OAM and SAA

In This Chapter

This chapter provides information about the Operations, Administration and Management (OAM) and Service Assurance Agent (SAA) commands available in the CLI for troubleshooting services.

Topics in this chapter include:

OAM Overview

Delivery of services requires a number of operations occur properly and at different levels in the service delivery model. For example, operations such as the association of packets to a service, VC-labels to a service and each service to a service tunnel must be performed properly in the forwarding plane for the service to function properly. In order to verify that a service is operational, a set of in-band, packet-based Operation, Administration, and Maintenance (OAM) tools is required, with the ability to test each of the individual packet operations.

For in-band testing, the OAM packets closely resemble customer packets to effectively test the customer's forwarding path, but they are distinguishable from customer packets so they are kept within the service provider's network and not forwarded to the customer.

The suite of OAM diagnostics supplement the basic IP ping and traceroute operations with diagnostics specialized for the different levels in the service delivery model. There are diagnostics for MPLS LSPs, SDPs, services and VPLS MACs within a service.

Two-Way Active Measurement Protocol

Two-Way Active Measurement Protocol (TWAMP) provides a standards-based method for measuring the round-trip IP performance (packet loss, delay and jitter) between two devices. TWAMP uses the methodology and architecture of One-Way Active Measurement Protocol (OWAMP) to define a way to measure two-way or round-trip metrics.

There are four logical entities in TWAMP: the control-client, the session-sender, the server, and the session-reflector. The control-client and session-sender are typically implemented in one physical device (the "client") and the server and session-reflector in a second physical device (the "server") with which the two-way measurements are being performed. The router acts as the server.

The control-client and server establish a TCP connection and exchange TWAMP-Control messages over this connection. When the control-client wants to start testing, the client communicates the test parameters to the server. If the server agrees to conduct the described tests, the test begin as soon as the client sends a Start-Sessions message. As part of a test, the session-sender sends a stream of UDP-based test packets to the session-reflector, and the session reflector responds to each received packet with a response UDP-based test packet. When the session-sender receives the response packets from the session-reflector, the information is used to calculate two-way delay, packet loss, and packet delay variation between the two devices.

LSP Diagnostics: LSP Ping and Trace

The router LSP diagnostics are implementations of LSP ping and LSP trace based on RFC 4379, Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures. LSP ping provides a mechanism to detect dataplane failures in MPLS LSPs. LSP ping and LSP trace are modeled after the ICMP echo request/reply used by ping and trace to detect and localize faults in IP networks.

For a given LDP FEC or RSVP P2P LSP, LSP ping verifies whether the packet reaches the egress label edge router (LER), while in LSP trace mode, the packet is sent to the control plane of each transit label switched router (LSR) which performs various checks to see if it is actually a transit LSR for the path.

The downstream mapping TLV is used in lsp-ping and lsp-trace to provide a mechanism for the sender and responder nodes to exchange and validate interface and label stack information for each downstream of an LDP FEC or an RSVP LSP and at each hop in the path of the LDP FEC or RSVP LSP.

Two downstream mapping TLVs are supported. The original Downstream Mapping (DSMAP) TLV defined in RFC 4379 and the new Downstream Detailed Mapping (DDMAP) TLV defined in RFC 6424.

When the responder node has multiple equal cost next-hops for an LDP FEC prefix, the downstream mapping TLV can further be used to exercise a specific path of the ECMP set using the path-destination option. The behavior in this case is described in the ECMP sub-section below.

LSP Ping/Trace for an LSP Using a BGP IPv4 Label Route

This feature adds support of the target FEC stack TLV of type BGP Labeled IPv4 /32 Prefix as defined in RFC 4379.

The new TLV is structured as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4 Prefix                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Length |                Must Be Zero                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 18: Target FEC Stack TLV for a BGP Labeled IPv4 Prefix

The user issues a LSP ping using the existing CLI command and specifying a new type of prefix:

oam lsp-ping bgp-label prefix *ip-prefix/mask* [**src-ip-address** *ip-address*] [**fc** *fc-name* [**profile** {**in**|**out**}]] [**size** *octets*] [**ttl** *label-ttl*] [**send-count** *send-count*] [**timeout** *timeout*] [**interval** *interval*] [**path-destination** *ip-address* [**interface** *if-name* | **next-hop** *ip-address*]] [**detail**]

The path-destination option is used for exercising specific ECMP paths in the network when the LSR performs hashing on the MPLS packet.

Similarly, the user issues a LSP trace using the following command:

oam lsp-trace bgp-label prefix *ip-prefix/mask* [**src-ip-address** *ip-address*] [**fc** *fc-name* [**profile** {**in**|**out**}]] [**max-fail** *no-response-count*] [**probe-count** *probes-per-hop*] [**size** *octets*] [**min-ttl** *min-label-ttl*] [**max-ttl** *max-label-ttl*] [**timeout** *timeout*] [**interval** *interval*] [**path-destination** *ip-address* [**interface** *if-name* | **next-hop** *ip-address*]] [**detail**]

The following are the procedures for sending and responding to an LSP ping or LSP trace packet:

1. The next-hop of a BGP label route for a core IPv4 /32 prefix is always resolved to an LDP FEC or an RSVP LSP. Thus the sender node encapsulates the packet of the echo request message with a label stack which consists of the LDP/RSVP outer label and the BGP inner label.

If the packet expires on an RSVP or LDP LSR node which does not have context for the BGP label IPv4 /32 prefix, it validates the outer label in the stack and if the validation is successful it replies the same way as it does today when it receives an echo request message for an LDP FEC which is stitched to a BGP IPv4 label route. In other words it replies with return `code 8 Label switched at stack-depth <RSC>`.

2. An LSR node which is the next-hop for the BGP label IPv4 /32 prefix as well as the LER node which originated the BGP label IPv4 prefix have full context for the BGP IPv4 target FEC stack and can thus perform full validation of it.

3. If the BGP IPv4 label route is stitched to an LDP FEC, the egress LER for the resulting LDP FEC will not have context for the BGP IPv4 target FEC stack in the echo request message and replies with return `code 4 Replying router has no mapping for the FEC at stack- depth <RSC>`. This is the same behavior as that of an LDP FEC which is stitched to a BGP IPv4 label route when the echo request message reaches the egress LER for the BGP prefix.

Note that only BGP label IPv4 /32 prefixes are supported since these are usable as tunnels on the 7x50 platform. BGP label IPv6 /128 prefixes are not currently usable as tunnels on the 7x50 platform and as such are not supported in LSP ping/trace.

ECMP Considerations

When the responder node has multiple equal cost next-hops for an LDP FEC or a BGP label IPv4 prefix, it replies in the Downstream Mapping TLV with the downstream information of the outgoing interface which is part of the ECMP next-hop set for the prefix.

Note however that when BGP label route is resolved to an LDP FEC (of the BGP next-hop of the BGP label route), ECMP can exist at both the BGP and LDP levels. The following selection of next-hop is performed in this case:

1. For each BGP ECMP next-hop of the label route, a single LDP next-hop is selected even if multiple LDP ECMP next-hops exist. Thus, the number of ECMP next-hops for the BGP IPv4 label route will be equal to the number of BGP next-hops.

2. ECMP for a BGP IPv4 label route is only supported at PE router (BGP label push operation) and not at ABR/ASBR (BGP label swap operation). Thus at an LSR, a BGP IPv4 label route will be resolved to a single BGP next-hop which itself is resolved to a single LDP next-hop.

3. LSP trace will return one downstream mapping TLV for each next-hop of the BGP IPv4 label route. Furthermore, it will return exactly the LDP next-hop the data path programmed for each BGP next-hop.

The following description of the behavior of LSP ping and LSP trace makes a reference to a FEC in a generic way and which can represent an LDP FEC or a BGP IPv4 label route. In addition the reference to a downstream mapping TLV means either the DSMAP TLV or the DDMAP TLV.

1. If the users initiates an lsp-trace or lsp-ping of the FEC without the **path-destination** option specified, then the sender node will not include multi-path information in the Downstream Mapping TLV in the echo request message (multipath type=0). In this case, the responder node will reply with a Downstream Mapping TLV for each outgoing interface which is part of the ECMP next-hop set for the FEC. Note however the sender node will select the first Downstream Mapping TLV only for the subsequent echo request message with incrementing TTL.

2. If the user initiates an lsp-ping of the FEC with the **path-destination** option specified, then the sender node will not include the Downstream Mapping TLV. However, the user can use the **interface** option, part of the same **path-destination** option, to direct the echo request message at the sender node to be sent out a specific outgoing interface which is part of an ECMP path set for the FEC.

3. If the user initiates an lsp-trace of the FEC with the **path-destination** option specified but configured not to include a downstream mapping TLV in the MPLS echo request message using the CLI command **downstream-map-tlv** {**none**}, then the sender node will not include the Downstream Mapping TLV. However, the user can use the **interface** option, part of the same **path-destination** option, to direct the echo request message at the sender node to be sent out a specific outgoing interface which is part of an ECMP path set for the FEC.

4. If the user initiates an lsp-trace of the FEC with the **path-destination** option specified, then the sender node will include the multipath information in the Downstream Mapping TLV in the echo request message (multipath type=8). The **path-destination** option allows the user to exercise a specific path of a FEC in the presence of ECMP. This is performed by having the user enter a specific address from the 127/8 range which is then inserted in the multipath type 8 information field of the Downstream Mapping TLV. The CPM code at each LSR in the path of the target FEC runs the same hash routine as the data path and replies in the Downstream Mapping TLV with the specific outgoing interface the packet would have been forwarded to if it did not expire at this node and if DEST IP field in the packet’s header was set to the 127/8 address value inserted in the multipath type 8 information.. This hash is based on:

    a. The {incoming port, system interface address, label-stack} when the **lsr-load-balancing** option of the incoming interface is configured to **lbl-only**. In this case the 127/8 prefix address entered in the **path-destination** option is not used to select the outgoing interface. All packets received with the same label stack will map to a single and same outgoing interface.

    b. The {incoming port, system interface address, label-stack, SRC/DEST IP fields of the packet} when the **lsr-load-balancing** option of the incoming interface is configured to **lbl-ip**. The SRC IP field corresponds to the value entered by the user in the **src-ip-**

address option (default system IP interface address). The DEST IP field corresponds to the 127/8 prefix address entered in the **path-destination** option. In this case, the CPM code will map the packet, as well as any packet in a sub-range of the entire 127/8 range, to one of the possible outgoing interface of the FEC.

c. The {SRC/DEST IP fields of the packet} when the **lsr-load-balancing** option of the incoming interface is configured to **ip-only**. The SRC IP field corresponds to the value entered by the user in the **src-ip-address** option (default system IP interface address). The DEST IP field corresponds to the 127/8 prefix address entered in the **path-destination** option. In this case, the CPM code will map the packet, as well as any packet in a sub-range of the entire 127/8 range, to one of the possible outgoing interface of the FEC.

d. In all above cases, the user can use the interface option, part of the same **path-destination** option, to direct the echo request message at the sender node to be sent out a specific outgoing interface which is part of an ECMP path set for the FEC.

e. Note that if the user enabled the **system-ip-load-balancing hash** option (**config>system>system-ip-load-balancing**), then the LSR hashing is modified by applying the system IP interface, with differing bit-manipulation, to the hash of packets of all three options (**lbl-only**, **lbl-ip, ip-only**). This system level option enhances the LSR packet distribution such that the probability of the same flow selecting the same ECMP interface index or LAG link index at two consecutive LSR nodes is minimized.

Lsp-ping and lsp-trace over Unnumbered IP Interface

Lsp-ping and p2mp-lsp-ping operate over a network using unnumbered links without any changes. Lsp-trace, p2mp-lsp-trace and ldp-treetrace are modified such that the unnumbered interface is properly encoded in the downstream mapping (DSMAP/DDMAP) TLV.

In a RSVP P2P or P2MP LSP, the upstream LSR encodes the downstream router-id in the “Downstream IP Address” field and the local unnumbered interface index value in the “Downstream Interface Address” field of the DSMAP/DDMAP TLV as per RFC 4379. Both values are taken from the TE database.

In a LDP unicast FEC or mLDP P2MP FEC, the interface index assigned by the peer LSR is not readily available to the LDP control plane. In this case, the alternative method described in RFC 4379 is used. The upstream LSR sets the Address Type to IPv4 Unnumbered, the Downstream IP Address to a value of 127.0.0.1, and the interface index is set to 0. If an LSR receives an echo-request packet with this encoding in the DSMAP/DDMAP TLV, it will bypass interface verification but continue with label validation.

Downstream Detailed Mapping (DDMAP) TLV

The DDMAP TLV provides with exactly the same features as the existing DSMAP TLV. plus the enhancements to trace the details of LSP stitching and LSP hierarchy. The latter is achieved using a new sub-TLV of the DDMAP TLV called the FEC stack change sub-TLV. The following are the structures of these two objects as defined in RFC 6424.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               MTU             | Address Type  |    DS Flags   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Downstream Address (4 or 16 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Downstream Interface Address (4 or 16 octets)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Return Code  | Return SubCode|        Sub-tlv length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                      List of Sub TLVs                         .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 19: DDMAP TLV

The DDMAP TLV format is derived from the DSMAP TLV format. The key change is that variable length and optional fields have been converted into sub-TLVs. The fields have the same use and meaning as in RFC 4379.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Operation Type | Address type  | FEC-tlv length|  Reserved     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Remote Peer Address (0, 4 or 16 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                          FEC TLV                              .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 20: FEC Stack Change Sub-TLV

The operation type specifies the action associated with the FEC stack change. The following operation types are defined.

```
Type #     Operation
------     ---------
```

```
        1          Push
        2          Pop
```

More details on the processing of the fields of the FEC stack change sub-TLV are provided later in this section.

The user can configure which downstream mapping TLV to use globally on a system by using the following command:

configure test-oam mpls-echo-request-downstream-map {**dsmap** | **ddmap**}

This command specifies which format of the downstream mapping TLV to use in all LSP trace packets and LDP tree trace packets originated on this node. The Downstream Mapping (DSMAP) TLV is the original format in RFC 4379 and is the default value. The Downstream Detailed Mapping (DDMAP) TLV is the new enhanced format specified in RFC 6424.

This command applies to LSP trace of an RSVP P2P LSP, a MPLS-TP LSP, or LDP unicast FEC, and to LDP tree trace of a unicast LDP FEC. It does not apply to LSP trace of an RSVP P2MP LSP which always uses the DDMAP TLV.

The global DSMAP/DDMAP setting impacts the behavior of both OAM LSP trace packets and SAA test packets of type **lsp-trace** and is used by the sender node when one of the following events occurs:

1. An SAA test of type **lsp-trace** is created (not modified) and no value is specified for the per-test **downstream-map-tlv** {**dsmap**|**ddmap**|**none**} option. In this case the SAA test **downstream-map-tlv** value defaults to the global **mpls-echo-request-downstream-map** value.

2. An OAM test of type **lsp-trace** test is executed and no value is specified for the per-test **downstream-map-tlv** {**dsmap**|**ddmap**|**none**} option. In this case, the OAM test **downstream-map-tlv** value defaults to the global **mpls-echo-request-downstream-map** value.

A consequence of the rules above is that a change to the value of **mpls-echo-request-downstream-map** option does not affect the value inserted in the downstream mapping TLV of existing tests.

The following are the details of the processing of the new DDMAP TLV:

1. When either the DSMAP TLV or the DDMAP TLV is received in an echo request message, the responder node will include the same type of TLV in the echo reply message with the proper downstream interface information and label stack information.

2. If an echo request message without a Downstream Mapping TLV (DSMAP or DDMAP) expires at a node which is not the egress for the target FEC stack, the responder node always includes the DSMAP TLV in the echo reply message. This can occur in the following cases:

a. The user issues a LSP trace from a sender node with a **min-ttl** value higher than 1 and a **max-ttl** value lower than the number of hops to reach the egress of the target FEC stack. This is the sender node behavior when the global configuration or the per-test setting of the DSMAP/DDMAP is set to DSMAP.

b. The user issues a LSP ping from a sender node with a **ttl** value lower than the number of hops to reach the egress of the target FEC stack. This is the sender node behavior when the global configuration of the DSMAP/DDMAP is set to DSMAP.

c. The behavior in (a) is changed when the global configuration or the per-test setting of the Downstream Mapping TLV is set to DDMAP. The sender node will include in this case the DDMAP TLV with the Downstream IP address field set to the all-routers multicast address as per Section 3.3 of RFC 4379. The responder node then bypasses the interface and label stack validation and replies with a DDMAP TLV with the correct downstream information for the target FEC stack.

3. A sender node never includes the DSMAP or DDMAP TLV in an lsp-ping message.

Using DDMAP TLV in LSP Stitching and LSP Hierarchy

In addition to performing the same features as the DSMAP TLV, the new DDMAP TLV addresses the following scenarios:

1. Full validation of an LDP FEC stitched to a BGP IPv4 label route. In this case, the LSP trace message is inserted from the LDP LSP segment or from the stitching point.

2. Full validation of a BGP IPv4 label route stitched to an LDP FEC. This includes the case of Full validation of a BGP IPv4 label route stitched to an LDP FEC. This includes the case of explicit configuration of the LDP-BGP stitching in which the BGP label route is active in Route Table Manager (RTM) and the case of a BGP IPv4 label route resolved to the LDP FEC due to the IGP route of the same prefix active in RTM. In this case, the LSP trace message is inserted from the BGP LSP segment or from the stitching point.

3. Full validation of an LDP FEC which is stitched to a BGP LSP and stitched back into an LDP FEC. In this case, the LSP trace message is inserted from the LDP segments or the or from the stitching points.

4. Full validation of an LDP FEC tunneled over an RSVP LSP using LSP trace.

In order to properly check a target FEC which is stitched to another FEC (stitching FEC) of the same or a different type, or which is tunneled over another FEC (tunneling FEC), it is necessary for the responding nodes to provide details about the FEC manipulation back to the sender node. This is achieved via the use of the new FEC stack change sub-TLV in the Downstream Detailed Mapping TLV (DDMAP) defined in RFC 6424.

When the user configures the use of the DDMAP TLV on a trace for an LSP that does not undergo stitching or tunneling operation in the network, the procedures at the sender and responder nodes are the same as in the case of the existing DSMAP TLV.

This feature however introduces changes to the target FEC stack validation procedures at the sender and responder nodes in the case of LSP stitching and LSP hierarchy. These changes pertain to the processing of the new FEC stack change sub-TLV in the new DDMAP TLV and the new return `code 15 Label switched with FEC change`. The following is a description of the main changes which are a superset of the rules described in Section 4 of RFC 6424 to allow greater scope of interoperability with other vendor implementations.

Responder Node Procedures

1. As a responder node, the 7x50 will always insert a global return code return code of either 3 `Replying router is an egress for the FEC at stack-depth <RSC>` or 14 `See DDMAP TLV for Return Code and Return Subcode`.

2. When the responder node inserts a global return code of 3, it will not include a DDMAP TLV.

3. When the responder node includes the DDMAP TLV, it inserts a global return `code 14 See DDMAP TLV for Return Code and Return Subcode` and:

    a. On a success response, include a return code of 15 in the DDMAP TLV for each downstream which has a FEC stack change TLV.

    b. On a success response, include a return `code 8 Label switched at stack-depth <RSC>` in the DDMAP TLV for each downstream if no FEC stack change sub-TLV is present.

    c. On a failure response, include an appropriate error return code in the DDMAP TLV for each downstream.

4. A tunneling node indicates that it is pushing a FEC (the tunneling FEC) on top of the target FEC stack TLV by including a FEC stack change sub-TLV in the DDMAP TLV with a FEC operation type value of PUSH. It also includes a return `code 15 Label switched with FEC change`. The downstream interface address and downstream IP address fields of the DDMAP TLV are populated for the pushed FEC. The remote peer address field in the FEC stack change sub-TLV is populated with the address of the control plane peer for the pushed FEC. The Label stack sub-TLV provides the full label stack over the downstream interface.

5. A node that is stitching a FEC indicates that it is performing a POP operation for the stitched FEC followed by a PUSH operation for the stitching FEC and will thus include two FEC stack change sub-TLVs in the DDMAP TLV in the echo reply message. It also includes and a

return `code 15 Label switched with FEC change`. The downstream interface address and downstream address fields of the DDMAP TLV are populated for the stitching FEC. The remote peer address field in the FEC stack change sub-TLV of type POP is populated with a null value (0.0.0.0). The remote peer address field in the FEC stack change sub-TLV of type PUSH is populated with the address of the control plane peer for the tunneling FEC. The Label stack sub-TLV provides the full label stack over the downstream interface.

6. If the responder node is the egress for one or more FECs in the target FEC Stack, then it must reply with no DDMAP TLV and with a return `code 3 Replying router is an egress for the FEC at stack-depth <RSC>`. RSC must be set to the depth of the topmost FEC. This operation is iterative in a sense that at the receipt of the echo reply message the sender node will pop the topmost FEC from the target stack FEC TLV and resend the echo request message with the same TTL value as explained in (5) below. The responder node will thus perform exactly the same operation as described in this step until all FECs are popped or until the topmost FEC in the target FEC stack TLV matches the tunneled or stitched FEC. In the latter case, processing of the target FEC stack TLV follows again steps (1) or (2).

Sender Node Procedures

1. If the echo reply message contains the return `code 14 See DDMAP TLV for Return Code and Return Subcode` and the DDMAP TLV has a return `code 15 Label switched with FEC change`, the sender node adjusts the target FEC Stack TLV in the echo request message for the next value of the TTL to reflect the operation on the current target FEC stack as indicated in the FEC stack change sub-TLV received in the DDMAP TLV of the last echo reply message. In other words, one FEC is popped at most and one or more FECs are pushed as indicated.

2. If the echo reply message contains the return `code 3 Replying router is an egress for the FEC at stack-depth <RSC>`, then:

   a. If the value for the label stack depth specified in the Return Sub-Code (RSC) field is the same as the depth of current target FEC Stack TLV, then the sender node considers the trace operation complete and terminates it. A 7x50 responder node will cause this case to occur as per step (6) of the responder node procedures.

   b. If the value for the label stack depth specified in the Return Sub-Code (RSC) field is different from the depth of the current target FEC Stack TLV, the sender node must continue the LSP trace with the same TTL value after adjusting the target FEC stack TLV by removing the top FEC. Note this step will continue iteratively until the value for the label stack depth specified in the Return Sub-Code (RSC) field is the same as the depth of current target FEC Stack TLV and in which case step (a) is performed. A 7x50 responder node will cause this case to occur as per step (6) of the responder node procedures.

c. If a DDMAP TLV with or without a FEC stack change sub-TLV is included, then the sender node must ignore it and processing is performed as per steps (a) or (b) above. A 7x50 responder node will not cause this case to occur but a third party implementation may do.

3. As a sender node, the 7x50 can accept an echo-reply message with the global return code of either 14 (with DDMAP TLV return code of 15 or 8), or15 and process properly the FEC stack change TLV as per step (1) of the sender node procedures.

4. If an LSP ping is performed directly to the egress LER of the stitched FEC, there is no DDMAP TLV included in the echo request message and thus the responder node, which is the egress node, will still reply with return `code 4 Replying router has no mapping for the FEC at stack- depth <RSC>`. This case cannot be resolved with this feature.

5. Note the following limitation when a BGP IPv4 label route is resolved to an LDP FEC which itself is resolved to an RSVP LSP all on the same node. This 2-level LSP hierarchy is not supported as a feature on the SROS but user is not prevented from configuring it. In that case, user and OAM packets are forwarded by the sender node using two labels (T-LDP and BGP). The LSP trace will fail on the downstream node with return `code 1 Malformed echo request received` since there is no label entry for the RSVP label.

LDP Tree Trace: End-to-End Testing of Paths in an LDP ECMP Network

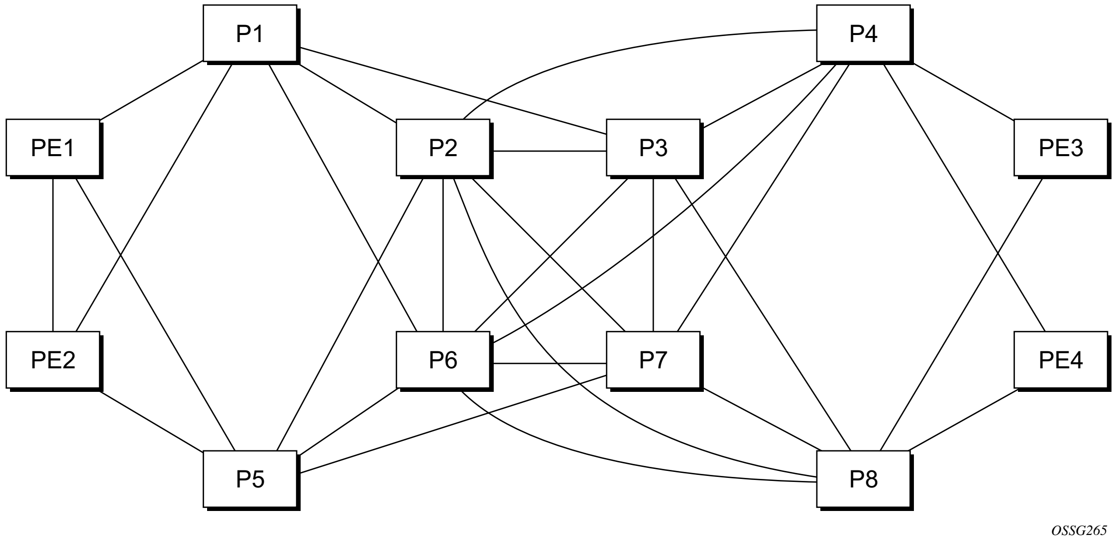

Figure 21: Network Resilience Using LDP ECMP

Figure 21 depicts an IP/MPLS network which uses LDP ECMP for network resilience. Faults that are detected through IGP and/or LDP are corrected as soon as IGP and LDP re-converge. The impacted traffic will be forwarded on the next available ECMP path as determined by the hash routine at the node that had a link failure.

However, there are faults which the IGP/LDP control planes may not detect. These faults may be due to a corruption of the control plane state or of the data plane state in a node. Although these faults are very rare and mostly due to misconfiguration, the LDP Tree Trace OAM feature is intended to detect these “silent” data plane and control plane faults. For example, it is possible that the forwarding plane of a node has a corrupt Next Hop Label Forwarding Entry (NHLFE) and keeps forwarding packets over an ECMP path only to have the downstream node discard them. This data plane fault can only be detected by an OAM tool that can test all possible end-to-end paths between the ingress LER and the egress LER. A corruption of the NLHFE entry can also result from a corruption in the control plane at that node.

LDP ECMP Tree Building

When the LDP tree trace feature is enabled, the ingress LER builds the ECM tree for a given FEC (egress LER) by sending LSP trace messages and including the LDP IPv4 Prefix FEC TLV as well as the downstream mapping TLV.In order to build the ECMP tree, the router LER inserts an IP address range drawn from the 127/8 space. When received by the downstream LSR, it will use this range to determine which ECMP path is exercised by any IP address or a sub-range of addresses within that range based on its internal hash routine. When the MPLS echo reply is received by the router LER, it will record this information and proceed with the next echo request message targeted for a node downstream of the first LSR node along one of the ECMP paths. The sub-range of IP addresses indicated in the initial reply will be used since the objective is to have the LSR downstream of the router LER pass this message to its downstream node along the first ECMP path.

The following figure illustrates the behavior through the following example adapted from RFC 4379, *Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures*:

```
PE1 ---- A ----- B ----- C ------ G ----- H ---- PE2
         \       \---- D ------/       /
             \      \--- E------/       /
              -- F --------------------/
```

LSR A has two downstream LSRs, B and F, for PE2 FEC. PE1 receives an echo reply from A with the Multipath Type set to 4, with low/high IP addresses of 127.1.1.1->127.1.1.255 for downstream LSR B and 127.2.1.1->127.2.1.255 for downstream LSR F. PE1 reflects this information to LSR B. B, which has three downstream LSRs, C, D, and E, computes that 127.1.1.1->127.1.1.127 would go to C and 127.1.1.128-> 127.1.1.255 would go to D. B would then respond with 3 Downstream Mappings: to C, with Multipath Type 4 (127.1.1.1->127.1.1.127); to D, with Multipath Type 4 (127.1.1.127->127.1.1.255); and to E, with Multipath Type 0.

The router supports multipath type 0 and 8, and up to a maximum of 36 bytes for the multipath length and supports the LER part of the LDP ECMP tree building feature.

A user configurable parameter sets the frequency of running the tree trace capability. The minimum and default value is 60 minutes and the increment is 1 hour.

The router LER gets the list of FECs from the LDP FEC database. New FECs will be added to the discovery list at the next tree trace and not when they are learned and added into the FEC database. The maximum number of FECs to be discovered with the tree building feature is limited to 500. The user can configure FECs to exclude the use of a policy profile.

Periodic Path Exercising

The periodic path exercising capability of the LDP tree trace feature runs in the background to test the LDP ECMP paths discovered by the tree building capability. The probe used is an LSP ping message with an IP address drawn from the sub-range of 127/8 addresses indicated by the output of the tree trace for this FEC.

The periodic LSP ping messages continuously probes an ECMP path at a user configurable rate of at least 1 message per minute. This is the minimum and default value. The increment is 1 minute. If an interface is down on a router LER, then LSP ping probes that normally go out this interface will not be sent.

The LSP ping routine updates the content of the MPLS echo request message, specifically the IP address, as soon as the LDP ECMP tree trace has output the results of a new computation for the path in question.

LSP Ping for RSVP P2MP LSP (P2MP)

Note: For more information about P2MP refer to the 7750 SR OS MPLS Guide.

The P2MP LSP ping complies to RFC 6425, *Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS) - Extensions to LSP Ping*.

An LSP ping can be generated by entering the following OAM command:

```
oam p2mp-lsp-ping lsp-name [p2mp-instance instance-name [s2l-dest-addr
ip-address [...up to 5 max]]] [fc fc-name [profile {in | out}]] [size
octets] [ttl label-ttl] [timeout timeout] [detail]
```

The echo request message is sent on the active P2MP instance and is replicated in the data path over all branches of the P2MP LSP instance. By default, all egress LER nodes which are leaves of the P2MP LSP instance will reply to the echo request message.

The user can reduce the scope of the echo reply messages by explicitly entering a list of addresses for the egress LER nodes that are required to reply. A maximum of 5 addresses can be specified in a single execution of the **p2mp-lsp-ping** command. If all 5 egress LER nodes are 7750 nodes, they will be able to parse the list of egress LER addresses and will reply. Note however that RFC 6425 specifies that only the top address in the P2MP egress identifier TLV must be inspected by an egress LER. When interoperating with other implementations, an 7750 egress LER will respond if its address is anywhere in the list. Furthermore, if another vendor implementation is the egress LER, only the egress LER matching the top address in the TLV may respond.

If the user enters the same egress LER address more than once in a single p2mp-lsp-ping command, the head-end node displays a response to a single one and displays a single error warning message for the duplicate ones. When queried over SNMP, the head-end node issues a single response trap and issues no trap for the duplicates.

The **timeout** parameter should be set to the time it would take to get a response from all probed leaves under no failure conditions. For that purpose, its range extends to 120 seconds for a p2mp-lsp-ping from a 10 second lsp-ping for P2P LSP. The default value is 10 seconds.

A 7750 head-end node displays a “Send_Fail” error when a specific S2L path is down only if the user explicitly listed the address of the egress LER for this S2L in the **ping** command.

Similarly, a 7750 head-end node displays the timeout error when no response is received for an S2L after the expiry of the timeout timer only if the user explicitly listed the address of the egress LER for this S2L in the **ping** command.

The user can configure a specific value of the **ttl** parameter to force the echo request message to expire on a 7750 branch node or a bud LSR node. The latter replies with a downstream mapping TLV for each branch of the P2MP LSP in the echo reply message. Note however that a maximum of 16 downstream mapping TLVs can be included in a single echo reply message. It also sets the multipath type to zero in each downstream mapping TLV and will thus not include any egress address information for the reachable egress LER nodes for this P2MP LSP.

If a 7750 ingress LER node receives the new multipath type field with the list of egress LER addresses in an echo reply message from another vendor implementation, it will ignore but will not cause an error in processing the downstream mapping TLV.

If the ping expires at an LSR node which is performing a re-merge or cross-over operation in the data path between two or more ILMs of the same P2MP LSP, there will be an echo reply message for each copy of the echo request message received by this node.

The output of the command without the **detail** parameter specified provides a high-level summary of error codes and/or success codes received.

The output of the command with the **detail** parameter specified shows a line for each replying node as in the output of the LSP ping for a P2P LSP.

The display is delayed until all responses are received or the timer configured in the timeout parameter expired. No other CLI commands can be entered while waiting for the display. A control-C (^C) command will abort the ping operation.

LSP Trace for RSVP P2MP LSP

The P2MP LSP trace complies to RFC 6425. An LSP trace can be generated by entering the following OAM command:

```
oam p2mp-lsp-trace lsp-name p2mp-instance instance-name s2l-dest-address
ip-address [fc fc-name [profile {in|out}]] [size octets] [max-fail no-
response-count] [probe-count probes-per-hop] [min-ttl min-label-ttl]
[max-ttl max-label-ttl] [timeout timeout] [interval interval] [detail]
```

The LSP trace capability allows the user to trace the path of a single S2L path of a P2MP LSP. Its operation is similar to that of the **p2mp-lsp-ping** command but the sender of the echo reply request message includes the downstream mapping TLV to request the downstream branch information from a branch LSR or bud LSR. The branch LSR or bud LSR will then also include the downstream mapping TLV to report the information about the downstream branches of the P2MP LSP. An egress LER does not include this TLV in the echo response message.

The **probe-count** parameter operates in the same way as in LSP trace on a P2P LSP. It represents the maximum number of probes sent per TTL value before giving up on receiving the echo reply message. If a response is received from the traced node before reaching maximum number of probes, then no more probes are sent for the same TTL. The sender of the echo request then increments the TTL and uses the information it received in the downstream mapping TLV to start sending probes to the node downstream of the last node which replied. This continues until the egress LER for the traced S2L path replied.

Since the command traces a single S2L path, the timeout and interval parameters keep the same value range as in LSP trace for a P2P LSP.

The P2MP LSP Trace makes use of the Downstream Detailed Mapping (DDMAP) TLV. The following excerpt from RFC 6424 details the format of the new DDMAP TLV:

```
      0                     1                     2
3
        0 1  2  3 4 5  6 7  8 9  0 1  2 3  4 5  6  7 8 9  0 1  2 3  4  5 6  7 8  9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |            MTU                      | Address Type    |    DS Flags        |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |            Downstream Address (4 or 16 octets)                             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |        Downstream Interface Address (4 or 16 octets)                      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Return Code  | Return Subcode    |        Sub-tlv Length                 |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     .
.
     .                 List of Sub-TLVs
.
     .
.
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 22: Downstream Detailed Mapping TLV

Figure 22 depicts Downstream Detailed Mapping TLV entered in the path-destination belongs to one of the possible outgoing interface of the FEC.

The Downstream Detailed Mapping TLV format is derived from the Downstream Mapping (DSMAP) TLV format. The key change is that variable length and optional fields have been converted into sub-TLVs. The fields have the same use and meaning as in RFC 4379.

Similar to p2mp-lsp-ping, an LSP trace probe results on all egress LER nodes eventually receiving the echo request message but only the traced egress LER node will reply to the last probe.

Also any branch LSR node or bud LSR node in the P2MP LSP tree may receive a copy of the echo request message with the TTL in the outer label expiring at this node. However, only a branch LSR or bud LSR which has a downstream branch over which the traced egress LER is reachable must respond.

When a branch LSR or BUD LSR node responds to the sender of the echo request message, it sets the global return code in the echo response message to RC=14 - "See DDMAP TLV for Return Code and Return Sub-Code" and the return code in the DDMAP TLV corresponding to the outgoing interface of the branch used by the traced S2L path to RC=8 - "Label switched at stack-depth <RSC>".

Since a single egress LER address, for example an S2L path, can be traced, the branch LSR or bud LSR node will set the multipath type of zero in the downstream mapping TLV in the echo response message as no egress LER address need to be included.

LSP Trace Behavior When S2L Path Traverses a Re-Merge Node

When a 7750 LSR performs a re-merge of one or more ILMs of the P2MP LSP to which the traced S2L sub-LSP belongs, it may block the ILM over which the traced S2L resides. This causes the trace to either fail or to succeed with a missing hop.

The following is an example of this behavior.

S2L1 and S2L2 use ILMs which re-merge at node B. Depending of which ILM is blocked at B, the TTL=2 probe will either yield two responses or will timeout.

```
S2L1 = ACBDF (to leaf F)
S2L2 = ABDE (to leaf E)

   A
 /  \
B -- C
|
D
| \
F  E
```

- Tracing S2L1 when ILM on interface C-B blocked at node B:

  For TTL=1, A gets a response from C only as B does not have S2L1 on the ILM on interface A-B.

  For TTL=2, assume A gets first the response from B which indicates a success. It then builds the next probe with TTL=3. B will only pass the copy of the message arriving on interface A-B and will drop the one arriving on interface C-B (treats it like a data packet since it does not expire at node B). This copy will expire at F. However F will return a "DSMappingMismatched" error because the DDMAP TLV was the one provided by node B in TTL=2 step. The trace will abort at this point in time. However, A knows it got a second response from Node D for TTL=2 with a "DSMappingMismatched" error.

  If A gets the response from D first with the error code, it waits to see if it gets a response from B or it times out. In either case, it will log this status as **multiple replies received per probe** in the last probe history and aborts the trace.
- Tracing S2L2 when ILM on interface A-B blocked at node B:

  For TTL=1, B responds with a success. C does not respond as it does not have an ILM for S2L2.

  For TTL=2, B drops the copy coming on interface A-B. It receives a copy coming on interface B-C but will drop it as the ILM does not contain S2L2. Node A times out. Next, node A generates a probe with TTL=3 without a DDMAP TLV. This time node D will respond with a success and will include its downstream DDMAP TLV to node E. The rest of the path will be discovered correctly. The traced path for S2L2 will look something like: A-B-(*)-D-E.

A 7750 ingress LER detects a re-merge condition when it receives two or more replies to the same probe, such as the same TTL value. It displays the following message to the user regardless if the trace operation successfully reached the egress LER or was aborted earlier:

```
“Probe returned multiple responses. Result may be inconsistent.”
```

This warning message indicates to the user the potential of a re-merge scenario and that a p2mp-lsp-ping command for this S2L should be used to verify that the S2L path is not defective.

The 7750 ingress LER behavior is to always proceed to the next ttl probe when it receives an OK response to a probe or when it times out on a probe. If however it receives replies with an error return code, it must wait until it receives an OK response or it times out. If it times out without receiving an OK reply, the LSP trace must be aborted.

The following are possible echo reply messages received and corresponding ingress LER behavior:

- One or more error return codes + OK: display OK return code. Proceed to next ttl probe. Display warning message at end of trace.
- OK + One or more error return codes: display OK return code. Proceed to next ttl probe right after receiving the OK reply but keep state that more replies received. Display warning message at end of trace.
- OK + OK: should not happen for re-merge but would continue trace on 1st OK reply. This is the case when one of the branches of the P2MP LSP is activating the P2P bypass LSP. In this case, the head-end node will get a reply from both a regular P2MP LSR which has the ILM for the traced S2L and from an LSR switching the P2P bypass for other S2Ls. The latter does not have context for the P2MP LSP being tunneled but will respond after doing a label stack validation.
- One error return code + timeout: abort LSP trace and display error code. Ingress LER cannot tell the error is due to a re-merge condition.
- More than one error return code + timeout: abort LSP trace and display first error code. Display warning message at end of trace.
- Timeout on probe without any reply: display “*” and proceed to next ttl probe.

SDP Diagnostics

The router SDP diagnostics are SDP ping and SDP MTU path discovery.

SDP Ping

SDP ping performs in-band uni-directional or round-trip connectivity tests on SDPs. The SDP ping OAM packets are sent in-band, in the tunnel encapsulation, so it will follow the same path as traffic within the service. The SDP ping response can be received out-of-band in the control plane, or in-band using the data plane for a round-trip test.

For a uni-directional test, SDP ping tests:

- Egress SDP ID encapsulation
- Ability to reach the far-end IP address of the SDP ID within the SDP encapsulation
- Path MTU to the far-end IP address over the SDP ID
- Forwarding class mapping between the near-end SDP ID encapsulation and the far-end tunnel termination

For a round-trip test, SDP ping uses a local egress SDP ID and an expected remote SDP ID. Since SDPs are uni-directional tunnels, the remote SDP ID must be specified and must exist as a configured SDP ID on the far-end router SDP round trip testing is an extension of SDP connectivity testing with the additional ability to test:

- Remote SDP ID encapsulation
- Potential service round trip time
- Round trip path MTU
- Round trip forwarding class mapping

SDP MTU Path Discovery

In a large network, network devices can support a variety of packet sizes that are transmitted across its interfaces. This capability is referred to as the Maximum Transmission Unit (MTU) of network interfaces. It is important to understand the MTU of the entire path end-to-end when provisioning services, especially for virtual leased line (VLL) services where the service must support the ability to transmit the largest customer packet.

The Path MTU discovery tool provides a powerful tool that enables service provider to get the exact MTU supported by the network's physical links between the service ingress and service termination points (accurate to one byte).

Service Diagnostics

Alcatel-Lucent’s Service ping feature provides end-to-end connectivity testing for an individual service. Service ping operates at a higher level than the SDP diagnostics in that it verifies an individual service and not the collection of services carried within an SDP.

Service ping is initiated from a router to verify round-trip connectivity and delay to the far-end of the service. Alcatel-Lucent’s implementation functions for both GRE and MPLS tunnels and tests the following from edge-to-edge:

- Tunnel connectivity
- VC label mapping verification
- Service existence
- Service provisioned parameter verification
- Round trip path verification
- Service dynamic configuration verification

VPLS MAC Diagnostics

While the LSP ping, SDP ping and service ping tools enable transport tunnel testing and verify whether the correct transport tunnel is used, they do not provide the means to test the learning and forwarding functions on a per-VPLS-service basis.

It is conceivable, that while tunnels are operational and correctly bound to a service, an incorrect Forwarding Information Base (FIB) table for a service could cause connectivity issues in the service and not be detected by the ping tools. Alcatel-Lucent has developed VPLS OAM functionality to specifically test all the critical functions on a per-service basis. These tools are based primarily on the IETF document draft-stokes-vkompella-ppvpn-hvpls-oam-xx.txt, *Testing Hierarchical Virtual Private LAN Services*.

The VPLS OAM tools are:

- MAC Ping — Provides an end-to-end test to identify the egress customer-facing port where a customer MAC was learned. MAC ping can also be used with a broadcast MAC address to identify all egress points of a service for the specified broadcast MAC.
- MAC Trace — Provides the ability to trace a specified MAC address hop-by-hop until the last node in the service domain. An SAA test with MAC trace is considered successful when there is a reply from a far-end node indicating that they have the destination MAC address on an egress SAP or the CPM.
- CPE Ping — Provides the ability to check network connectivity to the specified client device within the VPLS. CPE ping will return the MAC address of the client, as well as the SAP and PE at which it was learned.
- MAC Populate — Allows specified MAC addresses to be injected in the VPLS service domain. This triggers learning of the injected MAC address by all participating nodes in the service. This tool is generally followed by MAC ping or MAC trace to verify if correct learning occurred.
- MAC Purge — Allows MAC addresses to be flushed from all nodes in a service domain.

MAC Ping

For a MAC ping test, the destination MAC address (unicast or multicast) to be tested must be specified. A MAC ping packet can be sent through the control plane or the data plane. When sent by the control plane, the ping packet goes directly to the destination IP in a UDP/IP OAM packet. If it is sent by the data plane, the ping packet goes out with the data plane format.

In the control plane, a MAC ping is forwarded along the flooding domain if no MAC address bindings exist. If MAC address bindings exist, then the packet is forwarded along those paths (if they are active). Finally, a response is generated only when there is an egress SAP binding to that MAC address. A control plane request is responded to via a control reply only.

In the data plane, a MAC ping is sent with a VC label TTL of 255. This packet traverses each hop using forwarding plane information for next hop, VC label, etc. The VC label is swapped at each service-aware hop, and the VC TTL is decremented. If the VC TTL is decremented to 0, the packet is passed up to the management plane for processing. If the packet reaches an egress node, and would be forwarded out a customer facing port, it is identified by the OAM label below the VC label and passed to the management plane.

MAC pings are flooded when they are unknown at an intermediate node. They are responded to only by the egress nodes that have mappings for that MAC address.

MAC Trace

A MAC trace functions like an LSP trace with some variations. Operations in a MAC trace are triggered when the VC TTL is decremented to 0.

Like a MAC ping, a MAC trace can be sent either by the control plane or the data plane.

For MAC trace requests sent by the control plane, the destination IP address is determined from the control plane mapping for the destination MAC. If the destination MAC is known to be at a specific remote site, then the far-end IP address of that SDP is used. If the destination MAC is not known, then the packet is sent unicast, to all SDPs in the service with the appropriate squelching.

A control plane MAC traceroute request is sent via UDP/IP. The destination UDP port is the LSP ping port. The source UDP port is whatever the system gives (note that this source UDP port is really the demultiplexor that identifies the particular instance that sent the request, when correlating the reply). The source IP address is the system IP of the sender.

When a traceroute request is sent via the data plane, the data plane format is used. The reply can be via the data plane or the control plane.

A data plane MAC traceroute request includes the tunnel encapsulation, the VC label, and the OAM, followed by an Ethernet DLC, a UDP and IP header. If the mapping for the MAC address is known at the sender, then the data plane request is sent down the known SDP with the appropriate tunnel encapsulation and VC label. If it is not known, then it is sent down every SDP (with the appropriate tunnel encapsulation per SDP and appropriate egress VC label per SDP binding).

The tunnel encapsulation TTL is set to 255. The VC label TTL is initially set to the min-ttl (default is 1). The OAM label TTL is set to 2. The destination IP address is the all-routers multicast address. The source IP address is the system IP of the sender.

The destination UDP port is the LSP ping port. The source UDP port is whatever the system gives (note that this source UDP port is really the demultiplexor that identifies the particular instance that sent the request, when correlating the reply).

The Reply Mode is either 3 (i.e., reply via the control plane) or 4 (i.e., reply through the data plane), depending on the reply-control option. By default, the data plane request is sent with Reply Mode 3 (control plane reply).

The Ethernet DLC header source MAC address is set to either the system MAC address (if no source MAC is specified) or to the specified source MAC. The destination MAC address is set to the specified destination MAC. The EtherType is set to IP.

CPE Ping

The MAC ping OAM tool makes it possible to detect whether a particular MAC address has been learned in a VPLS.

The **cpe-ping** command extends this capability to detecting end-station IP addresses inside a VPLS. A CPE ping for a specific destination IP address within a VPLS will be translated to a MAC-ping towards a broadcast MAC address. Upon receiving such a MAC ping, each peer PE within the VPLS context will trigger an ARP request for the specific IP address. The PE receiving a response to this ARP request will report back to the requesting 7750 SR. It is encouraged to use the source IP address of 0.0.0.0 to prevent the provider's IP address of being learned by the CE.

CPE Ping for PBB Epipe

CPE ping has been supported for VPLS services since Release 3.0 of SR OS. It enables the connectivity of the access circuit between a VPLS PE and a CPE to be tested, even if the CPE is unmanaged, and therefor the service provider cannot run standardized Ethernet OAM to the CPE. The command "cpe-ping" for a specific destination IP address within a VPLS is translated into a MAC-ping towards a broadcast MAC address. All destinations within the VPLS context are reached by this ping to the broadcast the MAC address. At all these destinations, an ARP will be triggered for the specific IP address (with the IP destination address equals to the address from the request, mac-da equals to all1's, mac-sa equals to the CPM-mac-address and the IP source address, which is the address found in the request). The destination receiving a response will reply back to the requestor.

Release 10.0 extended the CPE ping command for local, distributed, and PBB Epipe services provisioned over a PBB VPLS. CPE ping for Epipe implements an alternative behavior to CPE ping for VPLS that enables fate sharing of the CPE ping request with the Epipe service. Any PE within the epipe service (the source PE) can launch the CPE ping. The source PE builds an arp request and encapsulates it to be sent in the epipe as if it came from a customer device by using its chassis MAC as the source MAC address. The ARP request then egresses the remote PE device as any other packets on the epipe. The remote CPE device responds to the ARP and the reply is transparently sent on the epipe towards the source PE. The source PE will then look for a match on its chassis MAC in the inner customer DA. If a match is found, the source PE device intercepts this response packet.

This method is supported regardless of whether the network uses SDPs or SAPs. It is configured using the existing **oam>cpe-ping** CLI command.

Note: This feature does not support IPv6 CPEs

Hardware Support

This feature supports IOM3 and above.

Any IOM3-supporting mode are subjected to the following check.

To launch cpe-ping on an Epipe, all of the following must be true:

1. All SAPs on the Epipe must be provisioned on slots that are mode-d compatible.
2. If bound to a PBB tunnel, all SAPs on the B-VPLS must be provisioned on slots that are mode-d compatible.
3. If the Epipe or the B-VPLS (in the case of PBB Epipe) uses SDP-bindings then the system configuration must be network-chassis-mode-d compatible.

MAC Populate

MAC populate is used to send a message through the flooding domain to learn a MAC address as if a customer packet with that source MAC address had flooded the domain from that ingress point in the service. This allows the provider to craft a learning history and engineer packets in a particular way to test forwarding plane correctness.

The MAC populate request is sent with a VC TTL of 1, which means that it is received at the forwarding plane at the first hop and passed directly up to the management plane. The packet is then responded to by populating the MAC address in the forwarding plane, like a conventional learn although the MAC will be an OAM-type MAC in the FIB to distinguish it from customer MAC addresses.

This packet is then taken by the control plane and flooded out the flooding domain (squelching appropriately, the sender and other paths that would be squelched in a typical flood).

This controlled population of the FIB is very important to manage the expected results of an OAM test. The same functions are available by sending the OAM packet as a UDP/IP OAM packet. It is then forwarded to each hop and the management plane has to do the flooding.

Options for MAC populate are to force the MAC in the table to type OAM (in case it already existed as dynamic or static or an OAM induced learning with some other binding), to prevent new dynamic learning to over-write the existing OAM MAC entry, to allow customer packets with this MAC to either ingress or egress the network, while still using the OAM MAC entry.

Finally, an option to flood the MAC populate request causes each upstream node to learn the MAC, for example, populate the local FIB with an OAM MAC entry, and to flood the request along the data plane using the flooding domain.

An age can be provided to age a particular OAM MAC after a different interval than other MACs in a FIB.

MAC Purge

MAC purge is used to clear the FIBs of any learned information for a particular MAC address. This allows one to do a controlled OAM test without learning induced by customer packets. In addition to clearing the FIB of a particular MAC address, the purge can also indicate to the control plane not to allow further learning from customer packets. This allows the FIB to be clean, and be populated only via a MAC Populate.

MAC purge follows the same flooding mechanism as the MAC populate.

A UDP/IP version of this command is also available that does not follow the forwarding notion of the flooding domain, but the control plane notion of it.

VLL Diagnostics

VCCV Ping

VCCV ping is used to check connectivity of a VLL in-band. It checks that the destination (target) PE is the egress for the Layer 2 FEC. It provides a cross-check between the data plane and the control plane. It is in-band, meaning that the VCCV ping message is sent using the same encapsulation and along the same path as user packets in that VLL. This is equivalent to the LSP ping for a VLL service. VCCV ping reuses an LSP ping message format and can be used to test a VLL configured over an MPLS and GRE SDP.

VCCV-Ping Application

VCCV effectively creates an IP control channel within the pseudowire between PE1 and PE2. PE2 should be able to distinguish on the receive side VCCV control messages from user packets on that VLL. There are three possible methods of encapsulating a VCCV message in a VLL which translates into three types of control channels:

1. Use of a Router Alert Label immediately above the VC label. This method has the drawback that if ECMP is applied to the outer LSP label (for example, transport label), the VCCV message will not follow the same path as the user packets. This effectively means it will not troubleshoot the appropriate path. This method is supported by the 7750 SR.

2. Use of the OAM control word as illustrated in Figure 23.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |0 0 0 1| FmtID |   Reserved    |         Channel Type          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 23: OAM Control Word Format

The first nibble is set to 0x1. The Format ID and the reserved fields are set to 0 and the channel type is the code point associated with the VCCV IP control channel as specified in the PWE3 IANA registry (RFC 4446). The channel type value of 0x21 indicates that the Associated Channel carries an IPv4 packet.

The use of the OAM control word assumes that the draft-martini control word is also used on the user packets. This means that if the control word is optional for a VLL and is not configured, the 7750 SR PE node will only advertise the router alert label as the CC capability in the Label Mapping message. This method is supported by the 7750 SR.

3. Set the TTL in the VC label to 1 to force PE2 control plane to process the VCCV message. This method is not guaranteed to work under all circumstances. For instance, the draft mentions some implementations of penultimate hop popping overwrite the TTL field. This method is not supported by the 7750 SR.

When sending the label mapping message for the VLL, PE1 and PE2 must indicate which of the above OAM packet encapsulation methods (for example, which control channel type) they support. This is accomplished by including an optional VCCV TLV in the pseudowire FEC Interface Parameter field. The format of the VCCV TLV is shown in Figure 24.

```
0                   1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 0
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |      0x0c      |       0x04     |   CC Types     |   CV Types     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 24: VCCV TLV

Note that the absence of the optional VCCV TLV in the Interface parameters field of the pseudowire FEC indicates the PE has no VCCV capability.

The Control Channel (CC) Type field is a bitmask used to indicate if the PE supports none, one, or many control channel types.

- 0x00 None of the following VCCV control channel types are supported
- 0x01 PWE3 OAM control word (see Figure 23)
- 0x02 MPLS Router Alert Label
- 0x04 MPLS inner label TTL = 1

If both PE nodes support more than one of the CC types, then a 7750 SR PE will make use of the one with the lowest type value. For instance, OAM control word will be used in preference to the MPLS router alert label.

The Connectivity Verification (CV) bitmask field is used to indicate the specific type of VCCV packets to be sent over the VCCV control channel. The valid values are:

0x00 None of the below VCCV packet type are supported.

0x01 ICMP ping. Not applicable to a VLL over a MPLS or GRE SDP and as such is not supported by the 7750 SR.

0x02 LSP ping. This is used in VCCV ping application and applies to a VLL over an MPLS or a GRE SDP. This is supported by the 7750 SR.

A VCCV ping is an LSP echo request message as defined in RFC 4379. It contains an L2 FEC stack TLV which must include within the sub-TLV type 10 “FEC 128 Pseudowire”. It also contains a field which indicates to the destination PE which reply mode to use. There are four reply modes defined in RFC 4379:

Reply mode, meaning:

1. Do not reply. This mode is supported by the 7750 SR.

2. Reply via an IPv4/IPv6 UDP packet. This mode is supported by the 7750 SR.

3. Reply with an IPv4/IPv6 UDP packet with a router alert. This mode sets the router alert bit in the IP header and is not be confused with the CC type which makes use of the router alert label. This mode is not supported by the 7750 SR.

4. Reply via application level control channel. This mode sends the reply message inband over the pseudowire from PE2 to PE1. PE2 will encapsulate the Echo Reply message using the CC type negotiated with PE1. This mode is supported by the 7750 SR.

The reply is an LSP echo reply message as defined in RFC 4379. The message is sent as per the reply mode requested by PE1. The return codes supported are the same as those supported in the 7750 SR LSP ping capability.

The VCCV ping feature is in addition to the service ping OAM feature which can be used to test a service between 7750 SR nodes. The VCCV ping feature can test connectivity of a VLL with any third party node which is compliant to RFC 5085.

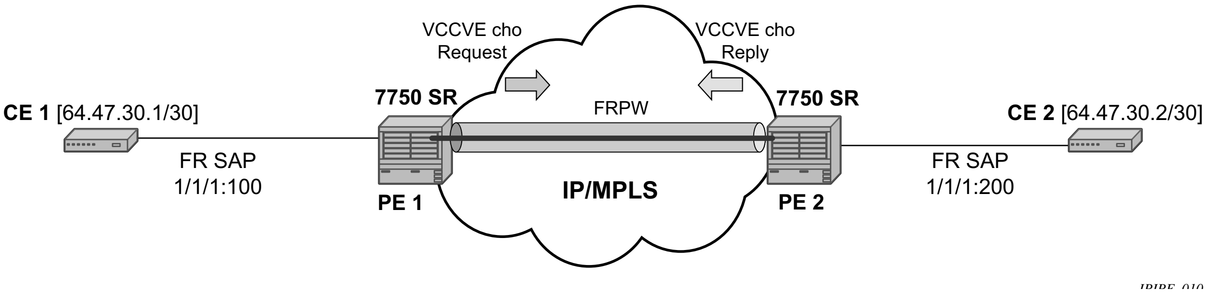

Figure 25: VCCV-Ping Application

VCCV Ping in a Multi-Segment Pseudowire

Figure 26 displays and example of an application of VCCV ping over a multi-segment pseudowire.

Pseudowire switching is a method for scaling a large network of VLL or VPLS services by removing the need for a full mesh of T-LDP sessions between the PE nodes as the number of these nodes grow over time. Pseudowire switching is also used whenever there is a need to deploy a VLL service across two separate routing domains.

In the network, a Termination PE (T-PE) is where the pseudowire originates and terminates. The Switching PE (S-PE) is the node which performs pseudowire switching by cross-connecting two spoke SDPs.

VCCV ping is extended to be able to perform the following OAM functions:

1. VCCV ping to a destination PE. A VLL FEC ping is a message sent by T-PE1 to test the FEC at T-PE2. The operation at T-PE1 and T-PE2 is the same as in the case of a single-segment pseudowire. The pseudowire switching node, S-PE1, pops the outer label, swaps the inner (VC) label, decrements the TTL of the VC label, and pushes a new outer label. The 7750 SR PE1 node does not process the VCCV OAM Control Word unless the VC label TTL expires. In that case, the message is sent to the CPM for further validation and processing. This is the method described in draft-hart-pwe3-segmented-pw-vccv.

Note that the originator of the VCCV ping message does not need to be a T-PE node; it can be an S-PE node. The destination of the VCCV ping message can also be an S-PE node.

VCCV trace to trace the entire path of a pseudowire with a single command issued at the T-PE. This is equivalent to LSP trace and is an iterative process by which T-PE1 sends successive VCCV ping messages while incrementing the TTL value, starting from TTL=1. The procedure for each iteration is the same as above and each node in which the VC label TTL expires checks the FEC and replies with the FEC to the downstream S-PE or T-PE node. The process is terminated when the reply is from T-PE2 or when a timeout occurs.

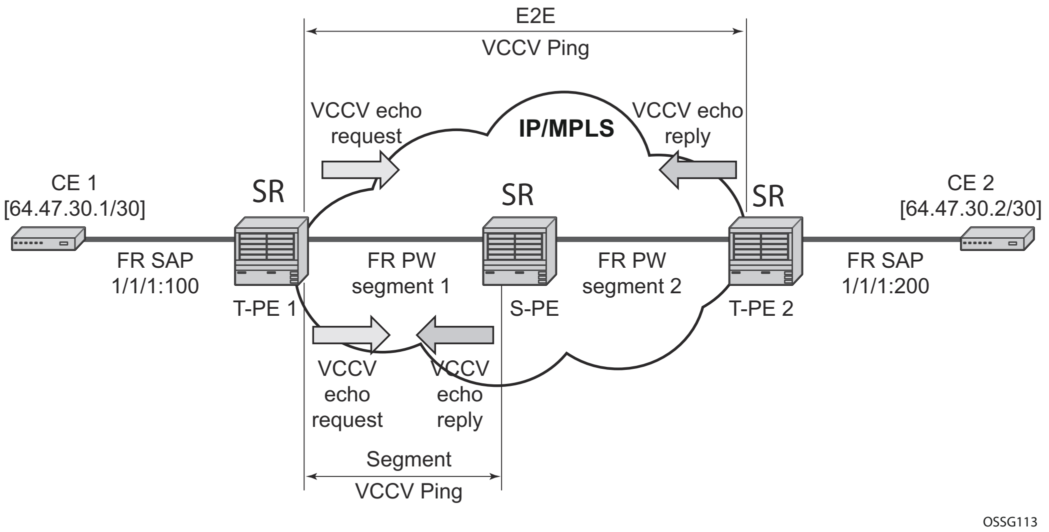

Figure 26: VCCV Ping over a Multi-Segment Pseudowire

Automated VCCV-Trace Capability for MS-Pseudowire

Although tracing of the MS-pseudowire path is possible using the methods explained in previous sections, these require multiple manual iterations and that the FEC of the last pseudowire segment to the target T-PE/S-PE be known a priori at the node originating the echo request message for each iteration. This mode of operation is referred to as a “ping” mode.

The automated VCCV-trace can trace the entire path of a pseudowire with a single command issued at the T-PE or at an S-PE. This is equivalent to LSP-trace and is an iterative process by which the ingress T-PE or T-PE sends successive VCCV-ping messages with incrementing the TTL value, starting from TTL=1.

The method is described in draft-hart-pwe3-segmented-pw-vccv, *VCCV Extensions for Segmented Pseudo-Wire*, and is pending acceptance by the PWE3 working group. In each iteration, the source T-PE or S-PE builds the MPLS echo request message in a way similar to VCCV Ping on page 160. The first message with TTL=1 will have the next-hop S-PE T-LDP session source address in the Remote PE Address field in the pseudowire FEC TLV. Each S-PE which terminates and processes the message will include in the MPLS echo reply message the FEC 128 TLV corresponding the pseudowire segment to its downstream node. The inclusion of the FEC TLV in the echo reply message is allowed in RFC 4379, *Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures*. The source T-PE or S-PE can then build the next echo reply message with TTL=2 to test the next-next hop for the MS-pseudowire. It will copy the FEC TLV it received in the echo reply message into the new echo request message. The process is terminated when the reply is from the egress T-PE or when a timeout occurs. If specified, the max-ttl parameter in the vccv-trace command will stop on SPE before reaching T-PE.

The results VCCV-trace can be displayed for a fewer number of pseudowire segments of the end-to-end MS-pseudowire path. In this case, the min-ttl and max-ttl parameters are configured accordingly. However, the T-PE/S-PE node will still probe all hops up to min-ttl in order to correctly build the FEC of the desired subset of segments.

Note that this method does not require the use of the downstream mapping TLV in the echo request and echo reply messages.

VCCV for Static Pseudowire Segments

MS pseudowire is supported with a mix of static and signaled pseudowire segments. However, VCCV ping and VCCV-trace is allowed until at least one segment of the MS pseudowire is static. Users cannot test a static segment but also, cannot test contiguous signaled segments of the MS-pseudowire. VCCV ping and VCCV trace is not supported in static-to-dynamic configurations.

Detailed VCCV-Trace Operation

In Figure 26 on page 164 a trace can be performed on the MS-pseudowire originating from T-PE1 by a single operational command. The following process occurs:

1. T-PE1 sends a VCCV echo request with TTL set to 1 and a FEC 128 containing the pseudo-wire information of the first segment (pseudowire1 between T-PE1 and S-PE) to S-PE for validation.

2. S-PE validates the echo request with the FEC 128. Since it is a switching point between the first and second segment it builds an echo reply with a return code of 8 and includes the FEC 128 of the second segment (pseudowire2 between S-PE and T-PE2) and sends the echo reply back to T-PE1.

3. T-PE1 builds a second VCCV echo request based on the FEC128 in the echo reply from the S-PE. It increments the TTL and sends the next echo request out to T-PE2. Note that the VCCV echo request packet is switched at the S-PE datapath and forwarded to the next downstream segment without any involvement from the control plane.

4. T-PE2 receives and validates the echo request with the FEC 128 of the pseudowire2 from T-PE1. Since T-PE2 is the destination node or the egress node of the MS-pseudowire it replies to T-PE1 with an echo reply with a return code of 3, (egress router) and no FEC 128 is included.

5. T-PE1 receives the echo reply from T-PE2. T-PE1 is made aware that T-PE2 is the destination of the MS pseudowire because the echo reply does not contain the FEC 128 and because its return code is 3. The trace process is completed.

Control Plane Processing of a VCCV Echo Message in a MS-Pseudowire

Sending a VCCV Echo Request

When in the ping mode of operation, the sender of the echo request message requires the FEC of the last segment to the target S-PE/T-PE node. This information can either be configured manually or be obtained by inspecting the corresponding sub-TLV's of the pseudowire switching point TLV. However, the pseudowire switching point TLV is optional and there is no guarantee that all S-PE nodes will populate it with their system address and the pseudowire-id of the last pseudowire segment traversed by the label mapping message. Thus the 7750 SR implementation will always make use of the user configuration for these parameters.

When in the trace mode operation, the T-PE will automatically learn the target FEC by probing one by one the hops of the MS-pseudowire path. Each S-PE node includes the FEC to the downstream node in the echo reply message in a similar way that LSP trace will have the probed node return the downstream interface and label stack in the echo reply message.

Receiving an VCCV Echo Request

Upon receiving a VCCV echo request the control plane on S-PEs (or the target node of each segment of the MS pseudowire) validates the request and responds to the request with an echo reply consisting of the FEC 128 of the next downstream segment and a return code of 8 (label switched at stack-depth) indicating that it is an S-PE and not the egress router for the MS-pseudowire.

If the node is the T-PE or the egress node of the MS-pseudowire, it responds to the echo request with an echo reply with a return code of 3 (egress router) and no FEC 128 is included.

Receiving an VCCV Echo Reply

The operation to be taken by the node that receives the echo reply in response to its echo request depends on its current mode of operation such as ping or trace.

In ping mode, the node may choose to ignore the target FEC 128 in the echo reply and report only the return code to the operator.

However, in trace mode, the node builds and sends the subsequent VCCV echo request with a incrementing TTL and the information (such as the downstream FEC 128) it received in the echo request to the next downstream pseudowire segment.

IGMP Snooping Diagnostics

MFIB Ping

The multicast forwarding information base (MFIB) ping OAM tool allows to easily verify inside a VPLS which SAPs would normally egress a certain multicast stream. The multicast stream is identified by a source unicast and destination multicast IP address, which are mandatory when issuing an MFIB ping command.

An MFIB ping packet will be sent through the data plane and goes out with the data plane format containing a configurable VC label TTL. This packet traverses each hop using forwarding plane information for next hop, VC label, etc. The VC label is swapped at each service-aware hop, and the VC TTL is decremented. If the VC TTL is decremented to 0, the packet is passed up to the management plane for processing. If the packet reaches an egress node, and would be forwarded out a customer facing port (SAP), it is identified by the OAM label below the VC label and passed to the management plane.

ATM Diagnostics

The ATM OAM ping allows operators to test VC-integrity and endpoint connectivity for existing PVCCs using OAM loopback capabilities.

If portId:vpi/vci PVCC does not exist, a PVCC is administratively disabled, or there is already a ping executing on this PVCC, then this command returns an error.

Because oam atm-ping is a dynamic operation, the configuration is not preserved. The number of oam atm-ping operations that can be performed simultaneously on a 7750 SR7450 ESS7710 SR is configurable as part of the general OAM MIB configuration.

An operator can specify the following options when performing an oam atm-ping:

> **end-to-end** – this option allows sending oam atm-ping towards the connection endpoint in the line direction by using OAM end-to-end loopback cells
>
> **segment** – this option allows sending oam atm-ping towards the segment termination point in the line direction by using OAM segment loopback cells.

The result of ATM ping will show if the ping to a given location was successful. It also shows the round-trip time the ping took to complete (from the time the ping was injected in the ATM SAR device until the time the ping response was given to S/W by the ATM SAR device) and the average ping time for successful attempts up to the given ping response.

An oam atm ping in progress will time-out if a PVCC goes to the operational status down as result of a network failure, an administrative action, or if a PVCC gets deleted. Any subsequent ping attempts will fail until the VC’s operational state changes to up.

To stop a ping in progress, an operator can enter “CTRL – C”. This will stop any outstanding ping requests and will return ping result up to the point of interruption (a ping in progress during the above stop request will fail).

MPLS-TP On-Demand OAM Commands

Ping and Trace tools for PWs and LSPs must be supported with both IP encapsulation and the MPLS-TP on demand CV channel for non-IP encapsulation (0x025).

MPLS-TP Pseudowires: VCCV-Ping/VCCV-Trace

For vccv-ping and vccv-trace commands:

- Sub-type **static** must be specified. This indicates to the system that the rest of the command contains parameters that are applied to a static PW with a static PW FEC.
- Add the ability to specify the non-IP ACH channel type (0x0025). This is known as the **non-ip control-channel**. This is the default for type static. GAL is not supported for PWs.
- If the ip-control-channel is specified as the encapsulation, then the IPv4 channel type is used (0x0021). In this case, a destination IP address in the 127/8 range is used, while the source address in the UDP/IP packet is set to the system IP address, or may be explicitly configured by the user with the src-ip-address option. This option is only valid if the IPv4 control-channel is specified.
- The reply mode are always assumed to be the same application level control channel type for type static.
- Allow an MPLS-TP global-id and node-id specified under the spoke-sdps with a given sdp-id:vc-id, used for MPLS-TP PW MEPs, or node-id (prefix) only for MIPs.
- The following CLI command description shows the options that are only allowed if the type static option is configured. All other options are blocked.
- As in the existing implementation, the downstream mapping and detailed downstream mapping TLVs (DSMAP/DDMAP TLVs) is not supported on PWs.

```
vccv-ping static <sdp-id:vc-id> [dest-global-id <global-id> dest-node-id <node-id>] [con-
trol-channel ipv4 | non-ip] [fc <fc-name> [profile {in|out}]] [size <octets>] [count <send-
count>] [timeout <timeout>] [interval <interval>] [ttl <vc-label-ttl>][src-ip-address <ip-
address>]
vccv-trace static <sdp-id:vc-id>  [size <octets>][min-ttl <min-vc-label-ttl>][max-ttl
<max-vc-label-ttl>][max-fail <no-response-count>][probe-count <probe-count>] [control-
channel ipv4 | non-ip] [timeout <timeout-value>][interval <interval-value>][fc <fc-name>
[profile {in|out}]][src-ip-address <ip-address>] [detail]
```

If the spoke-sdp referred to by sdp-id:vc-id has an MPLS-TP PW-Path-ID defined, then those parameters are used to populate the static PW TLV in the target FEC stack of the vccv-ping or vccv-trace packet. If a Global-ID and Node-ID are specified in the command, then these values are used to populate the destination node TLV in the vccv-ping or vccv-trace packet.

The global-id/node-id are only used as the target node identifiers if the vccv-ping is not end-to-end (for example, a TTL is specified in the vccv-ping/trace command and it is < 255); otherwise, the value in the PW Path ID is used. For vccv-ping, the dest-node-id may be entered as a 4-octet IP address <a.b.c.d> or 32-bit integer <1..4294967295>. For vccv-trace, the destination node-id and global-id are taken form the spoke-sdp context.

The same command syntax is applicable for SAA tests configured under configure saa test a type.

MPLS-TP LSPs: LSP-Ping/LSP Trace

For lsp-ping and lsp-trace commands:

- sub-type **static** must be specified. This indicates to the system that the rest of the command contains parameters specific to a LSP identified by a static LSP FEC.
- The 7x50 supports the use of the G-ACh with non-IP encapsulation or labeled encapsulation with IP de-multiplexing for both the echo request and echo reply for LSP-Ping and LSP-Trace on LSPs with a static LSP FEC (such as MPLS-TP LSPs).
- It is possible to specify the target MPLS-TP MEP/MIP identifier information for LSP Ping. If the target global-id and node-id are not included in the lsp-ping command, then these parameters for the target MEP ID are taken from the context of the LSP. The **tunnel-number** <tunnel-num> and **lsp-num** <lsp-num> for the far-end MEP are always taken from the context of the path under test.

```
lsp-ping static <lsp-name>
     [force]
     [path-type [active|working|protect]]
     [fc <fc-name> [profile {in|out}]]
     [size <octets>]
     [ttl <label-ttl>]
     [send-count <send-count>]
     [timeout <timeout>]
     [interval <interval>]
     [src-ip-address <ip-address>]
     [dest-global-id <dest-global-id> dest-node-id dest-node-id]
     [control-channel none | non-ip][detail]
lsp-trace static  <lsp-name>
     [force]
     [path-type [active|working|protect]
     [fc <fc-name> [profile {in|out}]]
     [max-fail <no-response-count>]
     [probe-count <probes-per-hop>]
     [size <octets>]
     [min-ttl <min-label-ttl>]
     [max-ttl <max-label-ttl>]
     [timeout <timeout>]
     [interval <interval>]
     [src-ip-address <ip-address>]
      [control-channel none | non-ip]
     [downstream-map-tlv <dsmap|ddmap>]
     [detail]
```

The following commands are only valid if the sub-type **static** option is configured, implying that lsp-name refers to an MPLS-TP tunnel LSP:

path-type. Values: active, working, protect. Default: active.

dest-global-id <global-id> **dest-node-id** <node-id>: Default: the **to** global-id:node-id from the LSP ID.

control-channel: If this is set to none, then IP encapsulation over an LSP is used with a destination address in the 127/8 range. The source address is set to the system IP address, unless the user specifies a source address using the src-ip-address option. If this is set to non-ip, then non-IP encapsulation over a G-ACh with channel type 0x00025 is used. This is the default for sub-type static. Note that the encapsulation used for the echo reply is the same as the encapsulation used for the echo request.

downstream-map-tlv: LSP Trace commands with this option can only be executed if the control-channel is set to none. The DSMAP/DDMAP TLV is only included in the echo request message if the egress interface is either a numbered IP interface, or an unnumbered IP interface. The TLV will not be included if the egress interface is of type **unnumbered-mpls-tp**.

For lsp-ping, the dest-node-id may be entered as a 4-octet IP address <a.b.c.d> or 32-bit integer <1..4294967295>. For lsp-trace, the destination node-id and global-id are taken form the spoke-sdp context.

The send mode and reply mode are always taken to be an application level control channel for MPLS-TP.

The **force** parameter causes an LSP ping echo request to be sent on an LSP that has been brought oper-down by BFD (LSP-Ping echo requests would normally be dropped on oper-down LSPs). This parameter is not applicable to SAA.

The LSP ID used in the LSP Ping packet is derived from a context lookup based on lsp-name and path-type (active/working/protect).

Dest-global-id and dest-node-id refer to the target global/node id. They do not need to be entered for end-to-end ping and trace, and the system will use the destination global id and node id from the LSP ID.

The same command syntax is applicable for SAA tests configured under **configure>saa>test**.

Mirroring for MPLS-TP

Bidirectional MPLS-TP spoke-sdps with a configured pw-path-id can transport a mirrored service. Note that mirror services are not supported on static PWs with an MPLS-TP pw-path-id bound to an SDP that uses an RSVP-TE LSP.

Control channel status signaling is supported with PW redundancy on spoke-sdps in a mirror context.

The following is an example of PW redundancy for a mirror service. In this case, MPLS-TP spoke-sdps are used.

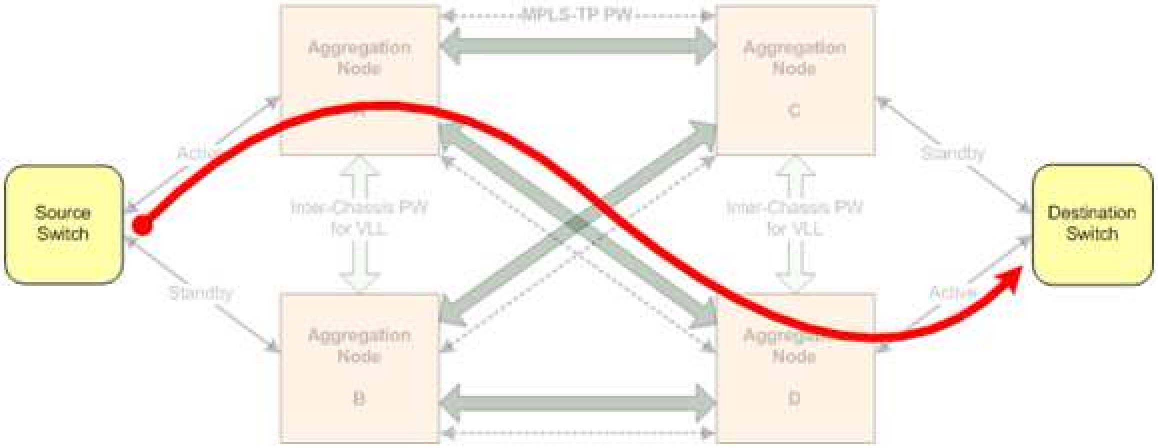

Figure 27: Mirroring with PW Redundancy using MPLS-TP

Note that mirroring traffic is usually unidirectional, flowing from "source" nodes (A or B) to "destination" nodes (C or D). However in case of MPLS-TP, the control channel status packets may flow in the reverse direction.

An example mirror service using MPLS-TP spoke-sdps is configured as follows:

Source Node A

```
debug mirror-source 20
   sap lag-2:100 ingress egress
   exit

mirror-dest 20 create
   endpoint "tx" create
   exit
   spoke-sdp 1000:20 endpoint "tx" create
      ingress
         vc-label 1000
         exit
      egress
         vc-label 2000
         exit
      control-word
      control-channel-status
         exit
      pw-path-id
         exit
      no shutdown
      exit
   spoke-sdp 2000:20 endpoint "tx" create
      ingress
         vc-label 1000
         exit
      egress
         vc-label 2000
         exit
      control-word
      control-channel-status
         exit
      pw-path-id
         exit
      no shutdown
      exit
no shutdown
exit
```

Destination Node C

```
mirror-dest 20 create
   endpoint "rx" create
      exit
   endpoint "tx" create
      exit
   remote-source
   spoke-sdp 1000:20 endpoint "rx" create  ! From node A
      ingress
         vc-label 2000
         exit
      egress
         vc-label 1000
         exit
      control-word
      control-channel-status
        exit
      pw-path-id
         exit
      no shutdown
      exit
   spoke-sdp 2001:20 endpoint "rx" create  ! From node B
      ingress
         vc-label 2000
            exit
      egress
         vc-label 1000
         exit
      control-word
      control-channel-status
        exit
      pw-path-id
         exit
      no shutdown
      exit

   spoke 3001:20 endpoint "rx" icb create  ! ICB from other destination node D
      ingress
         vc-label 2000
        exit
   egress
      vc-label 1000
         exit
   control-word
   control-channel-status
      exit
   pw-path-id
      exit
   no shutdown
   exit
sap lag-1:20 endpoint "tx" create
   exit
spoke 3000:20 endpoint "tx" icb create  ! ICB to other destination node D
   ingress
      vc-label 1000
         exit
   egress
      vc-label 2000
         exit
   control-word
   control-channel-status
      exit
   pw-path-id
      exit
   no shutdown
      exit
   no shutdown
   exit
no shutdown

exit
```

MPLS-TP Show Commands

Static MPLS Labels

The following new commands show the details of the static MPLS labels.

show>router>mpls-labels>label <start-label> [<end-label> [in-use|<label-owner>]]

show>router>mpls-labels>label-range

An example output is as follows:

```
*A:mlstp-dutA# show router mpls
mpls         mpls-labels
*A:mlstp-dutA# show router mpls label
label        label-range
*A:mlstp-dutA# show router mpls label-range
```

```
===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label     End Label       Aging          Total Available
-------------------------------------------------------------------------------
Static-lsp      32              16415           -              16364
Static-svc      16416           32799           -              16376
Dynamic         32800           131071          0              98268
===============================================================================
```

MPLS-TP Tunnel Configuration

These should show the configuration of a given tunnel.

show>router>mpls>tp-lsp

A sample output is as follows:

```
*A:mlstp-dutA# show router mpls tp-lsp
  - tp-lsp [<lsp-name>] [status {up|down}] [from <ip-address>|to <ip-address>]
    [detail]
  - tp-lsp [<lsp-name>] path [protect|working] [detail]
  - tp-lsp [<lsp-name>] protection

 <lsp-name>           : [32 chars max] - accepts * as wildcard char
 <path>               : keyword - Display LSP path information.
 <protection>         : keyword - Display LSP protection information.
 <up|down>            : keywords - Specify state of the LSP
 <ip-address>         : a.b.c.d
 <detail>             : keyword - Display detailed information.
*A:mlstp-dutA# show router mpls tp-lsp
path
protection
to <a.b.c.d>
<lsp-name>
 "lsp-32"  "lsp-33"  "lsp-34"  "lsp-35"  "lsp-36"  "lsp-37"  "lsp-38"  "lsp-39"
 "lsp-40"  "lsp-41"
status {up|down}
from <ip-address>
detail

*A:mlstp-dutA# show router mpls tp-lsp "lsp-
"lsp-32"  "lsp-33"  "lsp-34"  "lsp-35"  "lsp-36"  "lsp-37"  "lsp-38"  "lsp-39"
"lsp-40"  "lsp-41"
*A:mlstp-dutA# show router mpls tp-lsp "lsp-32"

===============================================================================
MPLS MPLS-TP LSPs (Originating)
===============================================================================
LSP Name                           To               Tun     Protect   Adm  Opr
                                                    Id      Path
-------------------------------------------------------------------------------
lsp-32                             0.0.3.234        32      No        Up   Up
-------------------------------------------------------------------------------
LSPs : 1
```

```
===============================================================================
*A:mlstp-dutA# show router mpls tp-lsp "lsp-32" detail

===============================================================================
MPLS MPLS-TP LSPs (Originating) (Detail)
===============================================================================
-------------------------------------------------------------------------------
Type : Originating
-------------------------------------------------------------------------------
LSP Name    : lsp-32
LSP Type    : MplsTp                           LSP Tunnel ID  : 32
From Node Id: 0.0.3.233+                       To Node Id     : 0.0.3.234
Adm State   : Up                               Oper State     : Up
LSP Up Time : 0d 04:50:47                      LSP Down Time  : 0d 00:00:00
Transitions : 1                                Path Changes   : 2

DestGlobalId: 42                               DestTunnelNum  : 32
```

MPLS-TP Path configuration.

This can reuse and augment the output of the current show commands for static LSPs. They should also show if BFD is enabled on a given path. If this referring to a transit path, this should also display (among others) the path-id (7 parameters) for a given transit-path-name, or the transit-path-name for a given the path-id (7 parameters)

show>router>mpls>tp-lsp>path

A sample output is as follows:

```
===============================================================================
*A:mlstp-dutA#  show router mpls tp-lsp path

===============================================================================
MPLS-TP LSP Path Information
===============================================================================
LSP Name      : lsp-32                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       32        32        AtoB_1          Up     Down
Protect                       2080      2080      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-33                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       33        33        AtoB_1          Up     Down
Protect                       2082      2082      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-34                           To            : 0.0.3.234
```

```
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       34        34        AtoB_1          Up     Down
Protect                       2084      2084      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-35                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       35        35        AtoB_1          Up     Down
Protect                       2086      2086      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-36                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       36        36        AtoB_1          Up     Down
Protect                       2088      2088      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-37                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       37        37        AtoB_1          Up     Down
Protect                       2090      2090      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-38                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       38        38        AtoB_1          Up     Down
Protect                       2092      2092      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-39                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       39        39        AtoB_1          Up     Down
Protect                       2094      2094      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-40                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
```

```
Working                       40        40        AtoB_1          Up     Down
Protect                       2096      2096      AtoC_1          Up     Up
===============================================================================
LSP Name      : lsp-41                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       41        41        AtoB_1          Up     Down
Protect                       2098      2098      AtoC_1          Up     Up

*A:mlstp-dutA#  show router mpls tp-lsp "lsp-32" path working

===============================================================================
MPLS-TP LSP Working Path Information
    LSP: "lsp-32"
===============================================================================
LSP Name      : lsp-32                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Working                       32        32        AtoB_1          Up     Down
===============================================================================
*A:mlstp-dutA#  show router mpls tp-lsp "lsp-32" path protect

===============================================================================
MPLS-TP LSP Protect Path Information
    LSP: "lsp-32"
===============================================================================
LSP Name      : lsp-32                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

-------------------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F         Admin  Oper
-------------------------------------------------------------------------------
Protect                       2080      2080      AtoC_1          Up     Up
===============================================================================
*A:mlstp-dutA#  show router mpls tp-lsp "lsp-32" path protect detail

===============================================================================
MPLS-TP LSP Protect Path Information
    LSP: "lsp-32" (Detail)
===============================================================================
LSP Name      : lsp-32                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

Protect path information
-------------------------------------------------------------------------------
Path Type     : Protect                          LSP Num       : 2
Path Admin    : Up                               Path Oper     : Up
Out Interface : AtoC_1                           Next Hop Addr : n/a
In Label      : 2080                             Out Label     : 2080
Path Up Time  : 0d 04:52:17                      Path Dn Time  : 0d 00:00:00
Active Path   : Yes                              Active Time   : 0d 00:52:56

MEP information
MEP State     : Up                               BFD           : cc
OAM Templ     : privatebed-oam-template          CC Status     : inService
                                                 CV Status     : unknown
Protect Templ : privatebed-protection-template   WTR Count Down: 0 seconds
RX PDU        : SF (1,1)                         TX PDU        : SF (1,1)
Defects       :
===============================================================================
*A:mlstp-dutA#  show router mpls tp-lsp "lsp-32" path working detail

===============================================================================
MPLS-TP LSP Working Path Information
    LSP: "lsp-32" (Detail)
===============================================================================
LSP Name      : lsp-32                           To            : 0.0.3.234
Admin State   : Up                               Oper State    : Up

Working path information
-------------------------------------------------------------------------------
Path Type     : Working                          LSP Num       : 1
Path Admin    : Up                               Path Oper     : Down
Down Reason   : ccFault ifDn
Out Interface : AtoB_1                           Next Hop Addr : n/a
In Label      : 32                               Out Label     : 32
Path Up Time  : 0d 00:00:00                      Path Dn Time  : 0d 00:53:01
Active Path   : No                               Active Time   : n/a

MEP information
MEP State     : Up                               BFD           : cc
OAM Templ     : privatebed-oam-template          CC Status     : outOfService
                                                 CV Status     : unknown
===============================================================================
*A:mlstp-dutA#
```

MPLS-TP Protection.

These should show the protection configuration for a given tunnel, which path in a tunnel is currently working and which is protect, and whether the working or protect is currently active.

show>router>mpls>tp-lsp>protection

A sample output is as follows:

```
*A:mlstp-dutA#  show router mpls tp-lsp protection

===============================================================================
MPLS-TP LSP Protection Information
Legend: W-Working, P-Protect,
===============================================================================
LSP Name                      Admin Oper  Path    Ingr/Egr      Act. Rx PDU
                              State State   State Label         Path Tx PDU
-------------------------------------------------------------------------------
lsp-32                        Up    Up    W Down      32/32     No   SF (1,1)
                                          P Up      2080/2080   Yes  SF (1,1)
lsp-33                        Up    Up    W Down      33/33     No   SF (1,1)
                                          P Up      2082/2082   Yes  SF (1,1)
```

```
lsp-34                        Up    Up    W Down      34/34     No   SF (1,1)
                                          P Up      2084/2084   Yes  SF (1,1)
lsp-35                        Up    Up    W Down      35/35     No   SF (1,1)
                                          P Up      2086/2086   Yes  SF (1,1)
lsp-36                        Up    Up    W Down      36/36     No   SF (1,1)
                                          P Up      2088/2088   Yes  SF (1,1)
lsp-37                        Up    Up    W Down      37/37     No   SF (1,1)
                                          P Up      2090/2090   Yes  SF (1,1)
lsp-38                        Up    Up    W Down      38/38     No   SF (1,1)
                                          P Up      2092/2092   Yes  SF (1,1)
lsp-39                        Up    Up    W Down      39/39     No   SF (1,1)
                                          P Up      2094/2094   Yes  SF (1,1)
lsp-40                        Up    Up    W Down      40/40     No   SF (1,1)
                                          P Up      2096/2096   Yes  SF (1,1)
lsp-41                        Up    Up    W Down      41/41     No   SF (1,1)
                                          P Up      2098/2098   Yes  SF (1,1)
-------------------------------------------------------------------------------
No. of MPLS-TP LSPs: 10
===============================================================================
```

BFD

The existing show>router>bfd context should be enhanced for MPLS-TP, as follows:

show>router>bfd>mpls-tp-lsp

Displays the MPLS –TP paths for which BFD is enabled.

show>router>bfd>session [src <ip-address> [dest <ip-address> | detail]]|[mpls-tp-path <lsp-id…> [detail]]

Should be enhanced to show the details of the BFD session on a particular MPLS-TP path, where lsp-id is the fully qualified lsp-id to which the BFD session is in associated.

A sample output is as follows:

```
*A:mlstp-dutA# show router bfd
  - bfd

      bfd-template    - Display BFD Template information
      interface       - Display Interfaces with BFD
      session         - Display session information

*A:mlstp-dutA# show router bfd bfd-template "privatebed-bfd-template"

===============================================================================
BFD Template privatebed-bfd-template
===============================================================================
Template Name           : privatebed-* Template Type             : cpmNp
Transmit Timer          : 10 msec      Receive Timer             : 10 msec
CV Transmit Interval    : 1000 msec
Template Multiplier     : 3            Echo Receive Interval     : 100 msec
```

```
Mpls-tp Association
privatebed-oam-template
===============================================================================
* indicates that the corresponding row element may have been truncated.
*A:mlstp-dutA# show router bfd session

===============================================================================
BFD Session
===============================================================================
Interface/Lsp Name            State                 Tx Intvl  Rx Intvl  Multipl
  Remote Address/Info         Protocols             Tx Pkts   Rx Pkts   Type
-------------------------------------------------------------------------------
wp::lsp-32                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-33                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-34                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-35                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-36                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-37                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-38                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-39                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-40                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
wp::lsp-41                    Down (1)              1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-32                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-33                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-34                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-35                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-36                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-37                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-38                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-39                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-40                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
pp::lsp-41                    Up (3)                1000      1000      3
    0::0.0.0.0                mplsTp                N/A       N/A       cpm-np
-------------------------------------------------------------------------------
No. of BFD sessions: 20
-------------------------------------------------------------------------------
wp = Working path   pp = Protecting path
===============================================================================
```

MPLS TP Node Configuration

Displays the Global ID, Node ID and other general MPLS-TP configurations for the node.

show>router>mpls>mpls-tp

A sample output is as follows:

```
*A:mlstp-dutA# show router mpls mpls-tp
  - mpls-tp


      oam-template    - Display MPLS-TP OAM Template information
      protection-tem* - Display MPLS-TP Protection Template information
      status          - Display MPLS-TP system configuration
      transit-path    - Display MPLS-TP Tunnel information

*A:mlstp-dutA# show router mpls mpls-tp oam-template

===============================================================================
MPLS-TP OAM Templates
===============================================================================
Template Name : privatebed-oam-template Router ID     : 1
BFD Template  : privatebed-bfd-template Hold-Down Time: 0 centiseconds
                                        Hold-Up Time  : 20 deciseconds
===============================================================================
*A:mlstp-dutA# show router mpls mpls-tp protection-template

===============================================================================
MPLS-TP Protection Templates
===============================================================================
Template Name  : privatebed-protection-template  Router ID      : 1
Protection Mode: one2one                         Direction      : bidirectional
Revertive      : revertive                       Wait-to-Restore: 300sec
Rapid-PSC-Timer: 10ms                            Slow-PSC-Timer : 5sec
===============================================================================
*A:mlstp-dutA# show router mpls mpls-tp status

===============================================================================
MPLS-TP Status
===============================================================================
Admin Status  : Up
Global ID     : 42                               Node ID       : 0.0.3.233
Tunnel Id Min : 1                                Tunnel Id Max : 4096
===============================================================================
*A:mlstp-dutA# show router mpls mpls-tp transit-path
  - transit-path [<path-name>] [detail]

 <path-name>          : [32 chars max]
 <detail>             : keyword - Display detailed information.
```

```
A:mplstp-dutC# show router mpls mpls-tp transit-path
  - transit-path [<path-name>] [detail]

 <path-name>          : [32 chars max]
 <detail>             : keyword - Display detailed information.


A:mplstp-dutC# show router mpls mpls-tp transit-path
<path-name>
 "tp-32"   "tp-33"   "tp-34"   "tp-35"   "tp-36"   "tp-37"   "tp-38"   "tp-39"
 "tp-40"   "tp-41"
detail

A:mplstp-dutC# show router mpls mpls-tp transit-path "tp-32"

===============================================================================
MPLS-TP Transit tp-32 Path Information
===============================================================================
Path Name     : tp-32
Admin State   : Up                               Oper State    : Up

------------------------------------------------------------------
Path        NextHop           InLabel   OutLabel  Out I/F
------------------------------------------------------------------
FP                            2080      2081      CtoB_1
RP                            2081      2080      CtoA_1
===============================================================================
A:mplstp-dutC# show router mpls mpls-tp transit-path "tp-32" detail

===============================================================================
MPLS-TP Transit tp-32 Path Information (Detail)
===============================================================================
Path Name     : tp-32
Admin State   : Up                               Oper State    : Up
-------------------------------------------------------------------------------
Path ID configuration
Src Global ID : 42                               Dst Global ID : 42
Src Node ID   : 0.0.3.234                        Dst Node ID   : 0.0.3.233
LSP Number    : 2                                Dst Tunnel Num: 32

Forward Path configuration
In Label      : 2080                             Out Label     : 2081
Out Interface : CtoB_1                           Next Hop Addr : n/a

Reverse Path configuration
In Label      : 2081                             Out Label     : 2080
Out Interface : CtoA_1                           Next Hop Addr : n/a
===============================================================================
A:mplstp-dutC#
```

MPLS-TP Interfaces

The existing show>router>interface command should be enhanced to display mpls-tp specific information.

The following is a sample output:

```
*A:mlstp-dutA# show router interface "AtoB_1"

===============================================================================
Interface Table (Router: Base)
===============================================================================
Interface-Name                   Adm         Opr(v4/v6)  Mode    Port/SapId
   IP-Address                                                    PfxState
-------------------------------------------------------------------------------
AtoB_1                           Down        Down/--     Network 1/2/3:1
   Unnumbered If[system]                                         n/a
-------------------------------------------------------------------------------
Interfaces : 1
```

Services using MPLS-TP PWs

The show>service command shuld be updated to display MPLS-TP specifc information such as the PW Path ID and control channel status signaling parameters.

The following is a sample output:

```
*A:mlstp-dutA# show service id 1 all

===============================================================================
Service Detailed Information
===============================================================================
Service Id        : 1                   Vpn Id            : 0
Service Type      : Epipe
Name              : (Not Specified)
Description       : (Not Specified)
Customer Id       : 1                   Creation Origin   : manual
Last Status Change: 12/03/2012 15:26:20
Last Mgmt Change  : 12/03/2012 15:24:57
Admin State       : Up                  Oper State        : Up
MTU               : 1514
Vc Switching      : False
SAP Count         : 1                   SDP Bind Count    : 1
Per Svc Hashing   : Disabled
Force QTag Fwd    : Disabled

-------------------------------------------------------------------------------
ETH-CFM service specifics
-------------------------------------------------------------------------------
Tunnel Faults     : ignore


-------------------------------------------------------------------------------
Service Destination Points(SDPs)
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
 Sdp Id 32:1  -(0.0.3.234:42)
-------------------------------------------------------------------------------
Description     : (Not Specified)
```

```
SDP Id             : 32:1                     Type              : Spoke
Spoke Descr     : (Not Specified)
VC Type            : Ether                    VC Tag            : n/a
Admin Path MTU     : 0                        Oper Path MTU     : 9186
Delivery           : MPLS
Far End            : 0.0.3.234:42
Tunnel Far End     : n/a                      LSP Types         : MPLSTP
Hash Label         : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label    : Disabled

Admin State        : Up                       Oper State        : Up
Acct. Pol          : None                     Collect Stats     : Disabled
Ingress Label      : 16416                    Egress Label      : 16416
Ingr Mac Fltr-Id   : n/a                      Egr Mac Fltr-Id   : n/a
Ingr IP Fltr-Id    : n/a                      Egr IP Fltr-Id    : n/a
Ingr IPv6 Fltr-Id  : n/a                      Egr IPv6 Fltr-Id  : n/a
Admin ControlWord  : Preferred                Oper ControlWord  : True
Admin BW(Kbps)     : 0                        Oper BW(Kbps)     : 0
Last Status Change : 12/03/2012 15:26:20      Signaling         : None
Last Mgmt Change   : 12/03/2012 15:24:57      Force Vlan-Vc     : Disabled
Endpoint           : N/A                      Precedence        : 4
PW Status Sig      : Enabled
Class Fwding State : Down
Flags              : None
Local Pw Bits      : None
Peer Pw Bits       : None
Peer Fault Ip      : None
Peer Vccv CV Bits  : None
Peer Vccv CC Bits  : None
Application Profile: None
Standby Sig Slave  : False
Block On Peer Fault: False

Ingress Qos Policy : (none)                   Egress Qos Policy : (none)
Ingress FP QGrp    : (none)                   Egress Port QGrp  : (none)
Ing FP QGrp Inst   : (none)                   Egr Port QGrp Inst: (none)

Statistics            :
I. Fwd. Pkts.      : 272969957                I. Dro. Pkts.     : 0
E. Fwd. Pkts.      : 273017433                E. Fwd. Octets    : 16381033352

-------------------------------------------------------------------------------
Control Channel Status
-------------------------------------------------------------------------------
PW Status          : enabled                  Refresh Timer     : 66 secs
Peer Status Expire : false                    Clear On Timeout  : true

-------------------------------------------------------------------------------
SDP-BIND PW Path Information
-------------------------------------------------------------------------------
AGI                : 1:1
SAII Type2         : 42:0.0.3.234:1
TAII Type2         : 42:0.0.3.233:1

-------------------------------------------------------------------------------
RSVP/Static LSPs
-------------------------------------------------------------------------------
Associated LSP List :
Lsp Name           : lsp-32
```

```
Admin State        : Up                       Oper State        : Up

*A:mlstp-dutA# show service id [1..4] all | match "Control Channel" pre-lines 1 post-lines
5
-------------------------------------------------------------------------------
Control Channel Status
-------------------------------------------------------------------------------
PW Status          : enabled                  Refresh Timer     : 66 secs
Peer Status Expire : false                    Clear On Timeout  : true

-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Control Channel Status
-------------------------------------------------------------------------------
PW Status          : enabled                  Refresh Timer     : 66 secs
Peer Status Expire : false                    Clear On Timeout  : true

-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Control Channel Status
-------------------------------------------------------------------------------
PW Status          : enabled                  Refresh Timer     : 66 secs
Peer Status Expire : false                    Clear On Timeout  : true

-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Control Channel Status
-------------------------------------------------------------------------------
PW Status          : enabled                  Refresh Timer     : 66 secs
Peer Status Expire : false                    Clear On Timeout  : true

-------------------------------------------------------------------------------
*A:mlstp-dutA#  show service id [1..4] all | match SDP-BIND pre-lines 1 post-lines 5
-------------------------------------------------------------------------------
SDP-BIND PW Path Information
-------------------------------------------------------------------------------
AGI                : 1:1
SAII Type2         : 42:0.0.3.234:1
TAII Type2         : 42:0.0.3.233:1

-------------------------------------------------------------------------------
SDP-BIND PW Path Information
-------------------------------------------------------------------------------
AGI                : 1:2
SAII Type2         : 42:0.0.3.234:2
TAII Type2         : 42:0.0.3.233:2

-------------------------------------------------------------------------------
SDP-BIND PW Path Information
-------------------------------------------------------------------------------
AGI                : 1:3
SAII Type2         : 42:0.0.3.234:3
TAII Type2         : 42:0.0.3.233:3

-------------------------------------------------------------------------------
SDP-BIND PW Path Information
-------------------------------------------------------------------------------
AGI                : 1:4
SAII Type2         : 42:0.0.3.234:4
```

```
TAII Type2         : 42:0.0.3.233:4
```

MPLS-TP DEBUG COMMANDS

The following command provides the debug command for an MPLS-TP tunnel:

tools>dump>router>mpls>tp-tunnel <lsp-name> [clear]

The following is a sample output:

```
A:mlstp-dutA# tools dump router mpls tp-tunnel
  - tp-tunnel <lsp-name> [clear]
  - tp-tunnel id <tunnel-id> [clear]

 <lsp-name>           : [32 chars max]
 <tunnel-id>          : [1..61440]
 <clear>              : keyword - clear stats after reading


*A:mlstp-dutA# tools dump router mpls tp-tunnel "lsp-
"lsp-32"  "lsp-33"  "lsp-34"  "lsp-35"  "lsp-36"  "lsp-37"  "lsp-38"  "lsp-39"
"lsp-40"  "lsp-41"
*A:mlstp-dutA# tools dump router mpls tp-tunnel "lsp-32"

 Idx: 1-32 (Up/Up): pgId 4, paths 2, operChg 1, Active: Protect
  TunnelId: 42::0.0.3.233::32-42::0.0.3.234::32
  PgState: Dn, Cnt/Tm: Dn 1/000 04:00:48.160 Up:3/000 00:01:25.840
  MplsMsg: tpDn 0/000 00:00:00.000, tunDn 0/000 00:00:00.000
           wpDn 0/000 00:00:00.000, ppDn 0/000 00:00:00.000
           wpDel 0/000 00:00:00.000, ppDel 0/000 00:00:00.000
           tunUp 1/000 00:00:02.070
  Paths:
   Work (Up/Dn): Lsp 1, Lbl 32/32, If 2/128 (1/2/3 : 0.0.0.0)
    Tmpl: ptc: , oam: privatebed-oam-template (bfd: privatebed-bfd-template(np)-10 ms)
    Bfd: Mode CC state Dn/Up handle 160005/0
    Bfd-CC (Cnt/Tm): Dn 1/000 04:00:48.160 Up:1/000 00:01:23.970
    State:  Admin Up (1::1::1)  port Up , if Dn ,  operChg 2
    DnReasons: ccFault ifDn

   Protect (Up/Up): Lsp 2, Lbl 2080/2080, If 3/127 (5/1/1 : 0.0.0.0)
    Tmpl: ptc: privatebed-protection-template, oam: privatebed-oam-template (bfd: pri-
vatebed-bfd-template(np)-10 ms)
    Bfd: Mode CC state Up/Up handle 160006/0
    Bfd-CC (Cnt/Tm): Dn 0/000 00:00:00.000 Up:1/000 00:01:25.410
    State:  Admin Up (1::1::1)  port Up , if Up ,  operChg 1

  Aps: Rx - 5, raw 3616, nok 0(), txRaw - 3636, revert Y
   Pdu: Rx - 0x1a-21::0101 (SF), Tx - 0x1a-21::0101 (SF)
   State: PF:W:L LastEvt pdu (L-SFw/R-SFw)
   Tmrs: slow
   Defects: None  Now: 000 05:02:19.130
  Seq  Event   state     TxPdu       RxPdu       Dir    Act          Time
  ===  ======  ========  ==========  ==========  =====  ====  ================
  000   start    UA:P:L    SF (0,0)    NR (0,0)  Tx-->  Work  000 00:00:02.080
  001     pdu    UA:P:L    SF (0,0)    SF (0,0)  Rx<--  Work  000 00:01:24.860
  002     pdu    UA:P:L    SF (0,0)    NR (0,0)  Rx<--  Work  000 00:01:26.860
  003     pUp        NR    NR (0,0)    NR (0,0)  Tx-->  Work  000 00:01:27.440
  004     pdu        NR    NR (0,0)    NR (0,0)  Rx<--  Work  000 00:01:28.760
  005     wDn    PF:W:L    SF (1,1)    NR (0,0)  Tx-->  Prot  000 04:00:48.160
  006     pdu    PF:W:L    SF (1,1)    NR (0,1)  Rx<--  Prot  000 04:00:48.160
  007     pdu    PF:W:L    SF (1,1)    SF (1,1)  Rx<--  Prot  000 04:00:51.080
```

The following command shows the free mpls tunnel IDs available between two values, start-range and end-range.

tools>dump>router>mpls>free-tunnel-id <start-range> <end-range>

The following command provides a debug tool to view control-channel-status signaling packets.

```
*A:bksim1611# /debug service id 700 sdp 200:700 event-type ?{config-change|oper-status-
change|neighbor-discovery|control-channel-status}

*A:bksim1611# /debug service id 700 sdp 200:700 event-type control-channel-status

*A:bksim1611#
1 2012/08/31 09:56:12.09 EST MINOR: DEBUG #2001 Base PW STATUS SIG PKT (RX):
"PW STATUS SIG PKT (RX)::
Sdp Bind 200:700 Instance 3
    Version          : 0x0
    PW OAM Msg Type  : 0x27
    Refresh Time     : 0xa
    Total TLV Length : 0x8
    Flags            : 0x0
    TLV Type         : 0x96a
    TLV Len          : 0x4
    PW Status Bits   : 0x0
"

2 2012/08/31 09:56:22.09 EST MINOR: DEBUG #2001 Base PW STATUS SIG PKT (RX):
"PW STATUS SIG PKT (RX)::
Sdp Bind 200:700 Instance 3
    Version          : 0x0
    PW OAM Msg Type  : 0x27
    Refresh Time     : 0xa
    Total TLV Length : 0x8
    Flags            : 0x0
    TLV Type         : 0x96a
    TLV Len          : 0x4
    PW Status Bits   : 0x0
"

3 2012/08/31 09:56:29.44 EST MINOR: DEBUG #2001 Base PW STATUS SIG PKT (TX):
"PW STATUS SIG PKT (TX)::
Sdp Bind 200:700 Instance 3
    Version          : 0x0
    PW OAM Msg Type  : 0x27
    Refresh Time     : 0x1e
    Total TLV Length : 0x8
    Flags            : 0x0
    TLV Type         : 0x96a
    TLV Len          : 0x4
    PW Status Bits   : 0x0
```

Ethernet Connectivity Fault Management (ETH-CFM)

The IEEE and the ITU-T have cooperated to define the protocols, procedures and managed objects to support service based fault management. Both IEEE 802.1ag standard and the ITU-T Y.1731 recommendation support a common set of tools that allow operators to deploy the necessary administrative constructs, management entities and functionality, Ethernet Connectivity Fault Management (ETH-CFM). The ITU-T has also implemented a set of advanced ETH-CFM and performance management functions and features that build on the proactive and on demand troubleshooting tools.

CFM uses Ethernet frames and is distinguishable by ether-type 0x8902. In certain cases the different functions will use a reserved multicast address that could also be used to identify specific functions at the MAC layer. However, the multicast MAC addressing is not used for every function or in every case. The Operational Code (OpCode) in the common CFM header is used to identify the type of function carried in the CFM packet. CFM frames are only processed by IEEE MAC bridges. With CFM, interoperability can be achieved between different vendor equipment in the service provider network up to and including customer premises bridges. The following table lists CFM-related acronyms used in this section.

IEEE 802.1ag and ITU-T Y.1731 functions that are implemented are available on the SR and ESS platforms.

This section of the guide will provide configuration example for each of the functions. It will also provide the various OAM command line options and show commands to operate the network. The individual service guides will provide the complete CLI configuration and description of the commands in order to build the necessary constructs and management points.

Acronym	Callout
1DM	One way Delay Measurement (Y.1731)
AIS	Alarm Indication Signal
CCM	Continuity check message
CFM	Connectivity fault management
DMM	Delay Measurement Message (Y.1731)
DMR	Delay Measurement Reply (Y.1731)
LBM	Loopback message
LBR	Loopback reply
LTM	Linktrace message
LTR	Linktrace reply

Acronym	Callout (Continued)
ME	Maintenance entity
MA	Maintenance association
MA-ID	Maintenance association identifier
MD	Maintenance domain
MEP	Maintenance association end point
MEP-ID	Maintenance association end point identifier
MHF	MIP half function
MIP	Maintenance domain intermediate point
OpCode	Operational Code
RDI	Remote Defect Indication
TST	Ethernet Test (Y.1731)
SLM	Synthetic Loss Message
SLR	Synthetic Loss Reply (Y.1731)

ETH-CFM Building Blocks

The IEEE and the ITU-T use their own nomenclature when describing administrative contexts and functions. This introduces a level of complexity to configuration, discussion and different vendors naming conventions. The SROS CLI has chosen to standardize on the IEEE 802.1ag naming where overlap exists. ITU-T naming is used when no equivalent is available in the IEEE standard. In the following definitions, both the IEEE name and ITU-T names are provided for completeness, using the format IEEE Name/ITU-T Name.

Maintenance Domain (MD)/Maintenance Entity (ME) is the administrative container that defines the scope, reach and boundary for faults. It is typically the area of ownership and management responsibility. The IEEE allows for various formats to name the domain, allowing up to 45 characters, depending on the format selected. ITU-T supports only a format of "none" and does not accept the IEEE naming conventions.

0 — Undefined and reserved by the IEEE.

1 — No domain name. It is the only format supported by Y.1731 as the ITU-T specification does not use the domain name. This is supported in the IEEE 802.1ag standard but not in currently implemented for 802.1ag defined contexts.

2,3,4 — Provides the ability to input various different textual formats, up to 45 characters. The string format (2) is the default and therefore the keyword is not shown when looking at the configuration.

Maintenance Association (MA)/Maintenance Entity Group (MEG) is the construct where the different management entities will be contained. Each MA is uniquely identified by its MA-ID. The MA-ID is comprised of the by the MD level and MA name and associated format. This is another administrative context where the linkage is made between the domain and the service using the **bridging-identifier** configuration option. The IEEE and the ITU-T use their own specific formats. The MA short name formats (0-255) have been divided between the IEEE (0-31, 64-255) and the ITU-T (32-63), with five currently defined (1-4, 32). Even though the different standards bodies do not have specific support for the others formats a Y.1731 context can be configured using the IEEE format options.

1 (Primary VID) — Values 0 — 4094

2 (String) — Raw ASCII, excluding 0-31 decimal/0-1F hex (which are control characters) form the ASCII table

3 (2-octet integer) — 0 — 65535

4 (VPN ID) — Hex value as described in RFC 2685, *Virtual Private Networks Identifier*

32 (icc-format) — Exactly 13 characters from the ITU-T recommendation T.50.

Note: When a VID is used as the short MA name, 802.1ag will not support VLAN translation because the MA-ID must match all the MEPs. The default format for a short MA name is an integer. Integer value 0 means the MA is not attached to a VID. This is useful for VPLS services on SR OS platforms because the VID is locally significant.

Maintenance Domain Level (MD Level)/Maintenance Entity Group Level (MEG Level) is the numerical value (0-7) representing the width of the domain. The wider the domain, higher the numerical value, the farther the ETH-CFM packets can travel.   It is important to understand that the level establishes the processing boundary for the packets. Strict rules control the flow of ETH-CFM packets and are used to ensure proper handling, forwarding, processing and dropping of these packets. To keep it simple ETH-CFM packets with higher numerical level values will flow through MEPs on MIPs on SAPs configured with lower level values. This allows the operator to implement different areas of responsibility and nest domains within each other. Maintenance association (MA) includes a set of MEPs, each configured with the same MA-ID and MD level used verify the integrity of a single service instance.

In the following example, a Y.1731 domain context and 802.1ag context are configured. The Y.1731 context can be identified by the **none** setting for the domain format.

```
configure eth-cfm domain 3 format none level 3
configure eth-cfm domain 4 format string name IEEE-Domain level 4

show eth-cfm domain
===============================================================================
CFM Domain Table
===============================================================================
Md-index   Level Name                                          Format
-------------------------------------------------------------------------------
3          3                                                   none
4          4     IEEE-Domain                                   charString
===============================================================================
```

The chassis does not support a domain format of **none** for the 802.1ag contexts. The domain index, the first numerical value, is not related to the level, even though in this example they do match.

The following example illustrates the creation of the association within the domain context. The association links the construct to the service using the value of the bridge-identifier. The value specified for the bridge-identifier is equivalent to the numerical value used to create the service.

```
config>eth-cfm# info
----------------------------------------------
        domain 3 format none level 3
            association 1 format icc-based name "123456789abcd"
                bridge-identifier 100
                exit
            exit
            association 2 format string name "Y1731ContextIEEEFormat"
                bridge-identifier 300
                exit
            exit
        exit
        domain 4 name "IEEE-Domain" level 4
            association 1 format string name "UpTo45CharactersForIEEEString"
```

```
                bridge-identifier 100
                exit
                ccm-interval 1
            exit
        exit
----------------------------------------------
*A:cses-E01>config>eth-cfm#  show eth-cfm association

================================================================================
CFM Association Table
================================================================================
Md-index   Ma-index   Name                     CCM-intrvl Hold-time Bridge-id
--------------------------------------------------------------------------------
3          1          123456789abcd            10         n/a       100
3          2          Y1731ContextIEEEFormat   10         n/a       300
4          1          UpTo45CharactersForIEEE* 1          n/a       100
================================================================================
```

* indicates that the corresponding row element may have been truncated.

This example show how to format the association within the domain to match the domain format, Y.1731 (domain 3/association 1) or 802.1ag (domain 4/association 1), and how the 802.1ag association format can be configured within a Y.1731 domain (domain 3/association 2). The mixed configuration represented by domain 3 association 2 may be of value in mixed Y.1731 and 802.1ag environments.

The CCM-interval is also specified within the association and has a default of 10 seconds unless specifically configured with another value. When the association is created and the MEP is a facility MEP the bridge-identifier is not to be included in the configuration since the facility MEP is not bound to a service. Facility MEPs are described in the 7750 SR OS Services Guide

Maintenance Endpoint (MEP)/MEG Endpoint (MEP) are the workhorses of ETH-CFM. A MEP is the unique identification within the association (0-8191). Each MEP is uniquely identified by the MA-ID, MEPID tuple. This management entity is responsible for initiating, processing and terminating ETH-CFM functions, following the nesting rules. MEPs form the boundaries which prevent the ETH-CFM packets from flowing beyond the specific scope of responsibility. A MEP has direction, **up** or **down**. Each indicates the directions packets will be generated; UP toward the switch fabric, **down** toward the SAP away from the fabric. Each MEP has an active and passive side. Packets that enter the active point of the MEP will be compared to the existing level and processed accordingly. Packets that enter the passive side of the MEP are passed transparently through the MEP.   Each MEP contained within the same maintenance association and with the same level (MA-ID) represents points within a single service.   MEP creation on a SAP is allowed only for Ethernet ports with NULL, q-tags, q-in-q encapsulations. MEPs may also be created on SDP bindings.

Maintenance Intermediate Point (MIP)/MEG Intermediate Point (MIP) are management entities between the terminating MEPs along the service path. These provide insight into the service path connecting the MEPs. MIPs only respond to Loopback Messages (LBM) and Linktrace Messages (LTM). All other CFM functions are transparent to these entities. Only one MIP is allowed per SAP or SDP binding. The creation of the MIPs can be done when the lower level domain is created (explicit) or manually (default). This is controlled by the use of the mhf-creation mode within the association under the bridge-identifier. MIP creation is supported on a SAP and SDP binding, not including Mesh SDP bindings. By default, no MIPs are created.

There are two locations in the configuration where ETH-CFM is defined. The domains, associations (including linkage to the service id), MIP creation method, common ETH-CFM functions and remote MEPs are defined under the top level **eth-cfm** command. It is important to note, when Y.1731 functions are required the context under which the MEPs are configured must follow the Y.1731 specific formats (domain format of none). Once these parameters have been entered, the MEP and possibly the MIP can be defined within the service under the SAP or SDP binding.

This is a general table that indicates the ETH-CFM support for the different services and SAP or SDP binding. It is not meant to indicate the services that are supported or the requirements for those services on the individual platforms.

Table 3: ETH-CFM Support Matrix

Service	Ethernet Connection	Down MEP	Up MEP	MIP	Virtual MEP
Epipe					No
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
VPLS					No
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
	Mesh-SDP	Yes	Yes	No	-
B-VPLS					Yes
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
	Mesh-SDP	Yes	Yes	No	-
I-VPLS					No
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
	Mesh-SDP	Yes	Yes	No	-

Service	Ethernet Connection	Down MEP	Up MEP	MIP	Virtual MEP
M-VPLS					No
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
	Mesh-SDP	Yes	Yes	No	-
PBB EPIPE					No
	SAP	Yes	Yes	Yes	-
	Spoke-SDP	Yes	Yes	Yes	-
	Mesh-SDP	Yes	Yes	No	-
IPIPE					No
	SAP	Yes	No	No	-
	Ethernet-Tunnel SAP	Yes	No	No	-
IES					No
	SAP	Yes	No	No	-
	Spoke-SDP (Interface)	Yes	No	No	-
	Subscriber Group-int SAP	Yes	No	No	-
VPRN					No
	SAP	Yes	No	No	-
	Spoke-SDP (Interface)	Yes	No	No	-
	Subscriber Group-int SAP	Yes	No	No	-
Note1	Ethernet-Tunnel (Control) SAP	Yes	No	No	-
	Ethernet-Tunnel (Path/Member)	Yes	Yes	No	-
	Ethernet-Ring (Data)	Yes	No	No	-

Note1: Ethernet-Tunnels and Ethernet-Rings are not configurable under all service types. Any service restrictions for MEP direction or MIP support will override the generic capability of the Ethernet-Tunnel or Ethernet-Ring MPs. Please check the applicable user guide for applicability

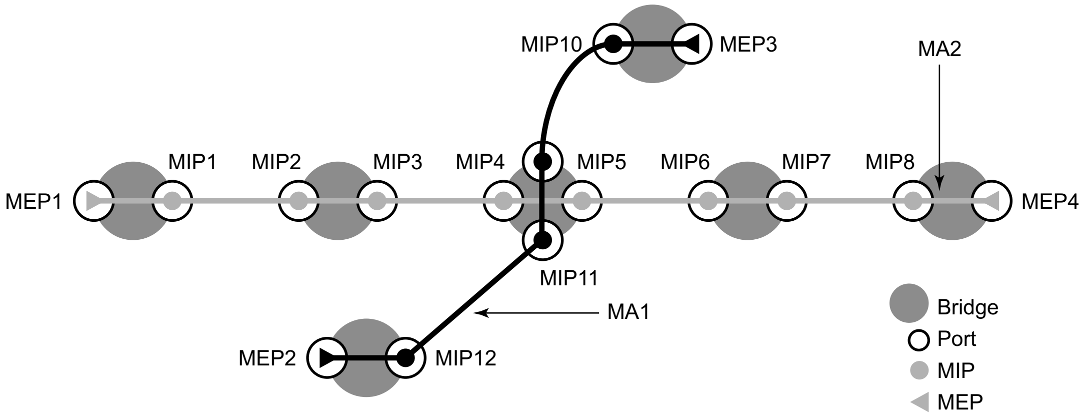

Figure 28: MEP and MIP

Figure 29 illustrates the usage of an EPIPE on two different nodes that are connected using ether SAP 1/1/2:100.31. The SAP 1/1/10:100.31 is an access port that is not used to connect the two nodes.

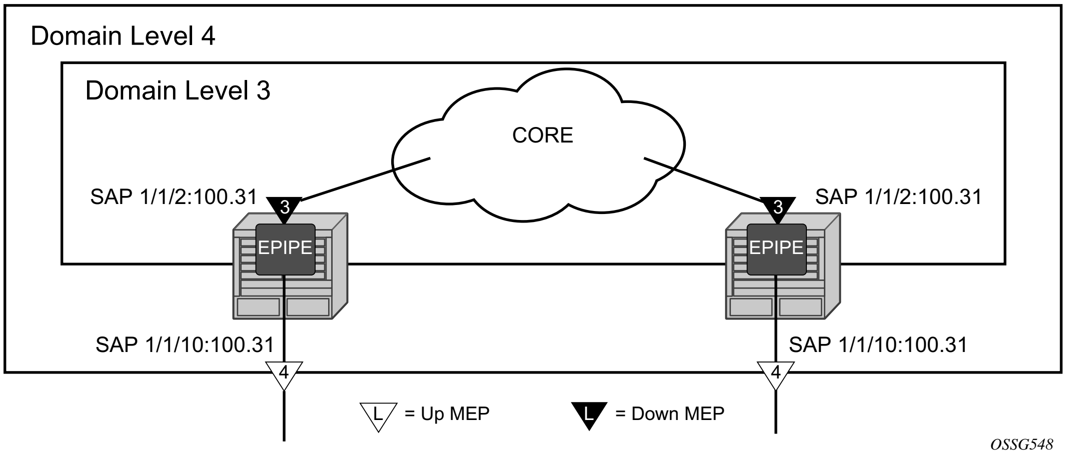

Figure 29: MEP Creation

```
NODE1
config>eth-cfm# info
----------------------------------------------
```

```
        domain 3 format none level 3
            association 1 format icc-based name "03-0000000101"
                bridge-identifier 100
                exit
            exit
        exit
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
        exit

*A:cses-E01>config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mep 111 domain 3 association 1 direction down
                       mac-address d0:0d:1e:00:01:11
                        no shutdown
                    exit
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

NODE 2
eth-cfm# info
----------------------------------------------
        domain 3 format none level 3
            association 1 format icc-based name "03-0000000101"
                bridge-identifier 100
                exit
            exit
        exit
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
        exit
----------------------------------------------
*A:cses-E02>config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mep 112 domain 3 association 1 direction down
                        mac-address d0:0d:1e:00:01:12
                        no shutdown
                    exit
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 102 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:02
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
*A:cses-E02>config>service>epipe#
```

Examining the configuration from NODE1, MEP 101 is configured with a direction of UP causing all ETH-CFM traffic originating from this MEP to generate into the switch fabric and out the mate SAP 1/1/2:100.31. MEP 111 uses the default direction of DOWN causing all ETH-CFM traffic that is generated from this MEP to send away from the fabric and only egress the SAP on which it is configured, SAP 1/1/2:100.31.

Further examination of the domain constructs reveal that the configuration properly uses domain nesting rules. In this case, the Level 3 domain is completely contained in a Level 4 domain.

The following display was taken from NODE1.

```
show eth-cfm cfm-stack-table
===============================================================================
CFM Stack Table Defect Legend:
R = Rdi, M = MacStatus, C = RemoteCCM, E = ErrorCCM, X = XconCCM, A = AisRx

===============================================================================
CFM SAP Stack Table
===============================================================================
Sap                Lvl Dir  Md-index   Ma-index   MepId  Mac-address     Defect
-------------------------------------------------------------------------------
1/1/2:100.31         3 Down          3          1  111 90:f3:01:01:00:02 ------
1/1/10:100.31       4   Up          4          1  101 d0:0d:1e:00:01:01 ------
===============================================================================
```

Figure 30 illustrates the creation of and explicit MIP.

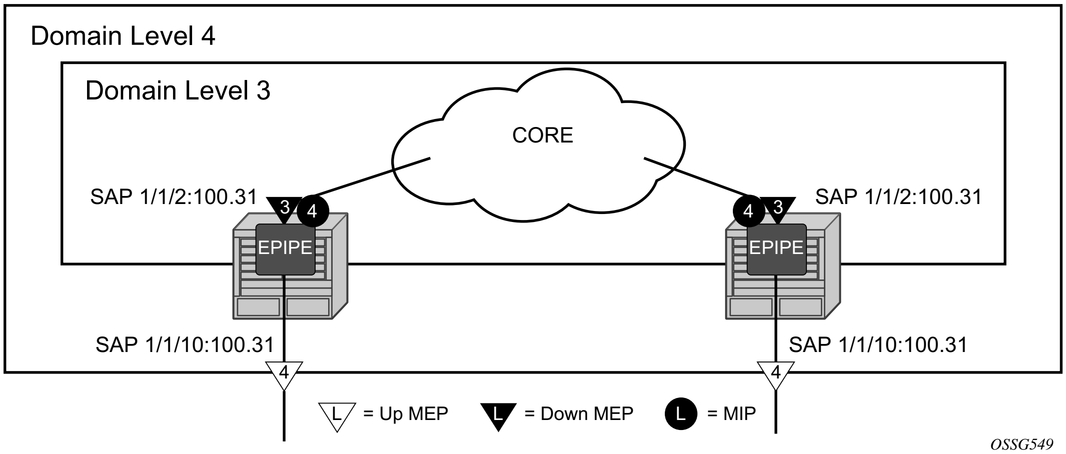

Figure 30: MIP Creation Example (NODE1)

```
NODE1
config>eth-cfm# info
----------------------------------------------
        domain 3 format none level 3
            association 1 format icc-based name "03-0000000101"
                bridge-identifier 100
                exit
            exit
        exit
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
     association 2 format icc-based name "04-MIP0000102"
                bridge-identifier 100
                    mhf-creation explicit
                exit
            exit
        exit

config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mep 111 domain 3 association 1 direction down
               mac-address d0:0d:1e:00:01:11
                        no shutdown
                    exit
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

NODE 2
eth-cfm# info
----------------------------------------------
        domain 3 format none level 3
            association 1 format icc-based name "03-0000000101"
                bridge-identifier 100
                exit
            exit
        exit
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
    association 2 format icc-based name "04-MIP0000102"
                bridge-identifier 100
                    mhf-creation explicit
                exit
            exit
        exit
----------------------------------------------

config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mep 112 domain 3 association 1 direction down
                        mac-address d0:0d:1e:00:01:12
                        no shutdown
                    exit
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 102 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:02
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
```

An addition of association 2 under domain four includes the **mhf-creation explicit** statement has been included. This means that when the level 3 MEP is assigned to the SAP 1/1/2:100.31 using the definition in domain 3 association 1, creating the higher level MIP on the same SAP. Since a

MIP does not have directionality “Both” sides are active. The service configuration and MEP configuration within the service did not change.

The following output is from Node 1.

```
show eth-cfm cfm-stack-table
===============================================================================
CFM Stack Table Defect Legend:
R = Rdi, M = MacStatus, C = RemoteCCM, E = ErrorCCM, X = XconCCM, A = AisRx

===============================================================================
CFM SAP Stack Table
===============================================================================
Sap                Lvl Dir  Md-index   Ma-index   MepId  Mac-address     Defect
-------------------------------------------------------------------------------
1/1/2:100.31         3 Down          3          1  111 d0:0d:1e:00:01:11 ------
1/1/2:100.31         4 Both          4          2  MIP 90:f3:01:01:00:02 ------
1/1/10:100.31        4   Up          4          1  101 d0:0d:1e:00:01:01 ------
===============================================================================
```

Figure 31 illustrates a simpler method that does not require the creation of the lower level MEP. The operator simply defines the association parameters and uses the **mhf-creation default** setting, then places the MIP on the SAP of their choice.

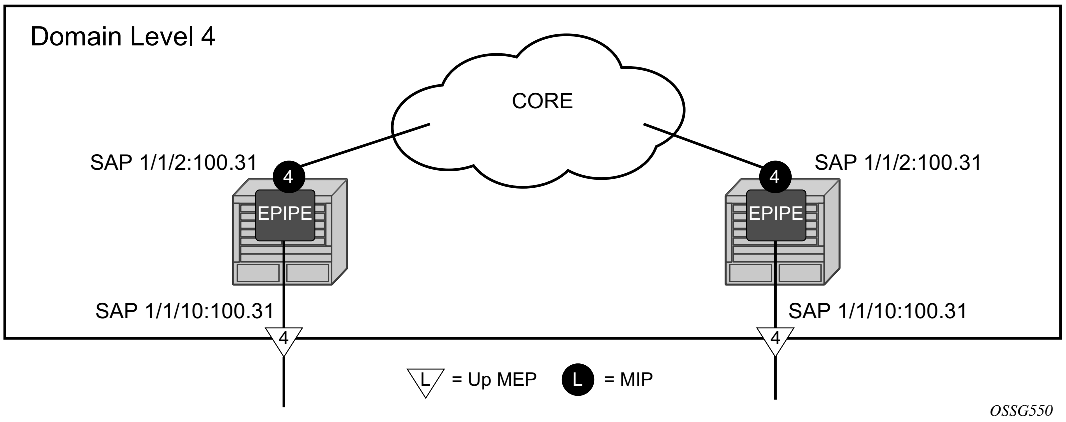

Figure 31: MIP Creation Default

```
NODE1
config>eth-cfm# info
----------------------------------------------
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
            association 2 format icc-based name "04-MIP0000102"
                bridge-identifier 100
                    mhf-creation default
                exit
            exit
        exit
----------------------------------------------

config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                     mip mac d0:0d:1e:01:01:01
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

# show eth-cfm cfm-stack-table
===============================================================================
CFM Stack Table Defect Legend:
R = Rdi, M = MacStatus, C = RemoteCCM, E = ErrorCCM, X = XconCCM, A = AisRx
===============================================================================
CFM SAP Stack Table
===============================================================================
Sap                Lvl Dir  Md-index   Ma-index   MepId  Mac-address     Defect
-------------------------------------------------------------------------------
1/1/2:100.31         4 Both          4          2  MIP d0:0d:1e:01:01:01 ------
1/1/10:100.31        4   Up          4          1  101 d0:0d:1e:00:01:01 ------
===============================================================================

NODE2
config>eth-cfm# info
----------------------------------------------
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
            exit
            association 2 format icc-based name "04-MIP0000102"
                bridge-identifier 100
                    mhf-creation default
                exit
            exit
        exit
----------------------------------------------

config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
```

```
                    mip mac d0:0d:1e:01:01:02
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 102 domain 4 association 1 direction up
                        mac-address d0:0d:1e:00:01:02
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

# show eth-cfm cfm-stack-table
===============================================================================
CFM Stack Table Defect Legend:
R = Rdi, M = MacStatus, C = RemoteCCM, E = ErrorCCM, X = XconCCM, A = AisRx

===============================================================================
CFM SAP Stack Table
===============================================================================
Sap                Lvl Dir  Md-index   Ma-index   MepId  Mac-address     Defect
-------------------------------------------------------------------------------
1/1/2:100.31         4 Both          4          2  MIP d0:0d:1e:01:01:02 ------
1/1/10:100.31        4   Up          4          1  102 d0:0d:1e:00:01:02 ------
===============================================================================
```

Figure 32 shows the detailed IEEE representation of MEPs, MIPs, levels and associations, using the standards defined icons.

SAPs support a comprehensive set of rules including wild cards to map packets to services. For example, a SAP mapping packets to a service with a port encapsulation of QinQ may choose to only look at the outer VLAN and wildcard the inner VLAN. SAP 1/1/1:100.* would map all packets arriving on port 1/1/1 with an outer VLAN 100 and any inner VLAN to the service the SAP belongs to. These powerful abstractions will extract inbound ETH-CFM PDUs only when there is an exact match to the SAP construct. In the case of the example when then an ETH-CFM PDU arrives on port 1/1/1 with a single VLAN with a value of 100 followed immediately with e-type (0x8902 ETH-CFM). Furthermore, the generation of the ETH-CFM PDUs that egress this specific SAP will be sent with only a single tag of 100. If the operator needs to extract ETH-CFM PDUs or generate ETH-CFM PDUs on wildcard SAPs Primary VLAN will be required.

Table 4 shows how packets that would normally bypass the ETH-CFM extraction would be extracted when Primary VLAN is configured. This assumes that the processing rules for MEPs and MIPs is met, E-type 0x8902, Levels and OpCodes.

Table 4: Extraction Comparison with Primary VLAN

Port Encapsulation	E-type	Ingress Tag(s)	Ingress SAP	No Primary VLAN ETH-CFM Extraction		With Primary VLAN (10) ETH-CFM Extraction	
				MEP	MIP	MEP	MIP
Dot1q	0x8902	10	x/y/z:*	No	No	Yes	Yes
Dot1q	0x8902	10.10	x/y/z:10	No	No	Yes	Yes
QinQ	0x8902	10.10	x/y/z:10.*	No	No	Yes	Yes
QinQ (Default Behavior)	0x8902	10.10	x/y/z:10.0	No	No	Yes	Yes
Null	0x8902	10	x/y/z	No	No	Yes	Yes

The mapping of the service data remains unchanged. The Primary VLAN function allows for one additional VLAN offset beyond the SAP configuration, up to a maximum of two VLANs in the frame. If a fully qualified SAP specifies two VLANs (SAP 1/1/1:10.10) and a primary VLAN of 12 is configured for the MEP there will be no extraction of ETH-CFM for packets arriving tagged 10.10.12. That exceeds the maximum of two tags.

The mapping or service data based on SAPs has not changed. ETH-CFM MPs functionality remains SAP specific. In instances where as service includes a specific SAP with a specified VLAN (1/1/1:50) and a wildcard SAP on the same port (1/1/1:*) it is important to understand how the ETH-CFM packets are handled. Any ETH-CFM packet with etype 0x8902 arriving with a single tag or 50 would be mapped to a classic MEP configured under SAP 1/1/1:50. Any packet arriving with an outer VLAN of 50 and second VLAN of 10 would be extracted by the 1/1/1:50 SAP and would require a Primary VLAN enabled MEP with a value of 10, assuming the operator would like to extract the ETH-CFM PDU of course. An inbound packet on 1/1/1 with an outer VLAN tag of 10 would be mapped to the SAP 1/1/1:*. If ETH-CFM extraction is required under SAP 1/1/1:* a Primary VLAN enabled MEP with a value of 10 would be required.

Obviously, the packet that is generated from a MEP or MIP with Primary VLAN enabled will include that VLAN. The SAP will encapsulate the Primary VLAN using the SAP encapsulation.

Primary VLAN support includes UP MEPs, DOWN MEPs and MIPs on Ethernet SAPs, including LAG for ePipe and VPLS services. There is no support for Primary VLAN configuration for vMEPs or MEPs on SDP binding. Classic MEPs, those without a primary VLAN enabled, and Primary VLAN enabled MEPs can co-exist under the same SAP. Classic MIPs and Primary VLAN enabled MIPs may also coexist. The enforcement of a single classic MIP per SAP continues to be enforced. However, the operator may configure multiple Primary VLAN enabled MIPs on the same SAP. MIPs in the Primary VLAN space must include the mhf-creation static under the association and must also include the specific VLAN on the MIP creation statement under the SAP. The **no** version of the **mip** command must include the entire statement including the VLAN information.

The eight MD Levels (0-7) are specific to context in which the Management Point (MP) is configured. This means the classic MPs have a discrete set of the levels from the Primary VLAN enabled space. Each Primary VLAN space has its own eight Level MD space for the specified Primary VLAN. Consideration must be given before allowing overlapping levels between customers and operators should the operator be provision a customer facing MP, like a MIP on a UNI. CPU Protection extensions for ETH-CFM are VLAN unaware and based on MD Level and the OpCode. Any configured rates will be applied to the Level and OpCode as a group.

There are two configuration steps to enable Primary VLAN. Under the bridging instance, contained within the association context (cfg>eth-cfm>domain>assoc>bridge) the VLAN information must be configured. Until this is enabled using the *primary-vlan-enable* option as part of the MEP creation step or the MIP statement (cfg>service>…>sap>eth-cfm>) the VLAN specified under the bridging instance remains inactive. This is to ensure backward interoperability.

Primary VLAN functions require a minimum of IOM3/IMM. There is no support for vpls-sap-templates. Sub second CCM intervals are not supported for Primary VLAN MEPs.

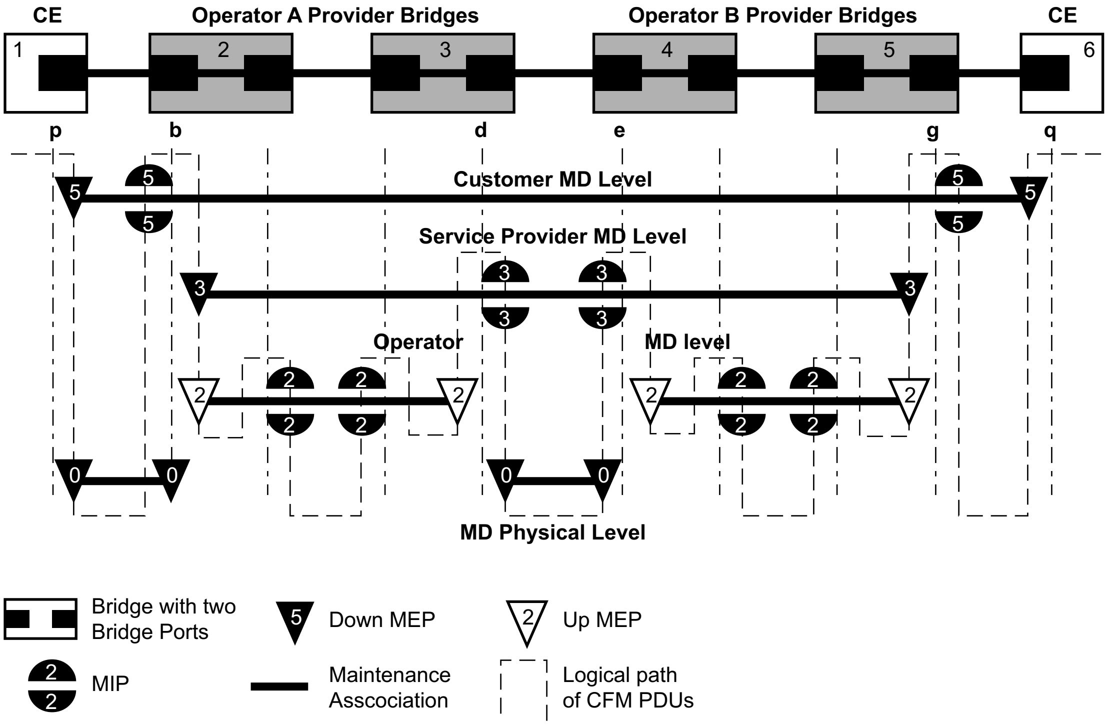

Figure 32: MEP, MIP and MD Levels

An operator may see the following INFO message (during configuration reload), or MINOR (error) message (during configuration creation) when upgrading to 11.0r4 or later if two MEPs are in a previously undetected conflicting configuration. The messaging is an indication that a MEP, the one stated in the message using format (domain <md-index> / association <ma-index> / mep <mep-id>), is already configured and has allocated that context. During a reload (INFO) a MEP that encounters this condition will be created but its state machine will be disabled. If the MINOR error occurs during a configuration creation this MEP will fail the creation step. The indicated MEP will need to be correctly re-configured.

```
INFO: ETH_CFM #1341 Unsupported MA ccm-interval for this MEP - MEP 1/112/21 conflicts with
sub-second config on this MA
MINOR: ETH_CFM #1341 Unsupported MA ccm-interval for this MEP - MEP 1/112/21 conflicts with
sub-second config on this MA
```

Loopback

A loopback message is generated by an MEP to its peer MEP or a MIP (Figure 33). The functions are similar to an IP ping to verify Ethernet connectivity between the nodes.

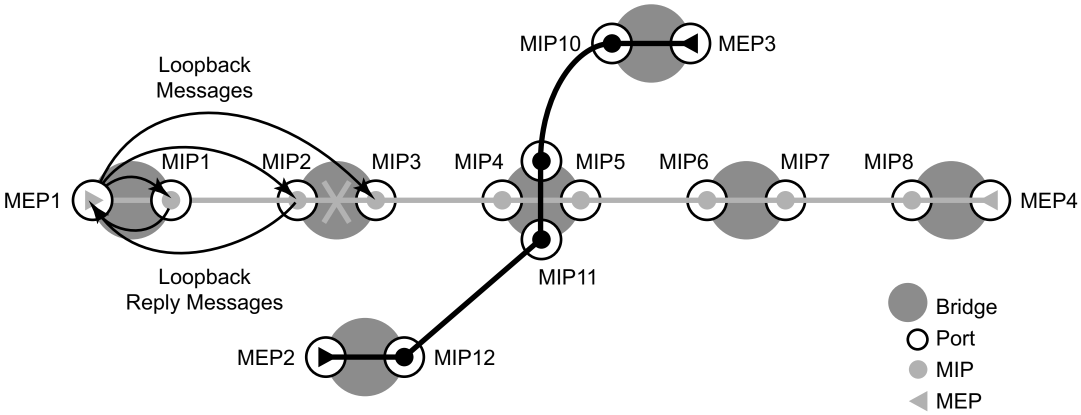

Figure 33: CFM Loopback

The following loopback-related functions are supported:

- Loopback message functionality on an MEP or MIP can be enabled or disabled.
- MEP — Supports generating loopback messages and responding to loopback messages with loopback reply messages.
- MIP — Supports responding to loopback messages with loopback reply messages when loopback messages are targeted to self.

- Displays the loopback test results on the originating MEP. There is a limit of ten outstanding tests per node.

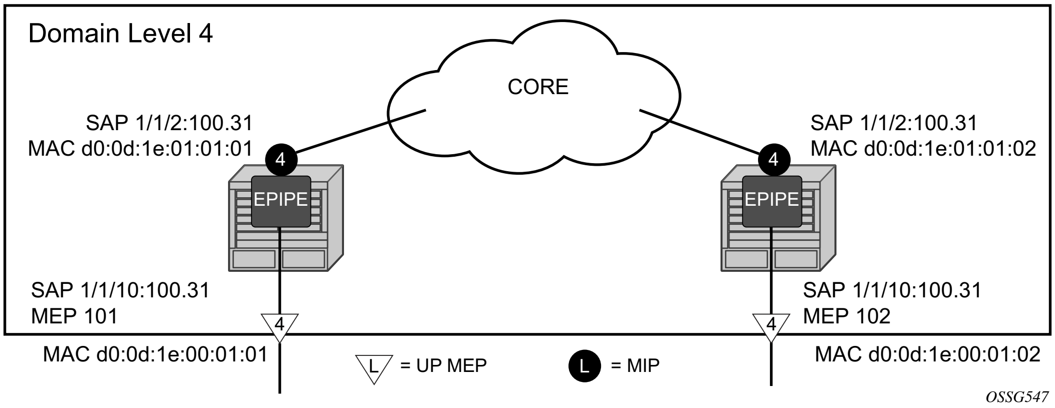

Figure 34: Loopback Configuration

```
# oam eth-cfm loopback d0:0d:1e:01:01:02 mep 101 domain 4 association
Eth-Cfm Loopback Test Initiated: Mac-Address: d0:0d:1e:01:01:02, out sap: 1/1/10:100.31
Sent 5 packets, received 5 packets [0 out-of-order, 0 Bad Msdu]

# oam eth-cfm loopback d0:0d:1e:00:01:02 mep 101 domain 4 association
Eth-Cfm Loopback Test Initiated: Mac-Address: d0:0d:1e:00:01:02, out sap: 1/1/10:100.31
Sent 5 packets, received 5 packets [0 out-of-order, 0 Bad Msdu]
```

Loopback Multicast

This on demand operation tool is used to quickly check the reachability of all MEPs within an Association.   A multicast address can be coded as the destination of an **oam eth-cm loopback** command. The specific class 1 multicast MAC address or the keyword “multicast” can be used as the destination for the loopback command. The class 1 ETH-CFM multicast address is in the format 01:80:C2:00:00:3x (where x = 0 - 7 and is the number of the domain level for the source MEP). When the “multicast” option is used, the class 1 multicast destination is built according to the local MEP level initiating the test.

Remote MEPs that receive this message, configured at the equivalent level, will terminate and process the multicast loopback message responding with the appropriate unicast loopback response (ETH-LBR).   Regardless of whether a multicast or unicast ETH-LBM is used, there is no provision in the standard LBR PDU to carry the MEP-ID of the responder. This means only the remote MEP MAC Address will be reported and subsequently displayed. MIPs will not respond to the multicast ETH-LBM. It is important to understand that although MIPs do not respond they perform the basic level and opcode check to determine whether they need to decode the packet. MIPs along the applicable path over which the LBM is sent that match the level and opcode will decode the packet, not respond and forward along the path.

Only a single on demand multicast ETH-LB may be run at any instance in time. When this test is in progress all other on demand unicast ETH-LB tests will be blocked. The MIB will store the first 1000 responses. Any additional responses received will not be stored in the MIB. It is important to check the scaling guides to ensure that the number of responders does not overwhelm the receive capability of the ETH-CFM application. One must consider all aspects of the configured ETH-CFM functions that are active.

MEP loopback stats are not updated as a result of this test being run. That means the received, out-of-order and bad-msdu counts are not affected by multicast loopback tests. The multicast loopback command is meant to provide immediate connectivity troubleshooting feedback for remote MEP reachability only.

Figure 35 displays a a four node hub and spoke network. The service has a level 4 UP MEP configured on one of the SAPs wherever the service exists.   In the example below, a multicast loopback is issued from the MEP 9, the level 4 MEP on the VPLS hub. Each of the remote UP MEPs will process the multicast loopback request, responding with the appropriate unicast LBR.

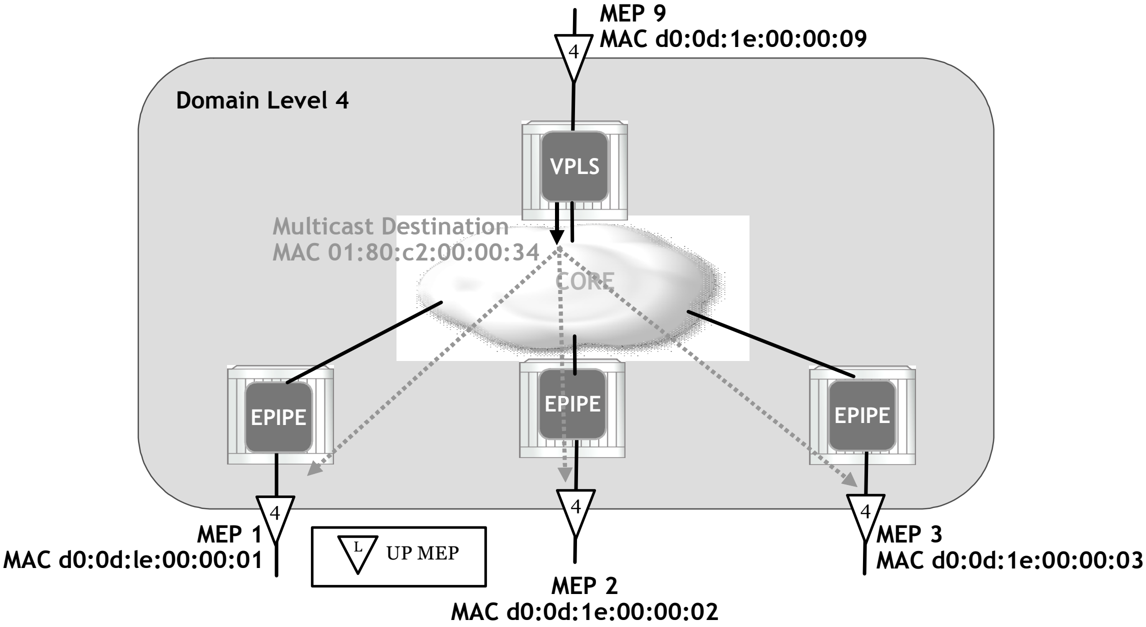

Figure 35: Multicast Loopback Message

The following output is shown from the VPLS hub containing MEP 9. The example displays a well behaved and fully connected service instance where all packets are responded to and in the proper order. The **SeqNum** in the **Mep Multicast Loopback Information** section is the sequence number that is set by the instantiating node as it generates the LBM. The **Rx Index** is the order in which the packets were received from the responding MEP.

```
oam eth-cfm loopback multicast mep 9 domain 14 association 1 send-count 5

Eth-Cfm Loopback Test Initiated: Mac-Address: multicast, out service: 1

MAC Address          Receive Order
-------------------------------------------------------------------------------
d0:0d:1e:00:00:01    1    2    3    4    5
d0:0d:1e:00:00:02    1    2    3    4    5
d0:0d:1e:00:00:03    1    2    3    4    5
Sent 5 multicast packets, received 15 packets

show eth-cfm mep 9 domain 14 association 1 loopback
===============================================================================
Eth-Cfm MEP Configuration Information
===============================================================================
Md-index           : 14                          Direction         : Up
Ma-index           : 1                           Admin             : Enabled
MepId              : 9                           CCM-Enable        : Disabled
SvcId              : 1
Description        : (Not Specified)
FngState           : fngReset                    ControlMep        : False
LowestDefectPri    : macRemErrXcon               HighestDefect     : none
```

```
Defect Flags       : None
Mac Address        : d0:0d:1e:00:00:09        ControlMep        : False
CcmLtmPriority     : 7
CcmTx              : 0                        CcmSequenceErr    : 0
Fault Propagation  : disabled                 FacilityFault     : n/a
MA-CcmInterval     : 10                       MA-CcmHoldTime    : 0ms
Eth-1Dm Threshold  : 3(sec)                   MD-Level          : 3
Eth-Ais:           : Disabled
Eth-Tst:           : Disabled

Redundancy:
     MC-LAG State   : n/a

CcmLastFailure Frame:
     None

XconCcmFailure Frame:
     None
-------------------------------------------------------------------------------
Mep Loopback Information
-------------------------------------------------------------------------------
LbRxReply          : 0                        LbRxBadOrder      : 0
LbRxBadMsdu        : 0                        LbTxReply         : 0
LbNextSequence     : 7                        LtNextSequence    : 1
LbStatus           : False                    LbResultOk        : True
DestIsMepId        : False                    DestMepId         : 0
DestMac            : 00:00:00:00:00:00        SendCount         : 0
VlanDropEnable     : True                     VlanPriority      : 7
Data TLV:
     None
-------------------------------------------------------------------------------
Mep Multicast Loopback Information
-------------------------------------------------------------------------------
MAC Address: d0:0d:1e:00:00:01   SeqNum: 2            Rx Index: 1
MAC Address: d0:0d:1e:00:00:01   SeqNum: 3            Rx Index: 2
MAC Address: d0:0d:1e:00:00:01   SeqNum: 4            Rx Index: 3
MAC Address: d0:0d:1e:00:00:01   SeqNum: 5            Rx Index: 4
MAC Address: d0:0d:1e:00:00:01   SeqNum: 6            Rx Index: 5
MAC Address: d0:0d:1e:00:00:02   SeqNum: 2            Rx Index: 1
MAC Address: d0:0d:1e:00:00:02   SeqNum: 3            Rx Index: 2
MAC Address: d0:0d:1e:00:00:02   SeqNum: 4            Rx Index: 3
MAC Address: d0:0d:1e:00:00:02   SeqNum: 5            Rx Index: 4
MAC Address: d0:0d:1e:00:00:02   SeqNum: 6            Rx Index: 5
MAC Address: d0:0d:1e:00:00:03   SeqNum: 2            Rx Index: 1
MAC Address: d0:0d:1e:00:00:03   SeqNum: 3            Rx Index: 2
MAC Address: d0:0d:1e:00:00:03   SeqNum: 4            Rx Index: 3
MAC Address: d0:0d:1e:00:00:03   SeqNum: 5            Rx Index: 4
MAC Address: d0:0d:1e:00:00:03   SeqNum: 6            Rx Index: 5
===============================================================================
```

The following output shows examples of some of the potential unexpected conditions that could be reported.   The dash “-” indicates that there was no LBR received for a particular sequence number from a specific peer. The asterisk “*” indicates that multiple LBRs were received with the same sequence number from a specific peer. The reversal of sequence numbers means that the order of LBR reception is out of sequence.

```
oam eth-cfm loopback multicast mep 9 domain 14 association 1 send-count 5

Eth-Cfm Loopback Test Initiated: Mac-Address: multicast, out service: 1

MAC Address          Receive Order
-------------------------------------------------------------------------------
d0:0d:1e:00:00:01    1    -    3    4    5
d0:0d:1e:00:00:02    1    2    *    4    5
d0:0d:1e:00:00:03    1    2    3    5    4
Sent 5 multicast packets, received 15 packets
```

Linktrace

A linktrace message is originated by an MEP and targeted to a peer MEP in the same MA and within the same MD level (Figure 36). Its function is similar to IP traceroute. Traces a specific MAC address through the service. The peer MEP responds with a linktrace reply message after successful inspection of the linktrace message. The MIPs along the path also process the linktrace message and respond with linktrace replies to the originating MEP if the received linktrace message that has a TTL greater than 1 and forward the linktrace message if a look up of the target MAC address in the Layer 2 FIB is successful. The originating MEP shall expect to receive multiple linktrace replies and from processing the linktrace replies, it can put together the route to the target bridge.

A traced MAC address is carried in the payload of the linktrace message, the target MAC. Each MIP and MEP receiving the linktrace message checks whether it has learned the target MAC address. In order to use linktrace the target MAC address must have been learned by the nodes in the network. If so, a linktrace message is sent back to the originating MEP. Also, a MIP forwards the linktrace message out of the port where the target MAC address was learned.

The linktrace message itself has a multicast destination address. On a broadcast LAN, it can be received by multiple nodes connected to that LAN. But, at most, one node will send a reply.

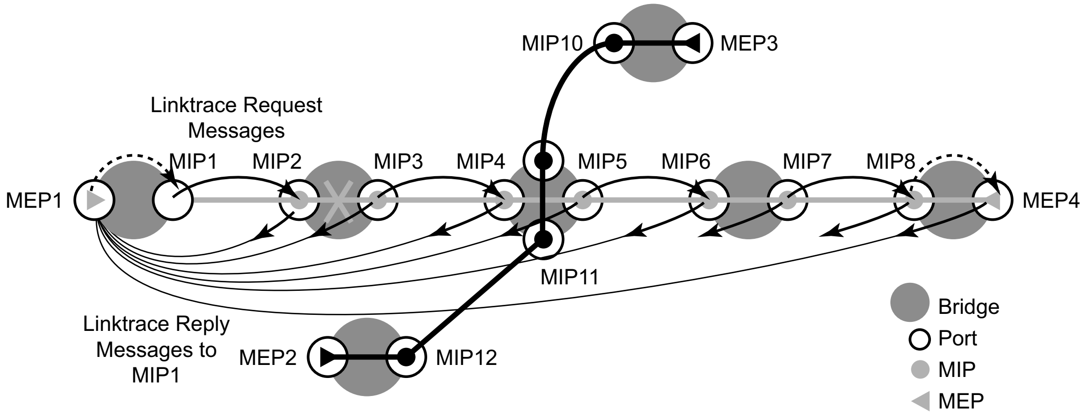

Figure 36: CFM Linktrace

The IEEE and ITU-T handle the linktrace reply slightly differently. An IEEE 802.1ag configured MEP requires the relay action field to be a valid non-zero integer. The ITU-T ignores the relay action field and will set the value to zero when responding to the LTM. In mixed 802.ag and

Y.1731 environments the operator may chose to configure a Y.1731 context with an IEEE domain format.

The following linktrace related functions are supported:

- Enable or disables linktrace functions on an MEP.
- MEP — Supports generating linktrace messages and responding with linktrace reply messages.
- MIP — Supports responding to linktrace messages with linktrace reply messages when encoded TTL is greater than 1, and forward the linktrace messages accordingly if a lookup of the target MAC address in the Layer 2 FIB is successful.
- Displays linktrace test results on the originating MEP. There is a limit of ten outstanding tests per node. Storage is provided for up to ten MEPs and for the last ten responses. If more than ten responses are received older entries will be overwritten.

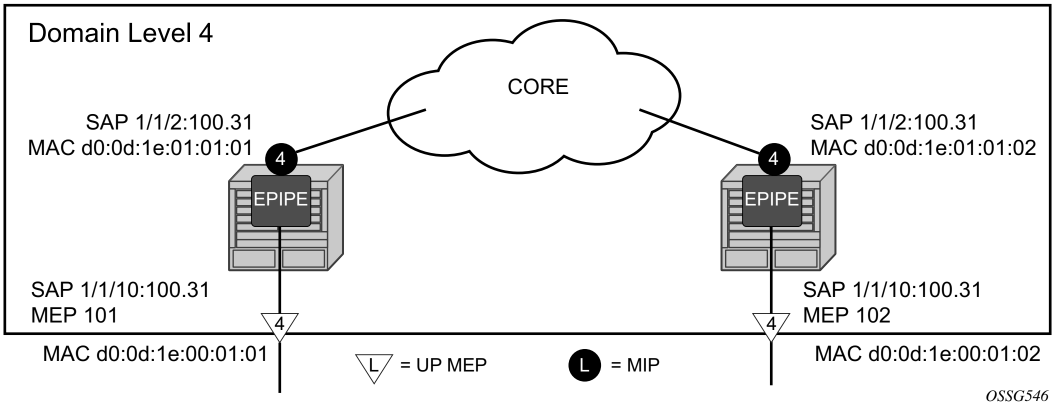

Figure 37: Linktrace Configuration

```
# oam eth-cfm linktrace d0:0d:1e:01:01:02 mep 101 domain 4 association 1

Index Ingress Mac          Egress Mac           Relay      Action
----- -------------------- -------------------- ---------- ----------
1     00:00:00:00:00:00    D0:0D:1E:01:01:01    n/a        forward
2     D0:0D:1E:01:01:02    00:00:00:00:00:00    n/a        none
----- -------------------- -------------------- ---------- ----------
No more responses received in the last 6 seconds.


# oam eth-cfm linktrace d0:0d:1e:00:01:02 mep 101 domain 4 association 1
```

```
Index Ingress Mac          Egress Mac           Relay      Action
----- -------------------- -------------------- ---------- ----------
1     00:00:00:00:00:00    D0:0D:1E:01:01:01    n/a        forward
2     D0:0D:1E:01:01:02    D0:0D:1E:00:01:02    n/a        terminate
----- -------------------- -------------------- ---------- ----------
No more responses received in the last 6 seconds.
```

Continuity Check (CC)

A Continuity Check Message (CCM) is a multicast frame that is generated by a MEP and multicast to all other MEPs in the same MA. The CCM does not require a reply message. To identify faults, the receiving MEP maintains an internal list of remote MEPs it should be receiving CCM messages from.

This list is based off of the remote-mepid configuration within the association the MEP is created in. When the local MEP does not receive a CCM from one of the configured remote MEPs within a pre-configured period, the local MEP raises an alarm.

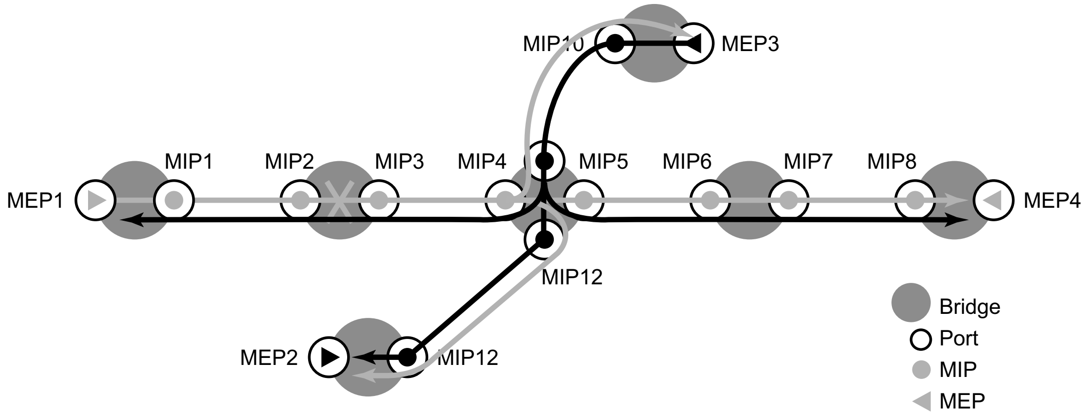

Figure 38: CFM Continuity Check

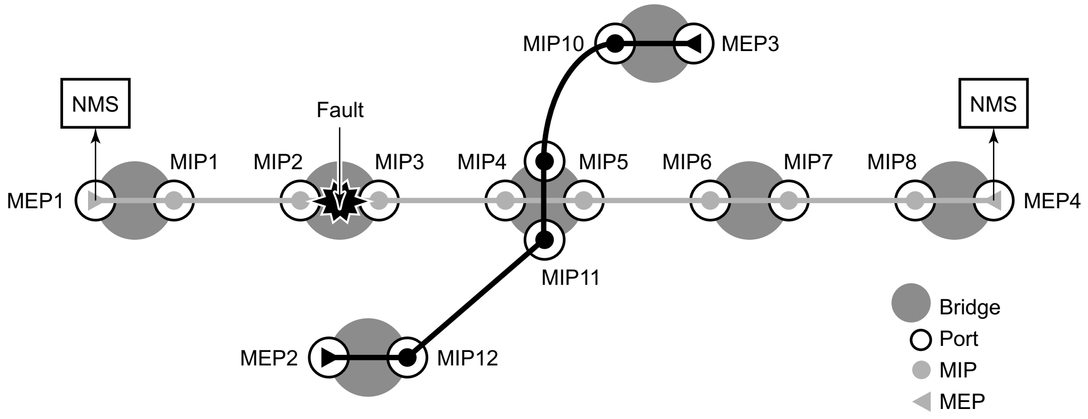

Figure 39: CFM CC Failure Scenario

An MEP may be configured to generate ETH-CC packet using a unicast destination Layer 2 MAC address. This may help reduce the overhead in some operational models where Down MEPs per peer are not available. For example, mapping an I-VPLS to a PBB core where a hub is responsible for multiple spokes is one of the applicable models. When ETH-CFM packets are generated from an I-context toward a remote I-context, the packets will traverse the B-VPLS context. Since many B-contexts are multipoint, any broadcast, unknown or multicast packet is flooded to all appropriate nodes in the B-context. When ETH-CC multicast packets are generated, all the I-VPLS contexts in the association must be configured with all the appropriate remote MEPids. If direct spoke to spoke connectivity is not part of the validation requirement, the operational complexity can be reduced by configuring unicast DA addressing on the "spokes" and continuing to use multicast CCM from the "hub". When the unicast MAC is learned in the forwarding DB, traffic will be scoped to a single node.

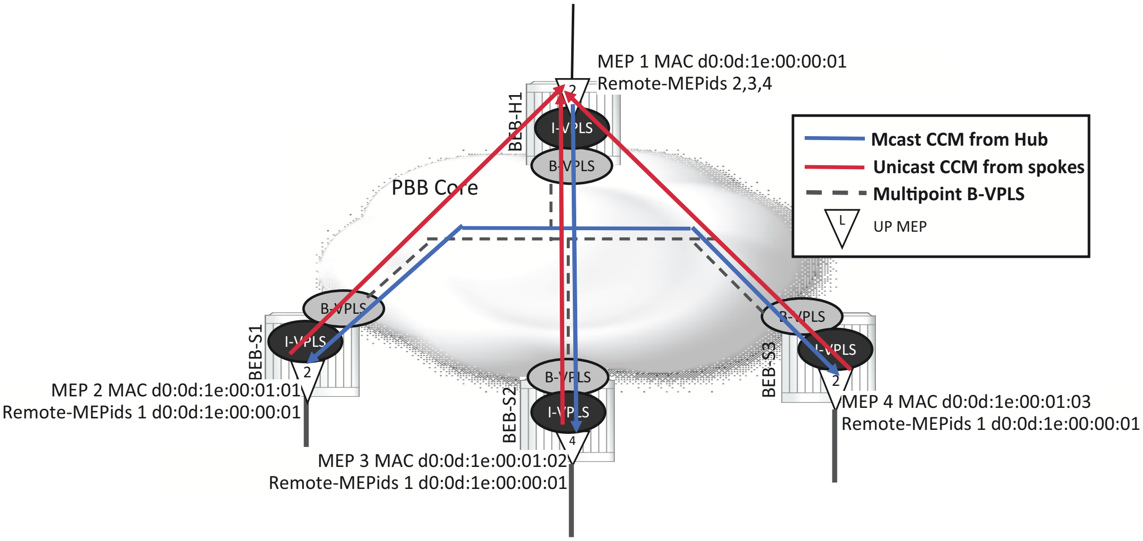
MEP 1 MAC d0:0d:1e:00:00:01
Remote-MEPids 2,3,4
BEB-H1
I-VPLS
B-VPLS
PBB Core
Mcast CCM from Hub
Unicast CCM from spokes
Multipoint B-VPLS
UP MEP
BEB-S1
B-VPLS
I-VPLS
MEP 2 MAC d0:0d:1e:00:01:01
Remote-MEPids 1 d0:0d:1e:00:00:01
BEB-S2
B-VPLS
I-VPLS
MEP 3 MAC d0:0d:1e:00:01:02
Remote-MEPids 1 d0:0d:1e:00:00:01
BEB-S3
B-VPLS
I-VPLS
MEP 4 MAC d0:0d:1e:00:01:03
Remote-MEPids 1 d0:0d:1e:00:00:01

Figure 40: Unicast CCM in Hub & Spoke Environments

Defect condition, reception, and processing will remain unchanged for both hub and spokes. When an ETH-CC defect condition is raised on the hub or spoke, the appropriate defect condition will be set and distributed throughout the association from the multicasting MEP. For example, should a spoke raise a defect condition or timeout, the hub will set the RDI bit in the multicast ETH-CC packet which is received on all spokes. Any local hub MEP defect condition will continue to be propagated in the multicast ETH-CC packet. Defect conditions will be cleared as per normal behavior.

The forwarding plane must be considered before deploying this type of ETH-CC model. A unicast packet will be handled as unknown when the destination MAC does not exist in local forwarding table. If a unicast ETH-CC packet is flooded in a multipoint context, it will reach all the appropriate I-contexts. This will cause the spoke MEPs to raise the "DefErrorCCM" condition because an ETH-CC packet was received from a MEP that has not been configured as part of the receiving MEPs database.

The remote unicast MAC address must be configured and is not automatically learned. A MEP cannot send both unicast and multicast ETH-CC packets. Unicast ETH-CC is only applicable to a local association with a single configured remote peer. There is no validation of MAC addresses for ETH-CC packets. The configured unicast destination MAC address of the peer MEP only replaces the multicast class 1 destination MAC address with a unicast destination.

Unicast CCM is not supported on any MEPs that are configured with sub second CCM-intervals.

The following functions are supported:

- Enable and disable CC for an MEP
- Configure and delete the MEP entries in the CC MEP monitoring database manually. It is only required to provision remote MEPs. Local MEPs shall be automatically put into the database when they are created.
- CCM transmit interval: 10ms, 100ms, 1s, 10s 60s, 600s. Default: 10s. Sub-second, or fast CC requires a ESS-7/ESS-12 and SR-7/SR-12 with a minimum SF/CPM-3, and with only a limited number supported on SF/CPM-1 and SF/CPM-2. When configuring MEPs with sub-second CCM intervals, bandwidth consumption must be taken into consideration. Each CCM PDU is approximately 100 bytes (800 bits). Taken individually, this is a small value. However, the bandwidth consumption increases rapidly as multiple MEPs are configured with 10ms timers, 100 packets per second.

  The following section describes some basic hierarchical considerations and the software requirements and configurations that need to be met when considering sub-second enabled MEPs.

  ç Down MEPs only

  ç Single peer only

  ç Any MD Level

    - As long as lower MD level MEPs are not CCM or ETH-APS enabled
      - G.8031 Ethernet-Tunnels enables OpCode39 Linear APS
      - G.8032 Ethernet-Rings enables OpCode 40 Ring APS
    - As long as lower MD levels MEPs are not receiving ETH-CCM or ETH-APS PDUs, even if they not locally enabled or configured to do so
      - The reception of the lower MD level ETH-CCM and ETH-APS PDUs will be processed by the sub second CCM enabled MEP, regardless of MD Level
      - All other ETH-CFM PDUs will be handled by the MEP at the MD level matching the PDU that has arrived, assuming one has been configured

  - ç Service MEPs (excluding Primary VLAN MEPs)
    - Ethernet SAPs configured on Port with any Ethernet Encapsulation (null, dot1q or QinQ)
  - ç Facility MEPs
    - Ethernet Port Based MEPs
    - Ethernet LAG Based MEPs
    - Ethernet QinQ Tunnel based MEPs (LAG+VLAN, PORT+VLAN)
    - Base Router IP Interfaces
  - ç Service MEPs and Facility MEPs can simultaneously execute sub second CCM enabled MEPs as these are considered different MEP families.
  - ç General processing rules for Service MEPs and Facility MEPs must be met regardless of the CCM interval. These are included here because of the impact misunderstanding could have on the CCM extraction.
    - All the above rules apply
    - MD level hierarchy must be ensured across different families
    - Facility MEPs are the first processing routine for ETH-CFM PDUs
    - VLAN encapsulation uniqueness must exist when processing the ETH-CFM PDU across the two families
      - Unique Example: An Ethernet Port Based Facility Down MEP configured on port 1/1/1 and Service Down MEP SAP 1/1/1:100 (dot1q encaps) are unique
      - Conflict Example: An Ethernet Port Based Facility Down MEP configured on port 1/1/1 and Service Down MEP SAP 1/1/1 (null encaps) are in conflict and cannot coexist. All ETH-CFM PDUs will arrive untagged and the Facility MEP takes precedence.
  - ç G.8031 (Ethernet-Tunnels) support both sub second and 1 second CCM intervals and optionally no CCM. When the MEP is created on a G.8031 Ethernet-Tunnel no other MEP that is any way connected to the G.8031 Ethernet-Tunnel can execute sub second CCM intervals.
    - Facility MEPs are not supported in conjunction with G.8031 (Ethernet-Tunnel MEPs)
  - ç G.8032 (Ethernet-Ring) support both sub second and 1 second CCM intervals and optionally no CCM.
    - Facility MEPs are supported and are considered the first processing route in this combination. G.8032 MEPs can be considered somewhat akin to service MEPs when determining their position in the hierarchy with regard to facility MEPs. This means that all the considerations for *General processing rules for Service MEPs and Facility MEPs* previously mentioned can be applied here.
- The size of the CCM PDU may be increased by configuring the optional Data TLV. This is accomplished by configuring the ccm-padding-size under the specific MEP. The

configured value represents the total length of the Data TLV that will be included with the other CCM PDU informational elements. The **no** form of this command removes the optional Data TLV from the CCM PDU. The operator must consider a CCM PDU is 83 byte size in length (75 base elements plus 8 bytes for port status and interface status). If the size of the optional TLV combined with the size of the CCM PDU exceeds 1500 bytes the packet will be dropped if the MTU is 1518/1522.

- CCM will declare a fault, when:
  - The CCM stops hearing from one of the remote MEPs for 3.5 times CC interval
  - Hears from a MEP with a LOWER MD level
  - Hears from a MEP that is not part of the local MEPs MA
  - Hears from a MEP that is in the same MA but not in the configured MEP list
  - Hears from a MEP in the same MA with the same MEP id as the receiving MEP
  - The CC interval of the remote MEP does not match the local configured CC interval
  - The remote MEP is declaring a fault
- An alarm is raised and a trap is sent if the defect is greater than or equal to the configured low-priority-defect value.
- Remote Defect Indication (RDI) is supported but by default is not recognized as a defect condition because the low-priority-defect setting default does not include RDI.

You can use the optional **ccm-tlv-ignore** command to ignore the reception of interface-status and port-status TLVs in the ETH-CCM PDU on Facility MEPs (Port, LAG, QinQ Tunnel and Router). No processing is performed on the ignored ETH-CCM TLVs values.

Any TLV that is ignored is reported as *absent* for that remote peer and the values in the TLV do not have an impact on the ETH-CFM state machine. This the same behavior as if the remote MEP never included the ignored TLVs in the ETH-CCM PDU. If the TLV is not properly formed, the CCM PDU will fail the packet parsing process, which will cause it to be discarded and a defect condition will be raised.

NODE1:

```
Config>eth-cfm# info
----------------------------------------------
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
                ccm-interval 1
                remote-mepid 102
            exit
        exit
----------------------------------------------
```

NODE2:

```
config>eth-cfm# info
----------------------------------------------
        domain 4 format none level 4
            association 1 format icc-based name "04-0000000102"
                bridge-identifier 100
                exit
                ccm-interval 1
                remote-mepid 101
            exit
        exit
----------------------------------------------
```

Common CCM attributes are defined within the association, including the list of remote peers and interval. Once this is complete, the MEP configured on the SAP within the service must enabled CCM and the priority of the packet can be set.

NODE1:

```
config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mip mac D0:0D:1E:01:01:01
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        ccm-enable
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
```

NODE2:

```
config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mip mac D0:0D:1E:01:01:02
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 102 domain 4 association 1 direction up
                        ccm-enable
                        mac-address d0:0d:1e:00:01:02
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
```

There are various display commands that are available to show the status of the MEP and the list of remote peers. The following illustrates the output from a few of these display commands, taken from NODE1.

No defect conditions are raised. The **Defect** column in the first display is clear and the **Defect Flags** is the second display is also clear.

```
show eth-cfm cfm-stack-table
===============================================================================
CFM Stack Table Defect Legend:
R = Rdi, M = MacStatus, C = RemoteCCM, E = ErrorCCM, X = XconCCM, A = AisRx

===============================================================================
CFM SAP Stack Table
===============================================================================
Sap                Lvl Dir  Md-index   Ma-index   MepId  Mac-address     Defect
-------------------------------------------------------------------------------
1/1/2:100.31         4 Both          4          2  MIP d0:0d:1e:01:01:01 ------
1/1/10:100.31        4   Up          4          1  101 d0:0d:1e:00:01:01 ------
===============================================================================

show eth-cfm mep 101 domain 4 association 1
===============================================================================
Eth-Cfm MEP Configuration Information
===============================================================================
Md-index           : 4                        Direction         : Up
Ma-index           : 1                        Admin             : Enabled
MepId              : 101                      CCM-Enable        : Enabled
IfIndex            : 35979264                 PrimaryVid        : 2031716
Description        : (Not Specified)
FngState           : fngReset                 ControlMep        : False
LowestDefectPri    : macRemErrXcon            HighestDefect     : none
Defect Flags       : None
Mac Address        : d0:0d:1e:00:01:01        ControlMep        : False
CcmLtmPriority     : 7
CcmTx              : 1639                     CcmSequenceErr    : 0
Fault Propagation  : disabled                 FacilityFault     : n/a
MA-CcmInterval     : 1                        MA-CcmHoldTime    : 0ms
```

```
Eth-1Dm Threshold  : 3(sec)                   MD-Level          : 4
Eth-Ais:           : Disabled
Eth-Tst:           : Disabled

Redundancy:
    MC-LAG State   : n/a

CcmLastFailure Frame:
    None

XconCcmFailure Frame:
    None
===============================================================================
```

The **all-remote-mepids** is the appropriate command to show the details for each configured peer, including the MAC address.

```
show eth-cfm mep 101 domain 4 association 1 all-remote-mepids
============================================================================
Eth-CFM Remote-Mep Table
============================================================================
R-mepId Rx CC  Rx Rdi Port-Tlv If-Tlv Peer Mac Addr     CCM status since
----------------------------------------------------------------------------
102     True   False  Up       Up     d0:0d:1e:00:01:02 02/02/2011 13:37:42
============================================================================
```

CCM Grace Period

When an ISSU operation or soft reset function is invoked, the ETH-Vendor Specific Message (ETH-VSM) PDU is used to announce a grace period to a remote CCM enabled peer which are administratively enabled. This Multicast Class 1 DA announcement includes the start of a grace period, the new remote timeout value of 90s and the completion of the grace process. Those MEPs configured with unicast destination MAC addresses will still receive the CCM messages as unicast.

At the start of the operation, a burst of three packets will be sent over a three second window in order to reduce the chances that a remote peer may miss the backoff announcement. This grace announcement will include an indication that the local node that is undergoing a maintenance operation that could possibly delay the announcement of CCM messages at the configured interval.

Three evenly spaced ETH-VSM messages will be sent during the interval advertised in the ETH-VSM message. This means that the ETH-VSM message will be sent every 10 seconds to all appropriate remote peers.   Reception of this packet refreshes the timeout calculation. The local node undergoing the maintenance operation will also delay the CCM timeout by the announced ETH-VSM interval. This local interval will be reset when any ETH-CC PDU is received on the MEP. An optional TLV is included in AIS packets to extend timeout values for active AIS conditions.

At the end of the maintenance operation there will be a burst of three more messages over a 10 second window that will indicate that the maintenance operation has completed. Once the first of these messages has been received the receiving peer will transition back to the ETH-CCM message and associated interval as the indication for the remote timeout (3.5*ccm-interval+hold if any).

CCM message will continue to be sent during this process but loss of the CCM packets during this 10s window will not affect the remote peer timeout. The only change to the CCM processing is which timer to use during the maintenance operation. During the operation, the value used is that announced as part of the ETH-VSM message. Outside a maintenance window the standard CCM-interval*3.5 + any configured hold time is used. Since CCM messages are sent during this time other faults and failures can still be conveyed and acted upon. These include AIS, Interface status settings, etc. Only the remote peer timeout (defRemoteCCM) is affected by the ETH-VSM announcement.

The grace announcement using ETH-VSM will continue until the upgrade or reset is completed. During an IOM soft reset ETH-CFM will not determine which peers are affected by a soft reset of a specific IOM. All remote peers will receive the ETH-VSM with the grace period announcement until the soft reset is completed. This means that all remote MEPs, regardless of location on the local node will enter a grace.

Clearing the IOM does not invoke the organizational specific TLV with the grace period announcement.

This is a value added function that is applicable to only nodes that implement support for ALU's approach for announcing grace using ETH-VSM. As specified in the standards, when a node does not support a specific optional function the message will be ignored and the no processing will be performed.

This feature is enabled by default. A system wide command is available to disable this transmission of these grace messages. Entering the no grace-tx-enable in the configuration under the **eth-cfm>system** context will prevent the grace announcements. If this configuration option is change from enable to disable while grace is being announced the three grace stop messages will be transmitted. Changing the state of this configuration option from disable to enable will only affect future ISSU and soft reset functions. It will have no affect on any ISSU or soft reset function that is active at the time this command was enabled.

CCM Hold Timers

In some cases the requirement exists to prevent a MEP from entering the defRemoteCCM defect, remote peer timeout, from more time than the standard 3.5 times the CCM-interval. Both the IEEE 802.1ag standard and ITU-T Y.1731 recommendation provide a non-configurable 3.5 times the CCM interval to determine a peer time out. However, when sub second CCM timers (10ms/100ms) are enabled the carrier may want to provide additional time for different network segments to converge before declaring a peer lost because of a timeout. In order to maintain compliance with the specifications the `ccm-hold-timer down <delay-down>` option has been introduced to artificially increase the amount of time it takes for a MEP to enter a failed state should the peer time out. This timer is only additive to CCM timeout conditions. All other CCM defect conditions, like defMACStatus, defXconCCM, and so on, will maintain their existing behavior of transitioning the MEP to a failed state and raising the proper defect condition without delay.

When the **ccm-hold-timer down** *delay-down* option is configured the following calculation is used to determine the remote peer time out (3.5 times the CCM-Interval + ccm-hold-timer delay-down).

This command is configured under the association. Only sub second CCM enabled MEPs support this hold timer. Ethernet-Tunnel Paths use a similar but slightly different approach and will continue to utilize the existing method. Ethernet-tunnels will be blocked from using this new hold timer.

It is possible to change this command on the fly without deleting it first. Simply entering the command with the new values will change to values without having to delete the command prior to the change.

It is possible to change the ccm-interval of a MEP on the fly without first deleting it. This means it is possible to change a sub second CCM enabled MEP to 1 second or above. The operator will be prevented from changing an association from a sub second CCM interval to a non-sub second CCM interval when a `ccm-hold-timer` is configured in that association. The `ccm-hold-timer` must be removed using the `no` option prior to allowing the transition from sub second to non-sub second CCM interval.

Alarm Indication Signal (ETH-AIS Y.1731)

Alarm Indication Signal (AIS) provides an Y.1731 capable MEP the ability to signal a fault condition in the reverse direction of the MEP, out the passive side. When a fault condition is detected the MEP will generate AIS packets at the configured client levels and at the specified AIS interval until the condition is cleared. Currently a MEP configured to generate AIS must do so at a level higher than its own. The MEP configured on the service receiving the AIS packets is required to have the active side facing the receipt of the AIS packet and must be at the same level the AIS, The absence of an AIS packet for 3.5 times the AIS interval set by the sending node will clear the condition on the receiving MEP.

AIS generation is also not subject to the low-priority-defect setting. AIS, when enabled, generates when the MEP enters any defect condition, including RDI.

AIS configuration has two components: receive and transmit. AIS reception is enabled when the command **ais-enable** is configured under the MEP. The transmit function is enabled when the **client-meg-level** is configured.

Alarm Indication Signal function is used to suppress alarms at the client (sub) layer following detection of defect conditions at the server (sub) layer. Due to independent restoration capabilities provided within the Spanning Tree Protocol (STP) environments, ETHAIS is not expected to be applied in the STP environment.

Transmission of frames with ETH-AIS information can be enabled or disabled on a MEP. Frames with ETH-AIS information can be issued at the client MEG Level by a MEP, including a Server MEP, upon detecting the following conditions:

- Signal failure conditions in the case that ETH-CC is enabled.
- AIS condition in the case that ETH-CC is disabled.

For a point-to-point ETH connection at the client (sub) layer, a client layer MEP can determine that the server (sub) layer entity providing connectivity to its peer MEP has encountered defect condition upon receiving a frame with ETH-AIS information. Alarm suppression is straightforward since a MEP is expected to suppress defect conditions associated only with its peer MEP.

For multipoint ETH connectivity at the client (sub) layer, a client (sub) layer MEP cannot determine the specific server (sub) layer entity that has encountered defect conditions upon receiving a frame with ETH-AIS information. More importantly, it cannot determine the associated subset of its peer MEPs for which it should suppress alarms since the received ETHAIS information does not contain that information. Therefore, upon reception of a frame with ETH-AIS information, the MEP will suppress alarms for all peer MEPs whether there is still connectivity or not.

Only a MEP, including a Server MEP, is configured to issue frames with ETH-AIS information. Upon detecting a defect condition the MEP can immediately start transmitting periodic frames with ETHAIS information at a configured client MEG Level. A MEP continues to transmit periodic frames with ETH-AIS information until the defect condition is removed. Upon receiving a frame with ETH-AIS information from its server (sub) layer, a client (sub) layer MEP detects AIS condition and suppresses alarms associated with all its peer MEPs. A MEP resumes alarm generation upon detecting defect conditions once AIS condition is cleared.

Specific configuration information required by a MEP to support ETH-AIS is the following:

- Client MEG Level — MEG level at which the most immediate client layer MIPs and MEPs exist.
- ETH-AIS transmission period — Determines transmission periodicity of frames with ETH-AIS information.
- Priority — Identifies the priority of frames with ETH-AIS information.
- Drop Eligibility — Frames with ETH-AIS information are always marked as drop ineligible.

A MIP is transparent to frames with ETH-AIS information and therefore does not require any information to support ETH-AIS functionality.

It is important to note that Facility MEPs do not support the generation of AIS to an explicitly configured endpoint. An explicitly configured endpoint is an object that contains multiple individual endpoints, as in pseudowire redundancy.

AIS is enabled under the service and has two parts, receive and transmit. Both of the components have their own configuration option. The **ais-enable** command under the SAP allows for the processing of received AIS packets at the MEP level. The **client-meg-level** command is the transmit portion that generates AIS if the MEP enter a fault state. AIS is independent of the **low-priority-defect** setting, so that any fault in the MEP causes AIS to be generated.

```
config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mip mac D0:0D:1E:01:01:01
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        ais-enable
                            client-meg-level 5
                        exit
                        ccm-enable
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
```

When MEP 101 enters a defect state, it starts to generate AIS out the passive side of the MEP, away from the fault. In this case, the AIS generates out sap 1/1/10:100.31 since MEP 101 is an up MEP on that SAP. The **Defect Flag** indicates that an RDI error state has been encountered and even though the **LowestDefectPri** setting is higher than the existing defect AIS is being transmitted. The **Eth-Ais Tx Counted** value is increasing, indicating that AIS is actively being sent.

```
# show eth-cfm mep 101 domain 4 association 1
===============================================================================
Eth-Cfm MEP Configuration Information
===============================================================================
Md-index           : 4                        Direction         : Up
Ma-index           : 1                        Admin             : Enabled
MepId              : 101                      CCM-Enable        : Disabled
IfIndex            : 35979264                 PrimaryVid        : 2031716
Description        : (Not Specified)
FngState           : fngReset                 ControlMep        : False
LowestDefectPri    : macRemErrXcon            HighestDefect     : none
Defect Flags       : bDefRDICCM
Mac Address        : d0:0d:1e:00:01:01        ControlMep        : False
CcmLtmPriority     : 7
CcmTx              : 2578                     CcmSequenceErr    : 0
Fault Propagation  : disabled                 FacilityFault     : n/a
MA-CcmInterval     : 1                        MA-CcmHoldTime    : 0ms
Eth-1Dm Threshold  : 3(sec)                   MD-Level          : 4
Eth-Ais:           : Enabled                  Eth-Ais Rx Ais:   : No
Eth-Ais Tx Priorit*: 7                        Eth-Ais Rx Interv*: 1
Eth-Ais Tx Interva*: 1                        Eth-Ais Tx Counte*: 288
Eth-Ais Tx Levels  : 5
Eth-Tst:           : Disabled

Redundancy:
    MC-LAG State   : n/a

CcmLastFailure Frame:
    None

XconCcmFailure Frame:
    None
===============================================================================
```

Test (ETH-TST Y.1731)

Ethernet test affords operators an Y.1731 capable MEP the ability to send an in service on demand function to test connectivity between two MEPs. The test is generated on the local MEP and the results are verified on the destination MEP. Any ETH-TST packet generated that exceeds the MTU will be silently dropped by the lower level processing of the node.

Specific configuration information required by a MEP to support ETH-test is the following:

- MEG level — MEG level at which the MEP exists
- Unicast MAC address of the peer MEP for which ETH-test is intended.
- Data - Optional element whose length and contents are configurable at the MEP.
- Priority — Identifies the priority of frames with ETH-Test information.
- Drop Eligibility — Identifies the eligibility of frames with ETHTest information to be dropped when congestion conditions are encountered.

A MIP is transparent to the frames with ETH-Test information and does not require any configuration information to support ETH-Test functionality.

Both nodes require the eth-test function to be enabled in order to successfully execute the test. Since this is a dual-ended test, initiate on sender with results calculated on the receiver, both nodes need to be check to see the results.

```
NODE1
config>service>epipe# info
----------------------------------------------
            sap 1/1/2:100.31 create
                eth-cfm
                    mip mac D0:0D:1E:01:01:01
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 101 domain 4 association 1 direction up
                        eth-test-enable
                        exit
                        mac-address d0:0d:1e:00:01:01
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
# oam eth-cfm eth-test d0:0d:1e:00:01:02 mep 101 domain 4 association 1 data-length 1000
# oam eth-cfm eth-test d0:0d:1e:00:01:02 mep 101 domain 4 association 1 data-length 1000
# oam eth-cfm eth-test d0:0d:1e:00:01:02 mep 101 domain 4 association 1 data-length 1000

NODE2
config>service>epipe# info
----------------------------------------------
```

```
            sap 1/1/2:100.31 create
                eth-cfm
                    mip mac D0:0D:1E:01:01:02
                exit
            exit
            sap 1/1/10:100.31 create
                eth-cfm
                    mep 102 domain 4 association 1 direction up
                        eth-test-enable
                        exit
                        mac-address d0:0d:1e:00:01:02
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

# show eth-cfm mep 102 domain 4 association 1 eth-test-results
===============================================================
Eth CFM ETH-Test Result Table
===============================================================
                                 Current        Accumulate
                  FrameCount     ErrBits        ErrBits
Peer Mac Addr     ByteCount      CrcErrs        CrcErrs
---------------------------------------------------------------
d0:0d:1e:00:01:01 3              0              0
                  3000           0              0
===============================================================
```

One-Way Delay Measurement (ETH-1DM Y.1731)

One-way delay measurement allows the operator the ability to check unidirectional delay between MEPs. An ETH-1DM packet is time stamped by the generating MEP and sent to the remote node. The remote node time stamps the packet on receipt and generates the results. The results, available from the receiving MEP, will indicate the delay and jitter. Jitter, or delay variation, is the difference in delay between tests. This means the delay variation on the first test will not be valid. It is important to ensure that the clocks are synchronized on both nodes to ensure the results are accurate. NTP can be used to achieve a level of wall clock synchronization between the nodes.

Note: accuracy relies on the nodes ability to timestamp the packet in hardware. Network elements that do not support this hardware time stamping, like the ESS-1 and SR-1, will display different results than hardware time stamp capable devices, like the SR-7/SR-12 and ESS-7/ESS-12.

Two-Way Delay Measurement (ETH-DMM Y.1731)

Two-way delay measurement is similar to one way delay measurement except it measures the round trip delay from the generating MEP. In this case wall clock synchronization issues will not influence the test results because four timestamps are used. This allows the remote nodes time to be removed from the calculation and as a result clock variances are not included in the results. The same consideration for first test and hardware based time stamping stated for one way delay measurement are applicable to two-way delay measurement.

Delay can be measured using one-way and two-way on demand functions. The two-way test results are available single-ended, test initiated, calculation and results viewed on the same node. There is no specific configuration under the MEP on the SAP in order to enabled this function. An example of an on demand test and results are below. The latest test result is stored for viewing. Further tests will overwrite the previous results. Delay Variation is only valid if more than one test has been executed.

```
oam eth-cfm two-way-delay-test d0:0d:1e:00:01:02 mep 101 domain 4 association 1

Two-Way-Delay-Test Response:
Delay 2955 microseconds        Variation 111 microseconds

# show eth-cfm mep 101 domain 4 association 1 two-way-delay-test
====================================================================
Eth CFM Two-way Delay Test Result Table
====================================================================
Peer Mac Addr         Delay (us)          Delay Variation (us)
--------------------------------------------------------------------
d0:0d:1e:00:01:02     2955                111
====================================================================
```

Synthetic Loss Measurement (ETH-SL)

Notes: Release 9.0R1 uses pre-standard OpCodes and will not interoperate with any other release or future release.

This synthetic loss measurement approach is a single-ended feature that allows the operator to run on-demand and proactive tests to determine “in”, “out” loss and “unacknowledged” packets. This approach can be used between peer MEPs in both point to point and multipoint services. Only remote MEP peers within the association and matching the unicast destination will respond to the SLM packet.

The specification uses various sequence numbers in order to determine in which direction the loss occurred. ALU has implemented the required counters to determine loss in each direction. In order to properly use the information that is gathered the following terms are defined;

- Count — The number of probes that are sent when the last frame is not lost. When the last frame(s) is/are lost, the count + unacknowledged equals the number of probes sent.
- Out-Loss (Far-end) — Packets lost on the way to the remote node, from test initiator to test destination
- In-Loss (Near-end) — Packet loss on the way back from the remote node to the test initiator.
- Unacknowledged — Number of packets at the end of the test that were not responded to.

The per probe specific loss indicators are available when looking at the on-demand test runs, or the individual probe information stored in the MIB. When tests are scheduled by Service Assurance Application (SAA) the per probe data is summarized and per probe information is not maintained. Any “unacknowledged” packets will be recorded as “in-loss” when summarized.

The on-demand function can be executed from CLI or SNMP. The on demand tests are meant to provide the carrier a means to perform on the spot testing. However, this approach is not meant as a method for storing archived data for later processing. The probe count for on demand SLM has a range of one to 100 with configurable probe spacing between one second and ten seconds. This means it is possible that a single test run can be up to 1000 seconds in length. Although possible, it is more likely the majority of on demand case will be run up to 100 probes or less at a one second interval. A node may only initiate and maintain a single active on demand SLM test at any given time. A maximum of one storage entry per remote MEP is maintained in the results table. Subsequent runs to the same peer will overwrite the results for that peer. This means when using on demand testing the test should be run and the results checked prior to starting another test.

The proactive measurement functions are linked to SAA. This backend provides the scheduling, storage and summarization capabilities. Scheduling may be either continuous or periodic. It also allows for the interpretation and representation of data that may enhance the specification. As an example, an optional TVL has been included to allow for the measurement of both loss and delay/jitter with a single test. The implementation does not cause any interoperability because the optional TVL will be ignored by equipment that does not support this. In mixed vendor environments loss measurement will continue to be tracked but delay and jitter will only report round trip times. It is important to point out that the round trip times in this mixed vendor environments will include the remote nodes processing time because only two time stamps will be included in the packet. In an environment where both nodes support the optional TLV to include time stamps unidirectional and round trip times will be reported. Since all four time stamps are included in the packet the round trip time in this case will not include remote node processing time. Of course, those operators that wish to run delay measurement and loss measurement at different frequencies are free to run both ETH-SL and ETH-DM functions. ETH-SL is not replacing ETH-DM. Service Assurance is only briefly discussed here to provide some background on the basic functionality. In order to completely understand how SAA functions please refer to the appropriate section of the user guide.

The ETH-SL packet format contains a test-id that will be internally generated and not configurable. The test-id will be visible for the on demand test in the display summary. It is possible a remote node processing the SLM frames will receive overlapping test-ids as a result of multiple MEPs measuring loss between the same remote MEP. For this reason, the uniqueness of the test is based on remote MEP-ID, test-id and Source MAC of the packet.

ETH-SL is applicable to up and down MEPs and as per the recommendation transparent to MIPs. There is no coordination between various fault conditions that could impact loss measurement. This is also true for conditions where MEPs are placed in shutdown state as a result of linkage to a redundancy scheme like MC-LAG. Loss measurement is based on the ETH-SL and not coordinated across different functional aspects on the network element. ETH-SL is supported on service based MEPs.

It is possible that two MEPs may be configured with the same MAC on different remote nodes. This will cause various issues in the FDB for multipoint services and is considered a misconfiguration for most services. It is possible to have a valid configuration where multiple MEPs on the same remote node have the same MAC. In fact, this is somewhat likely. In this release, only the first responder will be used to measure packet loss. The second responder will be dropped. Since the same MAC for multiple MEPs is only truly valid on the same remote node this should is an acceptable approach.

There is no way for the responding node to understand when a test is completed. For this reason a configurable “inactivity-timer” determines the length of time a test is valid. The timer will maintain an active test as long as it is receiving packets for that specific test, defined by the test-id, remote MEP Id and source MAC. When there is a gap between the packets that exceeds the inactivity-timer the responding node will respond with a sequence number of one regardless of what the sequence number was the instantiating node sent. This means the remote MEP believes the previous test has expired and these probes are part of a new test. The default for the inactivity-timer is 100 second and has a range of ten to 100 seconds.

The responding node will be limited to 1000 concurrent test SLM tests. Any test that attempts to involve a node that is already actively processing 1000 SLM tests will show up as “out loss” or “unacknowledged” packets on the node that instantiated the test because the packets will be silently discarded at the responder. It is important for the operator to understand this is silent and no log entries or alarms will be raised. It is also important to keep in mind that these packets are ETH-CFM based and the different platforms stated receive rate for ETH-CFM must not be exceeded.

Only the configuration is supported by HA. There will be no synchronization of data between active and standby. Any unwritten, or active tests will be lost during a switchover and the data will not be recoverable.

ETH-SL provides a mechanism for operators to proactively trend packet loss for service based MEPs.

Configuration Example

The following illustration shows the configuration required for proactive SLM test using SAA.

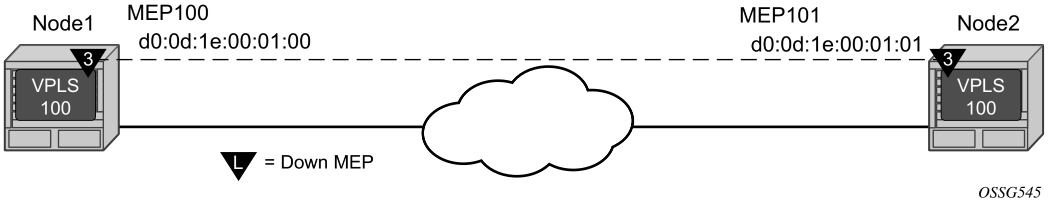

Figure 41: SLM Example

The output from the MIB is shown below as an example of an on-demand test. Node1 is tested for this example. The SAA configuration does not include the accounting policy required to collect the statistics before they are overwritten. NODE2 does not have an SAA configuration. NODE2 includes the configuration to build the MEP in the VPLS service context.

```
config>eth-cfm# info
----------------------------------------------
        domain 3 format none level 3
            association 1 format icc-based name "03-0000000100"
```

```
                bridge-identifier 100
                exit
                ccm-interval 1
                remote-mepid 101
            exit
        exit
----------------------------------------------

config>service>vpls# info
----------------------------------------------
            stp
                shutdown
            exit
            sap 1/1/3:100.100 create
            exit
            sap lag-1:100.100 create
                eth-cfm
                    mep 100 domain 3 association 1 direction down
                        ccm-enable
                        mac-address d0:0d:1e:00:01:00
                        no shutdown
                    exit
                exit
            exit
            no shutdown
----------------------------------------------

config>saa# info
----------------------------------------------
        test "slm1"
            type
                eth-cfm-two-way-slm d0:0d:1e:00:01:01 mep 100 domain 3
    association 1 count 100 timeout 1 interval 1
            exit
            continuous
            no shutdown
        exit
----------------------------------------------
```

The following sample output is meant to demonstrate the different loss conditions that an operator may see.   The total number of attempts is “99” is because the final probe in the test was not acknowledged.

```
# show saa slm1
Test Run: 183
Total number of attempts: 99
Number of requests that failed to be sent out: 0
Number of responses that were received: 48
Number of requests that did not receive any response: 50
Total number of failures: 50, Percentage: 50
 (in ms)            Min          Max      Average       Jitter
Outbound  :         -370         -362         -366        0.432
Inbound   :          363          371          367        0.308
Roundtrip :        0.000         5.93         1.38        0.496
Per test packet:
  Sequence     Outbound      Inbound    RoundTrip Result
```

```
         1       0.000       0.000       0.000 Out Loss
         2       0.000       0.000       0.000 Out Loss
         3       0.000       0.000       0.000 Out Loss
         4       0.000       0.000       0.000 Out Loss
…snip…
        46        -369         370        1.28 Response Received
        47        -362         363        1.42 Response Received
        48       0.000       0.000       0.000 In Loss
        49       0.000       0.000       0.000 In Loss
        50        -362         363        1.42 Response Received
        51        -362         363        1.16 Response Received
        52        -362         364        1.20 Response Received
        53        -362         364        1.18 Response Received
        54        -363         364        1.20 Response Received
…snip…
        96        -369         370        1.29 Response Received
        97        -369         370        1.30 Response Received
        98       0.000       0.000       0.000 Unacknowledged
        99       0.000       0.000       0.000 Unacknowledged
       100       0.000       0.000       0.000 Unacknowledged

===============================================================================
```

The following is an example of an on demand tests that and the associated output. Only single test runs are stored and can be viewed after the fact.

```
#oam eth-cfm two-way-slm-test d0:0d:1e:00:01:01 mep 100 domain 3 association 1 send-count
20 interval 1 timeout 1

Sending 20 packets to d0:0d:1e:00:01:01 from MEP 100/3/1 (Test-id: 588)

Sent 20 packets, 20 packets received from MEP ID 101, (Test-id: 588)
                 (0 out-loss, 0 in-loss, 0 unacknowledged)

# show eth-cfm mep 100 domain 3 association 1 two-way-slm-test
===============================================================================
Eth CFM Two-way SLM Test Result Table (Test-id: 588)
===============================================================================
Peer Mac Addr      Remote MEP       Count     In Loss    Out Loss         Unack
-------------------------------------------------------------------------------
d0:0d:1e:00:01:01         101          20           0           0             0
===============================================================================
```

ETH-CFM CoS Considerations

UP MEPs and Down MEPs have been aligned as of this release to better emulate service data. When an UP MEP or DOWN MEP is the source of the ETH-CFM PDU the priority value configured, as part of the configuration of the MEP or specific test, will be treated as the Forwarding Class (FC) by the egress QoS policy. If there is no egress QoS policy the priority value will be mapped to the CoS values in the frame. The discard ineligible by will be set. However, egress QoS Policy may overwrite this original value. The Service Assurance Agent (SAA) uses [fc {fc-name} [profile {in|out}]] to accomplish similar functionality.

UP MEPs and DOWN MEPs terminating an ETH-CFM PDU will use the received FC as the return priority for the appropriate response, again feeding into the egress QoS policy as the FC.

ETH-CFM PDUs received on the MPLS-SDP bindings will now properly pass the EXP bit values to the ETH-CFM application to be used in the response.

These are default behavioral changes without CLI options.

This does not include Ethernet Linktrace Response (ETH-LTR). The specification requires the highest priority on the bridge port should be used in response to an Ethernet Linktrace Message (ETH-LTM). This provides the highest possible chance of the response returning to the source. Operators may configure the linktrace response priority of the MEP using the ccm-ltm-priority. MIPs inherit the MEPs priority unless the mhf-ltr-priority is configured under the bridging instance for the association (config>eth-cfm>domain>assoc>bridge).

OAM Mapping

OAM mapping is a mechanism that enables a way of deploying OAM end-to-end in a network where different OAM tools are used in different segments. For instance, an Epipe service could span across the network using Ethernet access (CFM used for OAM), pseudowire (T-LDP status signaling used for OAM), and Ethernet access (E-LMI used for OAM). Another example allows an Ipipe service, where one end is Ethernet and the other end is Frame Relay or ATM.

In the SR OS implementation, the Service Manager (SMGR) is used as the central point of OAM mapping. It receives and processes the events from different OAM components, then decides the actions to take, including triggering OAM events to remote peers.

Fault propagation for CFM is by default disabled at the MEP level to maintain backward compatibility. When required, it can be explicitly enabled by configuration.

Fault propagation for a MEP can only be enabled when the MA is comprised of no more than two MEPs (point-to-point).

Fault propagation cannot be enabled for eth-tun control MEPs (MEPs configured under the eth-tun primary and protection paths). However, failure of the eth-tun (meaning both paths fail) will be propagated by SMGR because all the SAPs on the eth-tun will go down.

CFM Connectivity Fault Conditions

CFM MEP declares a connectivity fault when its defect flag is equal to or higher than its configured lowest defect priority. The defect can be any of the following depending on configuration:

- DefRDICCM
- DefMACstatus
- DefRemoteCCM
- DefErrorCCM
- DefXconCCM

The following additional fault condition applies to Y.1731 MEPs:

- Reception of AIS for the local MEP level

Setting the lowest defect priority to allDef may cause problems when fault propagation is enabled in the MEP. In this scenario, when MEP A sends CCM to MEP B with interface status down, MEP B will respond with a CCM with RDI set. If MEP A is configured to accept RDI as a fault, then it gets into a dead lock state, where both MEPs will declare fault and never be able to recover. The default lowest defect priority is DefMACstatus, which will not be a problem when interface status

TLV is used. It is also very important that different Ethernet OAM strategies should not overlap the span of each other. In some cases, independent functions attempting to perform their normal fault handling can negatively impact the other. This interaction can lead to fault propagation in the direction toward the original fault, a false positive, or worse, a deadlock condition that may require the operator to modify the configuration to escape the condition. For example, overlapping Link Loss Forwarding (LLF) and ETH-CFM fault propagation could cause these issues.

For the DefRemoteCCM fault, it is raised when any remote MEP is down. So whenever a remote MEP fails and fault propagation is enabled, a fault is propagated to SMGR.

CFM Fault Propagation Methods

When CFM is the OAM module at the other end, it is required to use any of the following methods (depending on local configuration) to notify the remote peer:

- Generating AIS for certain MEP levels
- Sending CCM with interface status TLV “down”
- Stopping CCM transmission

For using AIS for fault propagation, AIS must be enabled for the MEP. The AIS configuration needs to be updated to support the MD level of the MEP (currently it only supports the levels above the local MD level).

Note that the existing AIS procedure still applies even when fault propagation is disabled for the service or the MEP. For example, when a MEP loses connectivity to a configured remote MEP, it generates AIS if it is enabled. The new procedure that is defined in this document introduces a new fault condition for AIS generation, fault propagated from SMGR, that is used when fault propagation is enabled for the service and the MEP.

The transmission of CCM with interface status TLV must be done instantly without waiting for the next CCM transmit interval. This rule applies to CFM fault notification for all services.

Notifications from SMGR to the CFM MEPs for fault propagation should include a direction for the propagation (up or down: up means in the direction of coming into the SAP/SDP-binding; down means in the direction of going out of the SAP/SDP-binding), so that the MEP knows what method to use. For instance, an up fault propagation notification to a down MEP will trigger an AIS, while a down fault propagation to the same MEP can trigger a CCM with interface TLV with status down.

For a specific SAP/SDP-binding, CFM and SMGR can only propagate one single fault to each other for each direction (up or down).

When there are multiple MEPs (at different levels) on a single SAP/SDP-binding, the fault reported from CFM to SMGR will be the logical OR of results from all MEPs. Basically, the first fault from any MEP will be reported, and the fault will not be cleared as long as there is a fault in any local MEP on the SAP/SDP-binding.

Epipe Services

Down and up MEPs are supported for Epipe services as well as fault propagation. When there are both up and down MEPs configured in the same SAP/SDP-binding and both MEPs have fault propagation enabled, a fault detected by one of them will be propagated to the other, which in turn will propagate fault in its own direction.

CFM Detected Fault

When a MEP detects a fault and fault propagation is enabled for the MEP, CFM needs to communicate the fault to SMGR, so SMGR will mark the SAP/SDP-binding faulty but still oper-up. CFM traffic can still be transmitted to or received from the SAP/SDP-binding to ensure when the fault is cleared, the SAP will go back to normal operational state. Since the operational status of the SAP/SDP-binding is not affected by the fault, no fault handling is performed. For example, applications relying on the operational status are not affected.

If the MEP is an up MEP, the fault is propagated to the OAM components on the same SAP/SDP-binding; if the MEP is a down MEP, the fault is propagated to the OAM components on the mate SAP/SDP-binding at the other side of the service.

SAP/SDP-Binding Failure (Including Pseudowire Status)

When a SAP/SDP-binding becomes faulty (oper-down, admin-down, or pseudowire status faulty), SMGR needs to propagate the fault to up MEP(s) on the same SAP/SDP-bindings about the fault, as well as to OAM components (such as down MEPs and E-LMI) on the mate SAP/SDP-binding.

Service Down

This section describes procedures for the scenario where an Epipe service is down due to the following:

- Service is administratively shutdown. When service is administratively shutdown, the fault is propagated to the SAP/SDP-bindings in the service.
- If the Epipe service is used as a PBB tunnel into a B-VPLS, the Epipe service is also considered operationally down when the B-VPLS service is administratively shutdown or operationally down. If this is the case, fault is propagated to the Epipe SAP.
- In addition, one or more SAPs/SDP-bindings in the B-VPLS can be configured to propagate fault to this Epipe (see fault-propagation-bmac below). If the B-VPLS is operationally up but all of these entities have detected fault or are down, the fault is propagated to this Epipe's SAP.

Interaction with Pseudowire Redundancy

When a fault occurs on the SAP side, the pseudowire status bit is set for both active and standby pseudowires. When only one of the pseudowire is faulty, SMGR does not notify CFM. The notification occurs only when both pseudowire becomes faulty. The SMGR propagates the fault to CFM.

Since there is no fault handling in the pipe service, any CFM fault detected on an SDP binding is not used in the pseudowire redundancy's algorithm to choose the most suitable SDP binding to transmit on.

Ipipe Services

For Ipipe services, only down MEPs are supported on Ethernet SAPs.

CFM Detected Fault

When a MEP detects a fault and fault propagation is enabled for the MEP, CFM needs to communicate the fault to SMGR, so SMGR will mark the SAP/SDP-binding faulty but still oper-up. CFM traffic can still be transmitted to or received from the SAP/SDP-binding to ensure when the fault is cleared, the SAP will go back to normal operational state.

Because the MEP is a down MEP, the fault is always propagated to the OAM components on the mate SAP/SDP-binding at the other side of the service.

SAP/SDP-binding Failure (Including Pseudowire Status)

When a SAP/SDP-binding becomes faulty (oper-down, admin-down, or pseudowire status faulty), SMGR propagates the fault to OAM components on the mate SAP/SDP-binding.

Service Administratively Shutdown

When the service is administratively shutdown, SMGR propagates the fault to OAM components on both SAP/SDP-bindings.

Interaction with Pseudowire Redundancy

When the fault occurs on the SAP side, the pseudowire status bit is set for both active and standby pseudowires.

When only one of the pseudowire is faulty, SMGR does not notify CFM. The notification only occurs when both pseudowires become faulty. Then the SMGR propagates the fault to CFM. Since there is no fault handling in the PIPE service, any CFM fault detected on a SDP-binding is not used in the pseudowire redundancy's algorithm to choose the most suitable SDP-binding to transmit on.

VPLS Service

For VPLS services, on down MEPs are supported for fault propagation.

CFM Detected Fault

When a MEP detects a fault and fault propagation is enabled for the MEP, CFM communicate the fault to the SMGR. The SMGR will mark the SAP/SDP-binding as oper-down. Note that oper-down is used here in VPLS instead of "oper-up but faulty" in the pipe services. CFM traffic can be transmitted to or received from the SAP/SDP-binding to ensure when the fault is cleared, the SAP will go back to normal operational state.

Note that as stated in CFM Connectivity Fault Conditions on page 240, a fault is raised whenever a remote MEP is down (not all remote MEPs have to be down). When it is not desirable to trigger fault handling actions in some cases when a down MEP has multiple remote MEPs, operators can disable fault propagation for the MEP.

If the MEP is a down MEP, SMGR performs the fault handling actions for the affected service(s). Local actions done by the SMGR include (but are not limited to):

- Flushing MAC addresses learned on the faulty SAP/SDP-binding.
- Triggering transmission of MAC flush messages.
- Notifying MSTP/RSTP about topology change. If the VPLS instance is a management VPLS (mVPLS), all VPLS instances that are managed by the m VPLS inherits the MSTP/RSTP state change and react accordingly to it.
- If the service instance is a B-VPLS, and fault-propagation-bmac address(es) is/are configured for the SAP/SDP-binding, SMGR performs a lookup using the BMAC address(es) to find out which pipe services need to be notified, then propagates a fault to these services. There can be up to four remote BMAC addresses associated with an SAP/SDP-binding for the same B-VPLS.

SAP/SDP-Binding Failure (Including Pseudowire Status)

If the service instance is a B-VPLS, and an associated BMAC address is configured for the failed SAP/SDP-binding, the SMGR performs a lookup using the BMAC address to find out which pipe services will be notified and then propagate fault to these services.

Within the same B-VPLS service, all SAPs/SDP-bindings configured with the same fault propagation BMACs must be faulty or oper down for the fault to be propagated to the appropriate pipe services.

Service Down

When a VPLS service is down:

- If the service is not a B-VPLS service, the SMGR propagates the fault to OAM components on all SAP/SDP-bindings in the service.
- If the service is a B-VPLS service, the SMGR propagates the fault to OAM components on all SAP/SDP-bindings in the service as well as all pipe services that are associated with the B-VPLS instance.

Pseudowire Redundancy and Spanning Tree Protocol

A SAP or SDP binding that has a down MEP fault is made operationally down. This causes pseudowire redundancy or Spanning Tree Protocol (STP) to take the appropriate actions.

However, the reverse is not true. If the SAP or SDP binding is blocked by STP, or is not tx-active due to pseudowire redundancy, no fault is generated for this entity.

IES and VPRN Services

For IES and VPRN services, only down MEP is supported on Ethernet SAPs and spoke SDP bindings.

When a down MEP detects a fault and fault propagation is enabled for the MEP, CFM communicates the fault to the SMGR. The SMGR marks the SAP/SDP binding as operationally down. CFM traffic can still be transmitted to or received from the SAP/SDP-binding to ensure when the fault is cleared and the SAP will go back to normal operational state.

Because the SAP/SDP-binding goes down, it is not usable to upper applications. In this case, the IP interface on the SAP/SDP-binding go down. The prefix is withdrawn from routing updates to the remote PEs. The same applies to subscriber group interface SAPs.

When the IP interface is administratively shutdown, the SMGR notifies the down MEP and a CFM fault notification is generated to the CPE through interface status TLV or suspension of CCM based on local configuration.

Pseudowire Switching

When the node acts as a pseudowire switching node, meaning two pseudowires are stitched together at the node, the SMGR will not communicate pseudowire failures to CFM. Such features are expected to be communicated by pseudowire status messages, and CFM will run end-to-end on the head-end and tail-end of the stitched pseudowire for failure notification.

LLF and CFM Fault Propagation

LLF and CFM fault propagation are mutually exclusive. CLI protection is in place to prevent enabling both LLF and CFM fault propagation in the same service, on the same node and at the same time. However, there are still instances where irresolvable fault loops can occur when the two schemes are deployed within the same service on different nodes. This is not preventable by the CLI. At no time should these two fault propagation schemes be enabled within the same service.

802.3ah EFM OAM Mapping and Interaction with Service Manager

802.3ah EFM OAM declares a link fault when any of the following occurs:

- Loss of OAMPDU for a certain period of time
- Receiving OAMPDU with link fault flags from the peer

When 802.3ah EFM OAM declares a fault, the port goes into operation state down. The SMGR communicates the fault to CFM MEPs in the service.

OAM fault propagation in the opposite direction (SMGR to EFM OAM) is not supported.

Service Assurance Agent (SAA)

Service Application Agent (SAA) is a tool that allows operators to configure a number of different tests that can be used to provide performance information like delay, jitter and loss for services or network segments. The test results are saved in SNMP tables or summarized XML files. These results can be collected and reported on using network management systems.

SAA uses the resources allocated to the various OAM processes. These processes are not dedicated to SAA but shared throughout the system. Table 5 provides guidance on how these different OAM functions are logically grouped.

Table 5: SAA Test and Descriptions

Test	Description
Background	It is tasks configured outside of the SAA hierarchy that consume OAM task resources. Specifically, these include SDP-Keep Alive, Static route cpe-check, filter redirect-policy, ping-test, and vrrp policy host-unreachable. These are critical tasks that ensure the network operation and may affect data forwarding or network convergence.
SAA Continuous	It is configured SAA tests with the “continuous” key word, hence always scheduled.
SAA non-continuous	It is configured SAA tests that do not use the “continuous” key word, hence scheduled outside of the SAA application, requires the “oam saa start testname” to initiate the test run.
Non-SAA (Directed)	It is any task that does not include any configuration under SAA. These tests are SNMP or via the CLI that is used to troubleshoot or profile network condition. This would take the form “oam test-type” or ping/traceroute with the specific test parameters.

SAA test types are restricted to those that utilize a request response mechanism, single-ended tests. Dual-ended tests that initiate the test on one node but require the statistical gathering on the other node are not supported under SAA. As an example, Y.1731 defines two approaches for measuring frame delay and frame delay variation, single-ended and dual-ended. The single-ended approach is supported under SAA.

Post processing analysis of individual test runs can be used to determine the success or failure of the individual runs. The operator can set rising and lowering thresholds for delay, jitter, and loss. Exceeding the threshold will cause the test to have a failed result. A trap can be generated when the test fails. The operator is also able to configure a probe failure threshold and trap when these thresholds are exceeded.

Each supported test type has configuration properties specific to that test. Not all options, intervals, and parameters are available for all tests. Some configuration parameters, such as the sub second probe interval require specific hardware platforms.

The ETH-CFM SAA tests may be configured as “continuous”, meaning always scheduled. By default, all tests are configure in a waiting-to-start mode. This would require the operator to issue the “oam saa start testname” command to launch the test. When a test is executing the probe, spacing is be based on the interval parameter assuming there are no lost packets. In general, trace type tests will apply the timeout to each individual packet.   This is required because packet timeout may be required to move from one probe to the next probe. For those tests that do not require this type of behavior, typically ping functions, the probes will be sent at the specified probe interval and the timeout will only be applied at the end of the test if any probe has been lost during the run. When the timeout is applied at the end of the run, the test is considered complete when either all response have been received or the timeout expires at the end of the test run. For test marked as “continuous”, always scheduled, the spacing between the runs may be delayed by the timeout value when a packet is lost. The test run is complete when all probes have either been received back or the timeout value has expired.

In order to preserve system resources, specifically memory, the operator should only store summarized history results. By default, summary results are stored for tests configured with sub second probe intervals, or a probe count above 100 or is written to a file. By default, per probe information will be stored for test configured with an interval of one second or above counters, and probe counts of 100 or less and is not written to a file. The operator may choose to override these defaults using the **probe-history {keep|drop|auto}** option. The “auto” option sets the defaults above. The other options override the default retention schemes based on the operator requirements, per probe retention “keep” or summary only information “drop”. The probe data can be viewed using the “show saa test” command. If the per probe information is retained, this probe data is available at the completion of the test run. The summary data is updated throughout the test run. The overall memory system usage is available using the “show system memory-pools” command. The OAM entry represents the overall memory usage. This includes the history data stored for SAA tests. A “clear saa *testname*” option is available to release the memory and flush the test results.

The following example shows Y.1731 ETH-DMM packets to be sent from the local MEP 325, domain 12 and association 300 to destination MAC address d0:0d:1e:00:00:27. The tests will be scheduled as continuous and does not require an “oam saa start testname” to be issued by the operator. Each individual test run will contain 900 probes at 1 second intervals. This means each individual test run will be active for 15 minutes. If a packet is lost, the test will wait for the timeout (default 5s not shown) before closing one run and move to the next. If more than 10 probes are lost, the test will be marked as failed and a trap and log entry will be generated.

Test summary information and not per probe data is maintained for this test because the optional probe-history override is not configured. The summary information will be written to an XML file using the accounting-policy 1.

Example:

```
saa>test# info
----------------------------------------------
description "Two Way ETH-DDM To MEP 327 From MEP 325"
type
    eth-cfm-two-way-delay d0:0d:1e:00:00:27 mep 325 domain 12 association 300
    count 900 interval 1
exit
trap-gen
    probe-fail-enable
    probe-fail-threshold 10
exit
accounting-policy 1
continuous
no shutdown
```

SAA leverages the accounting record infrastructure. The sample configuration is included for completeness. For complete information on Accounting Policies consult the System Management Guide for the appropriate platform.

```
config>log# info
----------------------------------------------
file-id 1
    location cf3:
    rollover 60 retention 24
exit
accounting-policy 1
    description "SAA XML File"
    record saa
    collection-interval 15
    to file 1
    no shutdown
exit
```

SAA launched tests will maintain two most recent completed and one in progress test. The output below is the summary data from the test above.   Below, test run 18 and 19 have been completed and test run 20 is in progress. Once test run 20 is completed test run 18 data will be overwritten. It is important to ensure that the collection and accounting record process is configured in such a way to write the data to file before it is overwritten. Once the results are overwritten they are lost.

```
show saa "saa-dmm-1"

===============================================================================
SAA Test Information
===============================================================================
Test name                     : saa-dmm-1
Owner name                    : TiMOS CLI
Description                   : Two Way ETH-DDM To MEP 327 From MEP 325
Accounting policy             : 1
Continuous                    : Yes
Administrative status         : Enabled
Test type                     : eth-cfm-two-way-delay d0:0d:1e:00:00:27 mep
                                325 domain 12 association 300 count 900
                                interval 1
Trap generation               : probe-fail-enable probe-fail-threshold 10
```

```
Probe History                : auto (drop)
Test runs since last clear   : 3
Number of failed test runs   : 0
Last test result             : Success
-------------------------------------------------------------------------------
Threshold
Type        Direction Threshold  Value      Last Event          Run #
-------------------------------------------------------------------------------
Jitter-in   Rising    None       None       Never               None
            Falling   None       None       Never               None
Jitter-out  Rising    None       None       Never               None
            Falling   None       None       Never               None
Jitter-rt   Rising    None       None       Never               None
            Falling   None       None       Never               None
Latency-in  Rising    None       None       Never               None
            Falling   None       None       Never               None
Latency-out Rising    None       None       Never               None
            Falling   None       None       Never               None
Latency-rt  Rising    None       None       Never               None
            Falling   None       None       Never               None
Loss-in     Rising    None       None       Never               None
            Falling   None       None       Never               None
Loss-out    Rising    None       None       Never               None
            Falling   None       None       Never               None
Loss-rt     Rising    None       None       Never               None
            Falling   None       None       Never               None

===============================================================================
Test Run: 18
Total number of attempts: 900
Number of requests that failed to be sent out: 0
Number of responses that were received: 900
Number of requests that did not receive any response: 0
Total number of failures: 0, Percentage: 0
 (in ms)            Min          Max      Average       Jitter
Outbound  :        -29.3        -28.6        -28.9        0.000
Inbound   :         28.7         29.3         29.0        0.000
Roundtrip :        0.069        0.077        0.073        0.000
Per test packet:

Test Run: 19
Total number of attempts: 900
Number of requests that failed to be sent out: 0
Number of responses that were received: 900
Number of requests that did not receive any response: 0
Total number of failures: 0, Percentage: 0
 (in ms)            Min          Max      Average       Jitter
Outbound  :        -29.9        -29.3        -29.6        0.000
Inbound   :         29.3         30.0         29.7        0.001
Roundtrip :        0.069        0.080        0.073        0.001
Per test packet:

Test Run: 20
Total number of attempts: 181
Number of requests that failed to be sent out: 0
Number of responses that were received: 181
Number of requests that did not receive any response: 0
Total number of failures: 0, Percentage: 0
 (in ms)            Min          Max      Average       Jitter
```

```
Outbound  :        -30.0        -29.9        -30.0        0.001
Inbound   :         30.0         30.1         30.0        0.000
Roundtrip :        0.069        0.075        0.072        0.001
Per test packet:

===============================================================================
```

Any data not written to file will be lost on a CPU switch over.

There are a number of show commands to help the operator monitor the test oam tool set.

show test-oam oam-config-summary: Provides information about the configured tests.

show test-oam oam-perf: Provides the transmit (launched form me) rate information and remotely launched test receive rate on the local network element.

clear test-oam oam-perf: Provides the ability to clear the test oam performance stats for a current view of the different rates in the oam-perf command above.

monitor test-oam oam-perf: Makes use of the monitor command to provide time sliced performance stats for test oam functions.