the 6.4.3.R01 implementation of Ethernet OAM (802.1ag, version 8) and support for the ITU-T Y.1731 Recommendation.

Elements of Service OAM

- Maintenance End Points (MEPs) and Maintenance Intermediate Points (MIPs)
  - ➢ MEPs initiate OAM commands. MEPs prevent leakage between domains.
  - ➢ MIPs passively receive and respond to OAM frames.
- Maintenance Association (MA) is a logical connection between 2 or more MEPs.
- Point-to-point MA: logical sub-MA component only between 2 MEPs MA.
- Maintenance Domain: One or more MAs under the same administrative control.
- Maintenance Domain Levels: There are 8 levels defined in 802.1ag:
  - ➢ levels [5, 6, 7] are for operators,
  - ➢ levels [3, 4] are for service provider
  - ➢ levels [0, 1, 2] are for customers

  Multiple levels are supported for flexibility.
- CFM Mechanisms: continuity check, loopback, link trace

Connectivity Fault Management

Service OAM Connectivity Fault Management (CFM) consists of three types of messages that are used to help network administrators detect, verify, and isolate when a problem occurs in the network:

- **Continuity Check Messages (CCM).** These are multicast messages exchanged periodically by MEPs to detect loss of service connectivity between MEPs. These messages are also used by MEPs and MIPs to discover other MEPs within a domain.

  - ➢ CC messages flow within an MA—for example, a service (VLAN)—and are multicast within a VLAN. A MA can also monitor more than one VLAN and any MEP configured in the MA will inherit the MAID, MD Level, and Primary VID from its MA. Moreover, the selection of which of the MA's VIDs is the Primary VID can be overridden for a specific MEP.

- **Linktrace Messages (LTM).** These messages are transmitted by a MEP to trace the path to a destination maintenance point. The receiving maintenance point responds to LTMs with a linktrace reply (LTR). This mechanism is similar to the UDP Trace Route function. The transmission of linktrace messages is requested by an administrator.
- **Loopback Messages (LBM).** These messages are transmitted by a MEP to a specified MIP or MEP to determine whether or not the maintenance point is reachable. The receiving maintenance point responds to LBMs with a loopback reply (LBR). This mechanism is not used to discover a path to the destination; it is similar to the Ping function. The transmission of loopback messages is requested by an administrator.

Ethernet Service OAM Implementation Differences

The 6.4.3.R01 implementation of Ethernet Service OAM (802.1ag CFM) differs from the previous implementation as follows:

- A single MA will monitor more than one service (VLAN). While creating the MA, only the primary VID is required and the mapping between primary and non-primary VID(s) is stored in the VLAN Table.
- Transmitting the Sender ID TLV is a management action that is controlled by MD, MA, and Default-MD.
- Sender ID TLV now includes the Management Address and the Domain name of the TLV.
- The MD also controls the creating of the MIP. The MA mhfCreation object supersedes the mhfCreation object in the MD table. Both objects are controlled by the 'defer' value of the MA mhfCreation object.
- The Egress Identifier in the LTM frame is sent as LTM Egress Identifier TLV (type is 7).
- The last Egress Identifier and the next Egress Identifier in the LTR frame are sent as LTR Egress Identifier TLV (type is 8).
- LTR flags have changed. Refer to clause 21.9.1 in the standard for details.
- TLV types have changed for Data TLV (from 4 to 3) and Interface Status TLV (from 3 to 4).
- RDI condition has changed. Refer to clause 21.6.1.1 in the standard for details.
- The initiation mechanism for both Loopback and Linktrace messages has changed.

Interoperability with ITU-T Y.1731

Although this implementation of Service OAM supports both IEEE 802.1ag and ITU-T Y.1731 for CFM, the 802.1ag terminology and hierarchy is used for configuring CFM. The following table provides a mapping of 802.1ag terms to the equivalent ITU-T Y.1731 terms:

IEEE 802.1ag v8.1	ITU-T Y.1731
Maintenance Domain (MD)	Maintenance Entity (ME)
Maintenance Association (MA)	Maintenance Entity Group (MEG)
Maintenance Endpoint (MEP)	MEG Endpoint (MEP)
Maintenance Intermediate Point (MIP)	MEG Intermediate Point (MIP)
Maintenance Domain Level	MEG Level

Support for both the IEEE and ITU-T Ethernet CFM standards allows interoperability between OmniSwitch 802.1ag and Y.1731 CFM with the following considerations:

- The OmniSwitch MD format must be configured as “none”, which is now an option provided with the **ethoam domain** CLI command. For example:

```
→ ethoam domain MD1 format none
```

- When the MD format is “none”, the MD name is not encoded in the MAID and only the format, length, and name of the MD is encoded. Basically, MA in IEEE is analogous to MEG in ITU-T.
- ITU-T Y.1731 uses the “icc-based” format for a MEG, so the OmniSwitch MA format must also be configured to use the “icc-based” format. The **ethoam association** CLI command now includes an **icc-based** option. For example:

```
→ ethoam association MA1 format icc-based domain MD1 primary-vlan 100
```

- When the OmniSwitch MA is configured with the “icc-based” format, the MA name is automatically padded with zeros if the name specified is less than 13 characters.
- ITU-T Y.1731 does not include the LTM Egress Identifier TLV or the LTR Egress Identifier TLV. For compatibility with earlier implementations of ITU-T, IEEE does not require that these TLVs be present on received LTMs or LTRs. However, IEEE does require these TLVs to be transmitted in LTMs and LTRs.
- Even though IEEE 802.1ag and ITU-T Y.1731 have the same objective and use the same header format, they do not have the same field definitions.

ITU-T Y.1731 Performance Monitoring

The ITU-T Y.1731 Recommendation addresses the need to monitor performance to help enforce customer service level agreements (SLAs). Frame delay (latency) and frame delay variation (jitter) are important performance objectives, especially for those applications (such as voice) that cannot function with a high level of latency or jitter.

This implementation of Service OAM supports Ethernet frame delay measurement (ETH-DM). The ETH-DM feature allows for the configuration of on-demand OAM to measure frame delay and frame delay variation between endpoints.

Frame delay measurement is performed between peer MEPs (measurements to MIPs are not done) within the same MA. Although the OmniSwitch implementation of ETH-DM is compliant with ITU-T Y.1731, delay measurement can be performed for both ITU-T Y.1731 and IEEE 802.1ag MEPs.

Any MEP can initiate or reply to an ETH-DM request, depending on the type of delay measurement requested. There are two types of delay measurements supported: one-way and two-way.

One-Way Delay Measurement

- A MEP sends one-way delay measurement (1DM)) frames to a peer MEP. The sending MEP inserts the transmission time (TxTimeStampf) into the 1DM frame at the time the frame is sent.
- When a MEP receives a 1DM frame, the MEP calculates the one-way delay as the difference between the time at which the frame was received (t=RxTimeb) and the transmission time indicated by the frame timestamp (receive time minus transmission time). For example:

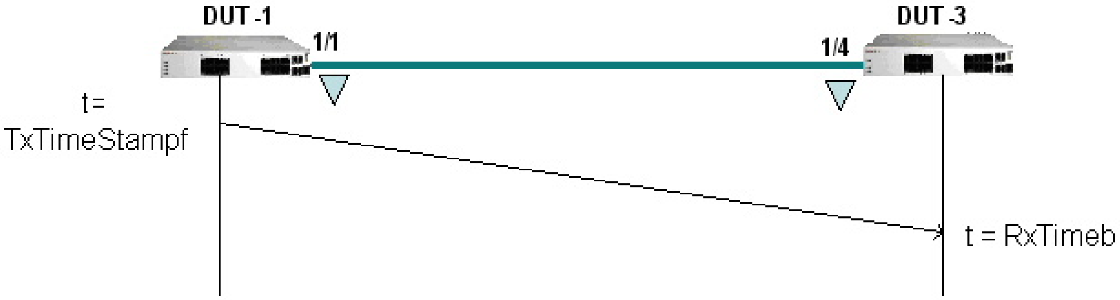
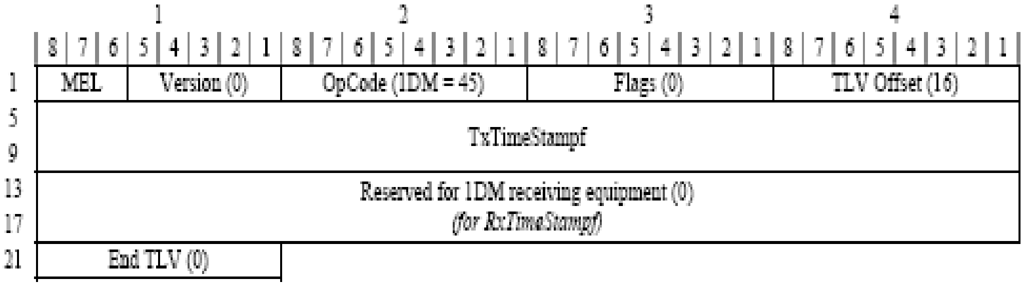


Figure 1. One-Way Frame Delay Measurement

- One-way delay measurement statistics are gathered and stored on the receiving MEP (the MEP that receives a 1DM request).

One-way delay measurement requires clock synchronization between the sending and receiving MEPs. Using NTP for clock synchronization is recommended. Two-way delay measurement does not require clock synchronization.

The PDU format for a one-way-delay frame is:



	1		2	3	4
	8 7 6	5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1
1	MEL	Version (0)	OpCode (1DM = 45)	Flags (0)	TLV Offset (16)
5 9	TxTimeStampf				
13 17	Reserved for 1DM receiving equipment (0) *(for RxTimeStampf)*				
21	End TLV (0)				

Figure 2. ETH One-Way-Delay PDU

Two-Way-Delay Measurement

- A MEP sends delay measurement message (DMM) frames to a peer MEP to request a two-way ETH-DM. The sending MEP inserts the transmission time (TxTimeStampf) into the DMM frame at the time the frame is sent.
- When a MEP receives a DMM frame, the MEP responds to the DMM with a delay message reply (DMR) frame that contains the following timestamps:
    - ➢ Timestamp copied from the DMM frame (TxTimeStampf).
    - ➢ Timestamp indicating when the DMM frame was received (t=RxTimeStampf).
    - ➢ Timestamp indicating the time at which the receiving MEP transmitted the DMR frame back to the sending MEP (t=TxTimeStampb).
- When a MEP receives a DMR frame, the MEP compares all the DMR timestamps with the time at which the MEP received the DMR frame (t=RxTimeb) to calculate the two-way delay.
- The two-way delay is the difference between the time the originating MEP sent a DMM request and the time at which the originating MEP received a DMR frame minus the time taken by the responding MEP to process the DMM request. For example:

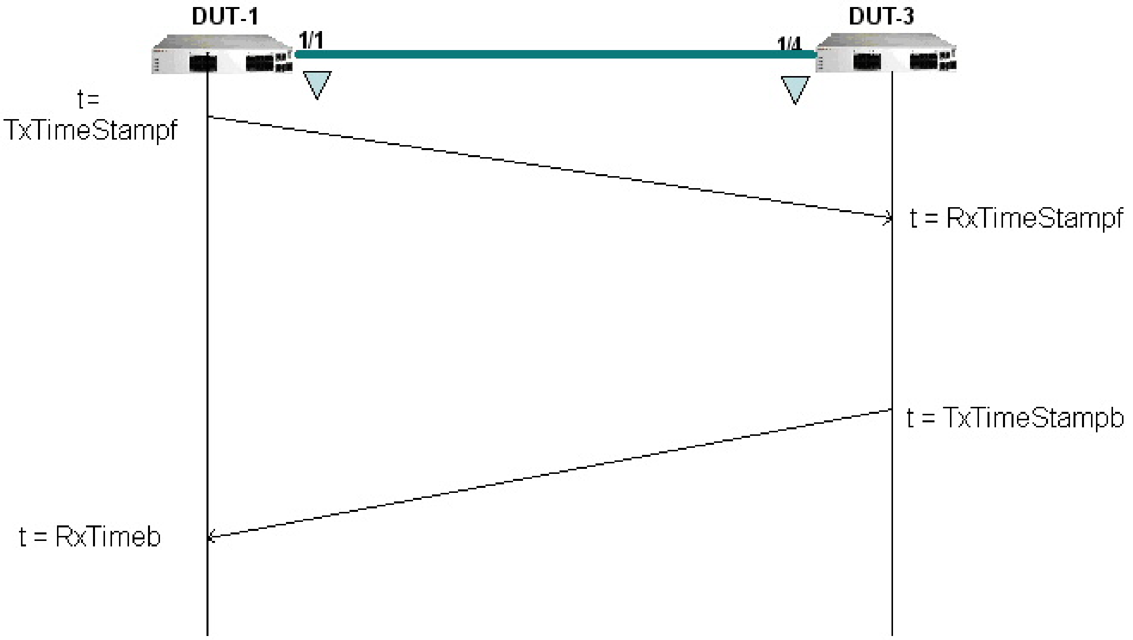
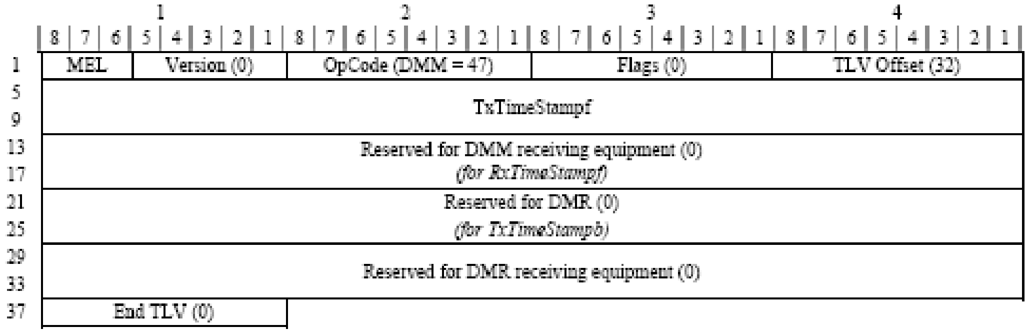


Figure 3. Two-Way Frame Delay Measurement

- Two-way delay measurement statistics are gathered and stored on the originating MEP (the MEP that initiates a DMM request).
- This method does not require clock synchronization between the transmitting and receiving MEPs.

Two-way delay measurement is an on-demand OAM performance measurement. To set up continuous two-way delay measurement, see the "Service Assurance Agent" section for more information.

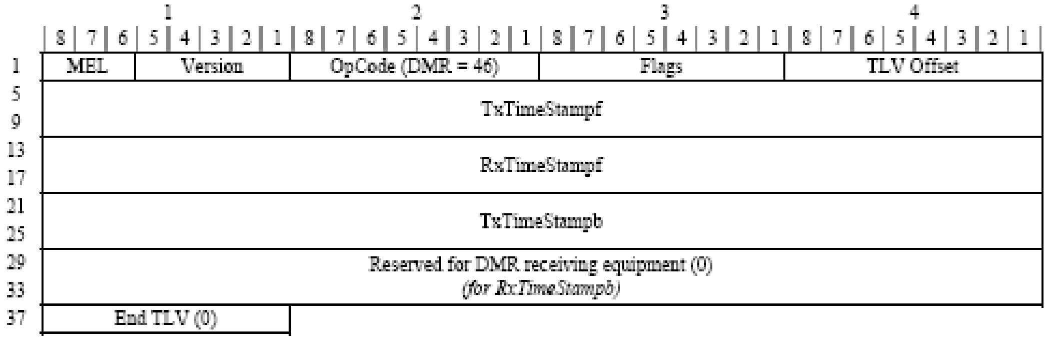
The PDU format for a two-way-delay request frame is:



	1		2	3	4
	8 7 6	5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1
1	MEL	Version (0)	OpCode (DMM = 47)	Flags (0)	TLV Offset (32)
5 9	TxTimeStampf				
13 17	Reserved for DMM receiving equipment (0) *(for RxTimeStampf)*				
21 25	Reserved for DMR (0) *(for TxTimeStampb)*				
29 33	Reserved for DMR receiving equipment (0)				
37	End TLV (0)				

Figure 4. ETH DMM PDU

The PDU format for a two-way-delay reply frame is:



	1		2	3	4
	8 7 6	5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1
1	MEL	Version	OpCode (DMR = 46)	Flags	TLV Offset
5 9	TxTimeStampf				
13 17	RxTimeStampf				
21 25	TxTimeStampb				
29 33	Reserved for DMR receiving equipment (0) *(for RxTimeStampb)*				
37	End TLV (0)				

Figure 5. ETH DMR PDU

Frame Delay Variation

The delay variation (jitter) for both one-way and two-way ETH-DM is determined by calculating the difference between the current delay measurement value and the previous delay measurement value. If a previous delay value is not available, which is the case when a DM request is first made, jitter is not calculated.

Other Guidelines

- Packet Loss measurement - Dual-ended packet loss measurement in either CCM messages or LMM/LMR messages is not supported. These functions rely on hardware packet generation and accurate ingress and egress service counters.
- Test ETH-Test - One way test signal is not supported. This function implies the generation of a test frame with test pattern (all-0, all-1, CRC) and verification of test frames.
- The following functions are also not supported:

  ETH-AIS
  ETH-LCK
  ETH-APS
  ETH-MCC
  ETH-EXM/EXR
  ETH-VXM/VXR

- Configuring a DOWN MEP on UNI port is not allowed. This is the same limitation that was in the previous 6.4.2.R01 release.
- LRU algorithm is not supported to record new entry in MIPCCM database by removing the oldest entry. This is same as in the previous 6.4.2.R01 release.
- This implementation does not provide the different control access for “OWNER” and “ADMINISTRATOR” as mentioned in the clause 12.1.4.1 of IEEE Std 802.1ag-2007.
- The one-way and two-way delay functions do not allow the configuration of the packet size. The standard does not specify this capability.
- ITU-T and IEEE handle Ethernet Loopback differently. AOS supports the IEEE 802.1ag Ethernet Loopback (i.e., Unicast ETH-LB not Multicast ETH-LB).
- Without time synchronization between the MEPs or without NTP running, the one-way frame delay measurement is not recommended. However, delay variation (jitter) measurement can be performed. Moreover, even if NTP is running the delay calculation is not accurate.

Configuration Examples

Example 1: Complete MA

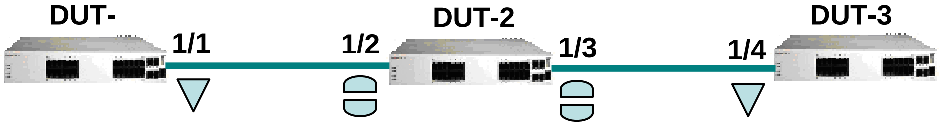


Figure 1: 3-box Linear Topology

DUT-1

```
→vlan 10-15
→vlan 10-15 802.1q 1/1
→ethoam vlan 11-15 primary-vlan 10
→ethoam domain MD format string level 3
→ethoam association MA format string domain MD primary-vlan 10
→ethoam association MA domain MD endpoint-list 10
→ethoam association MA domain MD endpoint-list 20
→ethoam endpoint 10 domain MD association MA direction down port 1/1
→ethoam endpoint 10 domain MD association MA admin-state enable
→ethoam endpoint 10 domain MD association MA ccm enable
```

DUT-2

```
→vlan 10-15
→vlan 10-15 802.1q 1/2
→vlan 10-15 802.1q 1/3
→ethoam vlan 11-15 primary-vlan 10
→ethoam domain MD format string level 3
→ethoam association MA format string domain MD primary-vlan 10
→ethoam association MA domain MD mhf default
```

DUT-3

```
→vlan 10-15
→vlan 10-15 802.1q 1/4
→ethoam vlan 10-14 primary-vlan 15
→ethoam domain MD format string level 3
→ethoam association MA format string domain MD primary-vlan 15
→ethoam association MA domain MD endpoint-list 10
→ethoam association MA domain MD endpoint-list 20
→ethoam endpoint 20 domain MD association MA direction down port 1/4 vlan 10
→ethoam endpoint 20 domain MD association MA admin-state enable
→ethoam endpoint 20 domain MD association MA ccm enable
```

Example 2: Inter-op with ITU-T

For interoperability with ITU-T, MD will support one more format, which is “none” and also MA will support another format, which is “ICC-based”. Figure 2 illustrates this example (it is important to note that the length of the MA name is 13):

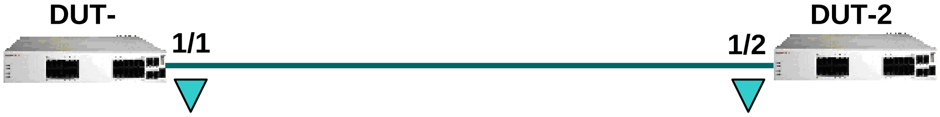


Figure 2: Inter-op with ITU-T

DUT-1

```
 → vlan 10-15
 → vlan 10-15 802.1q 1/1
 → ethoam vlan 11-15 primary-vlan 10
 → ethoam domain MD format none level 5
 → ethoam association AlcatelLucent format icc-based domain MD5 primary-vlan 10
 → ethoam association AlcatelLucent domain MD endpoint-list 10
 → ethoam association AlcatelLucent domain MD endpoint-list 20
 → ethoam endpoint 10 domain MD association AlcatelLucent direction down port 1/1
 → ethoam endpoint 10 domain MD association AlcatelLucent admin-state enable
 → ethoam endpoint 10 domain MD association AlcatelLucent ccm enable
```

DUT-2

```
 → vlan 10-15
 → vlan 10-15 802.1q 1/2
 → ethoam vlan 11-15 primary-vlan 10
 → ethoam domain MD format none level 5
 → ethoam association AlcatelLucent format icc-based domain MD primary-vlan 10
 → ethoam association AlcatelLucent domain MD endpoint-list 10
 → ethoam association AlcatelLucent domain MD endpoint-list 20
 → ethoam endpoint 20 domain MD association AlcatelLucent direction down port 1/2
 → ethoam endpoint 20 domain MD association AlcatelLucent admin-state enable
 → ethoam endpoint 20 domain MD association AlcatelLucent ccm enable
 →
```

Example 3: Sample Output – Performance Monitoring

```
 → ethoam one-way-delay target-endpoint 10 source-endpoint 12 domain MD association
   MA vlan-priority 4
 →
 → ethoam one-way-delay target-macaddress 00:e0:b1:6a:52:4c source-endpoint 12 domain
   MD association MA vlan-priority 4
 →
 → show ethoam one-way-delay domain MD association MA endpoint 10
     Legend: Jitter: - = undefined value

Remote Mac address       Delay (us)  Jitter (us)
--------------------+-------------+------------
   00:d0:95:ef:44:44          2369        1258
   00:d0:95:ef:66:88          5896         282
   00:d0:95:ef:88:88          2584          -
   00:d0:95:ef:66:55          2698        4782
```

```
→ show ethoam one-way-delay domain MD association MA endpoint 10 mac-address
  00:d0:95:ef:44:44
Legend: Jitter: - = undefined value
Remote Mac address        Delay (us)  Jitter (us)
--------------------+-------------+------------
   00:d0:95:ef:44:44           2369         1258
```

```
→ ethoam two-way-delay target-endpoint 10 source-endpoint 12 domain MD association
  MA vlan-priority 4
Reply from 00:0E:B1:6B:43:89 delay=2584us jitter=282us
```

```
→ ethoam two-way-delay target-macaddress 00:e0:b1:6a:52:4c source-endpoint 12 domain
  MD association MA vlan-priority 4
Reply form 00:E0:B1:6A:52:4C: delay=2584us jitter=282us
```

```
→ show ethoam two-way-delay domain MD association MA endpoint 10 mac-address
  00:d0:95:ef:44:44
Legend: Jitter: - = undefined value
      : RMEP-ID: 0 = two-way-delay was initiated with target mac-address

Remote Mac address  RMEP-ID      Delay (us)   Jitter (us)
-------------------+--------+--------------+------------
 00:d0:95:ef:44:44         12           2369        1258
```

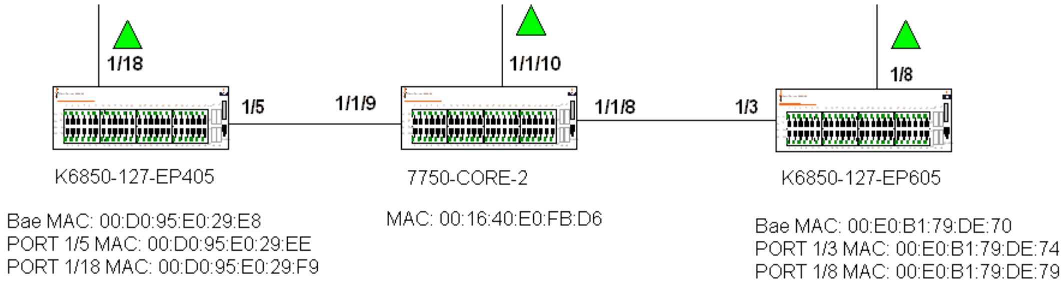
```
→ show ethoam two-way-delay domain MD association MA endpoint 10 remote-mep 0
Legend: Jitter: - = undefined value
      : RMEP-ID: 0 = two-way-delay was initiated with target mac-address

Remote Mac address  RMEP-ID      Delay (us)   Jitter (us)
-------------------+--------+--------------+------------
 00:d0:95:ef:66:88          0           5896          282
 00:d0:95:ef:88:88          0           2584         1856
```

```
→ show ethoam two-way-delay domain MD association MA endpoint 10 remote-mep 15
Legend: Jitter: - = undefined value
      : RMEP-ID: 0 = two-way-delay was initiated with target mac-address

Remote Mac address  RMEP-ID      Delay (us)   Jitter (us)
-------------------+--------+--------------+------------
 00:d0:95:ef:66:55         15           2736           -
```

```
→ show ethoam two-way-delay domain MD association MA endpoint 10
Legend: Jitter: - = undefined value
      : RMEP-ID: 0 = two-way-delay was initiated with target mac-address

Remote Mac address  RMEP-ID      Delay (us)   Jitter (us)
-------------------+--------+--------------+------------
 00:d0:95:ef:44:44         12           2369         1258
 00:d0:95:ef:66:88          0           5896          282
 00:d0:95:ef:88:88          0           2584         1856
 00:d0:95:ef:66:55         15           2736           -
```

Example 4: Interoperability with Alcatel-Lucent SR Series

This example shows an Ethernet OAM configuration consisting of two AOS OmniSwitch 6850 switches and one SR 7750 switch. In addition to the following illustration, a snapshot of the configuration on each switch is also provided.


1/18
1/1/10
1/8
1/5
1/1/9
1/1/8
1/3
K6850-127-EP405
7750-CORE-2
K6850-127-EP605
Bae MAC: 00:D0:95:E0:29:E8
PORT 1/5 MAC: 00:D0:95:E0:29:EE
PORT 1/18 MAC: 00:D0:95:E0:29:F9
MAC: 00:16:40:E0:FB:D6
Bae MAC: 00:E0:B1:79:DE:70
PORT 1/3 MAC: 00:E0:B1:79:DE:74
PORT 1/8 MAC: 00:E0:B1:79:DE:79

OmniSwitch 6850 Configuration Snapshot (K6850-127-EP405)

```
K6850-127EP405-> show configuration snapshot vlan
! VLAN :
vlan 1 enable name "VLAN 1"
ethernet-service svlan 100 name "VLAN 100"
vlan 172 1x1 stp disable flat stp disable name "VLAN 172"
vlan 172 port default 1/1
! VLAN SL:
! VLAN AGG:
! VLAN STACKING:
ethernet-service svlan 100 nni 1/5-6
ethernet-service sap-profile "CustomerA"
ethernet-service service-name "CustomerA" svlan 100
ethernet-service sap 10 service-name "CustomerA"
ethernet-service sap 10 sap-profile "CustomerA"
ethernet-service sap 10 uni 1/18
ethernet-service sap 10 cvlan 10
```

```
K6850-127-EP405> show configuration snapshot ethernet-oam
! Ethernet-OAM :
ethoam domain 4 format none level 2
ethoam domain 4 id-permission chassisid
ethoam association Metro.0000022 format icc-based domain 4 primary-vlan 100
ethoam association Metro.0000022 domain 4 mhf default
ethoam association Metro.0000022 domain 4 ccm-interval interval1s
ethoam association Metro.0000022 domain 4 endpoint-list 405
ethoam association Metro.0000022 domain 4 endpoint-list 605
ethoam association Metro.0000022 domain 4 endpoint-list 1220
ethoam endpoint 405 domain 4 association Metro.0000022 direction up port 1/18
primary-vlan 100
ethoam endpoint 405 domain 4 association Metro.0000022 admin-state enable
ethoam endpoint 405 domain 4 association Metro.0000022 ccm enable
K6850-127-EP405>
```

SR 7750 Configuration Snapshot (7750-Core-2)

```
#--------------------------------------------------
echo "Eth-CFM Configuration"
#--------------------------------------------------
    eth-cfm
        domain 4 format none level 2
            association 22 format icc-based name "Metro.0000022"
                bridge-identifier 100
                exit
                ccm-interval 1
                remote-mepid 405
                remote-mepid 605
            exit
        exit
    exit
#--------------------------------------------------
echo "Service Configuration"
#--------------------------------------------------
    service
        vpls 100 customer 1 create
            stp
                shutdown
            exit
            sap 1/1/8:100 create
                eth-cfm
                    mip
                exit
            exit
            sap 1/1/9:100 create
                eth-cfm
                    mip
                exit
            exit
            sap 1/1/10:100 create
                eth-cfm
                    mep 1220 domain 4 association 22 direction up
                        ccm-enable
                        no shutdown
                    exit
                exit
            exit
            no shutdown
        exit
```

OmniSwitch 6850 Configuration (K6850-128-EP605)

```
K6850-128-EP605> show configuration snapshot vlan
! VLAN :
vlan 1 enable name "VLAN 1"
ethernet-service svlan 100 name "VLAN 100"
vlan 172 enable name "VLAN 172"
vlan 172 port default 1/1
! VLAN SL:
! VLAN AGG:
! VLAN STACKING:
ethernet-service svlan 100 nni 1/5-6
ethernet-service sap-profile "CustomerA"
ethernet-service service-name "CustomerA" svlan 100
ethernet-service sap 10 service-name "CustomerA"
ethernet-service sap 10 sap-profile "CustomerA"
ethernet-service sap 10 uni 1/8
```

```
ethernet-service sap 10 cvlan 10
K6850-128-EP605> show configuration snapshot ethernet-oam
! Ethernet-OAM :
ethoam domain 4 format none level 2
ethoam domain 4 id-permission chassisid
ethoam association Metro.0000022 format icc-based domain 4 primary-vlan 100
ethoam association Metro.0000022 domain 4 mhf default
ethoam association Metro.0000022 domain 4 ccm-interval interval1s
ethoam association Metro.0000022 domain 4 endpoint-list 405
ethoam association Metro.0000022 domain 4 endpoint-list 605
ethoam association Metro.0000022 domain 4 endpoint-list 1220
ethoam endpoint 605 domain 4 association Metro.0000022 direction up port 1/8
primary-vlan 100
ethoam endpoint 605 domain 4 association Metro.0000022 admin-state enable
ethoam endpoint 605 domain 4 association Metro.0000022 ccm enable
K6850-128-EP605>
```

View all remote endpoints:

```
K6850-127-EP405> show ethoam remote-endpoint domain 4 association Metro.0000022
endpoint 405
Total number of remote MEPs = 2
Legends: PortStatusTlv: 1 = psBlocked, 2 = psUp, 0 = psNoTlv
         InterfaceStatusTlv: 1 = ifUp, 2 = ifDown, 4 = ifUnknown, 7 =
ifLowerLayerDown

RMEP-ID      RMEP        OkFailed      Mac Address       Port Status  I/f Status
RDI      Chassis ID     Chassis ID
            Status         Time                              Tlv         Tlv
Value      Subtype         Value
-------+--------------+-------------+------------------+-------------+----------+---
---+------------------+----------
   605   RMEP_OK           27319700   00:E0:B1:79:DE:70             2          1
false  LOCALLY_ASSIGNED   K6850-128-EP605

  1220   RMEP_OK           27318600   00:16:4D:E0:FB:D6             2          1
false  none               none
K6850-127-EP405>


*A:7750-core-2->>config# show eth-cfm mep 1220 domain 4 association 22 all-remote-
mepids

===========================================================================
Eth-CFM Remote-Mep Table
===========================================================================
R-mepId Rx CC  Rx Rdi Port-Tlv If-Tlv Peer Mac Addr     CCM status since
---------------------------------------------------------------------------
405     True   False  Up       Up     00:d0:95:e0:29:e8 09/09/2010 03:27:39
605     True   False  Up       Up     00:e0:b1:79:de:70 09/12/2010 07:14:24
===========================================================================


K6850-128-EP605> show ethoam remote-endpoint domain 4 association Metro.0000022
endpoint 605
Total number of remote MEPs = 2
Legends: PortStatusTlv: 1 = psBlocked, 2 = psUp, 0 = psNoTlv
         InterfaceStatusTlv: 1 = ifUp, 2 = ifDown, 4 = ifUnknown, 7 =
ifLowerLayerDown

RMEP-ID      RMEP        OkFailed      Mac Address       Port Status  I/f Status
RDI      Chassis ID     Chassis ID
            Status         Time                              Tlv         Tlv
Value      Subtype         Value
```

```
-------+--------------+-------------+-----------------+-------------+----------+---
---+------------------+----------
   405   RMEP_OK             39500   00:D0:95:E0:29:E8            2          1
false  LOCALLY_ASSIGNED   K6850-127-EP605

  1220   RMEP_OK             39400   00:16:4D:E0:FB:D6            2          1
false  none               none
K6850-128-EP605>
```

When port 1/18 of switch K2-127 goes down:

```
*A:7750-core-2->>config# show eth-cfm mep 1220 domain 4 association 22 all-remote-
mepids

===========================================================================
Eth-CFM Remote-Mep Table
===========================================================================
R-mepId Rx CC  Rx Rdi Port-Tlv If-Tlv Peer Mac Addr     CCM status since
---------------------------------------------------------------------------
405     True   False  Blocked  Down   00:d0:95:e0:29:e8 09/09/2010 03:27:39
605     True   True   Up       Up     00:e0:b1:79:de:70 09/12/2010 07:14:24
===========================================================================


K6850-128-EP605> show ethoam remote-endpoint domain 4 association Metro.0000022
endpoint 605
Total number of remote MEPs = 2
Legends: PortStatusTlv: 1 = psBlocked, 2 = psUp, 0 = psNoTlv
         InterfaceStatusTlv: 1 = ifUp, 2 = ifDown, 4 = ifUnknown, 7 =
ifLowerLayerDown

RMEP-ID      RMEP        OkFailed      Mac Address       Port Status  I/f Status
RDI      Chassis ID     Chassis ID
            Status         Time                              Tlv         Tlv
Value      Subtype         Value
-------+--------------+-------------+-----------------+-------------+----------+---
---+------------------+----------
   405   RMEP_OK             56100   00:D0:95:E0:29:E8            1          2
false  LOCALLY_ASSIGNED   K6850-127-EP605

  1220   RMEP_OK             56000   00:16:4D:E0:FB:D6            2          1
true   none               none
K6850-128-EP605>
```

References

- Chapter 45, “Ethernet OAM Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 43, “Configuring Ethernet OAM”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.22. SAA for ETHOAM

Service Assurance Agent (SAA) helps customers to provide service guarantees, increase network reliability by validating network performance, proactively identify network issues, and increase Return on Investment (ROI) by easing the deployment of new services. SAA uses active monitoring to generate traffic in a continuous, reliable, and predictable manner, thus enabling the measurement of network performance and health.

In addition to IP SAA, ETH-LB/DMM can be used to measure delay and jitter by sending out frames with DM information to the peer MEP and receiving frames with DM information from the peer MEP.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

This section contains configuration guidelines and additional information for the 6.4.3.R01 implementation of SAA for Ethernet OAM.

ETH-LB/DMM Timeout timer

This timer is started when response message with error code NO_ERROR is received from Ethoam CMM in reply of ETH-LB/DMM initiation request. The timer duration [(numPkts-1)*interPktDelay + 3] seconds (1sec timeout + 2sec to account for the IPC delay from Ethoam NI -> Ethoam CMM -> SAA CMM). If replies are received within this interval, the history statistics are updated; else the iteration is treated as failed.

Session timer

The session timer determines when the next iteration of the SAA will run. This timer is started when an SAA is started (i.e., the first iteration of the SAA is scheduled). For example, if an SAA is scheduled to start at 9:00pm and the current time is 6:00pm, the session timer will start at 9:00pm when the SAA is actually scheduled to start.

The timer value is configurable and measured in minutes. Its default value is 150 minutes.

Response timer

The response timer is started when SAA sends an ETH-LB/DMM initiation request to Ethoam CMM. The value of this timer is one second and specifies the maximum time for which SAA will wait for a reply from ETHOAM.

Scheduling of SAAs

A number of SAAs can be started at once, but only one SAA shall run at any instance of time. Subsequent SAAs shall be run one after the other. The SAA CMM task shall maintain a linked list to schedule the SAAs.

The SAAs shall be scheduled on a first come first serve basis, i.e. the SAA for which start request is received first shall be run first.

When a request to start an SAA is received, a node with the SAA information is inserted in the scheduler linked list. If this is the only node in the list, it is run immediately. If there are already some nodes in the scheduler linked list, then the node for this SAA is inserted at the tail of the linked list.

At the end of each iteration, the corresponding SAA node is removed from the scheduler linked list. If there are nodes remaining in the linked list, the one at the head is run. If there are no more nodes left, no action is taken.

Delay and Jitter Calculation

An aggregated record is kept for each SAA. This record retains the aggregated information of all the iterations that have been run. The record is updated after *n*th iteration as follows -

Minimum delay:

Minimum (delayOfIteration1, delayOfIteration2,......, delayOfIterationN)

Average delay:

[(avgDelayIteration1 * pktsRcvdInIteration1)+(avgDelayIteration2 * pktsRcvdInIteration2)+ . . . +(avgDelayIterationN * pktsRcvdInIterationN)] / (pktsRcvdInIteration1 + pktsRcvdInIteration2 + . . . + pktsRcvdInIterationN)

Maximum delay:

Maximum (delayOfIteration1, delayOfIteration2,......, delayOfIterationN)

Minimum jitter:

Minimum (jitterOfIteration1, jitterOfIteration2,...., jitterOfIterationN)

Average jitter:

[(avgJitterIteration1 * {pktsRcvdInIteration1 – 1}) + (avgJitterIteration2 * {pktsRcvdInIteration2 – 1}) + . . . + (avgJitterIterationN * {pktsRcvdInIterationN–1})] / ({pktsRcvdInIteration1–1}+{pktsRcvdInIteration2–1}+ . . . + {pktsRcvdInIterationN–1})

Maximum jitter:

Maximum (jitterOfIteration1, jitterOfIteration2,...., jitterOfIterationN)

Whenever a new iteration is run, the values are updated as:

Minimum delay:

Minimum (previousDelay, latestIterationDelay)

Average delay:

[(aggregatedDelay * totalNumPktsRcvd) + (latestIterationDelay * latestNumPktsRcvd)] / (totalNumPktsRcvd + latestNumPktsRcvd)

Maximum delay:

Maximum (previousDelay, latestIterationDelay)

Minimum jitter:

Minimum (previousJitter, latestIterationJitter)

Average jitter:

[{aggregatedJitter * (totalJitterSamplesRcvd)} + {latestIterationJitter * (latestNumPktsRcvd - 1)}] / {(totalJitterSamplesRcvd) + (latestPktsRcvd – 1)}

The totalJitterSamplesRcvd value is a count of the total number of jitter samples received. A jitter sample is obtained when at least two replies are received. For example, five replies received will provide four jitter samples in a single iteration.

Maximum jitter:

Maximum (previousJitter, latestIterationJitter)

For failed iterations, such as iterations in which no reply is received, the delay and jitter statistics are not updated. For iterations in which only one reply is received, only the delay statistics are updated, jitter statistics are not updated. In all other cases, for *n* replies received, *n* delays and *n*-1, jitter values are calculated and used for updating statistics.

Other Guidelines

- SAA for ETHOAM - Time-stamping is not available in hardware on all platforms. Therefore, time-stamping is done in software on all platforms.
- The SAA-ETHOAM operations use software based timestamps and hence do not provide precise measurement of network delay.
- Validation of PM family privileges is not supported in 6.4.3.R01 release. Hence, if a user only has bridging configuration privileges, the user is still able to configure SAA for IETH-LB or ETH-DMM.
- The network RTT includes the local software processing time in case the reply is received on a different NI.
- The behavior is undefined in case some SAAs are scheduled and the system time is changed. It is recommended that all the SAAs be rescheduled accordingly.
- Since the SAAs are identified by an alphanumeric name, it is not possible to support a range option in the SAA CLI commands.
- It is not possible to stop a SAA that has not started. Hence, scheduling an SAA to start and stop at a particular time is not possible.
- Advanced statistical measurements are not supported in the 6.4.3.R01 release.

Configuration Examples

Creating SAA:

```
→ saa saa1 description “saa for ip-ping”
→ saa saa2 description “saa for eth-lb” interval 160
→ saa saa3 description “saa for eth-dmm” interval 300
```

Configuring IP/ETH-LB/ETH-DMM SAA:

```
→ saa saa1 type ip-ping destination-ip 123.22.45.66 source-ip 123.35.42.125 type-of-
  service 5 inter-pkt-delay 1500 num-pkts 8 payload-size 1000
→ saa saa2 type ethoam-loopback target-endpoint 10 source endpoint 2 domain md1
  association ma1 vlan-priority 5 drop-eligible true inter-pkt-delay 500
→ saa saa3 type ethoam-two-way-delay target-endpoint 5 source endpoint 1 domain md2
  association ma2 vlan-priority 4 inter-pkt-delay 1000
```

Starting a SAA:

```
→ saa saa1 start
→ saa saa2 start at 2009-10-13,09:00:00.0
```

Stopping SAA:

```
→ saa saa1 stop
→ saa saa2 stop at 2009-10-13,10:00:00.0
```

Removing SAA:

```
→ no saa saa1
```

References

- Chapter 45, “Ethernet OAM Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 43, “Configuring Ethernet OAM”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.23. VPLS-MPLS

The OmniSwitch AOS implementation of MPLS provides the network architecture that is needed to provision a Virtual Private LAN Service (VPLS). VPLS allows multiple customer sites to transparently connect through a single bridging domain over an IP/MPLS-based network.

A VPLS-capable network consists of Customer Edges (CE), Provider Edges (PE), and a core MPLS network. The IP/MPLS core network interconnects the PEs but does not participate in the VPN functionality. Traffic is simply switched based on the MPLS labels.

MPLS forwarding is performed by routers called Label Switching Routers (LSRs). A Label Switched Path (LSP) is a path through one or more LSRs. There are two types of LSPs that are configurable using MPLS:

- **Signaled LSP**. The LSPs are set up using a signaling protocol, such as LDP. The signaling protocol allows the automatic assignment of labels from an ingress router to the egress router. Signaling is triggered by the ingress router, therefore configuration is only required on this router. A signaled LSP is confined to one gateway protocol area and, therefore, cannot cross an AS boundary.
- **Static LSPs**. A Static LSP specifies a statically defined path of LSRs. Configuration of label mappings and MPLS actions is required on each router that will participate in the static path. No signaling protocol, such as the Label Distribution Protocol (LDP), is required, and there is no dependence on a gateway protocol topology or local forwarding table. Static LSPs are able to cross an Autonomous System (AS) boundary.

In addition to static LSPs, a static Fast Reroute (FRR) feature is available that allows the configuration of backup static LSP tunnels. FRR uses these backup tunnels to provide alternate routes in the event an LSP goes down.

Platforms Supported

OmniSwitch 9000E

Guidelines

There were no changes or enhancements to the MPLS-VPLS feature in the 6.4.3.R01 release.

References

- Chapter 48, “MPLS LDP Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 49, “MPLS Static LSP and FRR Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 50, “Virtual Private LAN Service (VPLS) Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 45, “Configuring MPLS”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).
- Chapter 46, “Configuring VPLS”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.24. Internal DHCP Server

The Internal DHCP Server for the OmniSwitch is a port of the Vital QIP/QDHCP code (version 6.5). Compliant with RFC 2131, this implementation provides the ability to:

- Lease addresses for:
  - small offices
  - a management network
  - local phone services
- Enable or disable the DHCP server for the switch.
- Dynamically modify the DHCP configuration, using the **vi** editor, or through an accurately configured text file transferred to the switch.
- Restart the DHCP server.
- View the DHCP leases offered by the internal DHCP server.
- View the DHCP server statistics through the command line interface.
- Provide seamless DHCP server functionality during a takeover operation.
- Provide information used by the "out-of-the-box" Auto-configuration of an OmniSwitch.

The OmniSwitch DHCP server implementation makes use of the following policy, configuration, and server database files stored in the **/flash/switch** directory:

- **DHCP Template files**. The **dhcpd.conf.template** and **dhcpd.pcy.template** files contain the default configuration parameters and policy parameters respectively.
- **DHCP Policy file.** The **dhcpd.pcy** file initializes the global attributes for the DHCP server.
- **DHCP Configuration files.** The **dhcpd.conf** file is used to configure specific DHCP server settings on the switch, such as IP address ranges and options. The **dhcpd.conf.lastgood** file is a backup for the **dhcpd.conf** file.
- **DHCP Server Database file.** The **dhcpSrv.db** file is activated only during takeover and server restart of the DHCP server. It contains lease details of the client IP addresses.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

Consider the following when using the OmniSwitch Internal DHCP Server:

- Internal DHCP Server is supported only on the default VRF instance.
- Internal DHCP Server and DHCP Relay/Snooping are mutually exclusive; these features cannot run together in the default VRF instance.
- Only IPv4 is supported.
- Configuration on a per-VLAN basis is not supported.
- Configuring an IP interface is required.
- A maximum of 2K leases is recommended.
- If the **dhcpd.conf** file is edited, a restart of the DHCP Server (dhcp-server restart) is mandatory to reparse the modified configuration file.
- The larger the **dhcpd.conf** file size (file contains a lot of entries, such as with manual-dhcp), the more time it takes for the system to process the file.
- A DHCP server scope with Class A IP address subnets is not supported due to memory constraints.
- Configuring multiple DHCP server scopes for Class B IP address subnets will increase the amount of memory used. A single Class B scope reserves 2.3MB of memory.

Usage Notes

Use the following steps to initially configure the OmniSwitch Internal DHCP Server:

1. Copy the **dhcpd.conf.template** file to the **dhcpd.conf** and **dhcpd.conf.lastgood** files and the **dhcpd.pcy.template** file to **dhcpd.pcy**.
2. Edit the **dhcpd.conf** and **dhcpd.pcy** files as necessary.
3. Copy the **dhcpd.conf** and **dhcpd.pcy** files to the secondary CMM and any idle slots.
4. Reboot the switch.

Normal Operation and Switch Reboot

When the switch boots up or the DHCP Server is restarted, the DHCP Server application will parse both the **dhcpd.conf** and **dhcpd.pcy** files. The following scenarios may occur in regards to the configuring and processing of these files:

- If any error is detected during the parsing of the **dhcpd.conf** file, an error message is displayed indicating which keyword in the file triggered the error. When this occurs, the parsing of **dhcpd.conf** file will stop and the system will start parsing the **dhcpd.conf.lastgood** file, which is the configuration file the server will then use.
- If the **dhcpd.conf** file is parsed successfully, then the **dhcpd.conf.lastgood** file will be over written with the current configuration in the **dhcpd.conf** file. At this point, both files are identical as long as there is no change to the **dhcpd.conf** file contents.
- If any change is made to the **dhcpd.conf** file, restarting the DHCP Server (using the dhcp-server restart command) is required to parse the updated configuration. Restarting the server also automatically copies **dhcpd.conf**, **dhcpd.conf.lastgood** and **dhcpd.pcy** to the secondary CMM and other idle slots.
- If the DHCP Server is not restarted after any changes are made, these files are not synched to the CMM and other idle slots. In addition, there is no indication of this in the “show running-directory” command output.

- If both the **dhcpd.conf** and **dhcpd.conf.lastgood** files are corrupt, the DHCP Server will stop operating and an error message will flash on the switch console display. The error message displayed will indicate the necessary inputs required to correct the errors. For example, "invalid main key foo …". In this case, the errors must be fixed and the switch rebooted to continue DHCP Server operation.

Takeover

In the event of a CMM takeover, the probability for a DHCP Server failure is very low. The **dhcpd.conf**, **dhcpd.conf.lastgood** and **dhcpd.pcy** files are copied to the secondary CMM and to other NIs when the DHCP Server is restarted. As a result, any error rectified in the primary CMM configuration is applicable to the secondary CMM and other idle slots.

Types of Leases Allocated

- **Manual Bootp -** Allocates the address using the BootP config, where the MAC address is defined.

```
Syntax : manual-bootp 0f-ff-ff-ff-00-01 10.100.30.92 {
options list
      }
```

- **Automatic Bootp -** Allocates the address using the BootP config range, where the MAC address is not known.

```
Syntax : automatic-bootp range 10.252.0.200 10.252.0.206 {
option-list
      }
```

- **Automatic DHCP -** Allocates the address using the DHCP protocol, from a DHCP config range with an infinite lease (no lease specified).

```
Syntax : automatic-dhcp range 10.252.0.200 10.252.0.206 {
option-list
        }
```

- **Dynamic DHCP -** Defines the address using the DHCP protocol, from a DHCP config range for a specific lease time.

```
Syntax : dynamic-dhcp range 10.252.0.200 10.252.0.206 {
option-list
        }
```

- **Manual DHCP -** Allocates this address using the DHCP config, where the MAC address is defined.

```
Syntax : manual-dhcp 0f-ff-ff-ff-00-01 10.100.30.92 {
options list
       }
```

Configuration File Command Comparison

The following table provides a comparison of the configuration syntax used in the **dhcpd.conf** file with the commands used to configure the Cisco DHCP server:

Feature	Cisco	QIP	Supported in QIP	Remarks

IP address Exclusion	ip dhcp excluded-address 10.252.0.101 10.252.0.199	dynamic-dhcp range 10.252.0.1 10.252.0.100 userclass "pool0" {
		}
		dynamic-dhcp range 10.252.0.200 10.252.0.206 userclass "pool1" {
		}
			Yes	Achieved by breaking the scopes.

Address pool	ip dhcp pool 1
	network 172.16.0.0 /16
		subnet 10.252.0.0 netmask 255.255.0.0 {
		}
			Yes	No name can be associated with a pool

Address pool with secondary subnet	ip dhcp pool 1
	network 172.16.0.0 /16
	network 10.10.0.0 255.255.0.0 secondary
		There is no equivalent feature. However, there is the shared subnet feature, where a client request from any subnet in the shared network can be offered an address from any of the subnets in that shared network.
		Example:
		In the following example any client request which comes from any of the following subnets (10.100.x.0) would be assigned IP address from any of the subnet scopes.
		shared-network _10_100_20_0 {
		subnet 10.100.20.0 netmask 255.255.255.0 {
		manaul-dhcp 0f-ff-ff-ff-00-09 10.100.20.2 {
		option subnet-mask 255.255.255.0;
		option routers 192.168.32.1;
		option domain-name "quadritek.com";
		option dhcp-lease-time 4294967295;
		}
		}
		subnet 10.100.30.0 netmask 255.255.255.0 {
		manual-dhcp 0f-ff-ff-ff-00-01 10.100.30.92 {
		option subnet-mask 255.255.255.0;
		option routers 10.100.30.1;
		option domain-name "quadritek.com";
		option dhcp-lease-time 4294967295;
		}
		dynamic-dhcp range 10.100.30.81 10.100.30.85 {
		option subnet-mask 255.255.255.0;
		option domain-name "quadritek.com";
		option routers 10.100.30.1;
		option dhcp-lease-time 60000;
		}
		}
		subnet 10.100.50.0 netmask 255.255.255.0 {
		dynamic-dhcp range 10.100.50.51 10.100.50.55 userclass "macprefix" {
		option subnet-mask 255.255.255.0;
		option domain-name "quadritek.com";
		option routers 10.100.50.1;
		option dhcp-lease-time 60000;
		}
		}
		}
			Yes	When the primary subnet (172.16/16) is exhausted, the addresses are assigned from the secondary subnet (10.10/16)

	ip dhcp pool 1
	host 172.16.2.254
	hardware-address 02c7.f800.0422
	ieee802
	client-name Mars
		manual-dhcp 0f-ff-ff-ff-00-01 10.100.30.92 {
		}
			Yes	IP address is statically assigned instead of automatically assigned from a pool

Static mapping	More than one Static bindings are kept in a file which DHCP server reads	The static mappings are all defined in the dhcpd.conf file; there is no separate file.
		Example:
		manual-dhcp 0f-ff-ff-ff-00-01 10.100.30.92 {
		option subnet-mask 255.255.255.0;
		option routers 192.168.32.1;
		option domain-name "quadritek.com";
		option dhcp-lease-time 4294967295;
		}
		manual-dhcp 0f-ff-ff-ff-00-02 10.100.30.93 {
		option subnet-mask 255.255.255.0;
		option routers 192.168.32.1;
		option domain-name "quadritek.com";
		option dhcp-lease-time 4294967295;
		}
			No	

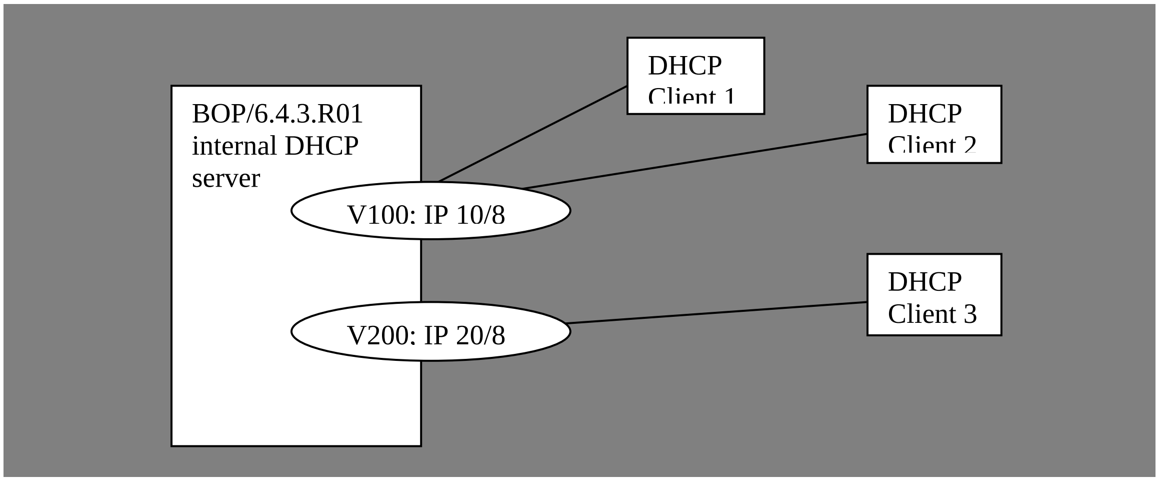
Address allocation using option-82	ip dhcp class CLASS1 relay agent information relay-information hex 01030a0b0c02050000000123	QIP has a similar feature that is going to be released in the next point release of the DHCP 5.6 server, but it is not GA at this time. This new point release must go through QA testing before it is released. There is no planned date for this testing at this time, however.	No Available in next QIP release	This feature is designed to allow the Cisco IOS DHCP server to use option 82 information to help determine which IP addresses to allocate to clients. The information sent via option 82 is used to identify on which port the DHCP request was received. This feature does not parse out the individual sub-options contained within option 82. Instead, the address allocation is done by matching a configured pattern byte by byte.
Domain name	domain-name cisco.com	option domain-name "quadritek.com";	Yes	
DNS	dns-server 172.16.1.102 172.16.2.102	option domain-name-servers 172.16.1.102, 172.16.2.102;	Yes	
Default router	default-router 172.16.1.100 172.16.1.101	option routers 10.100.30.1;	Yes	
Lease time	lease 30	option dhcp-lease-time 4294967295;	Yes	
NetBios	netbios-name-server 172.16.1.103 172.16.2.103	option netbios-name-servers 172.16.1.103, 172.16.2.103	Yes	
Netbios node type	netbios-node-type h-node	option netbios-node-type H-node; (choices are B-node, P-node, M-node, H-node)	Yes	
Ping attempts	ip dhcp ping packets 5	PingAttempts=5 in dhcpd.pcy file	Yes	Uses policy file instead of dhcpd.conf
Ping	ip dhcp ping timeout 850	PingDelay=850 (milliseconds)	Yes	Uses policy file instead of dhcpd.conf

Ignore all bootp packets	ip dhcp bootp ignore	QIP comment: There is no exact match here, but in VitalQIP, the DHCP server parameter SupportBootp, which defaults to TRUE, controls whether QIP writes Static bootp and automatic bootp objects into the dhcpd.conf file. When set to FALSE, the bootp object types are NOT included in the dhcpd.conf. Note that this is a VitalQIP parameter, NOT a DHCP server policy – it does not get pushed to the server in dhcpd.pcy, but rather controls what VitalQIP pushes in the .conf file.	No	

Configuration Examples

This implementation of the OmniSwitch Internal DHCP Server supports the following topologies:

Topology 1: DHCP Server Directly Connected to DHCP Clients (L3 model)



In this topology the clients and server are connected over a Layer 3 subnet. The DHCP server uses the incoming IP interface subnet on which the client packet is received to search the scope and assign an IP address to the clients. For example:

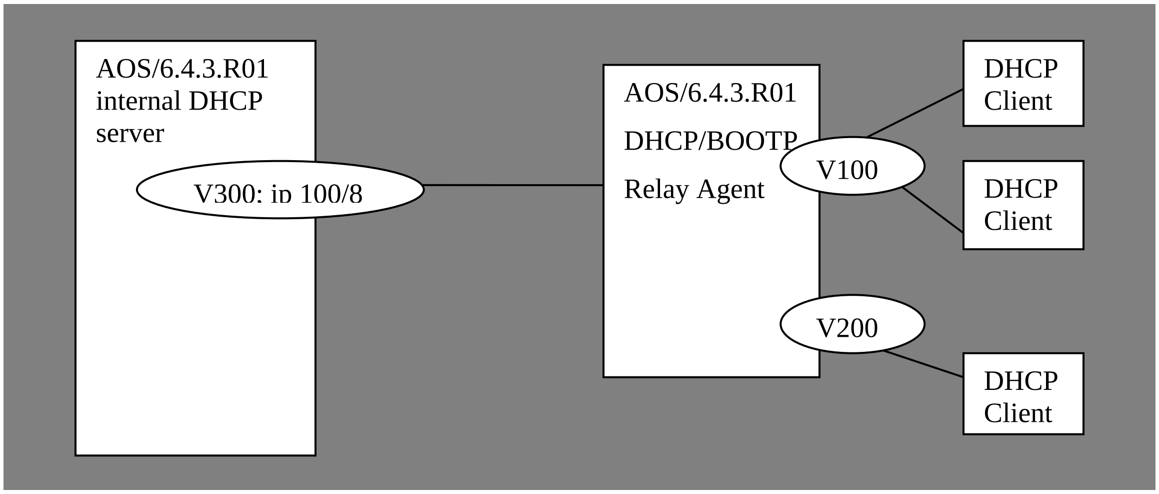
- DHCP Client 1 and Client 2 connect to the switch through VLAN 100, which is configured with the 10.0.0.1/8 IP interface. As a result, both clients will get a lease from the 10.0.0.0/8 scope.
- DHCP Client 3 connects to the switch through VLAN 200, which is configured with the 20.0.0.1/8 IP interface. As a result, Client 3 will get a lease from the 20.0.0.0/8 scope.

If multiple IP interfaces are configured for the same VLAN (IP multinetting), then the IP address subnet of the primary interface is used to determine the scope.

A DHCP client will not get a lease if:

- There are no IP interfaces configured.
- There is no scope defined for the IP interface subnet in which the client requests are received.

Topology 2: DHCP Server Connected to DHCP Clients via DHCP Relay Agent (L2/L3 model)



In this topology,

- The clients are connected to a relay agent via a L2 network.
- The relay agent routes the packets received from the clients to the server network.
- The server uses the giaddr value present in the DHCP discover packet to determine the scope from which an IP address is assigned to the client.

In order to fine tune the lease assignment, the relay agent can be configured to include option 82 header in the packet sent to the server. The option 82 header includes the circuit ID (similar to port), which the server can use to assign more granular leases to the clients.

Sample Configuration

The OmniSwitch Internal DHCP Server provides several **show** commands that display the server configuration and statistics. This subsection provides examples of such displays:

```
→ show dhcp-server leases

Total leases: 8

IP Address        MAC address        Lease Granted         Lease Expiry          Type
-----------------+------------------+---------------------+---------------------+-----------
200.0.1.1         00:00:01:b8:91:3f  DEC 15 14:10:59 2009  DEC 19 01:30:59 2009  DYNAMIC
200.0.1.2         00:00:01:b8:91:37  DEC 15 14:11:05 2009  DEC 19 01:31:05 2009  DYNAMIC
200.0.1.3         00:00:01:b8:91:3b  DEC 15 14:11:48 2009  DEC 19 01:31:48 2009  DYNAMIC
200.0.1.4         00:00:01:b8:91:3d  DEC 15 14:11:53 2009  DEC 19 01:31:53 2009  DYNAMIC
220.0.0.2         00:00:01:1d:4f:7e  DEC 15 14:11:45 2009  DEC 15 22:31:45 2009  DYNAMIC
220.0.0.3         00:00:01:5a:0b:76  DEC 15 14:12:00 2009  DEC 15 22:32:00 2009  DYNAMIC
220.0.0.4         00:00:01:1d:4f:7d  DEC 15 14:11:53 2009  DEC 15 22:31:53 2009  DYNAMIC
120.0.0.4         00:00:02:12:4f:8c  DEC 15 14:11:53 2009  DEC 15 23:31:53 2009  STATIC

→ show dhcp-server leases ip-address 200.0.1.2

IP Address        MAC address        Lease Granted         Lease Expiry          Type
-----------------+------------------+---------------------+---------------------+-----------
00:00:01:b8:91:37  DEC 15 14:11:05 2009  DEC 19 01:31:05 2009  DYNAMIC

→ show dhcp-server leases mac-address 00:00:01:1d:4f:7d
```

```
IP Address        MAC address        Lease Granted         Lease Expiry          Type
-----------------+------------------+---------------------+---------------------+-----------
220.0.0.4         00:00:01:1d:4f:7d  DEC 15 14:11:53 2009  DEC 15 22:31:53 2009  DYNAMIC

→ show dhcp-server leases type static

Total leases: 1

IP Address        MAC address        Lease Granted         Lease Expiry          Type
-----------------+------------------+---------------------+---------------------+-----------
120.0.0.4         00:00:02:12:4f:8c  DEC 15 14:11:53 2009  DEC 15 23:31:53 2009  STATIC

→ show dhcp-server statistics
General:
  DHCP Server Name          : mample.vitalqip.com,
  DHCP Server Status               : Enabled,
  Total Subnets Managed            : 7,
  Total Subnets Used               : 2,
  Total Subnets Unused             : 5,
  Total Subnets Full               : 0,
  DHCP Server System Up Time       : TUE DEC 15 14:10:27.9956
    Lease DB Sync:
      DB Sync time (in sec) : 60,
      Last sync time               : TUE DEC 15 14:21:34 2009,
      Next sync time               : TUE DEC 15 14:22:34 2009

→ show dhcp-server statistics packets
Packets:
  Total DHCP Discovers             : 12,
  Total DHCP Offers                : 12,
  Total DHCP Requests              : 16,
  Total DHCP Request Grants        : 10,
  Total DHCP Request Renews        : 6,
  Total DHCP Declines              : 0,
  Total DHCP Acks                  : 16,
  Total DHCP Nacks                 : 0,
  Total DHCP Releases              : 0,
  Total DHCP Informs               : 0,
  Total Bootp requests             : 0,
  Total Bootp response             : 0,
  Total Unknown packets            : 0

→ show dhcp-server statistics leases
Leases:
    Total:
      Leases Managed        : 1365,
      Leases used                  : 7,
      Leases unused                : 1358,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Static DHCP:
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Dynamic DHCP:
      Leases Managed               : 1365,
      Leases used                  : 7,
      Leases unused                : 1358,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Automatic DHCP:
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Static Bootp:
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Automatic Bootp                :
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0

→ show dhcp-server statistics subnets
Subnets:
    Subnet1:
      Subnet                : 200.0.0.0,
      Total                        : 1022,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 1022,
      Automatic DHCP               : 0,
      Static Bootp          : 0,
      Automatic Bootp              : 0
        Ranges:
          Start                    : 200.0.1.1,
          End                      : 200.0.2.255,
          Mask                     : 255.255.253.0,
          Type                     : 5
          Used                     : 4,
          Unused                   : 507,
          Pending                  : 0,
          Unavailable              : 0
    Subnet2:
      Subnet                : 220.0.0.0,
      Total                        : 508,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 508,
      Automatic DHCP               : 0,
      Static Bootp                 : 0,
      Automatic Bootp              : 0
        Ranges:
          Start                    : 220.0.0.2,
          End                      : 220.0.0.255,
       Mask                 : 255.255.255.0,
          Type                     : 5
             Unused                : 251,
     Used                   : 3,
     Pending                : 0,
     Unavailable            : 0
    Subnet3:
      Subnet                : 150.0.0.0,
      Total                        : 400,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 400,
      Automatic DHCP               : 0,
      Static Bootp                 : 0,
      Automatic Bootp              : 0
        Ranges:
          Range1:
            Start                  : 150.0.1.1,
            End                    : 150.0.1.100,
               Mask                : 255.255.255.0,
            Type                   : 5,
            Used                   : 0,
            Unused          : 100,
            Pending         : 0,
            Unavailable            : 0
          Range2:
            Start                  : 150.0.2.1,
            End                    : 150.0.2.100,
            Mask                   : 255.255.255.0,
            Type                   : 5,
            Unused          : 100,
            Used                   : 0,
            Pending         : 0,
            Unavailable             : 0
 Subnet4:
      Subnet                : 50.0.0.0,
      Total                 : 200,
      Static DHCP                   : 0,
      Dynamic DHCP          : 200,
      Automatic DHCP        : 0,
      Static Bootp          : 0,
      Automatic Bootp               : 0
        Ranges:
          Start                     : 50.0.1.1,
          End                       : 50.0.1.100,
          Mask                      : 255.255.255.0,
          Type                      : 5,
          Unused                    : 100,
          Used                      : 0,
          Pending                   : 0,
          Unavailable               : 0

→ show dhcp-server statistics all
General:
  DHCP Server Name          : mample.vitalqip.com,
  DHCP Server Status                : Enabled,
  Total Subnets Managed             : 7,
  Total Subnets Used                : 2,
  Total Subnets Unused              : 5,
  Total Subnets Full                : 0,
  DHCP Server System Up Time        : TUE DEC 15 14:10:27.9956
    Lease DB Sync:
      DB Sync time (in sec) : 60,
      Last sync time                : TUE DEC 15 14:21:34 2009,
      Next sync time                : TUE DEC 15 14:22:34 2009
Packets:
  Total DHCP Discovers              : 12,
  Total DHCP Offers                 : 12,
  Total DHCP Requests               : 16,
  Total DHCP Request Grants         : 10,
  Total DHCP Request Renews         : 6,
  Total DHCP Declines               : 0,
  Total DHCP Acks                   : 16,
  Total DHCP Nacks                  : 0,
  Total DHCP Releases               : 0,
  Total DHCP Informs                : 0,
  Total Bootp requests              : 0,
  Total Bootp response              : 0,
  Total Unknown packets             : 0
Leases:
    Total:
      Leases Managed        : 1365,
      Leases used                   : 7,
      Leases unused                 : 1358,
      Leases Pending                : 0,
      Leases unavailable            : 0
    Static DHCP:
      Leases Managed                : 0,
      Leases used                   : 0,
      Leases unused                 : 0,
      Leases Pending                : 0,
      Leases unavailable            : 0
    Dynamic DHCP:
      Leases Managed                : 1365,
      Leases used                   : 7,
      Leases unused                 : 1358,
      Leases Pending                : 0,
      Leases unavailable            : 0
    Automatic DHCP:
      Leases Managed                : 0,
      Leases used                   : 0,
      Leases unused                 : 0,
      Leases Pending                : 0,
      Leases unavailable           : 0
    Static Bootp:
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0
    Automatic Bootp                :
      Leases Managed               : 0,
      Leases used                  : 0,
      Leases unused                : 0,
      Leases Pending               : 0,
      Leases unavailable           : 0
Subnets:
    Subnet1:
      Subnet                : 200.0.0.0,
      Total                        : 1022,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 1022,
      Automatic DHCP               : 0,
      Static Bootp          : 0,
      Automatic Bootp              : 0
        Ranges:
          Start                    : 200.0.1.1,
          End                      : 200.0.2.255,
          Mask                     : 255.255.253.0,
          Type                     : 5
          Used                     : 4,
          Unused                   : 507,
          Pending                  : 0,
          Unavailable              : 0
    Subnet2:
      Subnet                : 220.0.0.0,
      Total                        : 508,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 508,
      Automatic DHCP               : 0,
      Static Bootp                 : 0,
      Automatic Bootp              : 0
        Ranges:
          Start                    : 220.0.0.2,
          End                      : 220.0.0.255,
       Mask                 : 255.255.255.0,
          Type                     : 5
             Unused                : 251,
     Used                   : 3,
     Pending                : 0,
     Unavailable            : 0
    Subnet3:
      Subnet                : 150.0.0.0,
      Total                        : 400,
      Static DHCP                  : 0,
      Dynamic DHCP                 : 400,
      Automatic DHCP               : 0,
      Static Bootp                 : 0,
      Automatic Bootp              : 0
        Ranges:
          Range1:
            Start                  : 150.0.1.1,
            End                    : 150.0.1.100,
               Mask                : 255.255.255.0,
            Type                   : 5,
            Used                   : 0,
            Unused          : 100,
            Pending         : 0,
            Unavailable            : 0
          Range2:
            Start                  : 150.0.2.1,
            End                    : 150.0.2.100,
            Mask                   : 255.255.255.0,
            Type                   : 5,
            Unused          : 100,
            Used                   : 0,
            Pending         : 0,
            Unavailable            : 0
 Subnet4:
      Subnet                : 50.0.0.0,
      Total                 : 200,
      Static DHCP                  : 0,
      Dynamic DHCP          : 200,
      Automatic DHCP        : 0,
      Static Bootp          : 0,
      Automatic Bootp              : 0
        Ranges:
          Start                    : 50.0.1.1,
          End                      : 50.0.1.100,
          Mask                     : 255.255.255.0,
           Type                    : 5,
          Unused                   : 100,
          Used                     : 0,
          Pending                  : 0,
          Unavailable              : 0
```

References

- Chapter 20, “DHCP Relay Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 24, “Configuring DHCP Server”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.25. DHCP Client

In the previous AOS implementation, a DHCP Client was only allowed on default VLAN 1 of the switch at boot up time. The 6.4.3.R01 release enhances this capability to support the following functionality:

- The DHCP Client is configurable on any one VLAN in any VRF instance, not just the default VLAN. This is done by configuring a DHCP Client IP interface for the designated VLAN.
- The **ip interface** command now includes a **dhcp-client** parameter to activate the DHCP Client functionality for the VLAN on which the interface is configured. In addition, new **release** and **renew** parameters were added to allow, when necessary, a manual release and renew of the DHCP Client IP address for the switch.
- A permanent DHCP client. The IP lease received by the switch is renewed and rebound according to RFC 2131. The renewal process for a permanent DHCP client keeps the entry active in the DHCP database and provides an indirect way of monitoring the availability of the switch.
- Configurable string for the option-60 field that is sent as part of the DHCP discover/request packets.
- DHCP option-12 support for retrieving the system name.

When the switch receives a valid IP address lease from a DHCP server:

- The IP address and the subnet mask (DHCP option 1) are assigned to the DHCP Client IP interface.
- A default static route is created according to DHCP option 3 (Router IP Address).
- The lease is periodically renewed and rebound according to the renew time (DHCP option 58) and rebind time (DHCP option 59) returned by the DHCP Server. If the lease cannot be renewed within the lease time (DHCP option 51) returned by the DHCP Server, the lease will be released. When not specified by the DHCP Server, a default lease time of 7 days is allocated.
- The system name is set according to the hostname (DHCP option 12) returned by the DHCP Server. This applies only when the default system name has not been changed. Once set by DHCP option 12, the system name can be saved in the boot.cfg, but will not change the running configuration status. Setting the system name in this manner is mainly used for the initial DHCP cycle, when the switch boots with no configured system name. Note that when an IP lease is released, the system name is not modified.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, and 9000E

Guidelines

- Previous commands for configuring a DHCP Client on the switch (**ip helper boot-up** and **ip helper boot-up enable**) have been deprecated and replaced with the **ip interface dhcp-client** command. However, the switch will accept a configuration file that contains the deprecated commands and will enable the DHCP client on the VLAN 1 interface. Upon the next write memory, the old CLI command will be replaced with the new **ip interface dhcp-client** command.
- In the previous implementation of this feature, the IP interface does not age out even after the IP address lease time expires. However, with the 6.4.3.R01 version, the DHCP Client will initiate the IP address lease extension period and update the IP address if there is any change. If the DHCP Client is not able to renew or get an IP address after the DHCP server assigned lease time expires, the switch will remove the default static route and will render the DHCP client interface inactive. At this point, a DHCP renew is required to start the DHCP client process again.
- The DHCP Client-enabled IP address serves as the primary IP address when multiple addresses are configured for a VLAN.

- Make sure the DHCP server is reachable through the DHCP Client VLAN.
- Setting the DHCP Server address lease time to 5 minutes is recommended.
- When a DHCP release is performed or the DHCP client interface is deleted, any default static route added for the client is also removed and the corresponding timers (such as release/renew timer) are cancelled. However, the system name will remain unchanged even if the name was updated using the DHCP client option-12 information.
- The previous implementation of this feature sent DHCP discover packets only from the first available slot. However, in the 6.4.3.R01 release, UDP relay will send DHCP packets on all available NIs. DHCP discover packets are flooded on all the ports that are members of the DHCP Client-enabled VLAN.

Reload and Takeover (dhcpClient.db file)

The following information is stored in the **dhcpClient.db** file located in the **/flash/switch** directory on the switch:

- DHCP server assigned IP
- VLAN information
- Subnet mask
- Router IP address
- Checksum value (validates the integrity of the file).

The **dhcpClient.db** file is used during a switch reload or CMM takeover. The IP address saved in this file is the address requested from the DHCP server in the event of a reload or takeover. The checksum value is used to validate the integrity of the file.

Whenever there is any change in the DHCP server assigned IP address, the **dhcpClient.db** file is updated with the new information and synchronized to the secondary CMM. This file is also synchronized at a periodic time interval of 5 minutes along with the DHCP snooping binding table. This synchronization shall not affect the switch synchronization status.

During takeover the new CMM shall use the **dhcpClient.db** file and try to acquire the same IP address again. The DHCP client shall try to send the BootP packets only after the NI is back up. For example, after a takeover:

- The DHCP client interface uses the **dhcpClient.db** file information to create the IP interface with a lease time of 10 minutes and try to acquire the same IP address.
- After successful renewal of the IP address, the lease time is modified as per the DHCP server assigned IP address.
- If the DHCP client is not able acquire the same IP address, the client will then try to get a new IP address after the switch-assigned DHCP lease time expires. Note that a trap message is sent whenever there is any change to the IP address.

Other Guidelines

- The IP address of a DHCP-Client interface is not configurable; this address is assigned through the DHCP Client process of requesting an IP address.
- DHCP Client only supports IPv4 addresses.
- When using this feature in a stack configuration, enable MAC Retention to ensure that the same IP address is obtained from the DHCP server after takeover.

- Do not configure the DHCP client interface on a switch where the interface will be the relay agent for the client VLAN.
- Although a DHCP Client is configurable for any VLAN in any VRF instance, only one VLAN DHCP Client per switch is allowed.

References

- Chapter 14, "IP Commands", and Chapter 20, "DHCP Relay Commands", *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 16, "Configuring IP", and Chapter 23, "Configuring DHCP Relay", *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

2.26. Out-of-the-Box Auto-Configuration

The Out-of-the-Box Auto-Configuration (automatic remote configuration download) capability automates and simplifies the deployment of large network installations eliminating the need for manual configuration of each device. It also ensures that each device is compliant with the centrally controlled device configuration policies and firmware revisions.

This feature is exclusive to the OmniSwitch and provides the ability to automate the following tasks for a newly deployed switch:

- OmniSwitch firmware upgrades.
- Configuration of a switch on bootup when the switch is first connected to the network.
- Download and installation of critical configuration bootup and image files.

These tasks are automated through the use of an instruction file. This file provides the necessary information a newly deployed OmniSwitch requires to download any necessary firmware upgrades or obtain a switch configuration without user intervention.

The auto-configuration download process automatically invokes the following actions necessary for a switch to gain network connectivity and access to the instruction file:

1. Configuration of a DHCP Client interface for VLAN 1 when the switch initially boots up.
2. Obtain an IP lease (IP address, mask, default gateway, and system name), the address of a TFTP file server, and the name of the instruction file from a reachable DHCP server.
3. Download the instruction file, which contains the information to obtain the configuration file, image files and/or script files from the given TFTP, FTP or SCP servers.
4. Download and apply the image and configuration file.
5. Reboot the switch with the upgraded image files and switch configuration file, or if no images or boot configuration is downloaded, scripted instructions are executed on the fly and the switch is made available remotely.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

The automatic configuration download process is triggered when the following occurs:

- There is no **boot.cfg** file found in the Working directory of the switch.
- The initialization process of the switch is complete and the network interfaces or ports are ready.
- There is connectivity with a DHCP server and a TFTP file server. This connectivity is mandatory.

Network Components

The following network components are required to support Out-of-the-Box Auto-Configuration:

- **DHCP server**. Any DHCP server can be used; this requirement is not based on a specific vendor. Using a Class C IP address pool is highly recommended; using Class B will consume a high amount of switch memory. For more information, see the **DHCP Server Configuration Example**.
- **Network gateway or router**. The switch must be able to reach the gateway that the DHCP server will provide to the switch. Depending on the topology, specific gateway configuration is required. This could be either UDP relay or VLAN configuration.
- **TFTP file server**. A reachable TFTP server is required to transfer the instruction file to the switch. Firmware and configuration files may also reside on this server.
- **Primary FTP/SFTP server**. Based on the instruction file, this server stores a certain number of files required to be transferred to the switch, such as firmware, configuration files, debug files, and a script file.
- **Secondary FTP/SFTP server (optional).** In case the primary FTP/TFTP server fails or the files are not available, auto remote config tries to download the files from the secondary server.

Download Files

The following types of files are used during the auto-configuration process:

Instruction file

This file is the initial file required for the automatic-configuration process to occur. The instruction file is stored in the TFTP server with the **.alu** extension and provides the names and location of servers and download files needed to complete the auto-configuration process.

When the switch obtains the instruction file, it compares the firmware version in the file with the version that is running on the switch. If both versions are the same, the firmware download is skipped. If the versions are different, the switch downloads the image files from the location provided in the instruction file.

This feature is able to distinguish the image files required for different platforms. Therefore, image files for different switch platforms can reside in the same firmware location.

For more information, see the **Instruction File Example.**

Firmware upgrade files

Firmware (image) files differ depending on the OmniSwitch platform. These files contain executable code, which provides support for the system, Ethernet ports, and network functions.

Configuration file

The configuration file (**boot.cfg**) is stored in the FTP/SFTP server. The auto-configuration process downloads this file from the server and reboots the switch. The configuration file is then applied to the switch after boot up. Any error in the configuration file will be reflected in the **.err** file in /**flash**. In the instruction file, the configuration file may have any name, but when the file is transferred to the switch, it will be placed in the **/flash/working** directory with the **boot.cfg** name.

Debug Configuration file

During the auto-configuration process, the debug configuration file is downloaded with the filename **AlcatelDebug.cfg**. The debug file is accessed to check the errors that occur during download process. All errors that occur during the automatic remote configuration download process are displayed on the switch command prompt and also stored in the switch log (**swlog.log** file). The switch is automatically rebooted when new firmware or configuration file has been downloaded or if the script file contains a reload command. However, if the debug file is the only file to be downloaded, then the switch is not automatically rebooted.

Script file

The script file is downloaded and stored with the same name in the **/flash/working** directory. During the auto-configuration process:

- The script file is downloaded and implemented.
- The DHCP client configured on VLAN 1 is removed.
- The commands in the script file are run on the switch in the order specified.

The script file contains the commands to be implemented on the switch after the configuration file, if available, is applied, or the script file can be used to configure the switch dynamically without a **boot.cfg** file. The following provides an example of a script file:

```
→vlan 100 enable name "VLAN 100"
→vlan 100 port default 1/1
→write memory
→reload working no rollback-timeout
```

The main purpose of script file is to configure a set of switches. In addition, the script file can have commands that are not desired in the configuration file, such as **write memory** and **reload working no rollback-timeout**. The auto-configuration process confirms that such command actions are completed.

Consider the following guidelines when using script files:

- A **write memory** command issued by a script file will override the **boot.cfg** file that was downloaded from the FTP/SFTP server. Therefore, make sure that a script file with such command is not downloaded along with a **boot.cfg** file.
- After the script file is downloaded, the auto-configuration process will not automatically reload the switch unless such a command exists in the script file.
- If both **write memory** and **reload working no rollback-timeout** are in the script file, after bootup, the switch will have only the configuration based on the script file contents.
- If **write memory** does not exist in the script file, only **reload working no rollback-timeout**, the switch state after bootup depends on whether or not the **boot.cfg** file was downloaded.
- If downloading the **boot.cfg** file was included in the instruction file, then the switch comes back up with the configuration file based on the **boot.cfg** file. However, if no **boot.cfg** file was downloaded or exists in **/flash/working**, when the switch comes back up the auto-configuration process will be initiated.

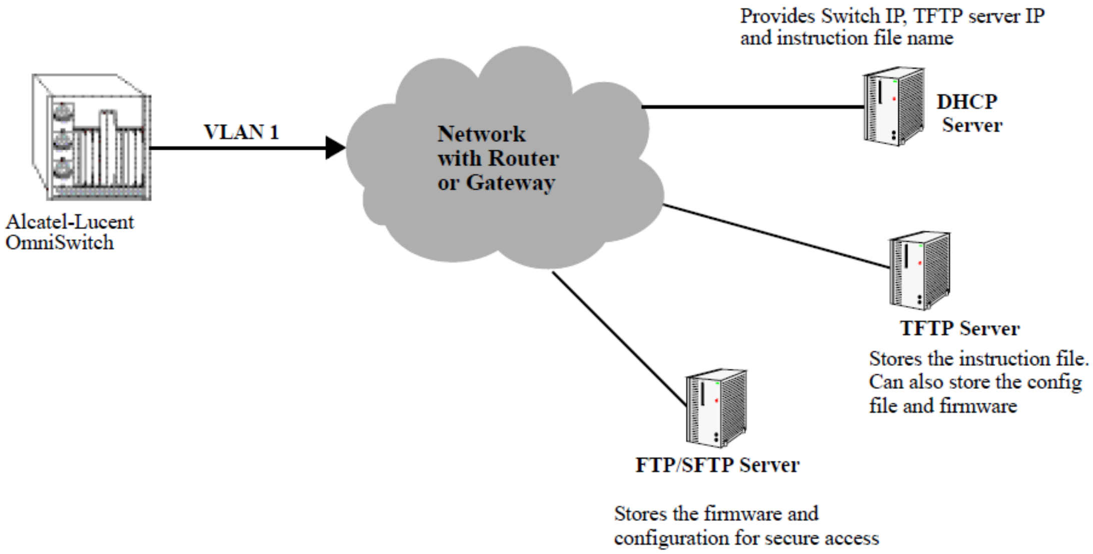
Manual Configuration

In the event auto-configuration fails and user-intervention is needed, the procedure for manual download and installation of component files is as follows:

1. Download the required files, which are present on the FTP/SFTP servers, or directly transfer files to the OmniSwitch using Zmodem.
2. Download the image files for firmware upgrade.
3. When the **boot.cfg** file is downloaded individually, the auto-configuration process takes over and initiates a **reload working no rollback-timeout**.
4. If a script file is downloaded along with the **boot.cfg** file, then the auto-configuration process runs the commands in the script file.
5. If required, perform a **write memory** and **reload working no-rollback timeout** command on the switch to install all the downloaded files and reload the switch.

Configuration Examples

The following illustration shows a basic auto-configuration network setup:


Provides Switch IP, TFTP server IP and instruction file name
VLAN 1
Network with Router or Gateway
DHCP Server
Alcatel-Lucent OmniSwitch
TFTP Server
Stores the instruction file. Can also store the config file and firmware
FTP/SFTP Server
Stores the firmware and configuration for secure access

DHCP Server Configuration Example

The DHCP server provides the following information to the switch:

- The IP address of the network gateway or router.
- The TFTP file server address.
- The name and location of the Instruction file.
- A dynamic IP address for the OmniSwitch (valid only for initial bootup process).

The following is an example of a DHCP server configuration for one subnet:

```
ddns-update-style interim;
ignore client-updates;
subnet 10.255.204.0 netmask 255.255.255.0 {
}
```

```
subnet 172.17.3.0 netmask 255.255.255.0 {
range 172.17.3.20 172.17.3.21;
option subnet-mask              255.255.255.0;
option broadcast-address        192.168.1.255;
option routers          172.17.3.254;
option domain-name              localdomain;
option tftp-server-name         10.255.204.100;
option bootfile-name          instruction_1_garuda.alu
option time-offset              -18000;
}
```

Instruction File Example

The instruction file provides the following information:

- Firmware version and file location.
- Configuration file name and location.
- Debug configuration file name and location.
- Script file name and location.
- Primary FTP/SFTP file server address / type / username.
- Secondary FTP/SFTP file server address / type / username.

The following is an example of an instruction file:

```
! Firmware version
Firmware version:OS_6_4_3_477_R01
Firmware location:/tftpboot/firmware
! Configuration file
Config filename:garuda_1.cfg
Config location:/tftpboot/configuration
! Debug file
Debug filename:None
Debug location:None
! Script file
Script location:None
Script filename:None
! Primary File Server
Primary server:10.255.204.100
Primary protocol:FTP
Primary user:tftptest
! Secondary File Server
Secondary server:None
Secondary protocol:None
Secondary user:None
```

Consider the following guidelines when configuring the instruction file:

- The instruction file is case sensitive and can contain only the keywords provided in the example above.
- The keywords can be placed in any order.
- If the *keyword*:*value* format is incorrect, the information on that line is discarded.
- The firmware version must be provided in the format as specified in the example.
- The pathnames provided must contain the complete path to the file location. Maximum length of the pathname is 255 characters, filename is 63 characters.
- If any file is not required, enter “None” for the keyword value. For example, if a debug configuration file download is not required, the following instruction file syntax is used:

```
Debug filename:None
Debug location:None
```

- The header line is the first line of the instruction file and begins with “!” character. Header line contents are logged to the switch log along with the other contents of the instruction file.
- The “!” character is also used to designate a comment line or to disable an option line. The header and comment lines begin with “!” character.
- For the FTP/TFTP server entries, the username and password are the same. The username should not exceed 16 characters. This is a limitation for the server, not the OmniSwitch.
- The instruction file must have the **.alu** extension. The instruction file is not downloaded if it does not include the **.alu** extension.
- If an error or failure occurs during the file transfer, the transfer process is retried up to 3 times.
- If file transfer and download fails, the automatic remote configuration process is stopped. This condition requires user intervention.
- All contents of the instruction file are stored in the switch log (**swlog.log**) file as evidence of the last auto-configuration download.

References

- Chapter 8, “Managing Automatic Remote Configuration Download”, *OmniSwitch AOS Release 6 Switch Management Guide* (060215-10, Rev. N).

2.27. RADIUS Service-Type Attribute

With RADIUS authentication, the Service-Type attribute in the Access Request Message is used by the RADUIS server to distinguish between different request types. This allows the server to forward the request to other servers (LDAP, Active Directory, etc.).

The Access Request packets sent from the OmniSwitch to the RADIUS server now contain a Service-Type attribute. The value of this attribute is set to **Framed-User (2)** when the authentication type is 802.1x supplicant, Captive Portal or ASA, and **Call-Check (10)** for MAC-based authentication or non-supplicant.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

The Service-Type attribute is not user-configurable and is automatically selected by the RADIUS client according to the authentication request type. No configuration changes are required on either the switch or the RADIUS server.

Access Request Packet Examples

Here are two RADIUS authentication packets: one with the Call-Check Service-Type and the other with the Framed-User Service-Type. The Service-Type field appears at the end of each example and is highlighted in green.

Packet 1 – Call-Check Service-Type

```
Frame 1 (118 bytes on wire, 118 bytes captured)
    Arrival Time: Feb  9, 2010 17:43:34.055829000
    [Time delta from previous captured frame: 0.000000000 seconds]
    [Time delta from previous displayed frame: 0.000000000 seconds]
    [Time since reference or first frame: 0.000000000 seconds]
    Frame Number: 1
    Frame Length: 118 bytes
    Capture Length: 118 bytes
    [Frame is marked: False]
    [Protocols in frame: eth:ip:udp:radius]
Ethernet II, Src: Alcatel-_9d:43:08 (00:d0:95:9d:43:08), Dst: DellComp_71:a2:54
(00:b0:d0:71:a2:54)
    Destination: DellComp_71:a2:54 (00:b0:d0:71:a2:54)
        Address: DellComp_71:a2:54 (00:b0:d0:71:a2:54)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory
default)
    Source: Alcatel-_9d:43:08 (00:d0:95:9d:43:08)
        Address: Alcatel-_9d:43:08 (00:d0:95:9d:43:08)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory
default)
    Type: IP (0x0800)
Internet Protocol, Src: 172.17.4.12 (172.17.4.12), Dst: 10.255.204.227
(10.255.204.227)
    Version: 4
    Header length: 20 bytes
    Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
        0000 00.. = Differentiated Services Codepoint: Default (0x00)
        .... ..0. = ECN-Capable Transport (ECT): 0
        .... ...0 = ECN-CE: 0
```

```
      Total Length: 104
      Identification: 0x6dfe (28158)
      Flags: 0x00
          0... = Reserved bit: Not set
          .0.. = Don't fragment: Not set
          ..0. = More fragments: Not set
      Fragment offset: 0
      Time to live: 63
      Protocol: UDP (0x11)
      Header checksum: 0x8587 [correct]
          [Good: True]
          [Bad : False]
      Source: 172.17.4.12 (172.17.4.12)
      Destination: 10.255.204.227 (10.255.204.227)
  User Datagram Protocol, Src Port: iad1 (1030), Dst Port: radius (1812)
      Source port: iad1 (1030)
      Destination port: radius (1812)
      Length: 84
      Checksum: 0xe444 [correct]
          [Good Checksum: True]
          [Bad Checksum: False]
  Radius Protocol
      Code: Access-Request (1)
      Packet identifier: 0x2 (2)
      Length: 76
      Authenticator: 7D67213F012248404FD762D73DE44717
      Attribute Value Pairs
          AVP: l=14  t=User-Name(1): 00B0D04CB64E
              User-Name: 00B0D04CB64E
          AVP: l=18  t=User-Password(2): Encrypted
              User-Password: \265\215\376H\r\212\250\b|\232\225\231\267<\307\226
          AVP: l=6  t=NAS-IP-Address(4): 172.17.4.12
              NAS-IP-Address: 172.17.4.12 (172.17.4.12)
          AVP: l=6  t=NAS-Port(5): 3014
              NAS-Port: 3014
          AVP: l=6  t=NAS-Port-Type(61): Async(0)
              NAS-Port-Type: Async (0)
          AVP: l=6  t=Service-Type(6): Call-Check(10)
              Service-Type: Call-Check (10)
```

Packet 2 – Framed-User Service-Type

```
  Frame 34 (171 bytes on wire, 171 bytes captured)
      Arrival Time: Feb  9, 2010 19:08:31.678438000
      [Time delta from previous captured frame: 0.253038000 seconds]
      [Time delta from previous displayed frame: 0.253038000 seconds]
      [Time since reference or first frame: 5097.622609000 seconds]
      Frame Number: 34
      Frame Length: 171 bytes
      Capture Length: 171 bytes
      [Frame is marked: False]
      [Protocols in frame: eth:ip:udp:radius:eap]
  Ethernet II, Src: Alcatel-_9d:43:08 (00:d0:95:9d:43:08), Dst: DellComp_71:a2:54
  (00:b0:d0:71:a2:54)
      Destination: DellComp_71:a2:54 (00:b0:d0:71:a2:54)
          Address: DellComp_71:a2:54 (00:b0:d0:71:a2:54)
          .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
          .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
      Source: Alcatel-_9d:43:08 (00:d0:95:9d:43:08)
          Address: Alcatel-_9d:43:08 (00:d0:95:9d:43:08)
          .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
          .... ..0. .... .... .... .... = LG bit: Globally unique address (factory
  default)
      Type: IP (0x0800)
```

```
Internet Protocol, Src: 172.17.4.12 (172.17.4.12), Dst: 10.255.204.227
(10.255.204.227)
    Version: 4
    Header length: 20 bytes
    Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
        0000 00.. = Differentiated Services Codepoint: Default (0x00)
        .... ..0. = ECN-Capable Transport (ECT): 0
        .... ...0 = ECN-CE: 0
    Total Length: 157
    Identification: 0x7473 (29811)
    Flags: 0x00
        0... = Reserved bit: Not set
        .0.. = Don't fragment: Not set
        ..0. = More fragments: Not set
    Fragment offset: 0
    Time to live: 63
    Protocol: UDP (0x11)
    Header checksum: 0x7edd [correct]
        [Good: True]
        [Bad : False]
    Source: 172.17.4.12 (172.17.4.12)
    Destination: 10.255.204.227 (10.255.204.227)
User Datagram Protocol, Src Port: iad1 (1030), Dst Port: radius (1812)
    Source port: iad1 (1030)
    Destination port: radius (1812)
    Length: 137
    Checksum: 0x1455 [correct]
        [Good Checksum: True]
        [Bad Checksum: False]
Radius Protocol
    Code: Access-Request (1)
    Packet identifier: 0xc (12)
    Length: 129
    Authenticator: 777E78540E80103051C73C9F36421AFF
    Attribute Value Pairs
        AVP: l=7  t=User-Name(1): user1
            User-Name: user1
        AVP: l=6  t=NAS-IP-Address(4): 172.17.4.12
            NAS-IP-Address: 172.17.4.12 (172.17.4.12)
        AVP: l=16  t=State(24): 5342522D434820343836357C3200
            State: 5342522D434820343836357C3200
        AVP: l=6  t=NAS-Port(5): 3014
            NAS-Port: 3014
        AVP: l=6  t=NAS-Port-Type(61): Unknown(1970496882)
            NAS-Port-Type: Unknown (1970496882)
        AVP: l=14  t=Calling-Station-Id(31): 00b0d04cb64e
            Calling-Station-Id: 00b0d04cb64e
        AVP: l=30  t=EAP-Message(79) Last Segment[1]
            EAP fragment
            Extensible Authentication Protocol
                Code: Response (2)
                Id: 2
                Length: 28
                Type: MD5-Challenge [RFC3748] (4)
                Value-Size: 16
                Value: D6D43717D14F8D076B0AEF51F983C7C3
                Extra data (6 bytes): 757365723100
        AVP: l=18  t=Message-Authenticator(80): 0939E4FCB0BB987730C82E8EE84C8BF1
            Message-Authenticator: 0939E4FCB0BB987730C82E8EE84C8BF1
        AVP: l=6  t=Service-Type(6): Framed-User(2)
            Service-Type: Framed-User (2)
```

References

- Chapter 34, "Managing Authentication Servers", *OmniSwitch AOS Release 6 Switch Management Guide* (060215-10, Rev. N).

2.28. IP Managed Interface

Currently, and by default, most applications that run on IP use the egress IP interface address as the source IP, while using a socket to communicate with a peer/ server. However, it may be desirable to have some applications use a specific source IP for the packets that are sent out using the socket.

This release provides the ability to configure a permanent source IP interface that is used when sending packets. The source IP interface can be the Loopback0 address or an existing IP interface on the switch. If a managed IP interface is not defined for an application, the application uses the egress IP interface address as the source IP.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

A source IP interface is configurable for the following applications within the specified VRF context:

Application	Default Source Interface	VRF Support
ASA Authentication Server		
LDAP-Server	Loopback0 if configured, otherwise outgoing interface	NO – Server can only be set in the default VRF
TACACS+	Outgoing interface	
RADIUS	Loopback 0 if configured, otherwise outgoing interface	YES – Can be configured with any VRF-ID (configuration only available in default VRF)
AAA Authentication Server		
RADIUS	Loopback 0 if configured, otherwise outgoing interface	YES – Can be configured with any VRF-ID (configuration only available in default VRF)
Switch Management Applications		
SNMP (includes traps)	Loopback 0 if configured, otherwise outgoing interface	NO – Servers/stations can only be set in the default VRF
SFLOW	Loopback 0 if configured, outgoing IP otherwise	
NTP	Loopback 0 if configured, otherwise outgoing interface	
SYSLOG	Outgoing interface	
DNS	Outgoing interface	
Switch Access and Utilities **(ping and traceroute command can specify a source address as an optional parameter)**		
Telnet	Outgoing interface	YES – Can be initiated in any VRF
FTP	Outgoing interface	NO – Can only be initiated in default VRF
SSH Includes scp sftp	Outgoing interface	YES – Can be initiated in any VRF
TFTP	Outgoing interface	NO – Can only be initiated in default VRF

CLI Commands and Examples

The following two new CLI commands were added to support this feature:

[no] ip managed-interface <*interface name*> **application [ldap-server] [tacacs] [radius] [snmp] [sflow] [ntp] [syslog] [dns] [telnet] [ftp] [ssh] [tftp] [all]**

show ip managed-interface

Here is a sample output from the new **show** command:

```
→ show ip managed-interface

Legend: "-" denotes no explicit configuration

 Application       Interface-Name
-----------------+---------------------
 tacacs            -
 sflow             -
 ntp               -
 syslog            -
 dns               -
 telnet            -
 ssh               -
 tftp              -
 ldap-server       -
 radius            vlan-172
 snmp              -
 ftp               -
```

The following CLI commands were deprecated with the introduction of the new **ip managed-interface** command:

aaa radius agent preferred [<*interface IP address***>|<non-loopback>|<default>]**

sflow agent ip <*ip-address*>

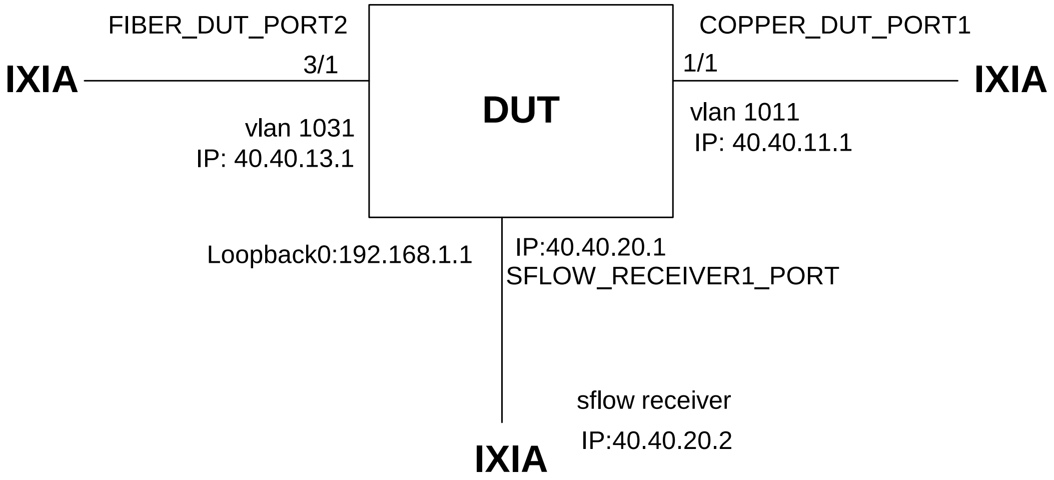
ntp src-ip preferred [<*interface IP address***>|<non-loopback>|<default>]**

snmp source ip preferred [default] [no-loopback0] [*ipv4address***]**

Configuration Examples

This section provides a sample configuration in which the source interface is configured and tested for the sFlow feature.

Topology



CLI Configuration

Configure VLANs, assign to a port and IP:

```
→vlan 1011 enable name "VLAN 1011"
→vlan 1011 port default 1/1
→ip interface "v1011" address 40.40.11.1 mask 255.255.255.0 vlan 1011 ifindex 2

→vlan 1020 enable name "VLAN 1020"
→vlan 1020 port default 1/13
→ip interface "v1020" address 40.40.20.1 mask 255.255.255.0 vlan 1020 ifindex 4

→vlan 1031 enable name "VLAN 1031"
→vlan 1031 port default 3/2
→ip interface "v1031" address 40.40.13.1 mask 255.255.255.0 vlan 1031 ifindex 3
```

Create loopback0 IP interface and managed interface for sFlow feature:

```
→ip interface "Loopback0" address 192.168.1.1
→ip managed-interface Loopback0 application sflow
```

Configure sFlow Receiver, Sampler and Poller:

```
→sflow receiver 1 name receiver1 address 40.40.20.2 udp-port 6475 packet-size 1400
 version 5 timeout 0
→sflow sampler 1 1/1 receiver 1 rate 1024 sample-hdr-size 128
→sflow sampler 1 3/2 receiver 1 rate 2048 sample-hdr-size 128
→sflow poller 1 1/1 receiver 1 interval 15
→sflow poller 1 3/2 receiver 1 interval 15
```

Verify the configuration:

```
→ show sflow receiver 1

 Receiver 1
 Name       = receiver1
 Address    = IP_V4  40.40.20.2
 UDP Port   = 6475
 Timeout    = No Timeout
 Packet Size= 1400
 DatagramVer= 5

→ >show sflow sampler

Instance  Interface  Receiver   Rate   Sample-Header-Size
----------------------------------------------------------
   1      1/ 1          1       1024          128
   1      3/ 2          1       2048          128

→ show sflow poller

Instance  Interface  Receiver   Interval(Secs)
----------------------------------------------
   1        1/ 1        1          15
   1        3/ 2        1          15
```

The managed interface IP and agent IP shows the configured IP address.

```
→ show ip managed-interface

Legend: "-" denotes no explicit configuration

 Application       Interface-Name
-----------------+----------------------
 tacacs             -
 sflow             Loopback0
 ntp                -
 syslog             -
 dns                -
 telnet             -
 ssh                -
 tftp               -
 ldap-server        -
 radius             -
 snmp               -
 ftp                -

→ show sflow agent

 Agent Version  = 1.3; Alcatel; 6.1.1
 Agent IP       = 192.168.1.1
```

Configuration Snapshot

```
→ show configuration snapshot ip vlan pmm
! VLAN :
vlan 1011 enable name "VLAN 1011"
vlan 1011 port default 1/1
vlan 1020 enable name "VLAN 1020"
vlan 1020 port default 1/13
vlan 1031 enable name "VLAN 1031"
vlan 1031 port default 3/2
! VLAN SL:
! IP :
debug ip set ipedrArpPoisonLrn 1
ip service all
ip interface "v1011" address 40.40.11.1 mask 255.255.255.0 vlan 1011 ifindex 2
ip interface "v1031" address 40.40.13.1 mask 255.255.255.0 vlan 1031 ifindex 3
ip interface "v1020" address 40.40.20.1 mask 255.255.255.0 vlan 1020 ifindex 4
ip interface "Loopback0" address 192.168.1.1
ip managed-interface Loopback0 application sflow
! VLAN AGG:
! Port mirroring :
sflow receiver 1 name receiver1 address 40.40.20.2 udp-port 6475 packet-size 1400
version 5 timeout 0
sflow sampler 1 1/1 receiver 1 rate 1024 sample-hdr-size 128
sflow sampler 1 3/2 receiver 1 rate 2048 sample-hdr-size 128
sflow poller 1 1/1 receiver 1 interval 15
sflow poller 1 3/2 receiver 1 interval 15
! VLAN STACKING:
```

References

- Chapter 14, “IP Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).

2.29. XNI-U12E

- Four Hi-Gig+ ports are used to interface to the Switching Fabric via a backplane connection.
- Other ports are configured to support twelve SFP+ pluggable transceivers, with each SFP+ port capable of running 10-Gig Ethernet traffic over multi-mode or single-mode fiber via LC-type connectors. If all 12 ports are active, this card will be 6:1 oversubscribed with single CMM.
- The total throughput of the module is designed to be 24 Gbps to the chassis fabric. Throughput will double to 48 Gbps in a chassis on an enhanced backplane with dual CMM equipped with 4 HIGIG+ links to the fabric. In this case the card will be 3:1 oversubscribed.

Platforms Supported

OmniSwitch 9800E, 9700E

Guidelines

The XNI-U12E blade can be mixed with other OmniSwitch 9000E blades with the following conditions:

- The XNI-U12E NI is only compatible with the OS9000E CMM. With the OS9000 CMM, the NI should remain DOWN.
- The XNI-U12E blade, BCM 56822, does not support the MPLS modules. If a chassis has the MPLS module, the XNI-U12E NI should remain DOWN.

References

- Chapter 5, “Network Interface (NI) Modules”, *OmniSwitch 9000/9000E Series Hardware Users Guide* (060211-10, Rev. K).

3. Existing Feature Guidelines

This section contains information and guidelines for the following OmniSwitch features that were introduced or enhanced in a previous release:

- **Spanning Tree**
- **Ethernet Ring Protection**
- **Ethernet Ring Protection**

Ethernet Ring Protection (ERP) is a protection switching mechanism for Ethernet ring topologies, such as multi-ring and ladder networks. This implementation of ERP is based on Recommendation ITU-T G.8032 and uses the ring Automatic Protection Switching (APS) protocol to coordinate the prevention of network loops within a bridged Ethernet ring.

Loop prevention is achieved by allowing the traffic to flow on all but one of the links within the protected Ethernet ring. This link is blocked and is referred to as the Ring Protection Link (RPL). When a ring failure condition occurs, the RPL is unblocked to allow the flow of traffic to continue through the ring.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- The total number of virtual ports supported for the ERP feature is 8K (OmniSwitch 9000E, 6850, and 6855) and 2.5K (OmniSwitch 6400)
- Maximum number of rings supported per node is 4. The ring ID is unique to a ring.
- ERP convergence of 50ms or less is only supported on the fiber 1Gig and 10 Gig ports and is not supported on the copper and combo ports. OmniSwitch 6855 stacking links can be used for the 10gig connections.

References

- Chapter 9, “Ethernet Ring Protection Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 10, “Configuring ERP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).
- General L3 Routing
- **General IPMS and Multicast Routing**.

3.1. Spanning Tree

The Alcatel-Lucent Spanning Tree implementation provides support for the Q2005 version of MSTP, the 802.1w Rapid Spanning Tree Algorithm and Protocol (RSTP) and the 802.1D Spanning Tree Algorithm and Protocol (STP). All three supported protocols ensure that there is always only one data path between any two switches for a given Spanning Tree instance to prevent network loops.

Q2005 (MSTP) is only available when the flat mode is active for the switch. The flat mode applies a single spanning tree instance across all VLAN port connections on a switch. MSTP allows the configuration of Multiple Spanning Tree Instances (MSTIs) in addition to the CST instance. Each MSTI is mapped to a set of VLANs. As a result, flat mode can support the forwarding of VLAN traffic over separate data paths.

802.1D STP and 802.1w RSTP are available in both the flat and 1x1 mode. However, when using 802.1D or 802.1w in the flat mode, the single spanning tree instance per switch algorithm applies.

Note that 802.1w is now the default Spanning Tree protocol for the switch regardless of which mode is active. This default value will apply to future releases as well.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- Maximum Number of STP Instances:
  - ➢ Flat Mode: STP/RSTP – 1 Instance, MSTP – 1 CIST and 16 MST Instances
  - ➢ 1x1 Mode: STP/RSTP –256 Instances
- Root bridge priority / path cost
- Default Spanning Tree mode is RSTP (802.1w)
- The bridge priority can be any value between 0 and 65535 for STP and RSTP in the 16-bit mode. By default, Spanning Tree follows the 16-bit path cost.
- The bridge priority can only be in multiples of 4096 in the 32-bit mode or in MSTP mode.
- MSTP can support 32 bit-mode per standard.
- Changing the STP protocol to MSTP will reset all priority and path cost of a bridge to the default values.
- Up to 128 Link Aggregates are supported with a maximum of 256 aggregated ports.
- The default port path costs for IEEE Std 802.1D-1998- 16 Bit are:

Port Speed	Path cost
10M	100
100M	19
1000 M	4
10000 M	3

- The default port path costs are for IEEE Std. 802.1Q-2005 32 Bit are:

Port Speed	Path cost

10M	2000000
100M	200000
1000 M	20000
10000 M	2000

- The default link aggregation path costs for 16 Bit are:

Linkagg speed	Linkagg size	Path cost
10M	2	60
	4	40
	8	30
100M	2	12
	4	9
	8	7
1000M	N/A	3
10000M	N/A	2

- The default link aggregation path costs for 32 Bit are:

LinkAgg speed	LinkAgg size	Path cost
10M	2	1200000
	4	800000
	8	600000
100M	2	120000
	4	80000
	8	60000
1000M	2	12000
	4	8000
	8	6000
10000M	2	1200
	4	800
	8	600

The path cost depends on the configured LinkAgg size and not on the active port size.

References

- Chapter 7, “Distributed Spanning Tree Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 7, “Using 802.1Q 2005 Multiple Spanning Tree”, and Chapter 8, “Configuring Spanning Tree Parameters”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

3.2. Ethernet Ring Protection

Ethernet Ring Protection (ERP) is a protection switching mechanism for Ethernet ring topologies, such as multi-ring and ladder networks. This implementation of ERP is based on Recommendation ITU-T G.8032 and uses the ring Automatic Protection Switching (APS) protocol to coordinate the prevention of network loops within a bridged Ethernet ring.

Loop prevention is achieved by allowing the traffic to flow on all but one of the links within the protected Ethernet ring. This link is blocked and is referred to as the Ring Protection Link (RPL). When a ring failure condition occurs, the RPL is unblocked to allow the flow of traffic to continue through the ring.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

- The total number of virtual ports supported for the ERP feature is 8K (OmniSwitch 9000E, 6850, and 6855) and 2.5K (OmniSwitch 6400)
- Maximum number of rings supported per node is 4. The ring ID is unique to a ring.
- ERP convergence of 50ms or less is only supported on the fiber 1Gig and 10 Gig ports and is not supported on the copper and combo ports. OmniSwitch 6855 stacking links can be used for the 10gig connections.

References

- Chapter 9, “Ethernet Ring Protection Commands”, *OmniSwitch CLI Reference Guide* (060218-10, Rev. N).
- Chapter 10, “Configuring ERP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).

3.3. General L3 Routing

This section provides general guidelines related to L3 routing.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

OmniSwitch 6850, 6855, 9000, 9000E Guidelines

- OS6850/6855 can learn up to 20K routes in software routing table, the CPU usage is far below the limit, but will cause the memory usage to be above the limitation and the memory will not be released even after all the configurations are deleted. Only a switch reboot will release the occupied memory.
- OS9000 can learn more than 40K routes in software, and still keep the CPU and memory usage below limit.
- OS9000E software routing table can support up to 96K IPv4 entries learnt through 2 gateways (without any traffic), or can support up to 70K entries learnt through 70 gateways (without any traffic).
- The OS9000E software routing table can support up to 36K IPv6 entries learnt through two gateways (without any traffic).
- One OS9000/9000E can support up to 65K BGP software routing table entries. OS9000 can release the memory when the BGP routes are withdrawn without rebooting. While learning 65K routes, the CPU and Memory usage are in normal range.
- Single unit or stack of OS6850 can support up 32 BGP peers, 65K BGP software routing table entries, but they can not release the memory when the BGP routes are withdrawn without rebooting. While learning 65K routes, the CPU usage is normal, but the Memory usages are normally above 94%.
- BGP graceful restart is supported on platform OS9000/9000E/OS6850

OmniSwitch 6400 Guidelines

- The OmniSwitch 6400 only supports the routing protocol RIP and RIPng.
- Up to 1024 RIP routes, 10 RIP peers, and 512 destinations with a maximum of 4 next hops RIP ECMP.

References

- Chapter 16, “Configuring IP”, Chapter 18 “Configuring IPv6”, Chapter 20 “Configuring RIP”, *OmniSwitch AOS Release 6 Network Configuration Guide* (060217-10, Rev. N).
- *OmniSwitch AOS Release 6 Advanced Routing Configuration Guide* (060216-10, Rev. N).

3.4. General IPMS and Multicast Routing

IP Multicast routing and switch control plane traffic are done in the software path, such as IGMP, MLD, PIM Join/Prune, Hello – DVMRP, PIM-DM, PIM-SSM, etc.

IP Multicast will be forwarded in Hardware mode in switched or routed networks, except register packets and DVMRP tunneling Packets.

Platforms Supported

OmniSwitch 6400, 6850, 6855, 9000, 9000E

Guidelines

L3 IPMC Replication

- Guidelines for the replication of multicast traffic when egressing multiple VLANs on multiple ports are as follows:
  - 1 egress port, 1021 flows can be replicated to at most 128 VLANs on that port.
  - 2 egress ports, 1021 flows can be replicated to at most 64 VLANs on those ports.
  - 12 egress ports, 170 flow can be replicated to at most 64 VLANs on those ports.
  - 24 egress ports, 85 flows can be replicated to at most 64 VLANS on those ports.
- Rule of thumb calculation for maximum number of supported simultaneous forwarded flows that can be replicated up to 64 VLANs is 2048 divided by the number of egress ports, if the number of egress ports is greater than 2 (e.g., 2048/12 == 170). The underlying hardware supports at most 24 ports on a single ASIC, so a 48-port NI will have 2 ASICs of 24 ports each.
- Number of (s,g) and (*,g) flows Global level:
  - OS6850,OS6855, OS6400, OS9000: 1021 flows.
  - OS9702: 4093 flows.

References

- N/A

4. Interoperability Test Results

This section provides the results of interoperability testing that was conducted between the specified OmniSwitch platforms and various Cisco and Juniper platforms. Note the following regarding these test results:

- An interoperability matrix is provided for each Cisco and Juniper platform tested. The matrix specifies the software features tested and the vendor model tested.
- The OmniSwitch platforms tested were running AOS Release 6.4.3.R01. The exception to this is the OmniSwitch 6800, which was running 6.3.1.R01.
- A “Yes” indicates that the interoperability test between the OmniSwitch and the other vendor platform was successful for the specified feature.
- An “N/S” indicates that interoperability between the OmniSwitch and the other vendor platform is not supported for the specified feature.
- An “N/A” indicates that the feature itself is not supported on the specified platform.

OmniSwitch Interoperability with Cisco 2800 Router, 12.4 (3g)

<table>
<tr><th>Feature</th><th>Cisco 2800
12.4 (3g)</th><th>OS9000</th><th>OS9000E</th><th>OS6850</th><th>OS6400</th><th>OS6855</th><th>OS6855-
U24X</th><th>OS6800</th></tr>
<tr><td>OSPF V2</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
<tr><td>RIP I</td><td colspan="8">Yes</td></tr>
<tr><td>RIP-II</td><td colspan="8">Yes</td></tr>
<tr><td>Static routes</td><td colspan="8">Yes</td></tr>
<tr><td>ISIS</td><td colspan="4">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>BGP 4</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
<tr><td>MPBGP ext. V6</td><td>N/S</td><td colspan="3">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>VRF</td><td>Yes</td><td>N/S</td><td>Yes</td><td>N/S</td><td>N/S</td><td>N/S</td><td>Yes</td><td>N/S</td></tr>
<tr><td>GRE</td><td colspan="8">Yes</td></tr>
<tr><td>GRE in VRF context</td><td>Yes</td><td>N/S</td><td>Yes</td><td>N/S</td><td>N/S</td><td>N/S</td><td>Yes</td><td>N/S</td></tr>
</table>

OmniSwitch Interoperability with Cisco 2600 Router, 12.3 (13) (fc2)

<table>
<tr><th>Feature</th><th>Cisco 2800
12.3 (13)
(fc2)</th><th>OS9000</th><th>OS9000E</th><th>OS6850</th><th>OS6400</th><th>OS6855</th><th>OS6855-
U24X</th><th>OS6800</th></tr>
<tr><td>IPV4</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet
Q-tag link</td><td colspan="8">Yes</td></tr>
<tr><td>OSPF V2</td><td colspan="4">Yes</td><td>N/A</td><td colspan="3">Yes</td></tr>
<tr><td>RIP I</td><td colspan="8">Yes</td></tr>
<tr><td>RIP-II</td><td colspan="8">Yes</td></tr>
<tr><td>Static routes</td><td colspan="8">Yes</td></tr>
<tr><td>ISIS</td><td>N/S</td><td colspan="3">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>BGP 4</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
<tr><td>MPBGP ext. V6</td><td colspan="4">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>VRF</td><td>Yes</td><td>N/S</td><td>Yes</td><td>N/S</td><td>N/S</td><td>N/S</td><td>Yes</td><td>N/S</td></tr>
<tr><td>GRE</td><td colspan="8">Yes</td></tr>
<tr><td>GRE in VRF context</td><td>Yes</td><td>N/S</td><td>Yes</td><td>N/S</td><td>N/S</td><td>N/S</td><td>Yes</td><td>N/S</td></tr>
<tr><td>OSPF V3</td><td colspan="4">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>RIPng</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>IPV6</td><td colspan="7">Yes</td><td>N/S</td></tr>
</table>

OmniSwitch Interoperability with Cisco 3750 Router, 12.2 (40) (SE)

<table>
<tr><th>Feature</th><th>Cisco 3750 12.2 (40) (SE)</th><th>OS9000</th><th>OS9000E</th><th>OS6850</th><th>OS6400</th><th>OS6855</th><th>OS6855-U24X</th><th>OS6800</th></tr>
<tr><td>Auto-negotiation on twisted pair 10/100/1000</td><td colspan="8">Yes</td></tr>
<tr><td>10/100 Q-tag link interoperability</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (STATIC)</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (STATIC) with 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (LACP)</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (LACP) 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>(1x1) Spanning Tree 802.1W (Rapid PVST)</td><td colspan="8">Yes</td></tr>
<tr><td>IPV4</td><td colspan="8">Yes</td></tr>
<tr><td>IPV6 neighbor discovery, ICMPV6</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>IPV6 static routing</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>RIP I</td><td colspan="8">Yes</td></tr>
<tr><td>RIP-II</td><td colspan="8">Yes</td></tr>
<tr><td>IPV6 MLD</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>IGMP V1/V2/V3</td><td colspan="8">Yes</td></tr>
<tr><td>PIM</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
</table>

OmniSwitch Interoperability with Cisco 2900XL Router, 12.0(5)WC11

<table>
<tr><th>Feature</th><th>Cisco 2900XL 12.0(5)WC11</th><th>OS9000</th><th>OS9000E</th><th>OS6850</th><th>OS6400</th><th>OS6855</th><th>OS6855-U24X</th><th>OS6800</th></tr>
<tr><td>10/100 Q-tag link</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (STATIC)</td><td colspan="8">Yes</td></tr>
<tr><td>Single Spanning Tree 802.1D</td><td colspan="8">Yes</td></tr>
</table>

OmniSwitch Interoperability with Cisco 6509 Router, IOS 12.2(33)SXI

<table>
<tr><th>Feature</th><th>Cisco 6509 IOS 12.2(33)SXI</th><th>OS9000</th><th>OS9000E</th><th>OS6850</th><th>OS6400</th><th>OS6855</th><th>OS6855-U24X</th><th>OS6800</th></tr>
<tr><td>Auto-negotiation on twisted pair 10/100/1000</td><td colspan="8">Yes</td></tr>
<tr><td>Auto-negotiation on Gigabit Ethernet</td><td colspan="8">Yes</td></tr>
<tr><td>10/100 Q-tag link interoperability</td><td colspan="8">Yes</td></tr>
<tr><td>Gigabit Q-tag link interoperability</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (STATIC)</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (STATIC) 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (LACP)</td><td colspan="8">Yes</td></tr>
<tr><td>Fast Ethernet Channel (LACP) 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>Giga Ethernet Channel (STATIC)</td><td colspan="8">Yes</td></tr>
<tr><td>Giga Ethernet Channel (STATIC) 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>Giga Ethernet Channel (LACP)</td><td colspan="8">Yes</td></tr>
<tr><td>Giga Ethernet Channel (LACP) 802.1Q</td><td colspan="8">Yes</td></tr>
<tr><td>(1x1) Spanning Tree 802.1D</td><td colspan="8">Yes</td></tr>
<tr><td>(1x1) Spanning Tree 802.1W</td><td colspan="8">Yes</td></tr>
<tr><td>Single Spanning Tree 802.1D</td><td colspan="8">Yes</td></tr>
<tr><td>Single Spanning Tree 802.1W</td><td colspan="8">Yes</td></tr>
<tr><td>RIPng</td><td colspan="7">Yes</td><td>N/A</td></tr>
<tr><td>Single Spanning Tree 802.1S</td><td colspan="8">Yes</td></tr>
<tr><td>IPV4</td><td colspan="8">Yes</td></tr>
<tr><td>IPV4 Static routing</td><td colspan="8">Yes</td></tr>
<tr><td>IPV6 neighbor discovery,ICMPV6</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>IPV6 static routing</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>RIP I</td><td colspan="8">Yes</td></tr>
<tr><td>RIP-II</td><td colspan="8">Yes</td></tr>
<tr><td>IGMP V1/V2/V3</td><td colspan="8">Yes</td></tr>
<tr><td>BGP 4</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
<tr><td>OSPF V2</td><td colspan="4">Yes</td><td>N/S</td><td colspan="3">Yes</td></tr>
<tr><td>OSPF V3</td><td colspan="4">Yes</td><td>N/S</td><td colspan="2">Yes</td><td>N/S</td></tr>
<tr><td>RIPng</td><td colspan="7">Yes</td><td>N/S</td></tr>
<tr><td>BGP4/RIP/OSPFV2 in VRF context</td><td>Yes</td><td>N/S</td><td>Yes</td><td>N/S</td><td>N/S</td><td>N/S</td><td>Yes</td><td>N/S</td></tr>
</table>

MVRP	Yes	NA

OmniSwitch Interoperability with Juniper M20, JUNOS 8.1R4.3

Feature	**Juniper M20 8.1R4.3**	**OS9000**	**OS9000E**	**OS6850**	**OS6400**	**OS6855**	**OS6855-U24X**	**OS6800**
10/100 Q-tag link	Yes							
IPV4	Yes							
OSPF V2	Yes				N/S	Yes		
RIP I	Yes							
RIP-II	Yes							
Static routes	Yes							
BGP 4	Yes				N/S	Yes		
BGP4/RIP/OSPFv2 in VRF context	Yes	N/S	Yes	N/S	N/S	N/S	Yes	N/S