Draft Recommendation Y.17ethoam

OAM Functions and Mechanisms for Ethernet based networks

Summary

<Mandatory material>

Ke ywords

<Optional>

TABLE OF CONTENTS

OAM Functions and Mechanisms for Ethernet based networks

1 Scope

2 References

The following ITU-T Recommendations and other references contain provisions, which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation

NOTE: The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[1] ITU-T Recommendation Y.1730 (2004), Requirements for OAM functions in Ethernet based networks

[2] ITU-T Recommendation I.610 (1999), *B-ISDN operation and maintenance principles and functions*.

[3] CCITT Recommendation M.20 (1992), *Maintenance philosophy for telecommunications networks*.

[4] ITU-T Recommendation G.805 (2000), *Generic functional architecture of transport networks*.

[5] ITU-T Recommendation G.8010 (2003), *Architecture of Ethernet Layer Networks*.

[6] ITU-T Recommendation G.8041 (2001) *Generic Framing Procedure (GFP)*.

[EDITOR'S NOTE] TO BE COMPLETED

3 Definitions

EDITOR'S NOTE: CHECK TERMS BELOW

This Recommendation introduces some functional architecture terminology that is required to discuss the network components associated with OAM. Relevant terms are defined below.

C-TAG: definition expected

S-TAG: definition expected

defect: Interruption of the capability of a transport entity (e.g. network connection) to transfer user or OAM information[2].

failure: Termination of the capability of a transport entity to transfer user or OAM information. A failure can be caused by a persisting defect[2].

ETH trail: a trail in the ETH layer

ETH link: a link in the ETH layer

link: A "topological component" which describes a fixed relationship between a "subnetwork" or "access group" and another "subnetwork" or "access group"[3].

trail: A "transport entity" which consists of an associated pair of "unidirectional trails" capable of simultaneously transferring information in opposite directions between their respective inputs and output[3]..

CE (customer edge device), which could be an Ethernet switch, a router or a host.

PE provider Edge device, which does not do any customer MAC switching, rather it encapsulates the customer traffic into a tunnel (e.g., IP, MPLS, ATM, FR, EOS).

Point to point Ethernet connection, is an end-to-end connection between two CEs.

[EDITOR'S NOTE] TO BE COMPLETED

4 Abbreviations

EDITOR'S NOTE: CHECK ABBREVIATIONS BELOW

This Recommendation uses the following abbreviations.

CE	Customer Edge device
DoS	Denial of Service
ETH	Ethernet
ETH-CC	Ethernet Continuity Check
ETH-LT	Ethernet Link-trace
ETH-LP	Ethernet Loopback
MAC	Media Access Control
ME	Maintenance Entity
MEP	Maintenance entity End Point
MIP	Maintenance entity Intermediate Point
NMS	Network Management System
OAM	Operation and Maintenance
PE	Provider Edge device
SLA	Service Level Agreement
TCP	Traffic Conditioning Point

[EDITOR'S NOTE] TO BE COMPLETED

5 Conventions

Maintenance Entity End Point (MEP) is a short name for an expanded ETH flow point that includes an ETH Segment flow termination function (that marks the end point of an ETH Maintenance Entity) and an ETH Diagnostic flow termination function. The ETH Maintenance Entity is capable to initiate fault management OAM like CC and RDI and terminate OAM like CC, AIS and RDI. The ETH Diagnostic flow termination function is capable to initiate and react to diagnostic OAM like loopback and link-trace. MEP is represented by square symbol as Figure 5-1.

EDITOR'S NOTE: THE FOLLOWING TWO FIGURES HAVE TO BE ALIGNED WITH THE FUNCTIONS AS DEFINED IN THIS RECOMMENDATION: E.G. TR SHOULD BE LINK TRACE, ETC

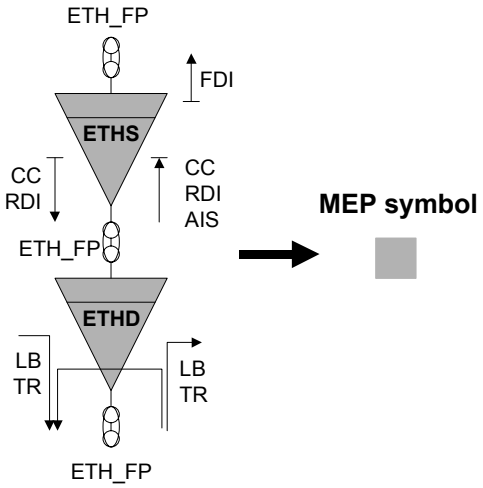


Figure 5-1: Maintenance entity End Point (MEP) symbol

Maintenance entity Intermediate Point (MIP) is a short name for an expanded ETH flow point including two ETH Diagnostic flow termination functions that are capable to react to diagnostic OAM like loopback and link-trace. MIP is represented by circle symbol as Figure 5-2.

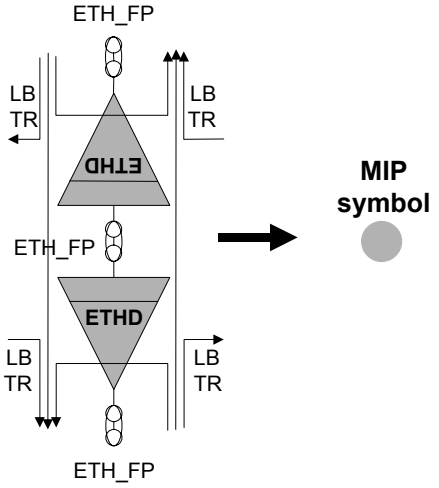


Figure 5- 2 – Maintenance entity Intermediate Point (MIP) symbol

Traffic Conditioning Point (TCP) is a short name for an expanded ETH flow point including an ETH traffic conditioning function like policing and shaping functions. TCP is represented by diamond symbol as Figure 5-3.

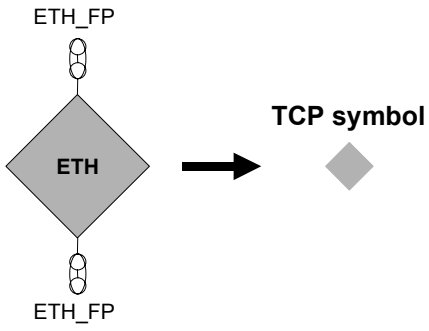


Figure 5-3 – Traffic Conditioning Point (TCP) symbol

6 Maintenance Entities (ME)

The maintenance entities in Ethernet networks are defined in Figures 23 and 24 of G.8010 [4] and in section 9 of Y.1730 [1]. The mapping of the maintenance entities as defined in both Recommendations and their relationship with the OAM flows defined in Y.1730 are shown in Table 6-1.

6.1 Relationship between Maintenance Entities and OAM Flows

Y.1730		G.8010
ME	**OAM flows**	**ME**
UNI-UNI (Customer)	UNI-UNI Flow	UNI_C to UNI-C ME
UNI-UNI (provider)	Transit Flow	UNI_N to UNI_N ME
Segment (PE-PE) intra-provider	Transit Flow	Intra Domain ME
Segment (PE-PE) inter-provider	Transit Flow Transit Link Flow	Inter Domain ME
Segment (any to any)	Transit Flow Transit Link Flow	
ETY Link OAM - UNI	UNI Link Flow	Access Link ME
ETY Link OAM - NNI	Transit Link Flow	Inter Domain ME

Table 6-1: MEs and OAM Flows

Y.1730 identifies different OAM flows which represent maintenance entities. The OAM flows can be inserted and extracted at the reference points, namely the flow points and termination flow points. The following OAM flows are identified:

- Customer UNI-UNI flow between reference points on the customer side of the UNI.
- Provider UNI-UNI flow between reference points on the provider side of the UNI
- Segment OAM flows:
  - Between flow points on the boundary of a provider network
  - Between flow points on the boundaries of two adjacent provider networks
  - Between any flow points as required
- ETY link OAM flow

Depending on the OAM flow, a provider may seek to limit it within its administrative boundary. For example, segment OAM flows between flow points on the boundary of a provider network may not be allowed to reach a customer network or another provider network. Similarly a segment OAM flow between flow points on the boundaries of two adjacent provider networks may not be allowed to reach a customer network or another provider network.

6.2 Relationship between VLANs and Maintenance Entities

Customer VLANs (C-VLAN and Service VLANs (S-VLAN) segregate the Maintenance Entities in their respective space. Same OAM mechanisms can be applied to each VLAN space. Thus, the same OAM Ethertype can be used for both C-VLAN and S-VLAN

CVLAN and SVLAN are one way to show Ethernet flow domain fragments. There are other ways. There is one OAM Ethertype for any Ethernet flow domain fragment.

EDITORS' NOTE: IMPROVE TEXT ON SEGREGATION HERE TO AVOID THE NEED TO MENTION C-VLAN AND S-VLAN IN THE REST OF THE DOCUMENT.

6.3 Representation of ME End and Intermediate Points and Traffic Conditioning Point in different Ethernet Network Scenarios

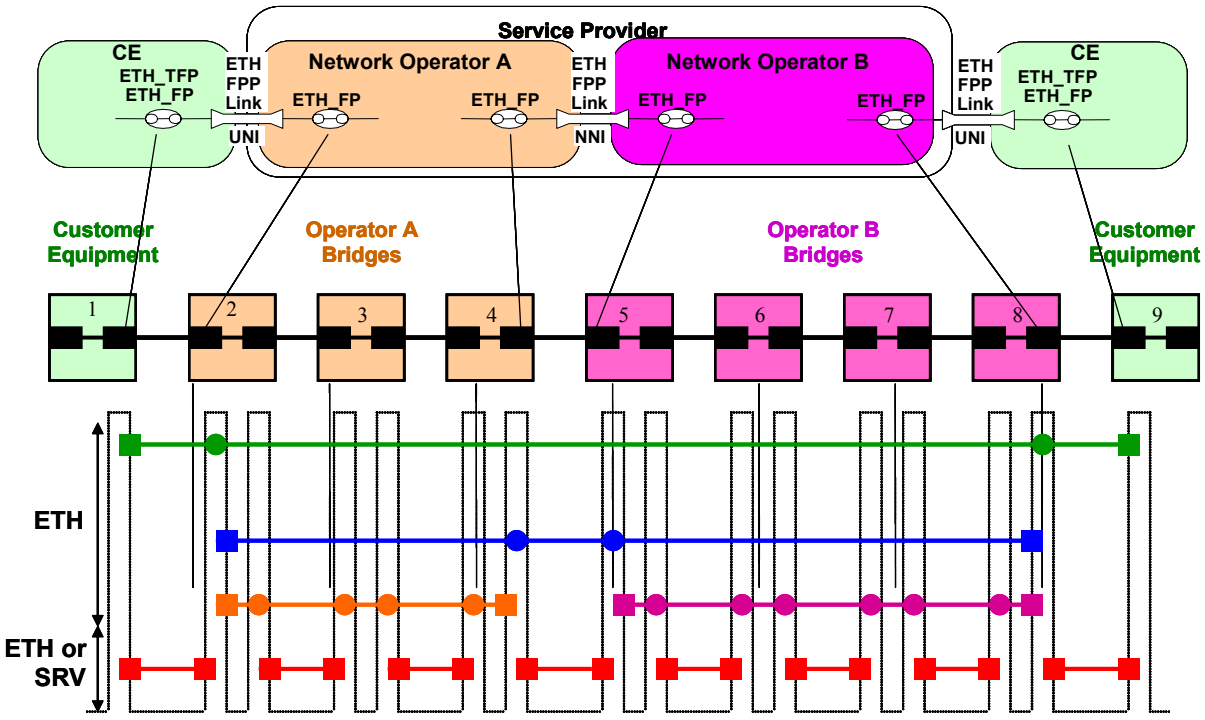


Figure 6-1: Example of ME End Point and ME Intermediate Point

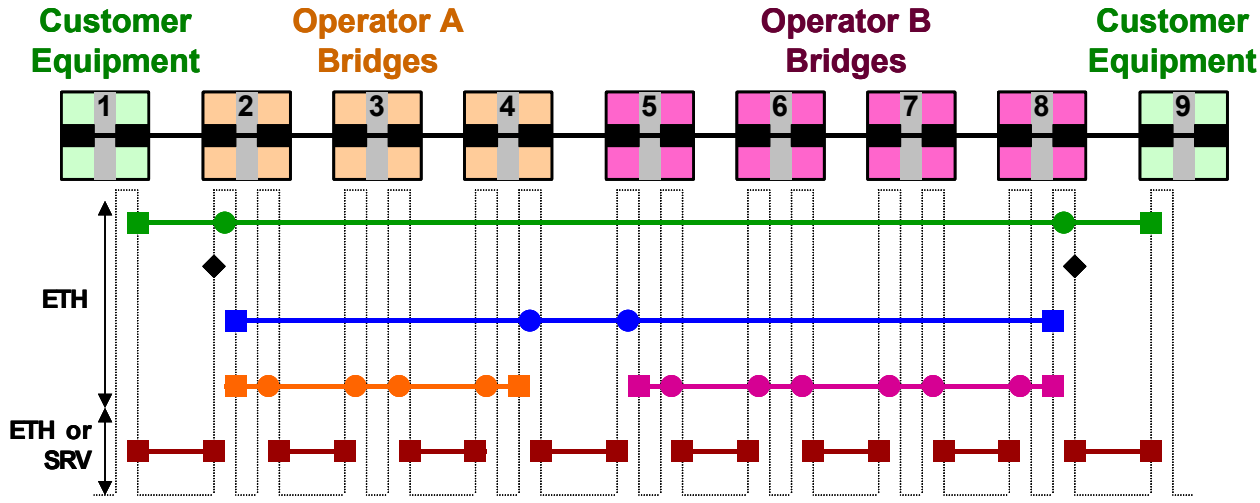


Figure 6-2 – Example of ETH MEs with their MEPs, MIPs and TCP

- p2p ETH connection between customer equipment 1 and 9 supported by a service provider and two network operators A and B

- green indicates a UNI-C to UNI-C ETH ME (customer) with MEPs in the interface ports facing the network in CE1 and CE9 and MIPs in the network interface ports facing the CEs (B2 and B8)
- blue indicates a UNI-N to UNI-N ETH ME (service provider) with MEPs at the edge of the network (B2,B8) and MIPs at the boundary of the two network operator domains (B4,B5)
- orange and mangenta indicate UNI-N to NNI ETH MEs (network operator) with MEPs at the edge of the operator networks (B2,B4 and B5,B8) and MIPs at each of the other interface ports
- brown indicates ETH link related MEs either realised as ETH MEs (sublayer monitoring) or as server layer MEs (inherent monitoring)
- black indicates location of unidirectional ETH TCPs; left TCP for direction CE1 to CE9 and right TCP for direction CE9 to CE1
- NOTE: the TCP must be located before any of the ETH MEPs in the interface port ➔ this is not correctly illustrated at the moment in the figure!! The black TCPs should be moved to the bottom of the figure if link is sublayer (ETH ME) monitored

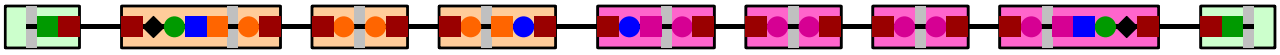


Figure 6-3 – Illustrating the order of MEPs, MIPs and TCP for example in Figure 6-2

- same p2p ETH connection as in previous figure, now represented in linear order; ETH_CI traffic units will pass through these MEPs, MIPs, TCPs in the presented order
- note: if link ME is an ETH ME (sublayer monitoring), then the TCPs must flip position with brown MEPs

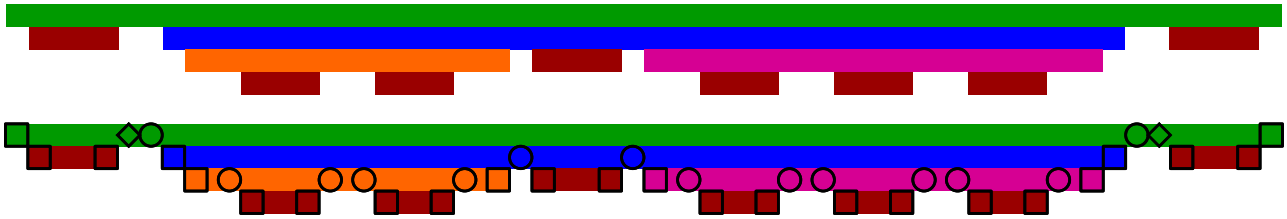


Figure 6-4 – Illustrating the stacking of ETH MEs or ETH and SRV MEs

- same p2p ETH connection as in previous two figures, now represented as a stack

6.3.1 MEP, MIP, TCP for Dual Relay Model and Bundling

6.3.1.1Single Integrated Provider Device

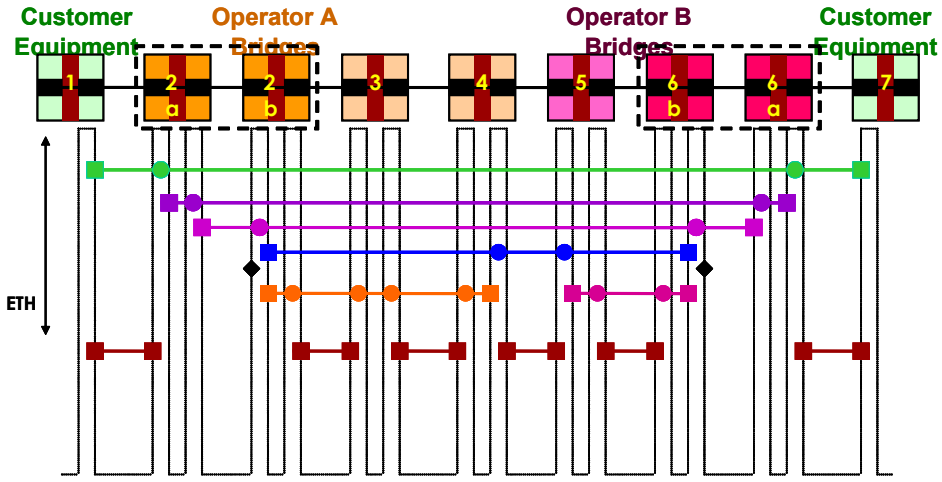


Figure 6-5: – Illustrating the location of MEPs, MIPs and TCPs in network model for the case of a one p2p connection across dual-relay provider devices

- Provider Device is represented as a dual relay model implemented with both relays. The first relay allows peering of customer L2CP protocols + multiplexing of multiple customer flows onto a single access link between the customer equipment 1 and provider bridge 2 (shown here as 2a and 2b).
- Due to the dual relay model, additional ME are introduced shown here in purple and pink between 2a and 6a. The Purple ME is associated with per customer VLAN at the provider equipment. The Pink ME is associated with per service instance (Service VLAN) that the provider applies to customer service frames.
- Between the dual relays, there are pseudo interfaces which correspond 1-to-1 with the Service VLAN or Provider Tag, which is expected to be inserted at second relay 2b.
- FFS: The positioning of the TCPs is for further study since TCPs can be positioned at customer access link level, per customer VLAN level and per provider VLAN (aka service) level.

6.3.1.2 Dual Relay Model with Single Relay as Provider Device

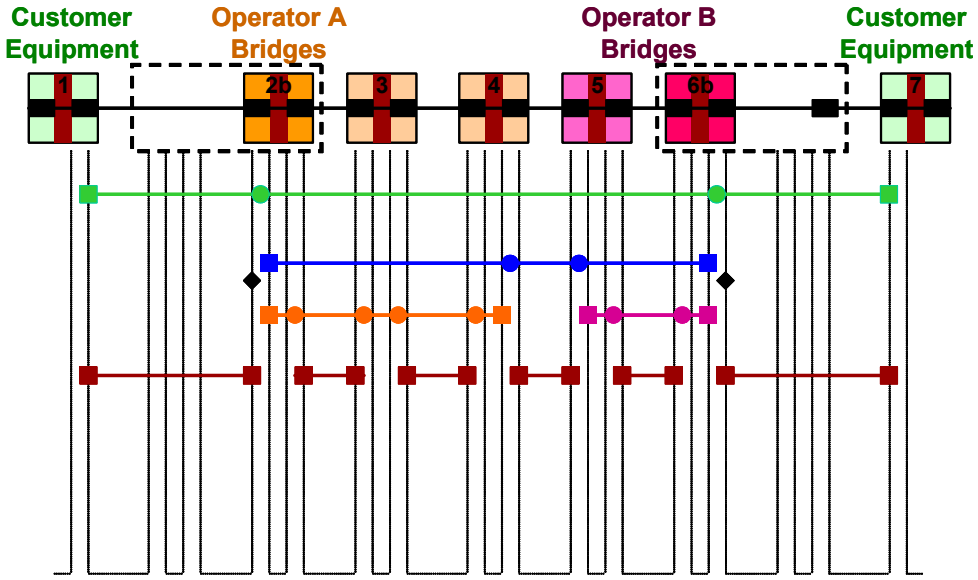


Figure 6-6: Illustrating the location of MEPs, MIPs and TCPs in network model for the case of a one p2p connection across dual-relay modelled provider device with a single relay

- Provider Device is represented as a dual relay model implemented with single relay (shown as 2b). In this case, the second relay does not allow peering of customer L2CP protocols + requires a single link for every service it supports across the customer device 1.
- Also customer device 1 is responsible for multiplexing multiple customer flows onto a single service link.
- Due to the provider using a single relay of dual relay model, additional ME that were introduced in Fig 1, are expected to be present at the customer device and are now shown here since customer is expected to manage his/her arrangements and its relationship with SP is limited to one single service instance ME marked here by Green ME between 1 and 7.
- The positioning of the TCP is clearer in this case and is shown in above figure. Customer is responsible for per customer VLAN level conditioning
- Provider Device is represented as a dual relay model implemented with single relay (shown as 2b). In this case, the second relay does not allow peering of customer L2CP protocols + requires a single link for every service it supports across the customer device 1.
- Also customer device 1 is responsible for multiplexing multiple customer flows onto a single service link.
- Due to the provider using a single relay of dual relay model, additional ME that were introduced in Fig 6-2, are expected to be present at the customer device and are now shown here since customer is expected to manage his/her arrangements and its relationship with SP is limited to one single service instance ME marked here by Green ME between 1 and 7.
- The positioning of the TCP is clearer in this case and is shown in above figure. Customer is responsible for per customer VLAN level conditioning

6.3.1.3 Dual Relay Model with Bundling for Single Integrated Provider Device

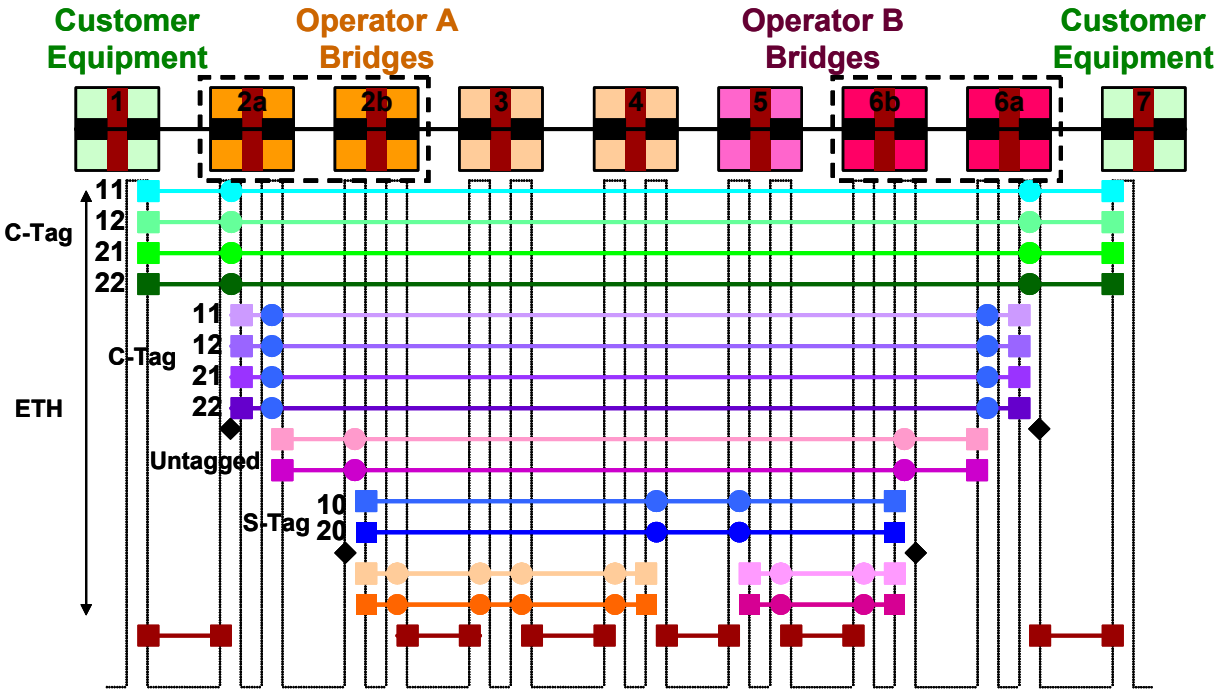


Figure 6-7: Illustrating the location of MEPs, MIPs and TCPs in network model for the case of a 2 p2p connection services with bundling across dual-relay provider devices

- Customer is shown using 4 customer VLANs (11, 12, 21, 22). It is also indicated that that customer signs up for 2 p2p connection services which the provider carries across the provider network using 2 provider VLANs (10 and 20). It is assumed that 2 customer VLANs (11 and 12) are mapped to provider VLAN 10 and other 2 customer VLANs (21 and 22) are mapped to provider VLAN 20.
- Additional ME are introduced in Figure 1 between 2a and 6a are replicated per customer VLAN and provider VLAN.
- MEs corresponding to the dual bridge pseudo interfaces which correspond 1-to-1 with the provider VLANs (10 and 20) are shown as untagged since frames from 1st relay e.g. 2a are expected to have no provider tag as they arrive at 2nd relay e.g. 2b.
- FFS: The positioning of the TCPs is for further study since TCPs can be positioned at customer access link level, per customer VLAN level and per provider VLAN (aka service) level.

6.3.1.4 Dual Relay Model with Bundling for Single Relay as Provider Device

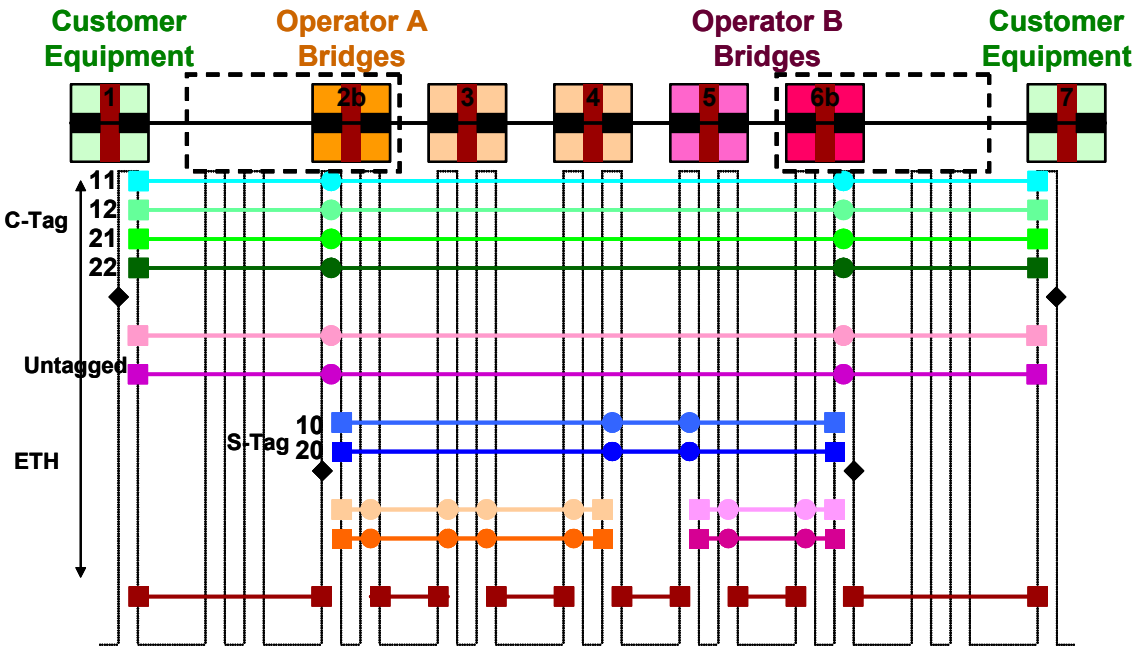


Figure 6-8: Illustrating the location of MEPs, MIPs and TCPs in network model for the case of a 2 p2p connection services with bundling across dual-relay provider devices with a single relay

- Due to the provider using a single relay of dual relay model, bundling is realized across the customer device 1 and 7.
- Additional ME is introduced at customer devices to highlight the responsibility of the customer for ME corresponding to per customer VLAN (shown here by 4 different green MEs between customer devices 1 and 7 for customer VLANs 11, 12, 21, and 22) and per service (shown here by 2 different purple MEs between customer devices 1 and 7) .
- The positioning of the TCP is clearer in this case and is shown in above figure. Customer is responsible for per customer VLAN level conditioning.

6.3.1.5 Dual Relay Model with all-to-one Bundling for Single Relay as Provider Device

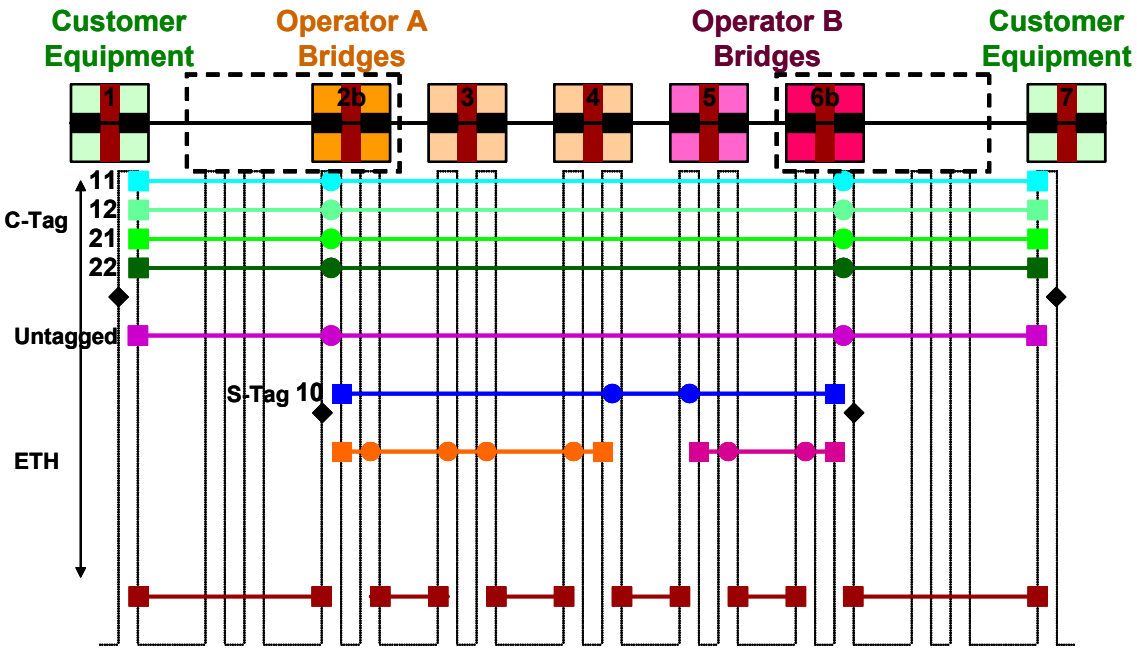


Figure 6-9: Illustrating the location of MEPs, MIPs and TCPs in network model for the case of a 1 p2p connection services with All-to-one bundling across dual-relay provider devices with a single relay

6.3.2 Network Demarcation Device (NDD) as an access Provider Device

An access scenario utilizing a Provider Network Demarcation Device (NDD) is illustrated below.

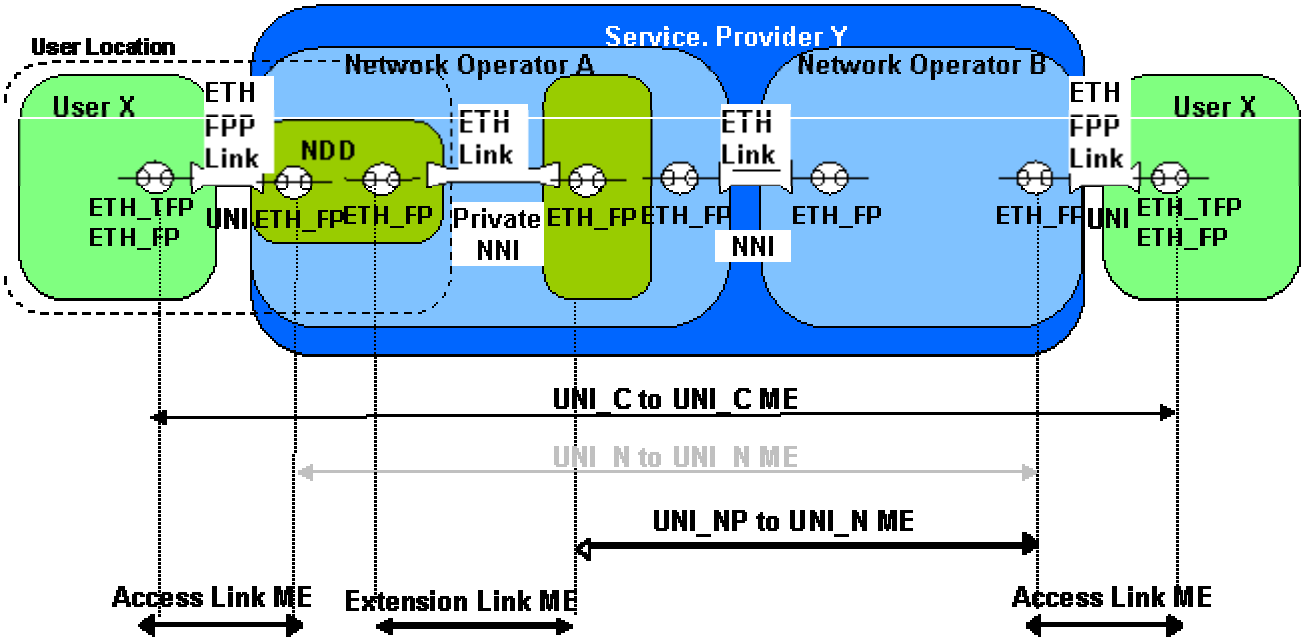


Figure 6.10: Illustrating the MEs in network model for the case of p2p connection services with Network Demarcation Device (NDD) as an access provider device

Deployment of this device in the network introduces the Extension Link ME (for the Private NNI ETH link) and the UNI_NP to UNI_N ME, which together form a subset of the previously defined UNI_N to UNI_N ME.

Figure 6-11 identifies the associated ME end points (MEPs) and ME intermediate points (MIPs) for this access scenario.

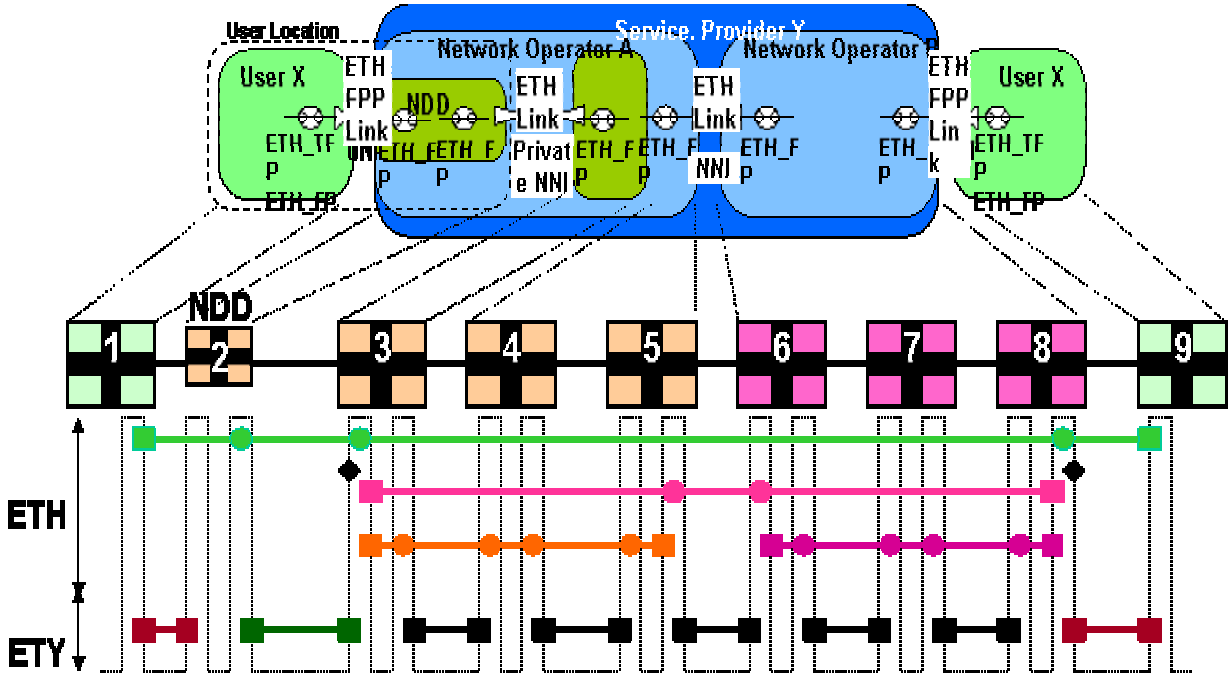


Figure 6.11: Illustrating the location of MEPs, MIPs in network model for the case of p2p connection services with Network Demarcation Device (NDD) as an access provider device

- A point-to-point ETH connection between customer equipment 1 and 9 is supported by service provider Y, consisting of a customer premise-located NDD connected to a metro transport network.
- Light green indicates a UNI-C to UNI-C ETH ME (customer) with MEPs in the interface ports facing the network in CE1 and CE9, and MIPs in the network interface ports facing the CEs (B2, B3 and B8).
- Dark green indicates at ETY layer indicates the Extension Link ETY ME. Brown represents an access link ME.
- Magenta indicates a UNI-NP to UNI-N ETH ME (metro network) with MEPs at the edge of the metro network (B3 and B8) and MIPs at each of the other interface ports in between.
- The UNI-N to UNI-N ETH ME (now shown in the figure) is realized by combination of Extension Link ETY ME and UNI-NP to UNI-N ETH ME.
- Orange and pink indicate intra-domain ETH MEs with MEPs at the edge of network operator A and B respectively (B3-B5 and B6-B8 respectively) and MIPs at each of the other interface ports in between.
- Black indicates ETH link MEs either realized as ETH MEs (sublayer monitoring) or as server layer MEs (inherent monitoring).

Other access scenarios are possible where the NDD could simply be a media converter (MC) with single flow-point. Figure 6.12 represents the scenario where the NDD is a Media Converter (MC) device where the Network Termination (NT) functionality is present in MC and Line Termination (LT) functionality is present in edge of provider domain.

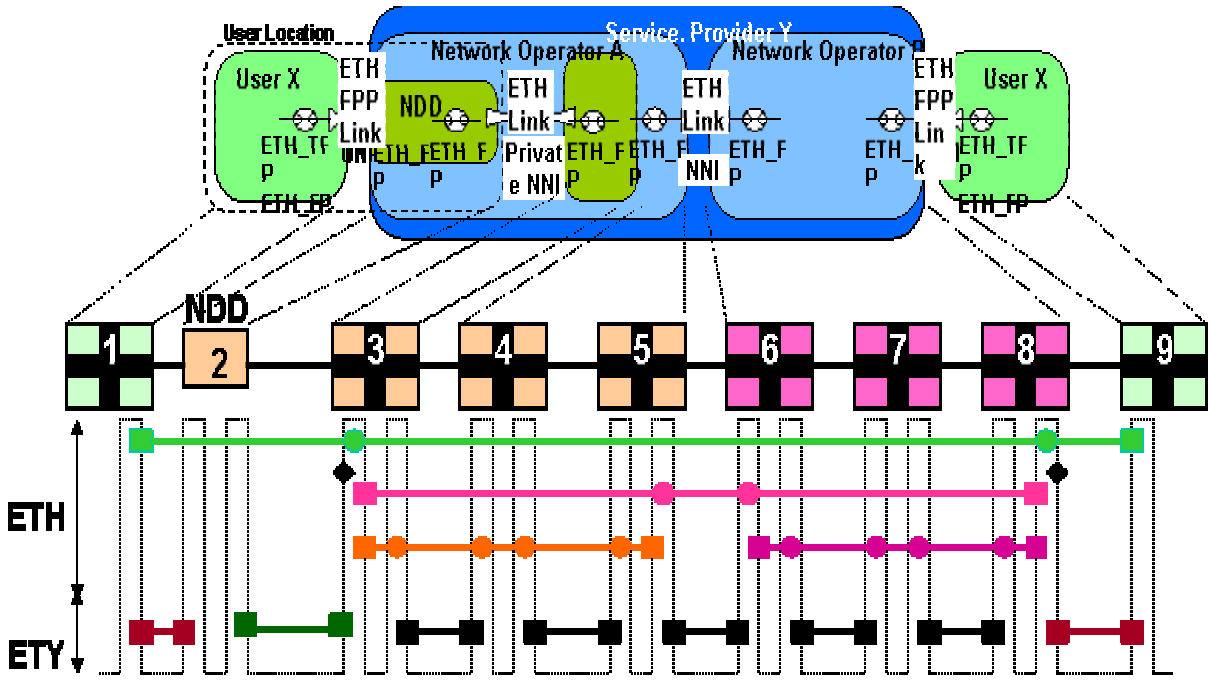


Figure 6.12: Illustrating the location of MEPs, MIPs in network model for the case of p2p connection services with Network Demarcation Device (NDD) as a NT provider device

- A point-to-point ETH connection between customer equipment 1 and 9 is supported by service provider Y. The NDD device here functions as a Media Converter (MC) where the MC realized a Network Termination (NT) device while the Line Termination (LT) functionality is integrated in the edge of network operator (B3).
- Light green indicates a UNI-C to UNI-C ETH ME (customer) with MEPs in the interface ports facing the network in CE1 and CE9, and MIPs in the network interface ports facing the CEs (B3 and B8).
- Dark green indicates at ETY layer indicates the Extension Link ETY ME. Brown represents an access link ME. This scenario requires some stitching between access link ME and Extension link ETY ME.
- Magenta indicates a UNI-NP to UNI-N ETH ME (metro network) with MEPs at the edge of the metro network (B3 and B8) and MIPs at each of the other interface ports in between.
- The UNI-N to UNI-N ETH ME (now shown in the figure) is realized by combination of Extension Link ETY ME and UNI-NP to UNI-N ETH ME.
- Orange and pink indicate intra-domain ETH MEs with MEPs at the edge of network operator A and B respectively (B3-B5 and B6-B8 respectively) and MIPs at each of the other interface ports in between.
- Black indicates ETH link MEs either realized as ETH MEs (sublayer monitoring) or as server layer MEs (inherent monitoring).

EDITOR'S NOTE: ANOTHER ACCESS SCENARIO WHERE SEPARATE MEDIA CONVERTERS (MC) WITH LT AND NT FUNCTIONALITIES ARE USED WILL BE INVISIBLE AT ETH LAYER. SUCH AN ARRANGEMENT MAY BE ADDRESSED IN

DRAFT RECOMMENDATION Y.17ETYOAM. THE DEPENDENCY OF SUCH AN ACCESS SCENARIO ON CURRENT RECOMMENDATION IS FFS.

CONTRIBUTIONS ARE INVITED, IF THERE IS INTERESTED IN THIS AREA.

6.4 ME Levels

ME Level defines the level at which an OAM flow, associated with an ME, operates. Since MEs can be nested, ME level is necessary to identify whether an OAM flow belongs to a level within the nesting. It is further useful to prevent the leak of OAM flows beyond the OAM Domains.
EDITOR'S NOTE: THE ABOVE TEXT HAS TO BE REWORDED. AS IT NOW , IT IS A CIRCULAR DEFINITION, I.E. USES THE SAME TERM TO DEFINE ITSELF
The number of ME levels needed are for further study.

Table 6-1 specifies an example of possible ME within the context of customer and provider domains, as mapped to Y.1730 and G.8010.

The mechanism used to assign the ME Levels is for further study. There are two options to assign ME Levels. The options are briefly described as follows:

- Absolute Assignment
- Flexible Assignment

NOTE: ONE OF THE TWO OPTIONS WILL BE CHOSEN TO DESCRIBE ME LEVELS

6.4.1 ME Level Absolute Assignment

For absolute assignment mechanism, each ME at a given level is pre-assigned a unique value. Since the exact number of possible levels is being studied, an example assignment is shown in Table 6.2

ME	ME Level	ME Level Indicator
UNI_C to UNI_C	6	255
UNI_N to UNI_N	5	253
Inter Domain ME	4	251
NNI ME	3	249
Access Link ME	2	247
Intra Domain ME	1	245

Table 6-2: Example of ME Level assignment

EDITORS' NOTE: ADD MORE EXAMPLES AND MOVE ALL OF THEM TO APPENDIX I

Table 6-2 highlights, as an example, how the ME levels can be assigned. It also indicates that some gaps in the absolute assignment can be left such that a finite set of MEs can be introduced in future if such a need arises. The finite set of MEs that can be introduced is determined by the gap between the absolute values. In Table 6.2, a gap of 1 is considered between MEs such that an additional level can be introduced between each two.

It may also be noted that the absolute values can be assigned either in the ascending order or descending order. However, this determination will be based on the filtering rules, which are for further study

6.4.2 ME Level Flexible Assignment

A possible solution for Flexible Assingment mechanism is a Stack-approach., where each MEP is associated with level 0 as it is shown in Figure 6-13

- When a MEP at ingress of a ME receives an OAM packets from outside the ME, it increments the level and passes the OAM packet along.
- When the MEP at egress of a ME receives an OAM packet from within the ME, it terminates the one with level 0 and passes other OAM packets by decrementing the level.

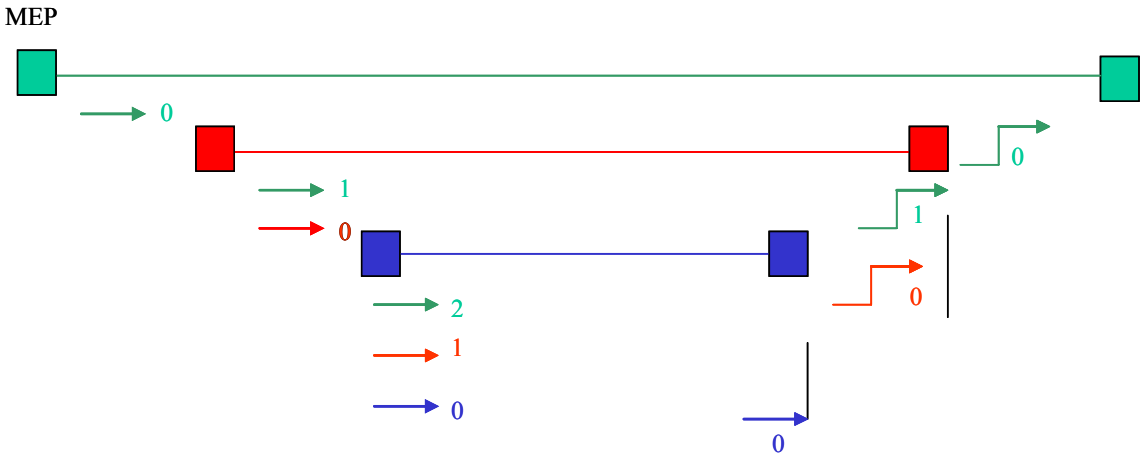
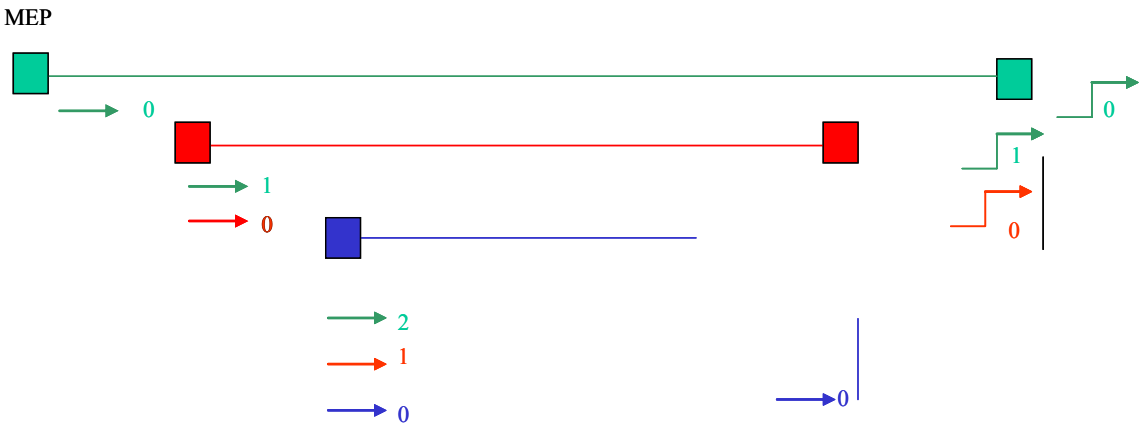


Figure 6-13: MEP Stack approach

- A limitation with this approach in that it can lead to OAM leaking when one MEP associated with a ME is faulty as it is shown in Figure 6-14:



- Limitation: This does not allow tapping at a certain level from within the network.

EDITOR'S NOTE: EXPLAIN THE MEANING OF THE ABOVE LIMITATION, IT IS NOT CLEAR

Figure 6-14: Limitation of the MEP stack approach

7 OAM functions for fault management

7.1 Continuity Check(CC)(keepalive)

Ethernet Continuity Check (CC) can be applied to detect connectivity and continuity faults across Maintenance Entities between a given pair of flow termination functions. It could also be used to detect the MAC addresses of ME end-points. Continuity failures could result due to hard or soft failures, with software failure, memory corruption, or mis-configurations being some soft failures. When used in context of a specific service instance, CC can be applied to detect connectivity failures across a given pair of flow termination functions that bound a ME. that support that common service instance.

EDITOR'S NOTE: NEED CLARIFICATION ON HIGHLIGHTED WORDS ABOVE AS IT IS CONFUSING NOW.

Although CC can be used to detect connectivity faults across a given pair of flow termination functions, i.e. any pair of flow points, it is particularly useful across a pair of edge flow points.

To detect connectivity failures with either a given set of flow point or all flow points meeting certain condition(s) within a boundary, CC OAM signal is generated and inserted in the ETH_FT_So and ETHS_FT_So functions. It is extracted and processed in the ETH_FT_Sk and ETHS_FT_Sk functions. CC is generated with either specific Unicast DAs or to a Multicast DA. Condition(s) could be that all edge flow points should receive this CC or all edge flow points participating in a service instance should receive this CC. Upon reception of the first CC from a particular flow point, the receiving flow point identifies continuity with sending flow point and expects to receive further periodic CCs. Once the receiving flow point stops receiving periodic CC from sending flow point, it detects that continuity to sending flow point is broken. Following detection of continuity failure, the detecting flow point may notify the operator, initiate fault verification followed by optional fault isolation step.

It may be noted that this mechanism has certain limitation in performing continuity failure detection. When a flow point starts participating in a network or within a network in a particular service instance for the first time, and if it has continuity failure with other flow point (s), the CC OAM frames will not reach those other flow point (s). Under such scenario, those other flow point (s) fail to detect continuity failures with this flow point. This scenario can be addressed by configuring for each flow point, a list of other flow points from which CC should be expected.

NOTE: CHANGE FLOW POINTS TO FUNCTIONS. ALSO MENTION THAT CONFIGURATION IS ALWAYS DONE AND THEN THIS MECHANISM CAN BE USED TO DETECT THE CONTINUITY …..REMARK THAT CONFIGURATION…..

EDITOR'S NOTE: NEED CLARIFICATION ON THE NOTE ABOVE When the flow termination function is present, CC will be present.

QUESTION: PERIODICITY OF CC CAN BE A CONFIGURABLE PARAMETER?

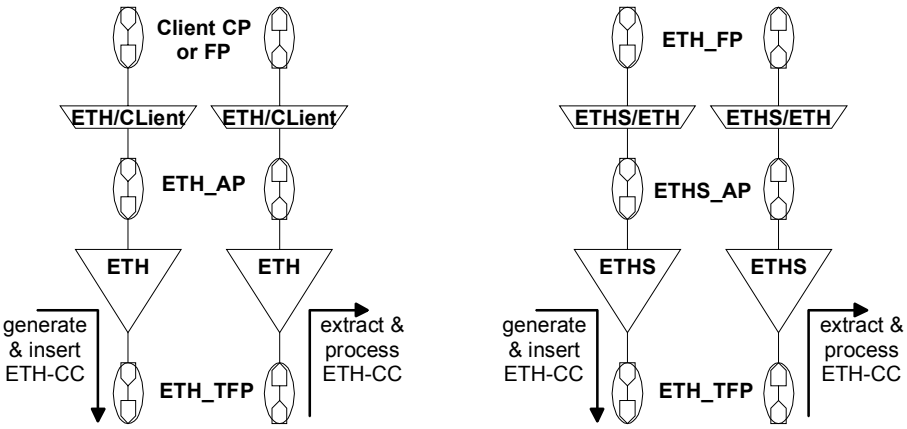


Figure 7.1-1 – Insertion/extraction & processing locations of ETH-CC OAM

In a mp connection with N endpoints there are N-1 ETH maintenance entities terminated by each ETHS_FT function. Each of these ETH maintenance entities is to be monitored for continuity and connectivity. An ETHS_FT_Sk function terminating those N-1 ETH maintenance entities should therefore expect to receive ETH-CC OAM from N-1 ETH_FT_So/ETHS_FT_So functions (Figure 7.1-2). If less than N-1 ETH-CC OAM frames are received the ETHS_FT_Sk should be able to state from which of the N-1 ETHS_FT_So functions it is not receiving the ETH-CC OAM frame(s). If it receives more than expected distinct id, then it can determine anomalies (about unexpected entities presence).

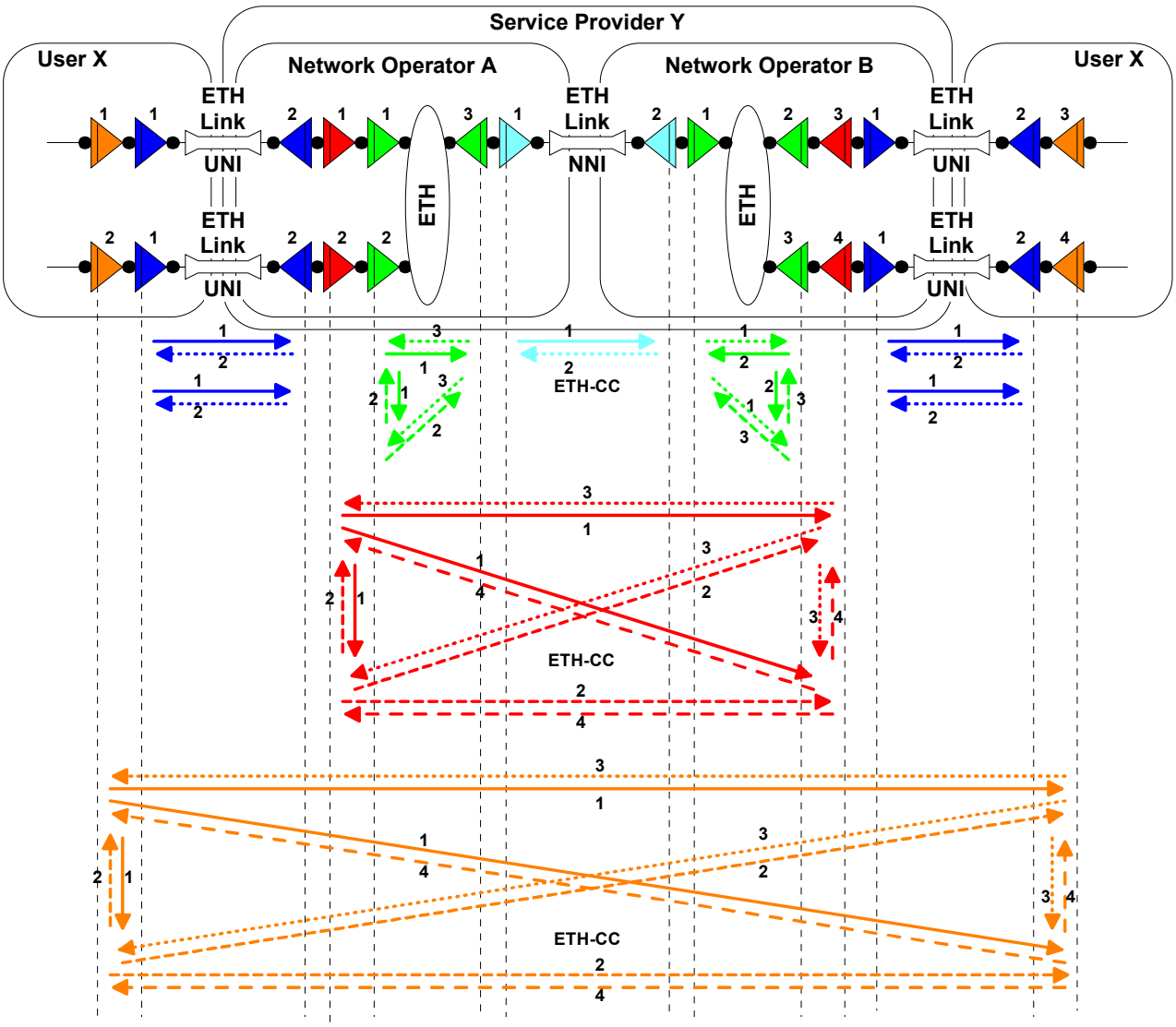


Figure 7.1-2 – ETH-CC in multi-operator ETH mp connection

NOTE: THIS DIAGRAM SHOULD MAKE CLEAR THAT CC MESSAGE STARTS WITH MULTICAST DA.

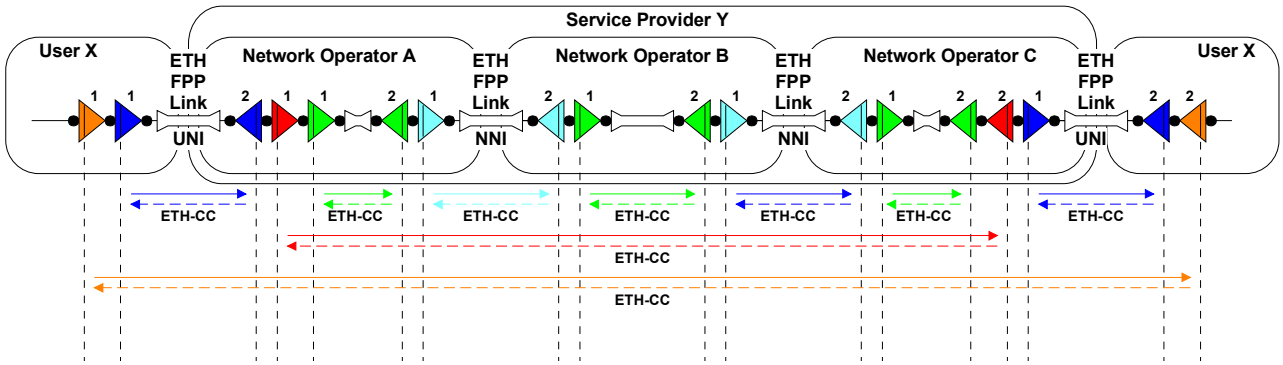


Figure 7.1-3 – ETH-CC in multi-operator ETH p2p connection

NOTE: Possible use is for loop detection using identifier information element.

EDITOR'S NOTES: CHECK THE TWO NOTES ABOVE AND DECIDE IF CONTRIBUTIONS ARE NECESSARY TO CLARIFY SO NOTES CAN BE DELETED

7.1.1 CC Transmission and Reception Process.

7.1.1.1. Transmission

Every active Maintenance Point sender function transmits a Continuity Check Message (CCM) as often as configured. Configured transmission intervals may range from 0.01 seconds to 655.35 seconds. The Lifetime TLV must be transmitted with a value of 2.5 times the transmission interval, so that another Maintenance Point can lose two Continuity Check Messages without declaring an error.

NOTE: IT IS TBD WHAT CONSTITUTES AN "ACTIVE" MAINTENANCE POINT. THE CONCEPT SHOULD ALLOW FOR A PORT THAT IS PHYSICALLY DOWN, OR FOR A PORT THAT IS BEING BROUGHT UP AND CHECKED OUT BEFORE/AFTER BEING PLACED IN SERVICE.

7.1.1.2 Reception

Every active Maintenance Point receives and catalogues Continuity Check Messages (CCM). Every CCM is examined to be sure that its Service Instance Identifier TLV matches that configured in the receiving Maintenance Point, and that its Maintenance Point Identifier TLV does not match that of the receiving Maintenance Point. The information in the Continuity Check Message is catalogued in the receiving Maintenance Point, indexed by the received MPID. Information saved includes the Lifetime TLV, so that the information can be timed out, the source MAC address and data path service identifier of the received Continuity Check Message, and the Bridge Port on which it was received. If the Lifetime TLV is 0, the catalogued information for the received MPID, if any, is discarded (or perhaps remembered as no longer active).

NOTE: THE USE OF LIFETIME TLV IS FOR FURTHER STUDY

When a Continuity Check Message is received at either a MEP or a MIP, the source MAC address, S-VLAN, ME Level, and ingress Bridge Port are recorded, indexed by MAC address, S-VLAN, and ME Level, in the Provider Bridge's CC Database.

NOTE: THE PROCESSING OF CC AT THE MIPs AND THE USE OF CC DATABASE IS FOR FURTHER STUDY. CONTRIBUTIONS ARE INVITED

7.2 Loopback

A Loopback function can be of two types:

- Intrusive Loopback
- Non-intrusive Loopback

7.2.1 Intrusive Loopback

Intrusive Loopback is used to place a remote flow point in a continuous Loopback such that all received frames would be looped back except OAM frames. Since this function results in Loopback

of data frames, the data path is impacted; it is therefore considered as Intrusive Loopback. Given the nature of this function, it is expected to be always point-to-point. Intrusive Loopback OAM frames, requesting start or termination of Loopback, are expected to be Unicast (with DA = address of remote network element). Moreover, the applicability of Intrusive Loopback is expected to be limited to EPL (Ethernet Private Line) service. This function is intended for out-of-service testing.

EDITOR'S NOTE: INPUT NEEDED HERE TO EXPLAIN THAT INTRUSIVE LOOPBACK IS A USEFUL TOOL BUT IT CAN BE DANGEROUS, IT SHOULD ONLY BE USED OUT-OF-SERVICE: INPUT EXPECTED FROM SHEZHAD (BT)

7.2.2 Non-Intrusive Loopback

Non-intrusive Loopback is used mainly to verify continuity with remote flow point (s). Non-intrusive Loopback is performed by sending OAM frames to remote flow point (s) and expecting a response back which verifies continuity and connectivity. Since the data frames are not looped back, and the data path is not impacted; this Loopback is considered as non-intrusive. As a result, this function can be used for in-service testing.

Though a Non-intrusive Loopback may be initiated at any time, it is particularly useful when verifying continuity or connectivity once a failure is detected. Non-intrusive Loopback request may be generated either:

- automatically following detection of continuity failure, where detection could be done using connectivity check (CC) function mentioned in 7.1, or
- On-demand via an operator initiated command, or
- Periodically.

Non-intrusive Loopback may be used for fault detection when used on a periodic basis. However, unlike CC mentioned in Section 7.1, Non-intrusive Loopback requires a response for each request. Response generation and response's handling by requestor require more processing in Non-intrusive Loopback as compared to CC. While a CC is suitable for detecting unidirectional connectivity failures, Non-intrusive Loopback can be used to detect bidirectional connectivity failures with single-ended maintenance entity.

A Non-intrusive Loopback can be of two types:

- Unicast Non-intrusive Loopback
- Multicast Non-intrusive Loopback

7.2.2.1 Unicast Non-intrusive Loopback

Unicast Non-intrusive Loopback request OAM frame is sent to a particular flow point (with DA = Unicast MAC address of destination network element). Upon reception of this request OAM frame, the destination flow point responds back with Non-intrusive Loopback response OAM frame (with DA = Unicast MAC address of requesting network element, learnt from request OAM frame). Other flow points that receive this request and/or response OAM frame forward these without processing.

Non-intrusive Loopback OAM signal is generated and inserted in the ETH_FT_So and ETHS_FT_So functions. It is extracted and processed in the ETH_FT_Sk and ETHS_FT_Sk functions. Refer to Figure 7.2-1.

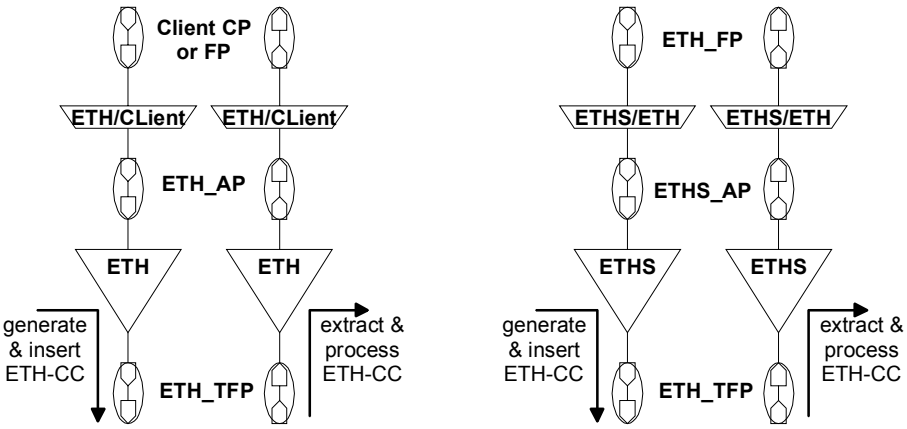


Figure 7.2-1 – Insertion/extraction & processing locations of Non-intrusive Loopback OAM

n

EDITORS' NOTE(REPLACE ETH-CC WITH NON-INTRUSIVE LOOPBACK IN FIGURES :CHECK WITH MAARTEN

EDITOR'S NOTE: NEED TO DECIDE IF REPLACE mp CONNECTION WITH MULTIPOINT CONNECTION??? OR PUT IN LIST OF ABBREVIATIONS BUT NEED CONSISTENCY IN TEXT

In an mp connection with N endpoints there are N-1 ETH maintenance entities terminated by an ETHS_FT function. Each of these ETH maintenance entities can be verified for continuity failures when continuity failures are detected using CC. An ETHS_FT_Sk function terminating those N-1 ETH maintenance entities can therefore expect to receive Non-intrusive Loopback OAM from N-1 ETH_FT_So functions (Figure 7.2-2). For Unicast Non-intrusive Loopback, ETHS_FT_Sk should receive OAM flow addressed to itself.

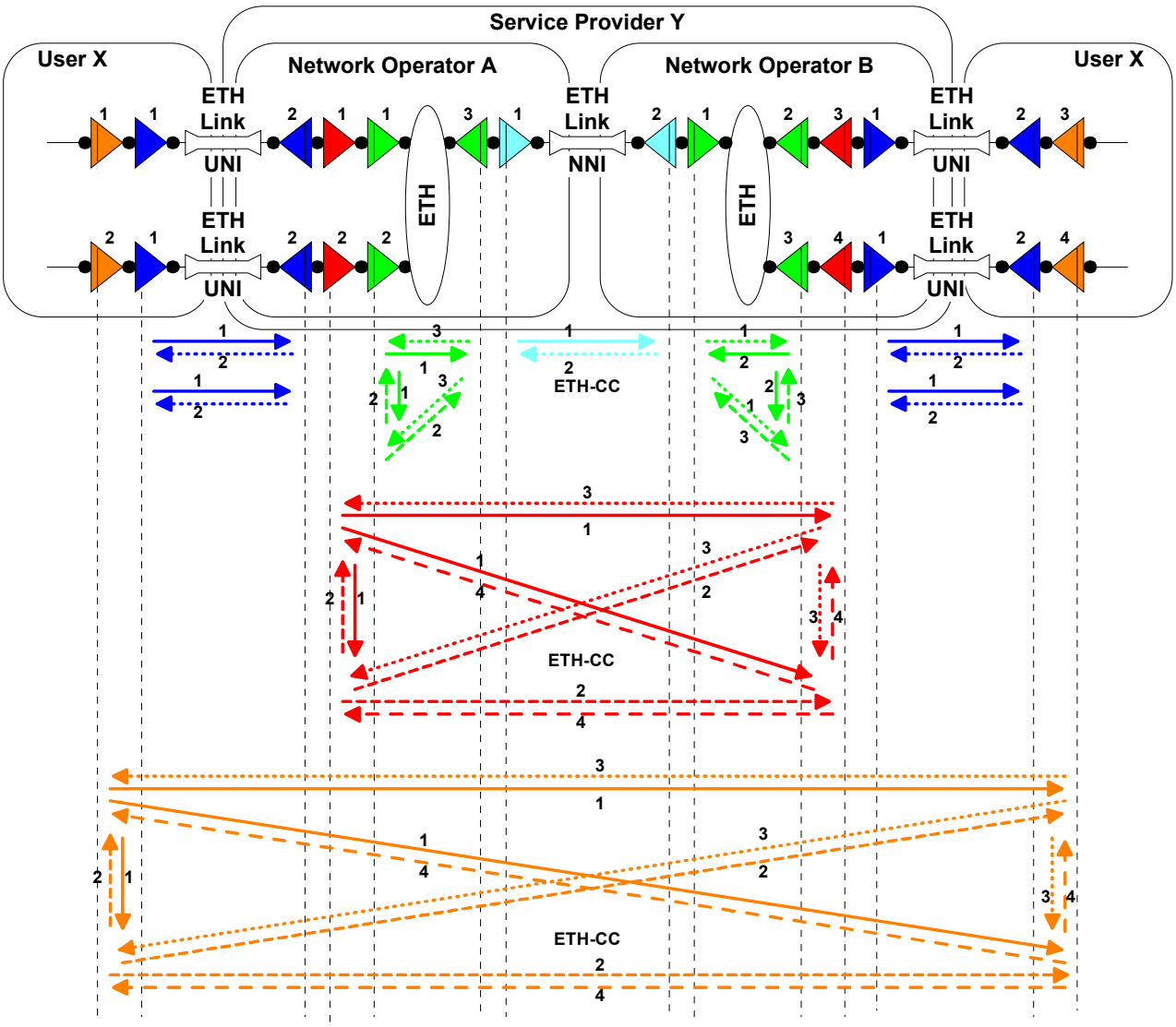


Figure 7.2-2 – Non-intrusive Loopback in multi-operator ETH mp connection

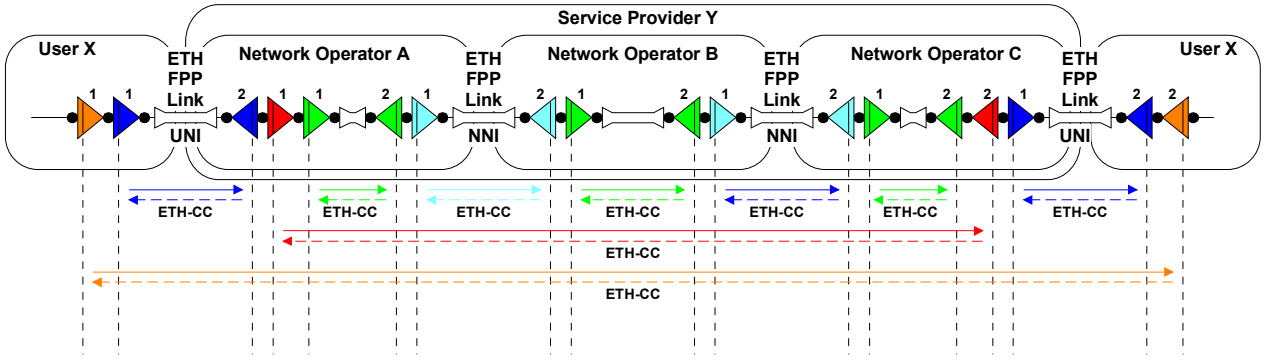


Figure 7.2-3 Non-intrusive Loopback in multi-operator ETH p2p connection

(Replace ETH-CC with Non-intrusive Loopback in the figure).

7.2.2.2 Non-intrusive Loopback Transmission and Reception Procedures

7.2.2.2.1Transmission

Loopback Messages are transmitted by operator command via the CLI or the Maintenance Point Control MIB. note: align with 7.2.2) The Transaction identifier transmitted is retained for at least 5 seconds after the Loopback Message is transmitted. The Transaction Identifier must be changed for every Loopback Message, and no Transaction Identifier from the same MEP may be repeated within one minute.

EDITOR'S NOTE: CHECK THE NOTE IN ABOVE PARAGRAPH, WHAT IS GOING TO BE ALIGNED WITH 7.2.2 AND WHO IS PROVIDING TEXT FOR THAT

7.2.2.2.2 Loopback Reception and Reply Transmission

Whenever a valid Loopback Message is received by a MIP or MEP receive function, Loopback Reply is generated and transmitted to the originating Maintenance Point. Every field in the Loopback Message is copied to the Loopback Reply with the following exceptions:

1. The source and destination MAC addresses are swapped.
2. The OpCode field is changed from Loopback Message to Loopback Reply.
3. The Checksum TLV is recalculated to reflect any changes to the message, such as the the OpCode field.
4. It is To Be Determined whether any other fields in the Loopback Message are to be interpreted by the receiver.

7.2.2.2.3 Loopback Reply Reception

When a Loopback Reply is received by a MIP receive function, or if the received Transaction ID is not in the list of transmitted Transaction Ids maintained by the MEP, the message is invalid. The MEP receive function may examine the TLVs returned in the Loopback Reply, and declare the message invalid if they do not match those sent in the corresponding Loopback Message.

7.2.2.3 Multicast Non-intrusive Loopback

Multicast Non-intrusive Loopback request OAM frame is sent to all functions supporting Non-intrusive Loopback meeting certain condition(s) within a boundary (with DA = Multicast DA). Condition(s) could be that all edge functions should receive this request OAM frame or that all edge functions participating in a service instance should receive this request OAM frame. Upon reception of this request OAM frame, the receiving function (s) that satisfy the above mentioned condition(s) respond back with a Unicast Non-intrusive Loopback response OAM frame (with DA = Unicast MAC address of requesting network element, learnt from request OAM frame). Other functions that do not meet these conditions receive this request and/or response OAM frame and forward without processing.

NOTE: THE APPLICATION OF MULTICAST NON-INTRUSIVE LOOPBACK CAN BE DANGEROUS WHILE USE IN SERVICE.

EDITOR'S NOTE: SOME TEXT IS EXPECTED TO EXPAND ON THE ABOVE NOTE

EDITOR'S NOTE: FOLLOWING THREE SUBSECTIONS ARE PROPOSED TO BE DELETED

7. 3 Link Trace (ETH-LT)

The main objectives of an Ethernet Link Trace (ETH-LT) are:

- Adjacent Relation Retrieval
- Fault Localization

7.3.1 Link Trace for Adjacent relation retrieval

EDITORS NOTE: TITLE SHOULD BE REVISED

ETH-LH can be used to identify adjacent MIP and/or MEP. In order to find the relationships, the identification of MIP and/or MEP such as MEP/MIP ID or MAC address, is required. In addition, the sequence of MIPs and/or MEPs in the tested links should be identified

7.3.2 Link Trace for Fault localization

ETH-LT can be applied for fault localization. When a fault (eg.in a link and/or bridge, etc) and/or a loop occurs, the sequence of MIPs and MEPs in the tested links will be changed from the expected one. The difference of both sequences provides information of the faulty element.

EDITORS NOTE: USAGE OF THE KIND OF INFORMATION THAT THE LINK TRACE GIVE FOR FAULT LOCALIZATION NEEDS TO BE CLARIFIED.

7.3.3 Link Trace Operation

7.3.3.1 Link Trace Origination

Link Trace Messages are transmitted by operator command via the Command Line Interface (CLI) or the Maintenance Point Control MIB. The Transaction Identifier of each Link Trace Message transmitted is retained for at least 5 seconds after the Link Trace Message is transmitted.

7.3.3.2 Link Trace Reception, Forwarding, and Replying

If a Link Trace Message is received by a Provider Bridge, and if the data frame targeted by the Link Trace Message would pass through a MEP or MIP on ingress or egress through the bridge, then the bridge must:

1. Examine the Link Trace Message's TTL TLV, and if 0, discard the Link Trace Message; else
2. Determine the information required to generate a Link Trace Reply;
3. If the data frame targeted by the Link Trace Message would pass through the Provider Bridge and out a single Egress Bridge Port, and if the Link Trace Message's TTL TLV was greater than 1 when received, then the Link Trace Message must be transmitted on the selected Egress Port. All fields and TLVs are transmitted exactly as received, except for the source MAC address and TTL TLV.
4. After a random time interval in the range 0-1 second, transmit a Link Trace Reply to the originating Maintenance Point.

If the data frame targeted by the Link Trace Message would not pass through a Maintenance Point or Loopback Point on ingress or egress through the Provider Bridge, then the bridge must pass the Link Trace Message through as normal data.

NOTE: TLVs for Link Trace Function are FFS

Figure 7.3-1 and Table 7.3-1 below show the packet handling and the packet information table by hop respectively

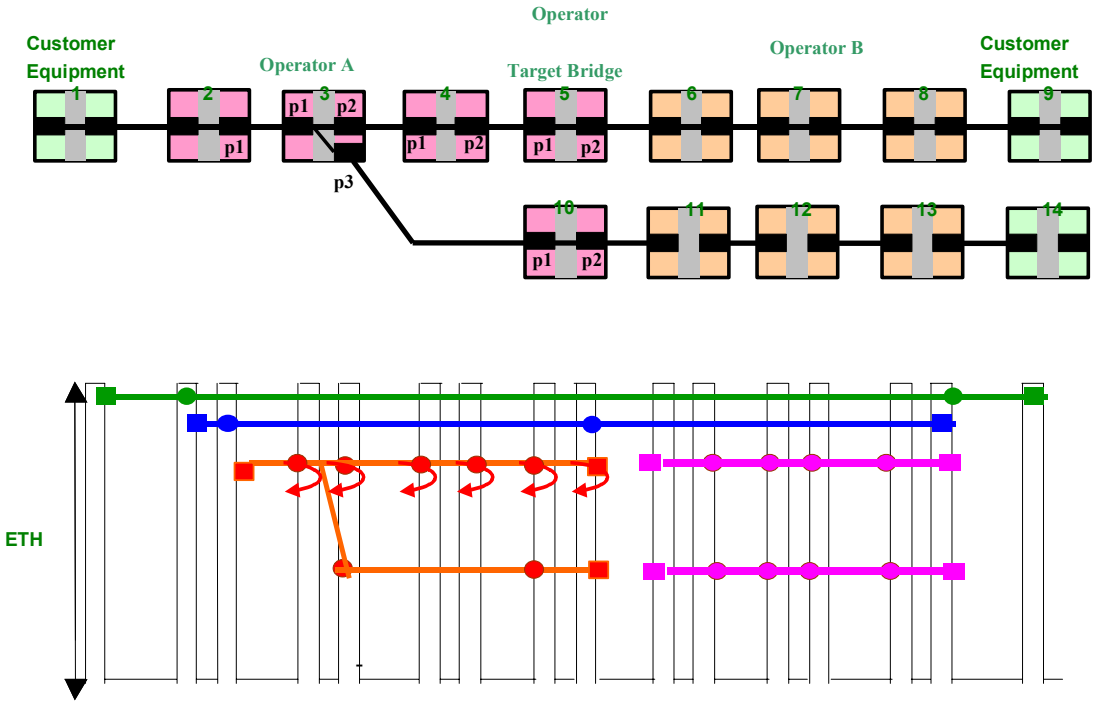


Figure 7.3-1: ETH-LT packet handling

Number of hops	Sent ME ID	Received ME ID
1	2/p1	3/p1
2	3/p1	3/p2
3	3/p2	4/p1
4	4/p1	4/p2
5	4/p2	5/p1
6	5/p1	5/p2

Table 7.3-1: Sorted ETH-LT packet information table by number of hops

7.4 ETH-AIS

ETH layer Alarm Indication Signal (ETH-AIS) can be used to notify higher layers about faults detected at lower layers such that the ETH-AIS can be used to suppress declaration of same fault at higher layers. This allows the fault to be reported to OSS (Operations Support Systems) or NMS

(Network Management Systems) by a single layer (at which the fault occurs and is detected) and not by all other higher layers.

Note: The current version describes applicability of ETH-AIS for point to point services offered across infrastructure where automatic reconfiguration mechanisms like STP are not used. Appendix III highlights some scenarios and issues associated with the multipoint services including when a service has only 2 endpoints.

EDITORS NOTE: MAARTEN WILL PROVIDE ATOMIC MODEL FIGURE FOR THIS SECTION AND MAARTEN AND DINESH NEED TO PROVIDE MATERIAL TO UPDATE APPENDIX III

Figure 7.4-1 shows, as an example, how a fault at the ETY layer, i.e. at Ethernet link between Operator A bridges 3 and 4, can be notified via ETH-AIS to higher level maintenance entities.

.

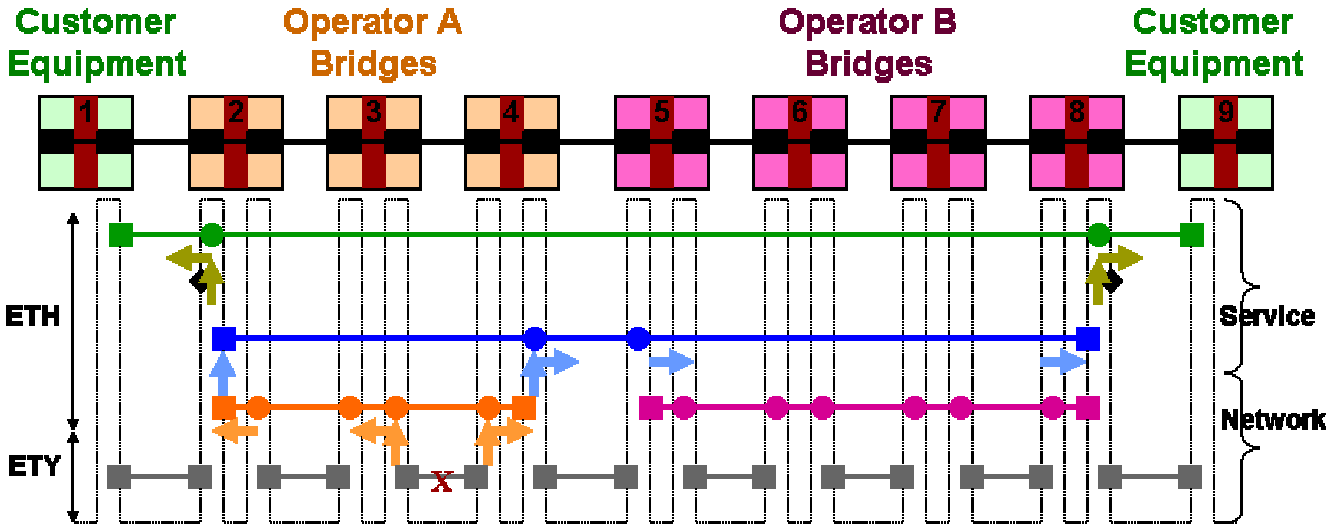


Figure 7.4-1: ETH-AIS mapped to MEs, MEPs and MIPs

- Black indicates ETH link MEs either realized as ETH MEs (sublayer monitoring) or as server layer MEs (inherent monitoring).
- Upon detection of this fault, the MEPs corresponding to failed link generate ETH_AIS (represented by orange arrows) which are forwarded by orange MIPs towards orange MEPs corresponding to orange ME.
- Upon receiving ETH_AIS, the orange MEPs generate higher level ETH_AIS (represented by blue arrows) which are forwarded by blue MIPs towards blue MEPs corresponding to blue ME.
- ETH_AIS promoted to blue ME remains transparent to the purple ME, where the purple ME is at a lower level compared to blue ME. In Figure 7.4-1, purple ME is shown as the same level as the orange ME.
- Similarly, upon receiving ETH_AIS, the blue MEPs generate higher level ETH_AIS (represented by green arrows), which are forwarded by green MIPs towards green MEPs corresponding to green ME.

It may be noted that in Figure 7.4-1, the green and blue MEs correspond to service level MEs while orange and black MEs correspond to network and/or facility level MEs. Therefore, it is conceivable that a network level failure could trigger ETH_AIS along the service level ME.

The following figures illustrate the ETH-AIS insertion and termination in a p2p connection for different fault locations e.g. UNI, operator A domain, inter-operator NNI, operator B domain, access link and/or extension link in NDD access scenario.

NOTE: FIGURES 7.4-2 TO 7.4-8 INDICATE A UNIDIRECTIONAL FLOW OF ETH-AIS, HOWEVER, IT MAY BE CHANGED

A Server layer or ETH sublayer MEP Sink function that detects a signal fail condition will insert ETH-AIS in its Srv/ETH_A_Sk[1] or ETHS/ETH_A_Sk function. A ETH sublayer MEP Sink function that detects ETH-AIS at its ME level will terminate the signal in its ETHS_FT_Sk function, detects dAIS, declares a signal fail condition and inserts in its ETHS/ETH_A_Sk function ETH-AIS (at the higher level).

The termination and re-generation of ETH-AIS within an ETH sublayer MEP Sink function provides some security by preventing internal MEP MAC addresses to be exposed outside a ME domain.

Note that a MIP function is transparent to ETH-AIS.

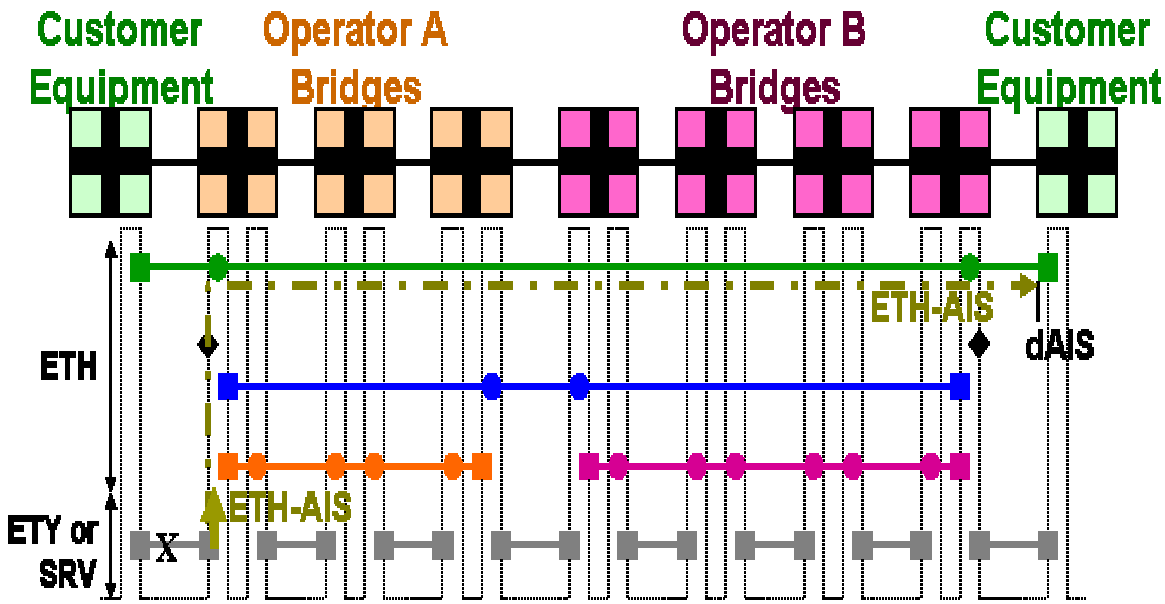


Figure 7.4-2 – ETH-AIS on p2p connection (failure on UNI)

NOTE - ETH-AIS incoming to a ME is marked as incoming and not processed within that ME domain.

At the egress, the marking is taken out. The method to mark incoming ETH-AIS depends on the ME level indication method that is selected. A stack pointer (relative level encoding) method will cause the pointer to be incremented/decremented. A MAC address (absolute level) encoding method will have the applicable level hard coded in the MAC address, and doesn't require further marking at ME in-/egress points.

---

[1] This Srv/ETH_A_Sk function will be part of a server layer's MEP.

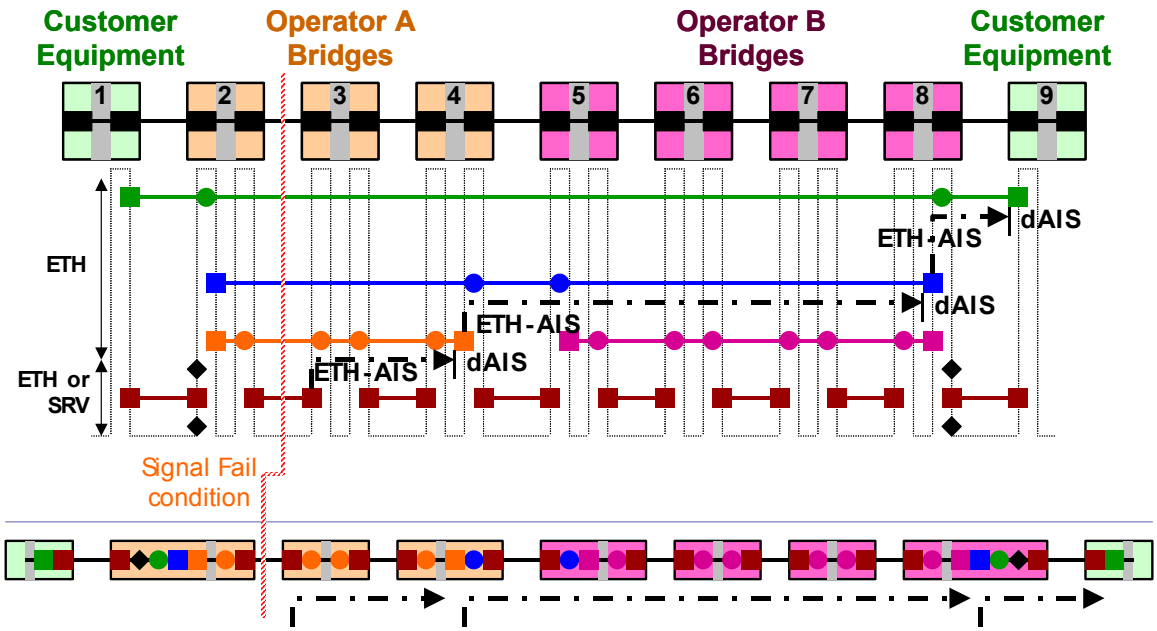


Figure 7.4-3 – ETH-AIS on p2p connection (failure in operator A domain)

EDITOR'S NOTE: FIGURE 7.4-3 WILL BE MADE CONSISTENT WITH FIGURE 7.4-2, DINESH WILL PROVIDE THE FIGURE

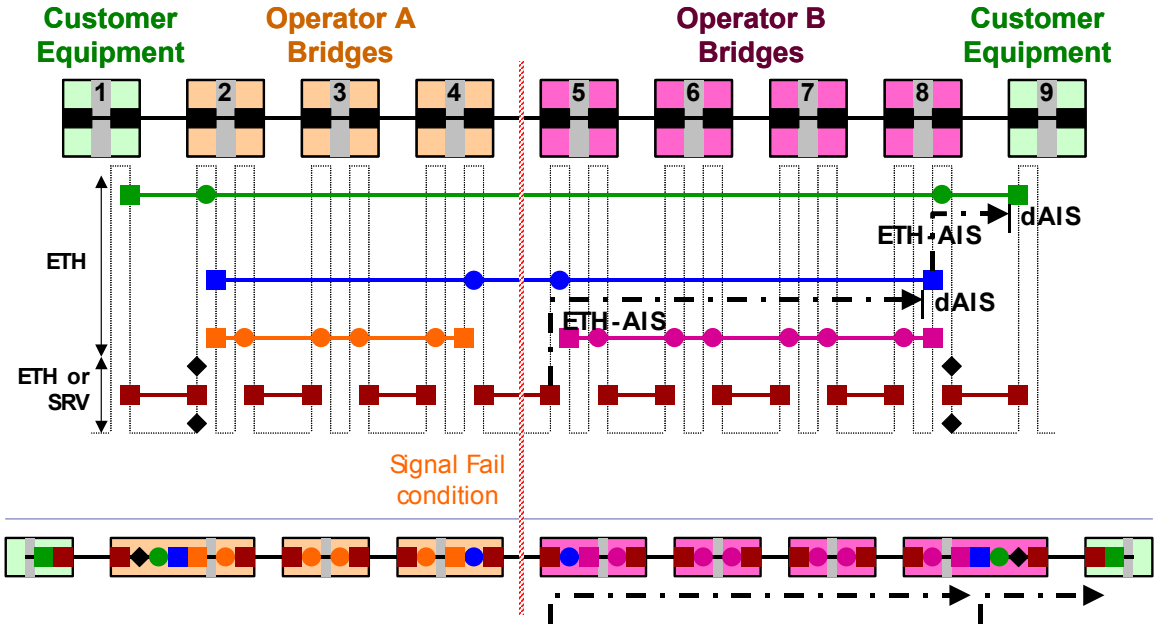


Figure 7.4-4 – ETH-AIS on p2p connection (failure on inter-provider NNI)

EDITOR'S NOTE: FIGURE 7.4-4 WILL BE MADE CONSISTENT WITH FIGURE 7.4-2

Interface ports with two or more MEP functions active will functionally terminate and re-generate ETH-AIS in each of the MEP Sink functions; see figure 4 **(EDITOR'S NOTE: CHECK THIS FIGURE NUMBER)**

The termination and re-generation of ETH-AIS may increase the recovery time of the higher level MEs after the fault is repaired. Care should be taken with its processing definitions.

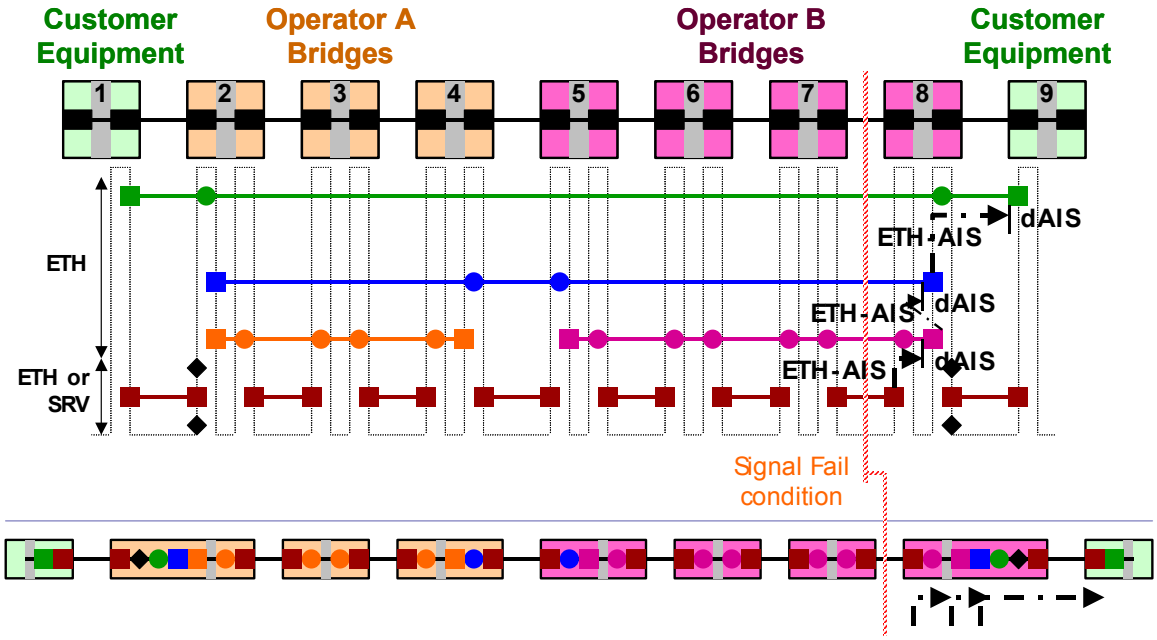


Figure 7.4-5 – ETH-AIS on p2p connection (failure in operator B domain)

EDITOR'S NOTE: FIGURE 7.4-5 WILL BE MADE CONSISTENT WITH FIGURE 7.4-2

When a FAILURE condition is detected in the final customer equipment's ingress port, the Server layer's MEP sink function will insert ETH-AIS, which will be terminated immediately in the next ETH layer MEP Sink function. If it is an ETHS_FT_Sk function then this function will also re-insert ETH-AIS to be forwarded through the customer domain towards the ETH flow termination. If it is an ETH_FT_Sk function (inside the LLC), then this MEP sink function will insert (if defined) client layer AIS.

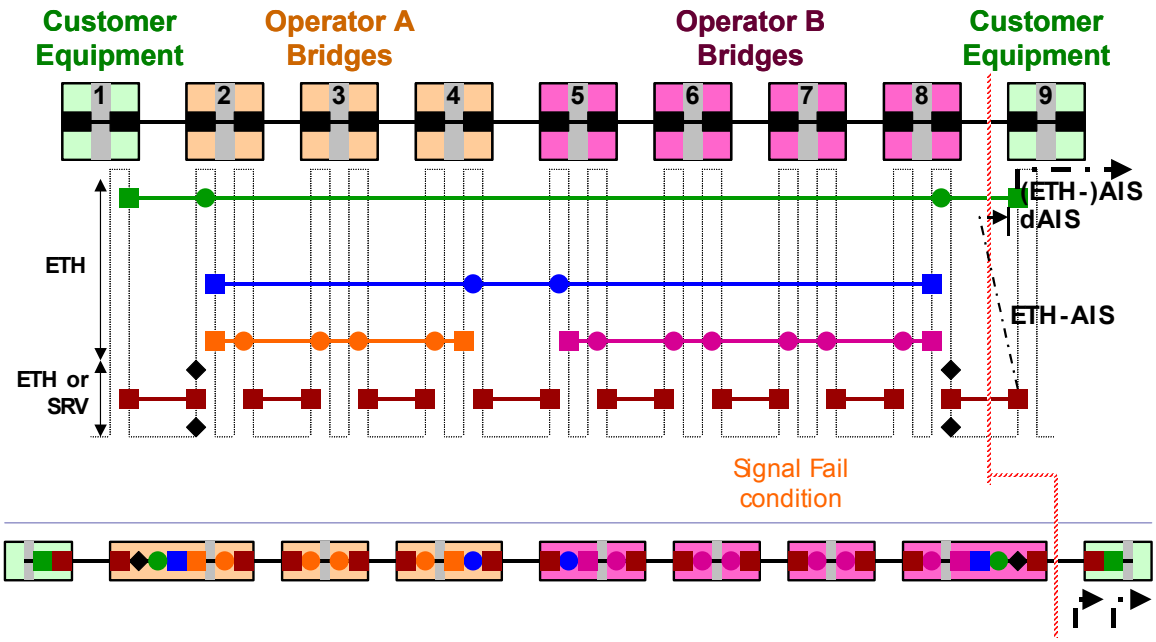


Figure 7.4-6 – ETH-AIS on p2p connections (failure on UNI)

EDITOR'S NOTE: FIGURE 7.4-6 WILL BE MADE CONSISTENT WITH FIGURE 7.4-2

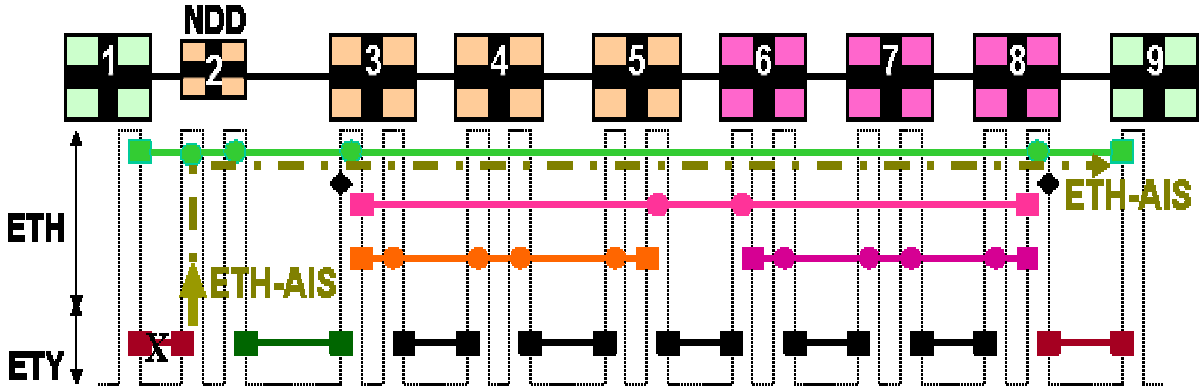


Figure 7.4-7 – ETH-AIS on p2p connections (failure on access link in NDD)

EDITOR'S NOTE: TEXT NEEDS TO BE ADDED FOR FIGURE 7.4-7. CONTRIBUTIONS ARE INVITED.

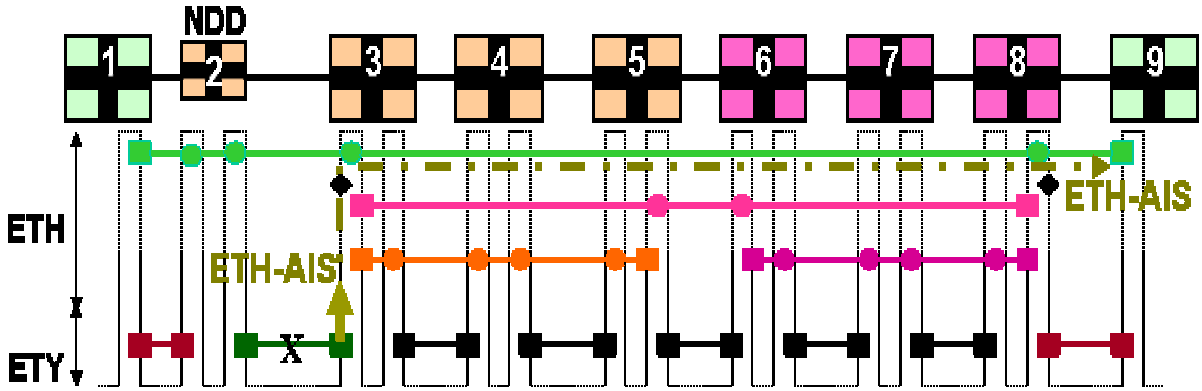


Figure 7.4-8 – ETH-AIS on p2p connections (failure on extension link in NDD)

EDITOR'S NOTE: TEXT NEEDS TO BE ADDED FOR FIGURE 7.4-8. CONTRIBUTIONS ARE INVITED.

7.5 ETH-RDI

The application of ETH layer Remote Defect Indication (ETH-RDI) is for further study.

Note: ETH layer is dependent upon an operational ETH link, where both transmit and receive directions are up. When either transmit or receive direction is down at a port of an ETH link, entire port and associated link is marked as operationally down. Determination of directionality in such case may be infeasible.

However, ETH-RDI may still be applicable for performance management and/or uni-directional network level failures. Another possible application is to differentiate between administrative shutdown and failure shutdown.

Application of ETH-RDI is therefore FFS

7.6 Test Signal Generation/Detection function

The test signal generation function generates test signal with specified throughput (bandwidth), frame size and frame transmission pattern. The detection function detects throughput (bandwidth), frame loss, frame disorders, bit errors, delay and delay variation.

7.6.1 Maintenance scenarios

This section shows some examples of maintenance scenarios for point-to-point, out-of-service case. NOTE: Multipoint application is FFS.

(a) Unidirectional measurement

An edge bridge generates a test signal and another edge bridge receives the test signal and measures the performance between these two edge bridges (Fig. 7.6-1).

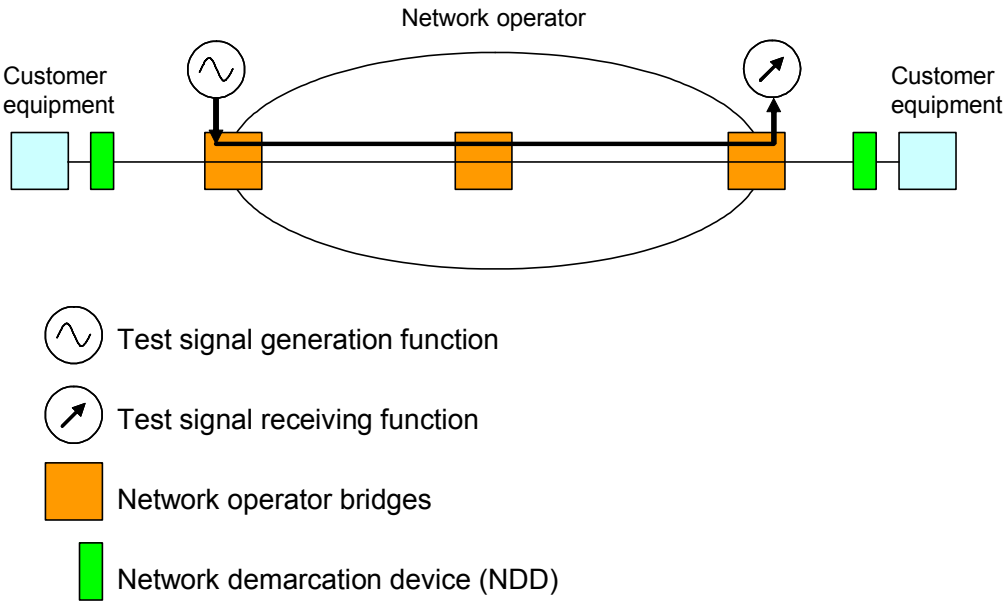


Figure 7.6-1 Unidirectional test.

(b) Bidirectional measurement

A core bridge or an edge bridge generates a test signal. Another edge bridge or a network demarcation device (NDD) is put into an intrusive loopback mode. The bridge generating the test signal sends the test signal towards the bridge in the intrusive loopback mode and receives the loopbacked test signal. Bidirectional (round trip) performance is measured with this (Figs. 7.6-2 and 7.6-3).

NOTE: REALIZING INTRUSIVE LOOPBACK AT HIGH BIT RATE MIGHT CAUSE EXCESSIVE COMPLEXITY.

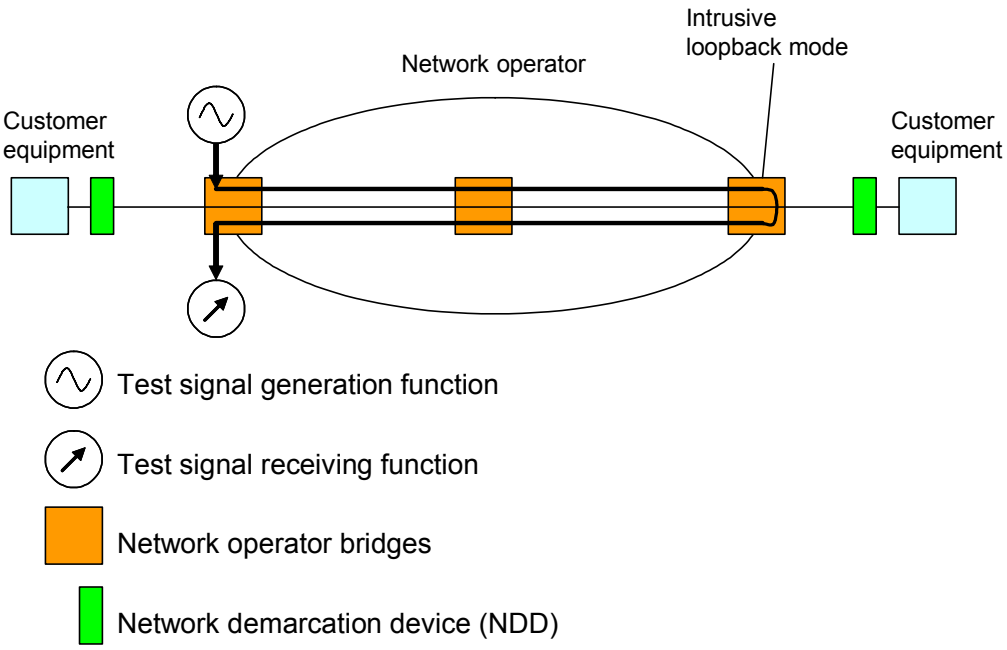


Figure 7.6-2 Bidirectional test (1).

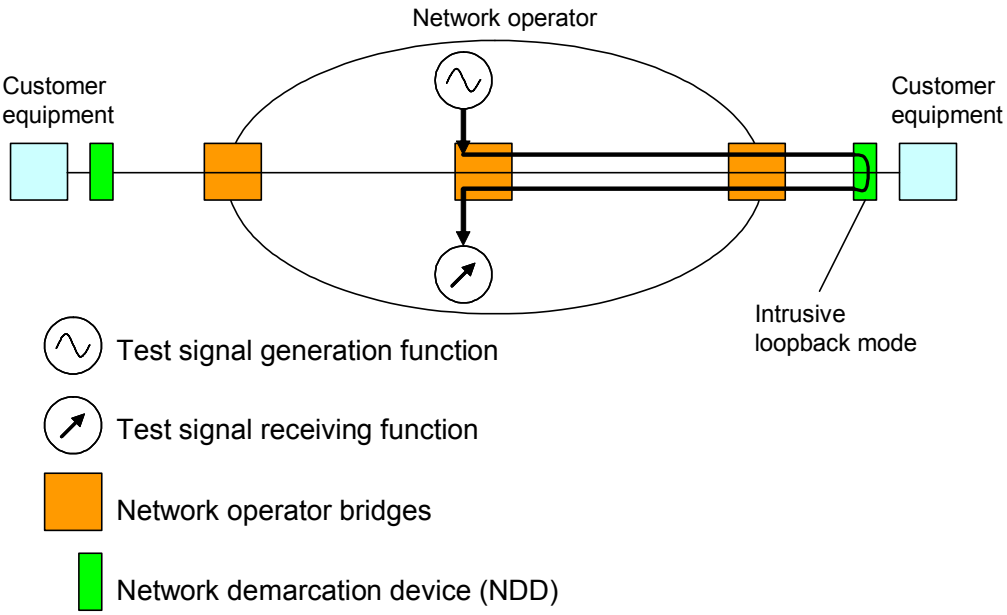


Figure 7.6-3 Bidirectional test (2).

8 OAM functions for performance management

8.1 Performance Parameters

Current discussions in Metro Ethernet Forum on performance parameters for Ethernet networks and services have focused on the following parameters, which are captured in an MEF working draft (January 2003)

- **Frame Loss (FL)**
  Difference between the number of service frames sent to ingress UNI and the number of service frames received at egress UNI. This is applied to Ethernet Virtual Connection (EVC) which corresponds to UNI_N to UNI_N ME.
- **Frame Delay (FD)**
  Frame delay can be specified in terms of round-trip delay, which is defined as the time elapsed since start of the transmission of the first bit of a frame by the source node until the reception of the last bit of the loop backed frame by the same source node, when the loop back is performed by the frame's destination node.
- **Frame Delay Variation (FDV)**
  Measure of the variations in the frame arrival pattern belonging to the same CoS instance compared to the arrival pattern at the ingress of the MEN.
- **Availability**
  Function of time the ME (associating service UNIs) is in available state. It is specified as a ratio of:
  Total Time ME is in Available State / Total Service Time

  where, **Total Service Time** is viewed as number of time intervals and **Available State** is viewed as interval when service meets FL, FD and FDV bounds. Unavailable state is encountered when at least one of the FL, FD or FDV measures exceed their bounds/thresholds during a time interval. These

bounds/thresholds are determined by the class of service (CoS).

Note 1:
The definition of Availability should be aligned with Y.1711 and/or Y.MPLSperf . The details of Availability are expected to be defined in a separate Recommendation being developed by Q.6/13 – Y.17EthPerf.

Note 2:

For sub rate or virtual services, the frame loss can be associated with both in-profile and out-of-profile service frames.

Additional performance parameters that may be taken into consideration include:

- **Errored Frame Seconds**
  Indicates if an error (e.g., frame error due to FCS or 8B/10B coding violation) has occurred within the second. This does not take into consideration errors when frames are received error free but are not delivered.
- **Service Status**
  Indicates if the service is in-service or out-of-service. In-service or out-of-service state can be based on **Available state** defined earlier.
- **Frame Throughput**
  Number of frames and/or bytes transmitted to network interface relative to CIR
- **Frame Tx**
  Number of frames transmitted out of the customer facing interface within the (previous) time interval (e.g. 1 second).
- **Frame Rx**
  Number of frames received from the customer facing interface within the (previous) time interval (e.g. 1 second).
- **Frame Drop**
  Number of frames dropped at the customer facing interface within the (previous) time interval (e.g. 1 second).
- **Loopback Status**
  Indicates whether the customer facing interface is in an intrusive Loopback state (potentially due to OAM interactions across Access Link ME).
- **Client Signal Fail**
  Indicates state of Access Link ME.
- **Unavailable Time**
  Number of time intervals (e.g. 1 second) when the service status is unavailable.

8.2 Measurement Mechanisms

Different measurement mechanisms are possible to perform performance measurements. One significant difference across these mechanisms is the level of accuracy of measurements. These mechanisms include:

- **Management plane statistical methods**
  Statistical methods use OAM frames to estimate data path behavior. Such methods are least accurate since they apply approximation to emulate data frames.
  The limitation lies in that the behavior of actual data frames may be quite different due to different addressing, processing, transient congestion conditions etc. Also, error conditions in networks typically happens in bursts thus statistical methods can likely miss those bursts and represent different results.

- **Management plane managed objects**
  Here OAM frames use data path managed objects to calculate performance parameters and are inserted and/or extracted via management plane. These methods are fairly accurate since they use data path statistics to measure data path performance.
  Their limitation lies in that since the insertion and extraction of these OAM frames is done via management plane, in-flight frames need to be accounted for. On the egress side of OAM frame, in-flight frames refers to data frames between accessing egress data path managed objects and actual transmission of OAM frame. On ingress side of OAM frame, in-flight frames refer to data frames between reception of OAM frame and subsequent accessing of ingress data plane managed objects. However, this limitation can be addressed by averaging such measurements across multiple time intervals.

- **Data path OAM frames**
  OAM frames use data path managed objects and are inserted and/or extracted via data plane. This method tends to be most accurate since it does not have the limitation associated with the in-flight frames.
  However, the current data path hardware/chips do not support the implementation of such methods since this requires Ethernet data path processing to include automatic insertion and/or extraction of OAM frames with data plane managed object values. Moreover, it would also require changes in hardware/chips to allow ingress and egress filtering rules across OAM frames to protect service provider administrative domains from unintended OAM frames.

  Of the three methods mentioned to measure performance the use of management plane managed objects mechanism seems to be the most suitable. The advantage of these mechanisms is that these require no changes in the existing hardware/chips and only require change in OAM client software that needs to be implemented. The steps involved in such measurement mechanism include:

  - Collection of managed object (s) information
  - Calculation of performance parameter (s)

8.2.1 Performance Management Collection Method

To collect managed object information, general or specific methods can be used. When a generic method is used, it can be applied to collect information across different managed objects e.g. using TLVs as information elements instead of specific information elements. However, when specific method with specific information elements is used, a separate method is needed per managed object or per set of managed objects.

Similarly, it is possible to use either a solicited or unsolicited collection method, where solicited method requires a response after an OAM request frame is sent while unsolicited methods does not require a response to an OAM frame. Some current examples of solicited and unsolicited methods include Loopback and Continuity Check respectively, though these are currently not used as performance management collection methods.

A generic method to send/receive data path managed object information can be used. This is similar to the variable request/response method used in IEEE 802.3ah [section 57.4.3.3/.4]. Also both solicited and unsolicited methods can be used and optionally extend the currently defined Loopback [section 7.2] and Continuity Check [section 7.1]. Note that this extension for PM will require additional processing and therefore should not be used for the measurement of delay.

EDITOR'S NOTE: INPUT ON ATOMIC FUNCTION MODEL FIGURE FROM MAARTEN EITHER FOR HERE OR FOR APPENDIX

8.2.2 Frame Loss Measurement

MEs which can support Frame Loss include:

- Service MEs for point-to-point service with dedicated UNIs

  - UNI_C to UNI_C
  - UNI_N to UNI_N
  - Access Link (UNI)
  - Inter-domain (NNI)
- Network MEs
  - Intra-domain
  - Inter-domain

8.2.3 Unsolicited Method

When applied across UNI_N to UNI_N ME, OAM frame is sent every N seconds (e.g. N=1) with **FramesTransmittedOK** value at ingress service UNI. Upon receiving this OAM frame, **FramesTransmittedOK** value is compared with **FramesReceivedOK** value at egress service UNI. Between two such consecutive OAM frames, the FL can be measured as:

Frame Loss = |CT2-CT1| - |CR2-CR1|,

where CT and CR are **FramesTransmittedOK** and **FramesReceivedOK** counts. Consecutive messages help in reducing error introduced by in-flight frames and lack of timing synchronization between sender and receiver. Within a measurement time interval, the Frame loss count can be averaged to improve the accuracy of this measurement.

NOTE: For measurement considerations with possible wrapping of CT/CR, refer to Appendix V

8.2.4 Solicited Method

Requestor sends OAM request frame to receiver every N seconds (e.g. N=1) with its managed objects (MOs) information and expects an OAM response frame with receiver's MOs information.

When applied across UNI_C to UNI_C ME, requestor sends **FramesTransmittedOK** value at egress service UNI and requests **FramesReceivedOK** value from receiver's ingress service UNI. Similarly, when applied across UNI_N to UNI_N ME, requestor sends **FramesReceivedOK** value at ingress service UNI and requests **FramesTransmittedOK** value from receiver's egress service UNI

Upon receiving the OAM request frame, receiver compares received MO information with its corresponding MO information and sends a response OAM frame back to requestor with requested MO information. When applied across UNI_C to UNI_C ME, receiver compares received **FramesTransmittedOK** value with **FramesReceivedOK** value and responds with its **FramesTransmittedOK** value. Similarly, when applied across UNI_N to UNI_N ME, receiver compares received **FramesReceivedOK** value with its **FramesTransmittedOK** value and responds with its **FramesTransmittedOK** value.

Upon receiving OAM response frame, requestor compares original sent value with received values, similar to receiver. It is possible that receiver returns the results of frame loss instead of MO information in response, however, if the MO information is returned, the performance collection method remains generic.

Between two such consecutive OAM frames, the FL can be measured as:

Frame Loss = |CT2-CT1| - |CR2-CR1|,

where CT and CR are **FramesTransmittedOK** and **FramesReceivedOK** counts. Consecutive messages help in reducing error introduced by in-flight frames and lack of timing synchronization between sender and receiver. Within a measurement time interval, the Frame loss count can be averaged to improve the accuracy of this measurement.

NOTE: For measurement considerations with possible wrapping of CT/CR, refer to Appendix V

The above method can be applied for measuring network level Frame Loss. The network level frame loss can be measured within the network independent of the services.

For non-dedicated point-to-point service types with multiplexed service UNI, where a UNI carries more than one service flow, it is possible to measure FL when data path MOs per service instance are supported.

8.2.5 Statistical Method

For multipoint-to-multipoint service type, statistical method across a pair of UNIs can be applied to estimate frame loss.

The requestor sends N OAM request frames to the recipient and receives M response frames back from the recipient such that M <= N. The data path frame loss can be estimated as:

Frame Loss = (N – M) per measurement time interval

As noted earlier, statistical methods are less accurate than proposed method in this contribution.

8.3 Frame Delay Measurement

Services supported include point-to-point and multipoint-to-multipoint between a given pair of UNIs.

MEs across which the frame delay can be measured are:

- Service MEs
  - UNI_C to UNI_C
  - UNI_N to UNI_N

Solicited Method – Loopback

This method measures round-trip or two-way frame delay. Requestor sends OAM request message with its timestamp to the receiver. Receiver replies copying the requestor's timestamp. At the requestor, the difference between the timestamps at the time of receiving the OAM response frame and original timestamp in the OAM response frame results in round trip frame delay.

8.4 Frame Delay Variation Measurement

Services supported include point-to-point and multipoint-to-multipoint between a given pair of UNIs.

MEs across which the frame delay variation can be measured are:

- Service MEs
  - –UNI_C to UNI_C
  - –UNI_N to UNI_N

Solicited Method – Loopback

This method measures round-trip or two-way frame delay per request and response frame. Within the period of observation, requestor keeps track of maximum frame delay (FD_{max}) and minimum frame delay (FD_{min}). Frame delay variation is calculated as:

Frame Delay Variation or Jitter = FD_{max} - FD_{min}

Information elements for FDV method in OAM Data mentioned in Y.17ethoam [3 - section 15.1] include:

- Sequence number
- Request Timestamp
    - FDV or Jitter = {FD (max) – FD (min)} per measurement time interval
    - Information elements for FD method in OAM Data
        - Sequence number
        - Request Timestamp

8.5 Availability Measurement

Services supported include point-to-point with at least dedicated UNIs.

MEs across which the frame delay variation can be measured are:

- Service MEs

–UNI_C to UNI_C

–UNI_N to UNI_N

Measurement Method

Measurement is based on FL, FD and FDV methods. Availability time interval (e.g. 24hr) can be divided into measurement time intervals (e.g. 1 minute). FL, FD and FDV are measured per measurement time interval. If any of the three measures crosses its corresponding thresholds, which are dependent on the service type, the measurement time interval is considered to be unavailable else it is considered to be available.

Availability = (# of available measurement time intervals)/(# of total measurement time intervals)x100%

NOTE: Mechanisms that can be used to measure availability are being proposed here but they will depend on the definition of availability and further details expected to be specified by Ethernet Traffic Management activities (q4/13 and q6/13).

8.6 Other Measurements

As per the unsolicited method explained before, the following parameters can be sent every time interval (e.g. 1 second) to the peer.

8.6.1 Errored Frame Seconds

ME: Access Link ME

Within 1 second, check if any increments in (aFrameCheckSequenceErrors, aAlignmentErrors, aFramesAbortedDueToXsColls, aFramesLostDueToIntMACXmitError, aCarrierSenseErrors, aFrameLostDueToIntMACRcvError)

If yes, declare that 1 second as errored frame second

8.6.2 Service Status

ME: UNI_C to UNI_C ME or UNI_N to UNI_N ME

Within the measurement time interval (e.g. 1 min), declare whether the service is up or down as per availability measurement, explained earlier

8.6.3 Frame Throughput

ME: UNI_N to UNI_N

Within the measurement time interval, aFramesTransmittedOK at egress UNI_N relative to CIR

8.6.4 Frame Tx

ME: Access Link ME

Within 1 second, aFramesTransmittedOK at egress UNI_N

8.6.5 Frame Rx

ME: Access Link ME

Within 1 second, aFramesReceivedOK at ingress UNI_N

8.6.6 Frame Drop

ME: Access Link ME

Within 1 second, ifInDiscards at ingress UNI_N and ifOutDiscards at egress UNI_N.

8.6.7 Loopback Status

ME: Access Link ME

aLoopbackStatus at UNI_N.

8.6.8 Client Signal Fail

ME: Access Link ME

aLinkStatus at UNI_N.

8.6.9 Unavailable time

ME: UNI_N to UNI_N

This is related to availability definition with the unavailable time intervals being counted within the observation period.

EDITOR'S NOTE: THE DISCOVERY MATERIAL WAS NOT CONTRIBUTED TO OR DISCUSSED DURING THE LAST THREE MEETINGS SO CONSULTED WITH RAPPORTEUR , IT IS PROPOSED TO DELETE MATERIAL AS BELOW

9 Information elements

EDITOR'S NOTE: THE COMMON INFORMATION ELEMENTS AND THE ONES EXCLUSIVE FOR CC HAVE BEEN UPDATED AT THE MEETING OF FEBRUARY 2004

9.1 Common Information Elements

- Addressing (DA, SA MAC)
- VLAN ID
- ME Level
- OAM EtherType
- Version
- OAM OpCode
- MPID
- ServiceID
- TransactionID

The Maintenance Point ID (MPID) is necessary as a change of hardware (e.g. an IF card or a bridge is removed/replaced) will imply a change in the MAC address. Each port has a corresponding MAC address.
The Service ID is necessary to identify a service instance. It is globally unique

Note that an MP represents either a MEP or a MIP.

Both the MPID and the Service ID may be TLVs.

Their corresponding position within the generic frame format is FFS.

The MPID is unique within a service instance.

The VLAN ID represents the data plane service instance identifier, when used.

EDITOR'S NOTE: SHOW SOME PROVISIONING MODEL IF POSSIBLE TO ADD INTO AN APPENDIX

EDITOR'S NOTE: COMPARE FUNCTION SPECIFIC INFORMATION ELEMENTS AND DELETE THE ONES THAT ARE COMMON TO ALL AS THEY ARE CAPTURED IN SUBSECTION CALLED COMMON INFORMATION ELEMENTS ABOVE

9.2 Specific Information Elements for Connectiviy Check

- Required CC Information Elements
  - Source MEP ID

Other Functionality/Information Elements are for further study:
e.g.
- CC Expiry Indication
- MEP Status - CC Activation/Deactivation Indicator

EDITOR'S NOTE: THE FOLLOWING TEXT REGARDING CC HAS BEEN DELETED AS THERE IS UPDATED (JUNE 2004) MATERIAL IN SECTION 7.1

9.3 Specific Information Elements for Non-intrusive Loopback

EDITOR'S NOTE: CHECK THE FOLLOWING INFORMATION ELEMENTS AS THEY WERE NOT REVISED FOR THE LAST TWO MEETINGS (JUNE 2004)

- Fault Detection
  - OAM Frame Identifier – (Detection of Loops, correlation)
  - Source Port Number – (Identification of specific source port, handle)
  - Destination Port Number – (Identification of specific target port, handle)?
  - Arbitrary data part – (Stress Test: could be used to test different MTUs, pad packet to full size, put worst case patterns)?
  - Note: Addressing used for this case is option (D) i.e. Unicast Destination MAC Address. Needs validation.
- Fault Localization
  - OAM Frame Identifier – (Detection of Loops, correlation)
  - Source Port Number – (Identification of specific source port, handle)
  - Destination Port Number – (Identification of specific target port, handle)?
  - Arbitrary data part – (Stress Test: could be used to test different MTUs, pad packet to full size, put worst case patterns)?
  - Note: Addressing used for this case is option (D) i.e. Unicast Destination MAC Address. Needs validation.
- Performance – for round-trip Delay and Jitter
  - OAM Frame Identifier – (Detection of Loops, correlation)
  - Source Port Number – (Identification of specific source port, handle)?
  - Destination Port Number – (Identification of specific target port, handle)?
  - Source Timestamp
  - Note: Addressing used for this case is option (D) i.e. Unicast Destination MAC Address. Needs validation.

Extra Elements:

- Response MAC Address
- Randomized Delay

NOTE: loopback will not be used for discovery purposes because of potential storms in DOS scenarios due to number of replies.

9. 4 Specific Information Elements for Link-Trace (Body)

- OAM Frame Identifier –
- Source Port Number – (Identification of specific source port, handle)
- Destination Port Number – (Identification of specific target port, handle)
- TLV for Checksum – (checksum for part that cannot be changed)
- Target MAC Address

- Source MAC Address
- Hop Count
- Others
  - Periodicity of Loopback – (when used proactively)
  - Randomized Delay - ?

9.5 Performance Monitoring Information Elements

EDITOR'S NOTE: THE FOLLOWING THREE SUBSECTIONS SHOULD BE CONSOLIDATED IN ONE AS THE PERFORMANCE MONITORING OAM FUNCTION SHOULD HAVE INFORMATION ELEMENTS THAT SERVE ALL OF THE PURPOSES

9.5.1 Information elements that can be applied to OAM Data for the Unsolicited Method

- Sequence number
- # of TLVs
- TLVs (Managed Object variable: **FramesTransmittedOK**, value length, value)

9.5.2 Information elements that can be applied to OAM Data for the Solicited Method

- Sequence number
- # of Transmit TLVs (value filled in by requestor, recipient simply copies it back in response)
- # of Request TLVs (value is filled in by recipient and sent back in response)
- TLVs (Managed Object variable: **FramesTransmittedOK & FramesReceivedOK**, value length, value)

9.5.3 Information elements for Frame Delay method in OAM Data

- Sequence number
- Request Timestamp

10 OAM frame formats

EDITOR'S NOTE: THIS SECTION WILL HAVE TO ALIGN WITH THE SECTION ON INFORMATION ELEMENTS.

10.1 Generic OAM Frame Format

A single generic format is defined for all Ethernet OAM messages as depicted blow. The VLAN tag is optional and if it is present, it indicates the service instance corresponding to the OAM message (e.g., only bridge nodes participating in that service instance will process/forward the OAM message). The OAM Ethernet Type has a value TBD and it indicates that the message is an OAM message for Service Provider use. It should be noted that all the OAM messages carry the same OAM Ethernet Type. It is recommended that another OAM Ethernet Type to be allocated for customer OAM usage (transparent to service provider nodes) so that there is clear differentiation between customer and provider OAM messages.

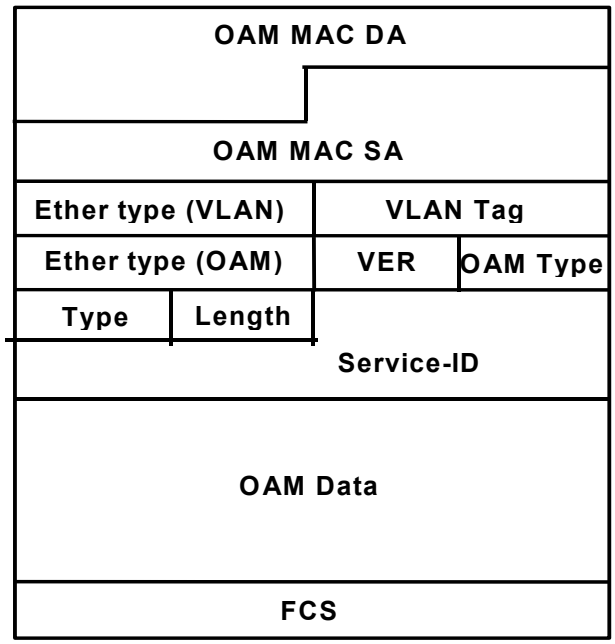


Figure 10-1: Generic OAM Message Format

EDITOR'S NOTE: THE GENERIC OAM MESSAGE FORMAT HAS TO BE ALIGNED WITH INFORMATION ELEMENTS. TO BE DECIDED IN NEXT MEETINGS

The fields for the generic OAM message format are defined as follows:

- **OAM Destination MAC Address**: This MAC address can be the unicast address of a bridge or a multicast address corresponding to a group of bridges or it can be a well-defined multicast address (e.g., "bridge all" multicast address)
- **OAM Source MAC Address**: This is the MAC address of the source bridged (a unique MAC address designated for OAM functionality) or it can be the MAC address of a bridge interface over which the OAM message is sourced.
- **VLAN Ether Type and Tag**: This is an optional field and it is present when the OAM message is related to a service instance. In this case, this VLAN tag designates the associated service instance.
- **OAM Ethernet Type**: This is a unique Ethernet Type that indicates Service Provider OAM messages. It is recommended to have another Ether Type for Customer OAM messages.

OAM Type: The OAM frame type defines the type of OAM frame. The OAM frame types that are defined in this recommendation are:

- **Intrusive Loopback Request (0x00)**
- **Intrusive Loopback Release (0x01)**
- **Intrusive Loopback Reply (0x02)**
- **Non-Intrusive Loopback Request (0x03)**
- **Non-Intrusive Loopback Reply (0x04)**
- **Ethernet Link Trace Request (0x05)**
- **Ethernet Link Trace Response (0x06)**
- **Continuty Check (0x07)**
- **Performance Monitoring Request (0x08)**
- **Performance Monitoring Reply (0x09)**
- **AIS (0x0A)**

- **RDI (0x0B)**
- **Vendor Specific (0xFF)**. The vendor specific op-code is provided to allow vendors or other organizations to extend OAM functions in proprietary ways.
- **Other op-codes may be defined in the future**. OAM frames with unexpected unknown op-codes MUST be silently discarded.

- **OAM Version**: The Version field identifies the OAM version number. Implementations conforming to this recommendation MUST use a value of 0x01 in this field. OAM frames with a different version MUST be silently discarded
- **Service ID**: TLV for the service instance identification
- **OAM Data**: This is a data field associated with the corresponding OAM type and sub-type and its format is dependent on the OAM type/sub-type fields. The OAM frame including OAM data portion must result in an Ethernet frame with valid length. Therefore, if necessary the OAM frame must be padded with zeros for a minimum size frame.

Annex A

EDITOR'S NOTE: THE MATERIAL OF THIS ANNEX IS TAKEN FROM WD 10 OF INTERIM MEETING OF Q.3/13 (NOVEMBER 2003) AND IT WAS AGREED TO INCLUDE IT IN THIS ANNEX FOR FURTHER CONSIDERATION

AIS/RDI MECHANISM FOR AN ETHERNET POINT-TO-POINT CONNECTION OVER A SINGLE SERVER LAYER (i.e. SDH or OTN)

G.8010 [4] Figure 18 (reproduced below) illustrates the architecture of an Ethernet point-to-point connection.

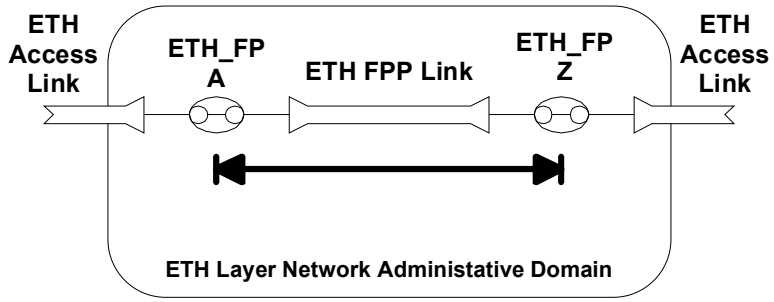


Figure 18/G.8010 – Point-to-point ETH connection (single link)

The representation of the corresponding maintenance areas is illustrated in G.8010 Figure 23 top right (reproduced below).

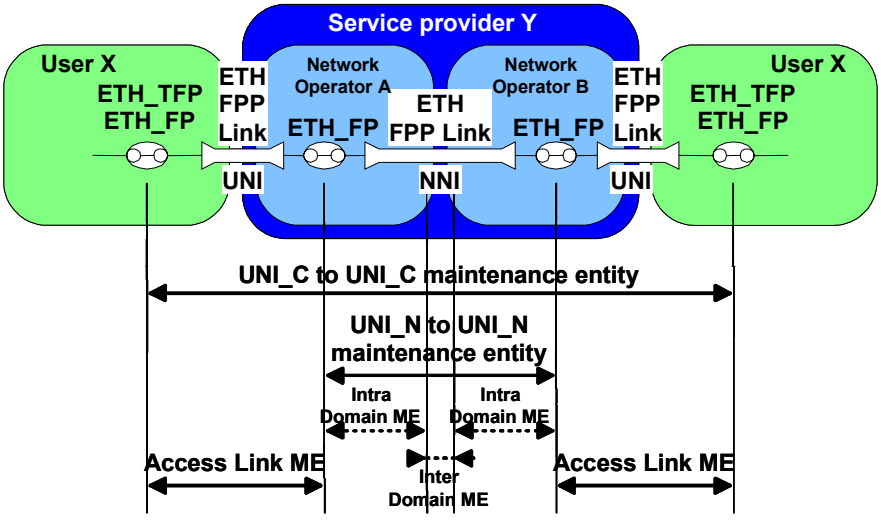


Figure 23 (top right)/G.8010 – Point-to-point ETH connection administrative domain associated maintenance entities

Of note is that there are no ETH flow points at the handoff between the two networks. So for the general case when there are multiple network operators and a single server layer, maintenance of the ETH layer is not possible within those operators' networks, and can only be performed via the server layer. That category of maintenance is called inherent monitoring, also discussed in [4].

To address the lack of AIS and network RDI functionality in EFM OAM, the issues are then:

a) How to convey an access link fault from one side of the network to the other.

b) How to convey a server layer fault to the access links.

In SG 15/Q.12, work has been progressing to define a Service Management Channel (SMC) to facilitate the provider edge NE-to-edge NE exchange of OAM information, as well as support for an intermediate provider NE to query OAM information from, and send test-related commands to, a provider edge NE.

Currently, the direction being taken in Q.12/15 proposes G.7041 [5] GFP-F Client Management Frames (CMF) for conveying the provider edge NE-to-edge NE OAM information, and a Path OH byte for the intermediate provider NE communications channel to a provider edge NE. G.7041 defines CMFs for conveying information about the client signal from an ingress edge NE to the egress edge NE. One of the defined CMF indications is Client Signal Fail (CSF). The figure below illustrates the GFP-F frame format for a CMF with a CSF indication.

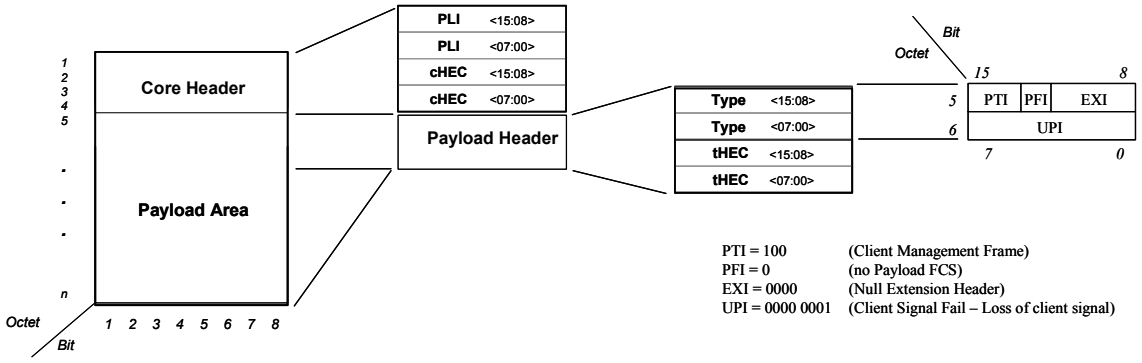


Figure A-1: GFP-F CMF CSF Frame Format

By using the EFM OAM Link Fault flag in conjunction with the GFP-F CMF CSF indication, the necessary AIS and network RDI mechanisms can be provided for an Ethernet point-to-point connection single server application.

A simplifying assumption can be made regarding the conditioning of the Ethernet access links on either side of the SDH/OTN transport network. For a dedicated point-to-point application, the access link is specific to a single service, and since an Ethernet service is bidirectional, a fault in either direction should result in the access link being conditioned as ‘failed'.

The following fault scenarios and accompanying figures illustrate the proposed interworking of the EFM OAM Link Fault flag with the GFP-F CMF CSF indication to appropriately condition the Ethernet access links. Only uni-directional faults are considered, the scenarios can be combined per the superposition principle to describe bi-directional faults.

Scenario 1

In the figure below a uni-directional fault occurs on the east access link on ingress to the carrier network.

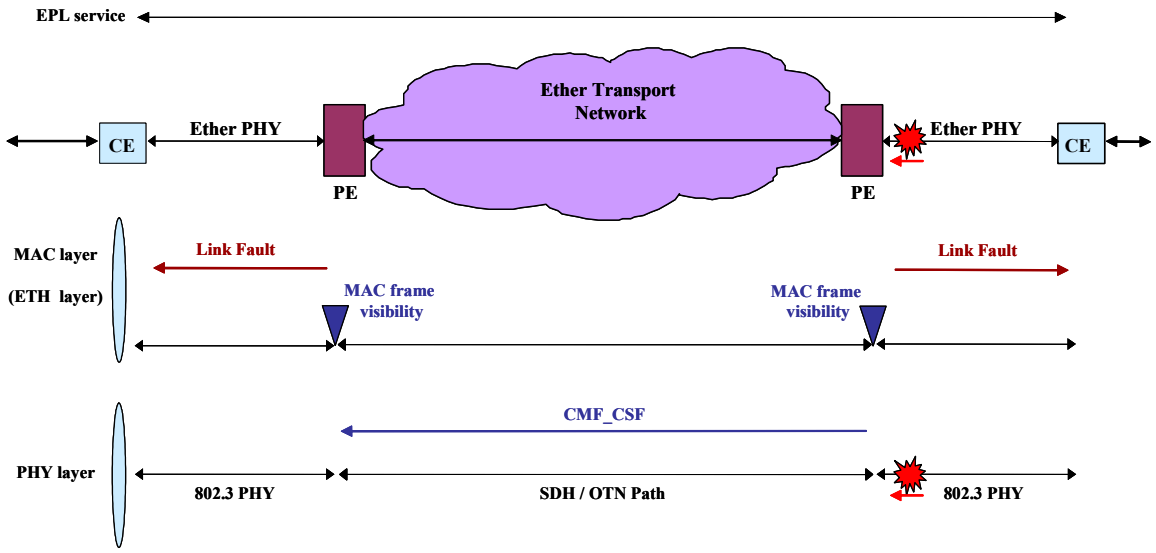


Figure A-2: Fault on Ingress

- The east PE detects the failure:
  - an 802.3ah OAM function sends Link Fault upstream, interspersed with Idles
  - a new function sends a GFP-F CMF CSF indication into the network
- The east CE detects Link Fault:
  - Idles are sent towards the network (and towards the enterprise)
- The west PE detects the GFP-F CMF CSF indication:
  - a new function translates it to 802.3ah Link Fault downstream, interspersed with Idles
- The west CE detects Link Fault:
  - Idles are sent towards the network (and towards the enterprise)

Scenario 2

In the figure below a uni-directional fault occurs westbound within the carrier network.

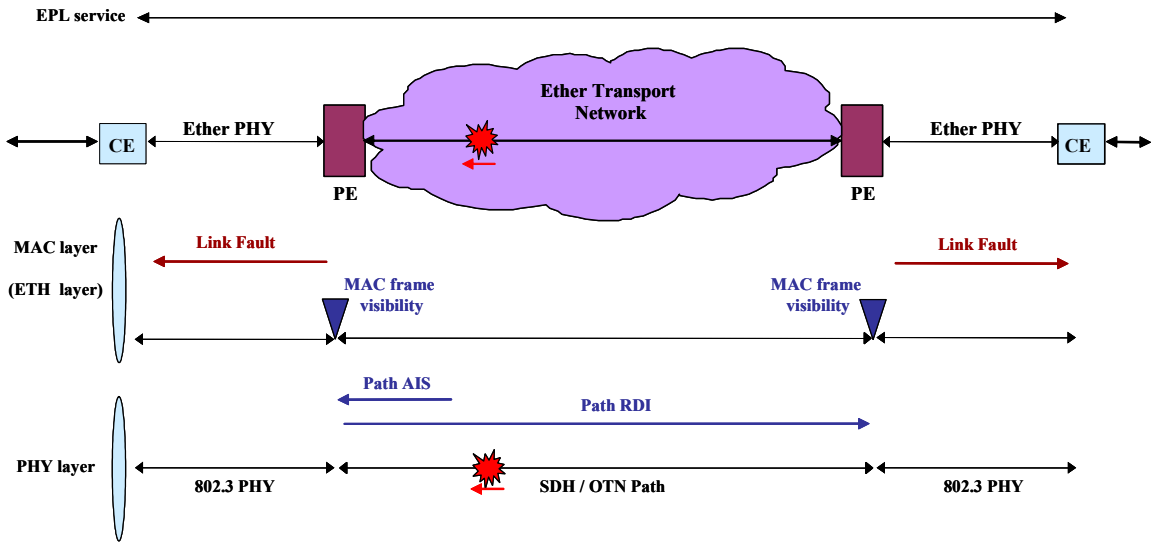


Figure A-3: Fault within Carrier Network

- An NE (or the west PE) in the carrier network detects the failure:
  - SDH Path AIS is generated downstream

- The west PE detects SDH Path AIS (or the fault directly):
  - a new function translates it to 802.3ah Link Fault downstream, interspersed with Idles
  - SDH Path RDI is generated back into the network
- The west CE detects Link Fault:
  - Idles are sent towards the network (and towards the enterprise)
- The east PE detects SDH Path RDI:
  - a new function translates it to 802.3ah Link Fault downstream, interspersed with Idles
- The east CE detects Link Fault:
  - Idles are sent towards the network (and towards the enterprise)

Scenario 3

In the figure below a uni-directional fault occurs on the west access link towards the enterprise network.

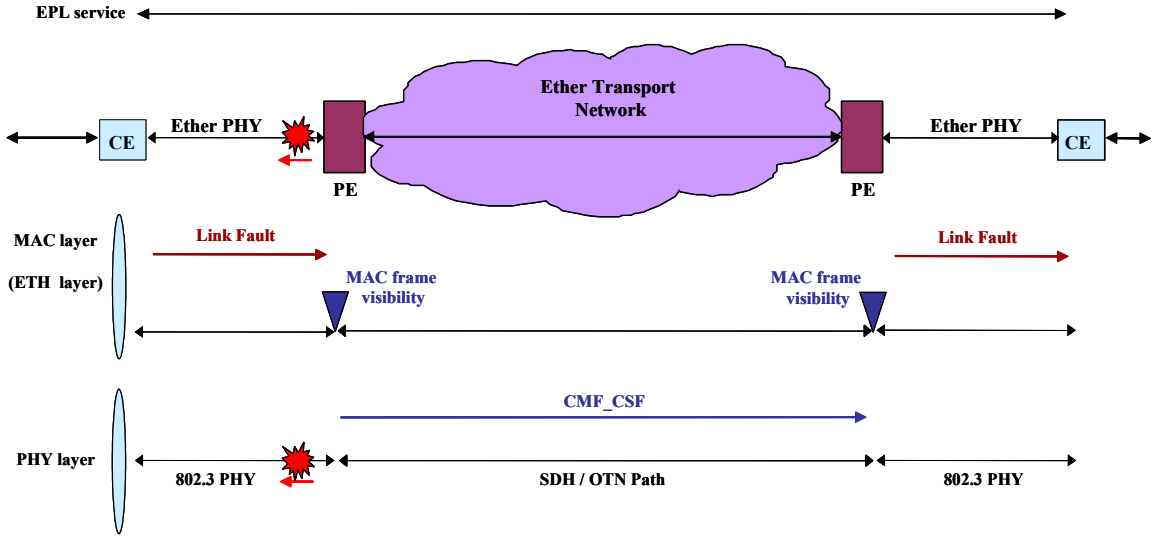


Figure A-4: Fault on Egress

- The west CE detects the failure:
  - an 802.3ah OAM function sends Link Fault upstream, interspersed with Idles
  - Idles are sent towards the enterprise
- The west PE detects Link Fault:
  - a new function translates it to a GFP-F CMF CSF indication into the network
  - Idles are sent towards the CE
- The east PE detects the GFP-F CMF CSF indication:
  - a new function translates it to 802.3ah Link Fault downstream, interspersed with Idles
- The east CE detects Link Fault:
  - Idles are sent towards the network (and towards the enterprise)

Summary

As a result, two maintenance signal translation functions, i.e. EFM OAM Link Fault flag and GFP-F CMF CSF indication can be used by the provider edge of a point-to-point single server application (i.e. SDH or OTN) in order to address network AIS and RDI mechanisms.

Appendix I

I.1 OAM Domains

Each provider can be associated with an administrative boundary, called OAM domain. A service may be carried across a single or multiple OAM domains.

As identified in Y.1730, network elements placed at the boundary of provider network serve as edge network elements and are associated with the ingress and egress of a network flow. When an edge network element of a provider performs hand-off of an ETH layer flow, while interacting with edge network element of another provider, that network element serves as an edge hand-off network element. Those network elements that are not associated with the ingress, egress or hand-off of a network flow serve as interior network elements.

It is also possible that a single provider network may have further administrative boundaries. Example is when a provider network consists of different operator networks. If this is the case, one could still identify edge, edge hand-off, and interior network elements within each such administrative boundary.

Ports on a network element in an OAM domain can be classified as interior or exterior to that OAM domain. Interior ports are those on which OAM frames, belonging to an OAM flow, are recognized and processed. Processing may result in either termination of OAM flow or relaying across other ports on the network element. Exterior ports are those on which OAM frames are not recognized and filtered. An edge network element has both interior and exterior ports to an OAM domain, while an interior network element has all its ports marked as interior ports to that OAM domain.

Within an OAM domain, OAM flows may be applicable between edge network elements only (edge hand-off network element is also an edge network element) or across all network elements (i.e. including all interior network element and edge network elements). OAM frames can be Unicast or Multicast frames. The difference between the two is based on the destination MAC address (DA). A Unicast OAM frame has a Unicast DA while a Multicast OAM frame has a Multicast DA. A Multicast OAM frame can associate itself to all edge networks elements or all network elements inside a domain based on its Multicast DA.

NOTE: Refer to G.8010 and G.805.

I.2 OAM Flows

Different OAM flows, as discussed in Section 6.1, can be identified by using an OAM flow identifier within the OAM frame. OAM flow identifier can assume the following values:

- $\text{UNI-UNI}_{\text{Customer}}$
  Customer UNI-UNI flow between reference points on the customer side of the UNI.
- $\text{UNI-UNI}_{\text{Provider}}$
  Provider UNI-UNI flow between reference points on the provider side of the UNI.
- $\text{Segment}_{\text{intra-provider}}$
  Segment OAM flow between flow points within the boundary of a provider network. This may include OAM flow between flow points on the boundary of a provider network or between any flow points within a provider network as required.
- $\text{Segment}_{\text{inter-provider}}$
  Segment OAM flow between flow points inside the boundaries of two or more provider networks. This may include OAM flow between flow points on the boundaries of two or more adjacent provider networks or between any flow points inside the boundaries of two or more provider networks, as required. Note: Under special cases, $\text{Segment}_{\text{inter-provider}}$ may be same as $\text{UNI-UNI}_{\text{Provider}}$.

- $UNI_{Segment}$
  OAM flow between reference points (i.e. TFP and FP) on the customer side and provider side of the UNI.
- $NNI_{Segment}$
  OAM flow between flow points on two edge hand-off network elements connected to each other. Each edge hand-off network element belongs to a different provider network.
- UNI_{Link}
  If the UNI is realized using a single ETY link, this OAM flow can be used for ETY link between customer and provider network.
- $Transit_{Link}$
  Any intermediate ETY link between network elements, this OAM flow can be used.

NOTE: Both UNI_{Link} and $Transit_{Link}$ can be based on IEEE 802.3ah. [However, the reference to IEEE 802.3ah may not be possible, until it becomes a standard, though it is close to being one]

NOTE: It is worth noting that though different OAM flows have been identified, not all will be applicable for all services and/or business models; especially, there may be some limitations within multiple provider scenarios.

An example of maintenance entities and ME levels can be seen in table I.1

Table I.1: Relationship between OAM flows and MEs

Y.1730	**G.8010**		**Examples**
ME	**ME**	**OAM flows**	**ME**
UNI-UNI (Customer)	UNI_C to UNI-C ME	UNI-UNI Flow	UNI-UNI (Customer)
UNI-UNI (provider)	UNI_N to UNI_N ME	Transit Flow	UNI-UNI (provider)
Segment (PE-PE) intra-provider	Intra Domain ME	Transit Flow	Segment (PE-PE) intra-provider
Segment (PE-PE) inter-provider	Inter Domain ME	Transit Flow Transit Link Flow	Segment (PE-PE) inter-provider
Segment (any to any)		Transit Flow Transit Link Flow	Segment (any to any)
ETY Link OAM - UNI	Access Link ME	UNI Link Flow	ETY Link OAM - UNI
ETY Link OAM - NNI	Inter Domain ME	Transit Link Flow	ETY Link OAM - NNI

Since the OAM Flows have a one-to-one correspondence with the MEs, the ME levels can be represented by octet values assigned to OAM Flow identifiers as follows:

(It is conceivable that value of OAM Flow Identifiers can be such that filtering can be done based on whether the OAM frame entering or exiting a domain have OAM Flow Identifier value smaller than minimum OAM Flow Identifier configured on the interior and or exterior port of a domain.)

- $UNI\text{-}UNI_{Customer}$ = 255 (0xFF)
- $UNI\text{-}UNI_{Provider}$ = 253 (0xFD)
- $Segment_{inter\text{-}provider}$ = 251 (0xFB)
- $NNI_{Segment}$ = 249 (0xF9)

- $UNI_{Segment}$ = 247 (0xF7)
- $Segment_{intra\text{-}provider}$ = 245 (0xF5)
- UNI_{Link} =
- $Transit_{Link}$ =

And, if the following minimum OAM flow Identifier values are configured across the different ports:

- NNI port = 249 (0xF9)
- UNI port = 247 (0xF7)
- Interior port = 245 (0xF5)

Filtering at edge network elements can be achieved such that OAM frames with OAM Flow identifier smaller than the minimum OAM Flow identifier are not allowed into or out of the OAM domain.

I.3 Fault Types

Two fault types are recognized in relationship with ETH OAM:

1) ETH Discontinuity
2) ETH Misconnection
3) ETH Link Faults

The two first fault types are shown in the following Figure I.3-1:

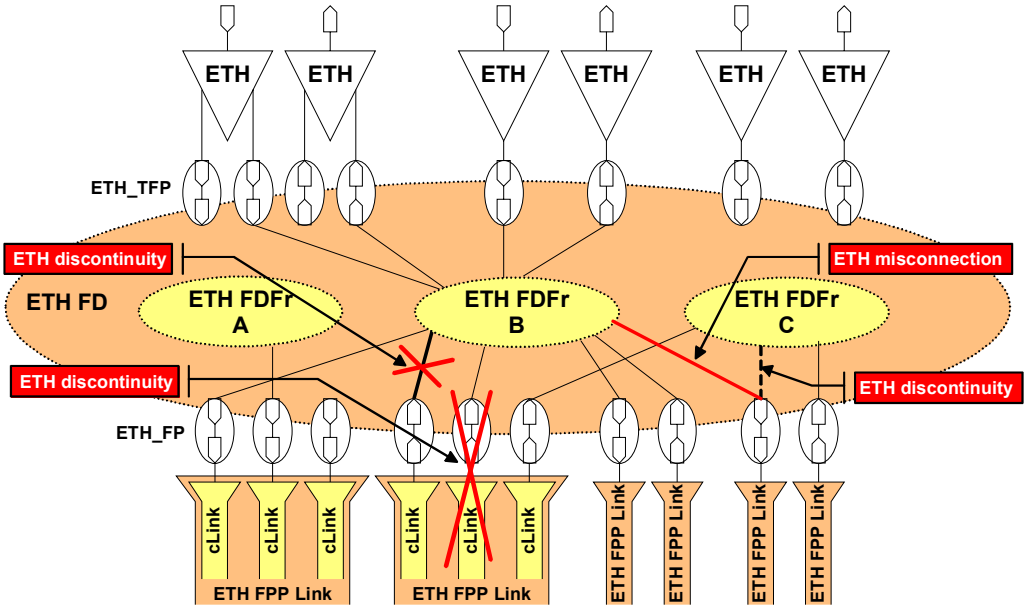


Figure I.3-1

ETH Discontinuity

Causes can be:

- Physical fault (i.e. fibre cut)
- Failure of a bridge
- Looping (customer or provider loops or due to the use of a wrong topology)

– Misconfiguration

ETH Misconnection

Can be caused by:

– Misconfiguration

Appendix II

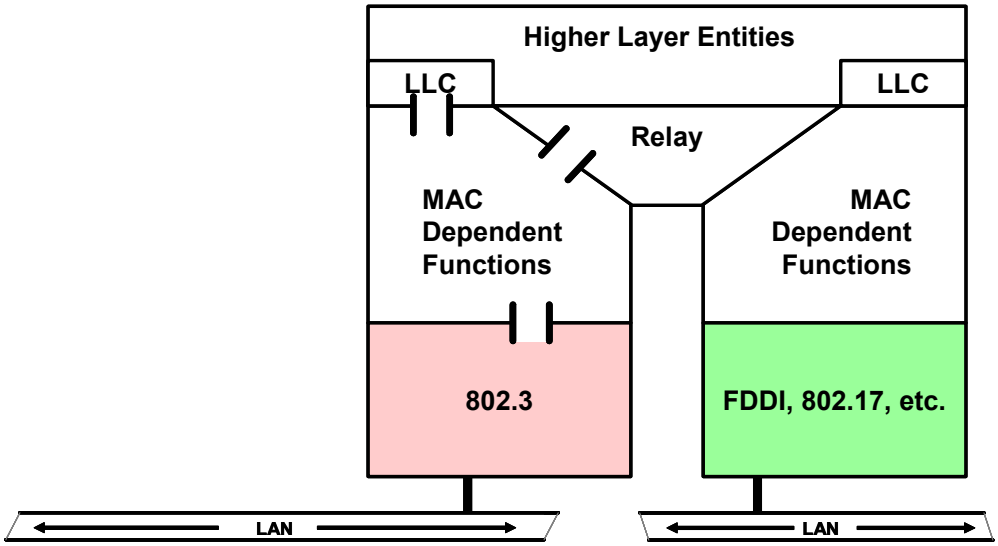


Figure II-1 The "Baggy Pants" diagram : IEEE Std. 802.1D-2003, Fig. 7-3

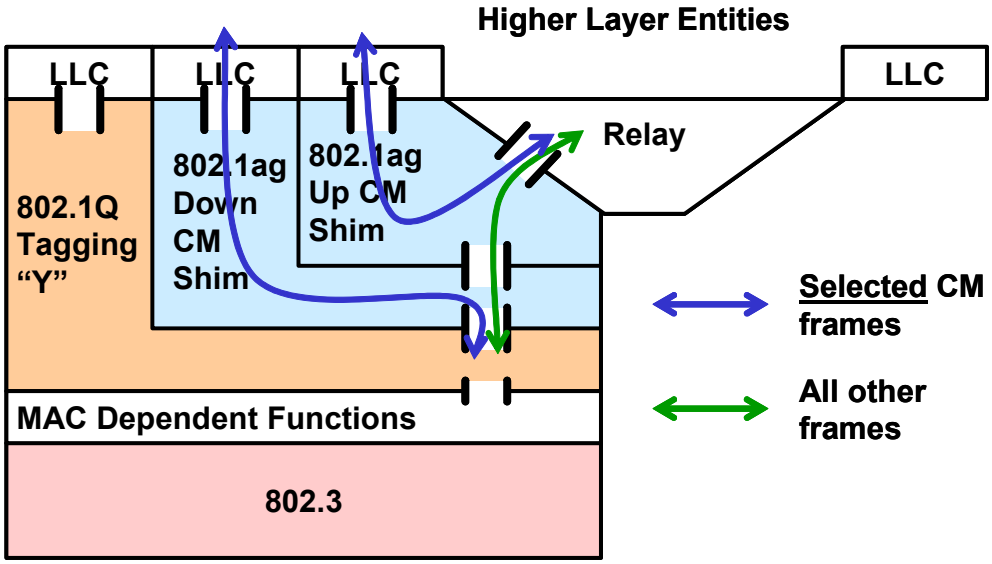


Figure II-2: The MAC stack: 802.1ag Connectivity Management Shim

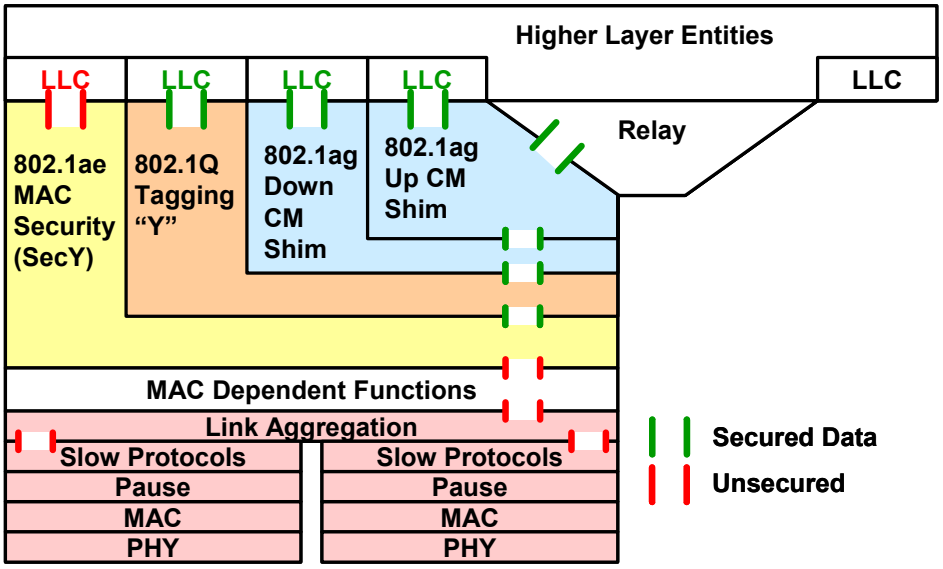


Figure II-3: The MAC stack: CM + Security + 802.3

The examples below relate to figure 6-2 of this Recommendation

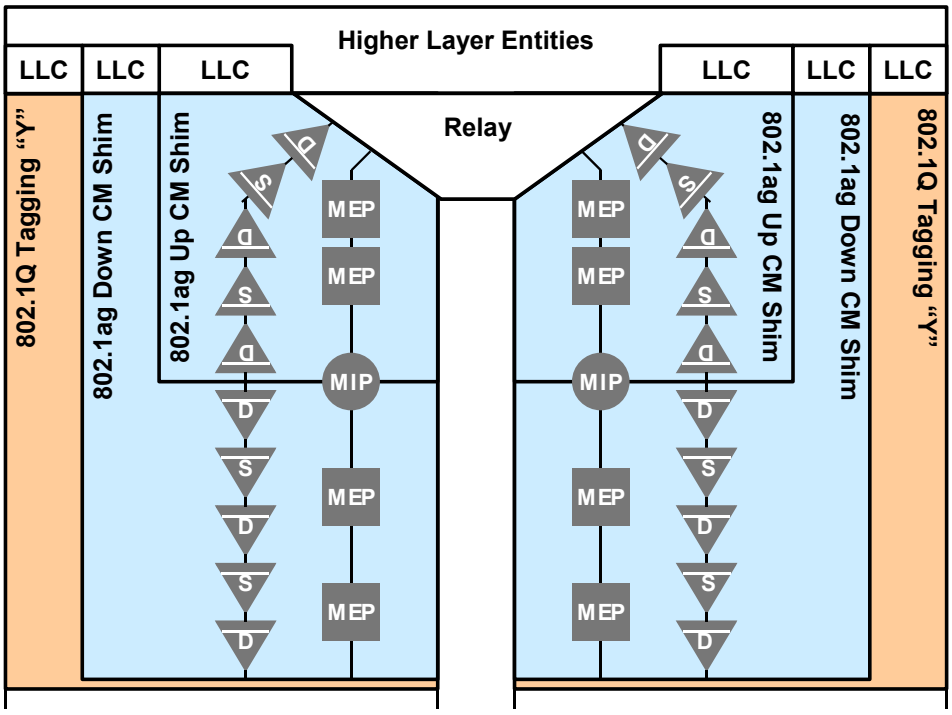


Figure II-4 – Illustrating the location of MEPs and MIPs in the 802.1ag Up and Down CM Shims in the IEEE 802.1 "baggy pants" model

- two representations are shown in parallel: the MEP and MIP shorthand symbolic representation and their associated atomic function representation
- both the up and down shims in an interface port to a bridge may support multiple MEPs; in this example each shim supports two MEPs

- the MIP is located at the junction of up and down shim as it has an ETH Diagnostic function in both the up and the down shim.
- the MEPs and MIP in an interface port have no pre-assigned knowledge of the ME level they will be operating at; this is represented by making them all colourless (i.e. grey)
- further baggy pants figures will present only MEPs and MIPs, not longer the associated atomic functions

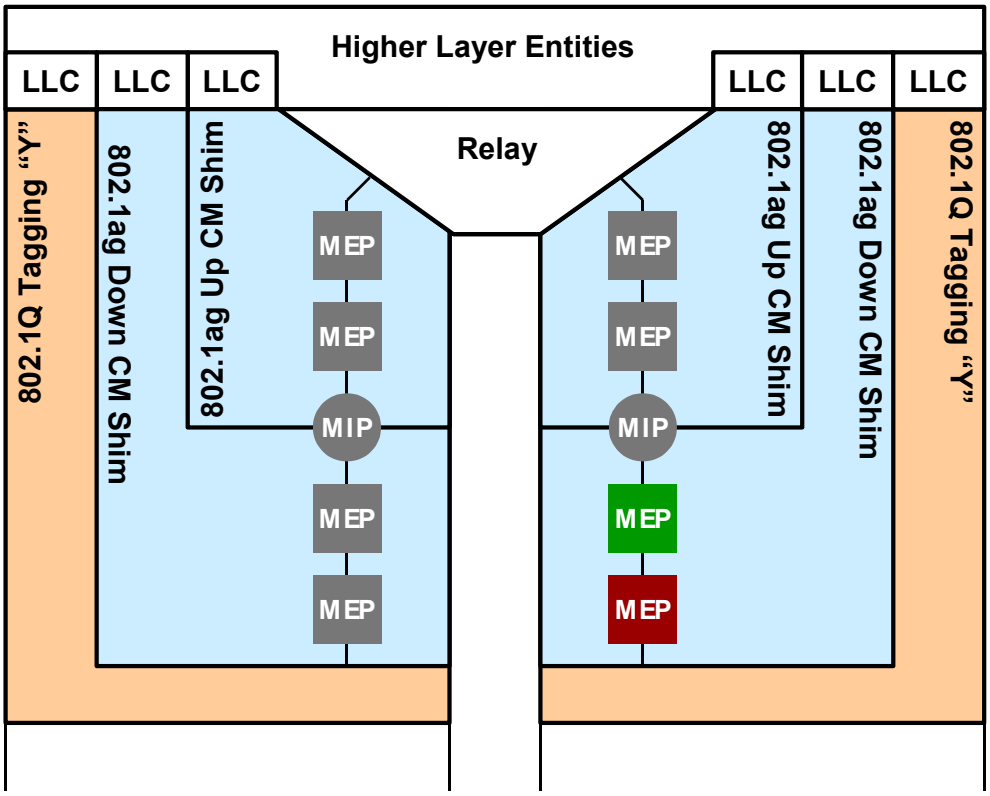


Figure II-5 – Illustrating the location of MEPs in the "baggy pants" model for CE1

- customer equipment number 1 has for the ETH connection of figure 6-2 two MEPs activated in the down shim to monitor the UNI-C to UNI-C connection and the CE1 to B2 link
- the other MEPs and MIPs are made transparent to represent that those are inactive for this connection

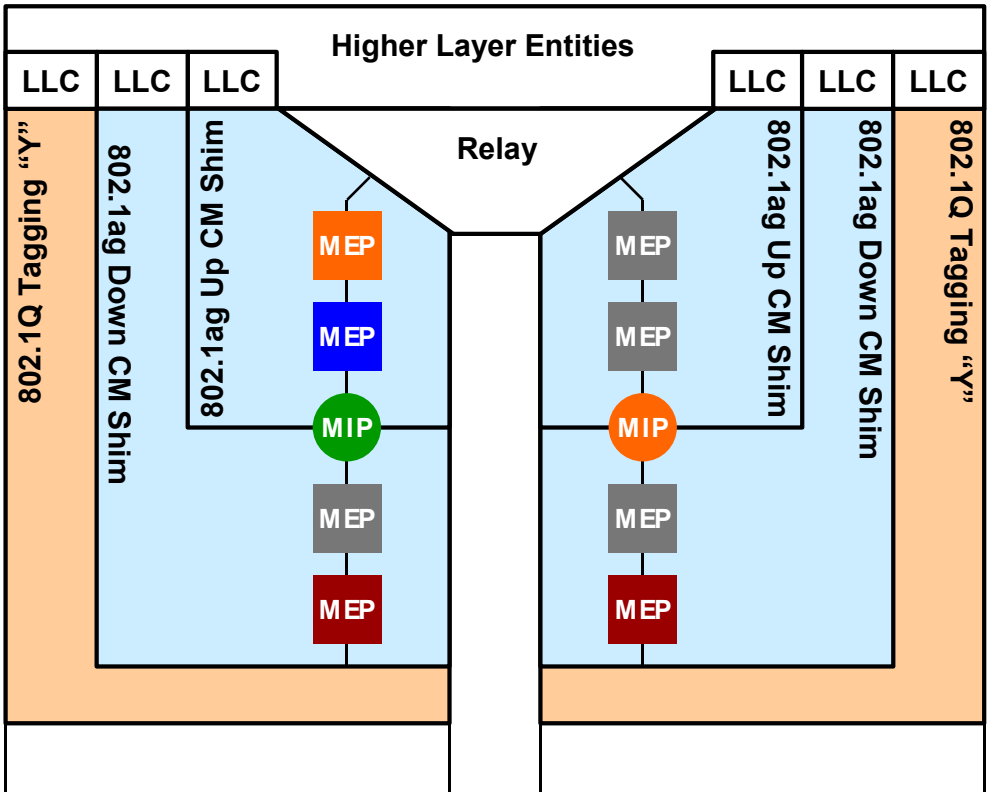


Figure II-6– Illustrating the location of MEPs in the "baggy pants" model for B2

- operator A bridge number 2 has at its customer equipment facing port three MEPs active; one to terminate the link ME, one to terminate the service provider's ETH ME (blue) and one to terminate the operator A's ETH ME
- this port has also an active MIP for use by the customer's UNI-C to UNI-C ETH ME
- the second interface port on this bridge has an active MIP (operator A's ETH ME) and an active MEP (B2 to B3 link)
- note that the choice of the active MEP in the two down shims is arbitrary; the other MEP could have been choosen as well. Equipment should be capable to change the MEP location hitless in order to support the addition of an ME level above or below an existing ME level

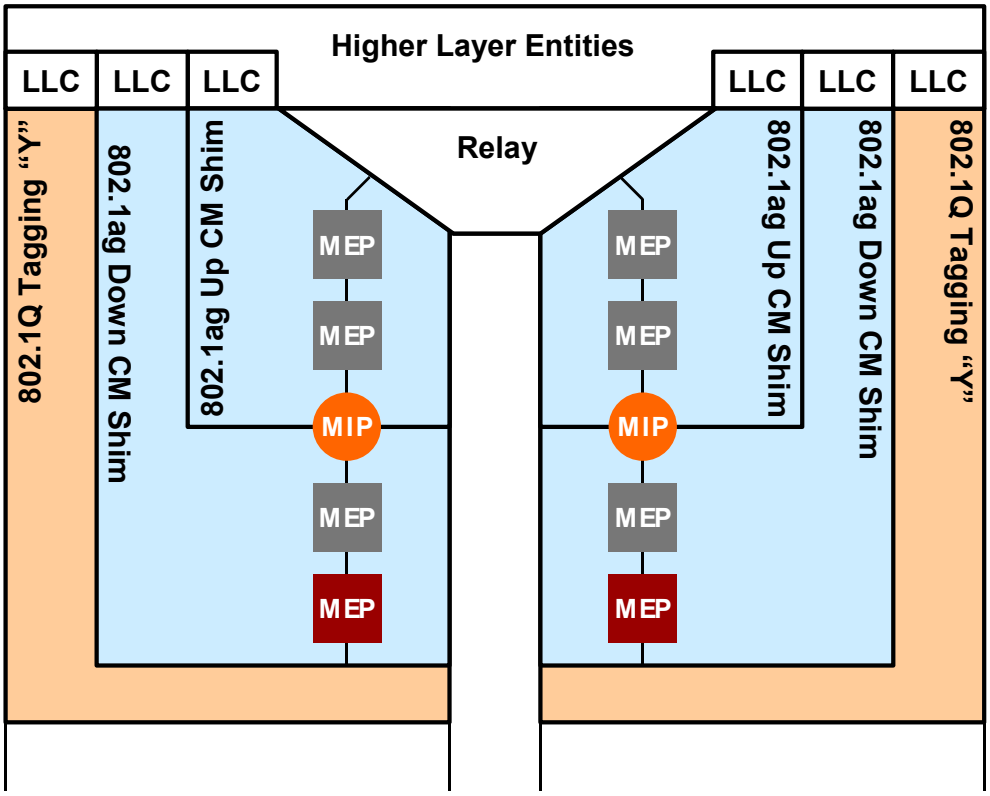


Figure II-7 – Illustrating the location of MEPs in the "baggy pants" model for B3

- operator A bridge number 3 has two ETH link related MEPs and two operator A ETH ME related MIPs active

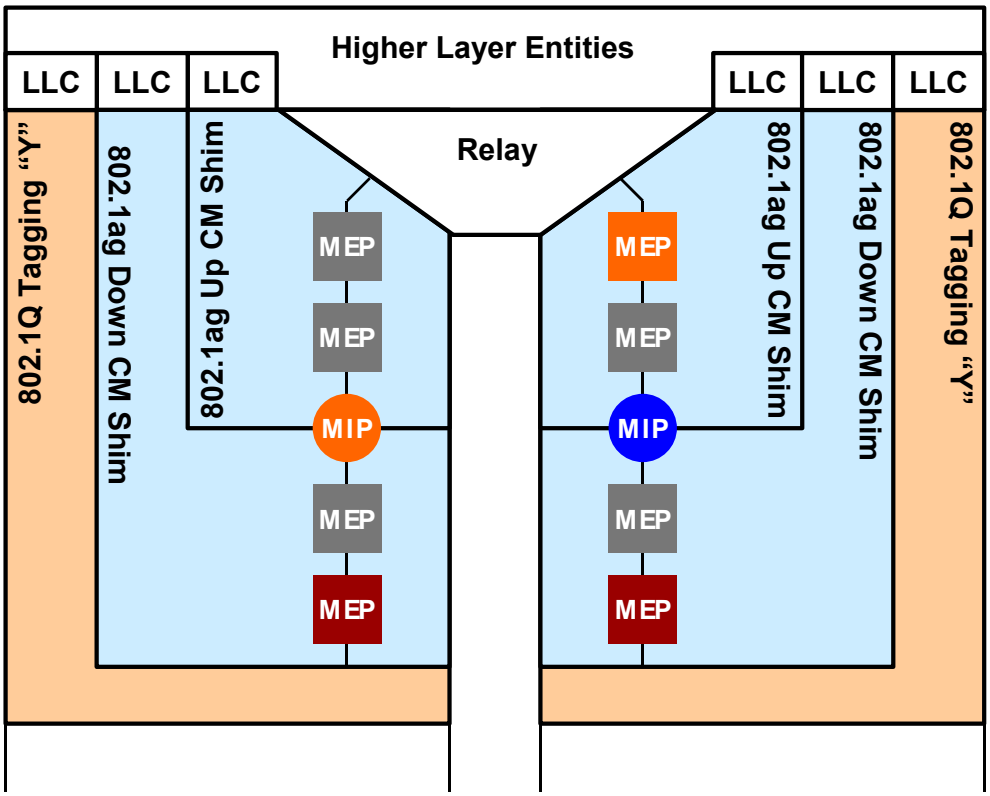


Figure II-8 – Illustrating the location of MEPs in the "baggy pants" model for B4

- operator A bridge number 4 has two ETH link related MEPs active, one in each interface port
- furthermore the operator A domain facing interface port (left) has its MIP active for use in the opertor A ETH ME level
- the operator B facing interface port (right) terminates operator A's ETH ME and has for that purpose a MEP in the up shim active
- as this right interface port is at a domain boundary, it has to support a MIP for the next higher ETH ME (service provider), to allow fault localization by this service provider (inside network of operator A, inside network of operator B or in the link between A and B)

The reader is assumed to be able to draw the MEP/MIP configurations for the other bridges at this point. Those are not shown therefore.

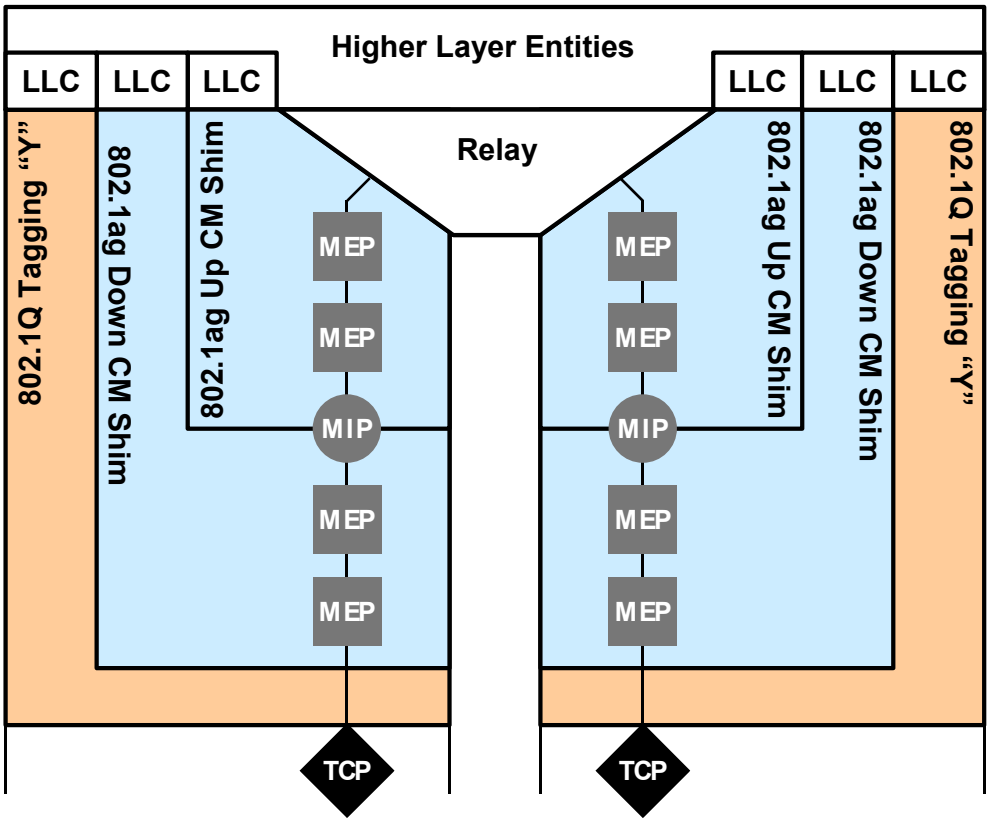


Figure II-9 – Illustrating the location of TCPs in the "baggy pants" model

- the interface port at the network side of a UNI will/may have a TCP that is located below the Down CM Shim in the baggy pants model. In this way the MEPs in the down shim will be able to register the effect (discarding) of the traffic conditioning and report this to the customer and service provider who share the responsibility for this UNI link

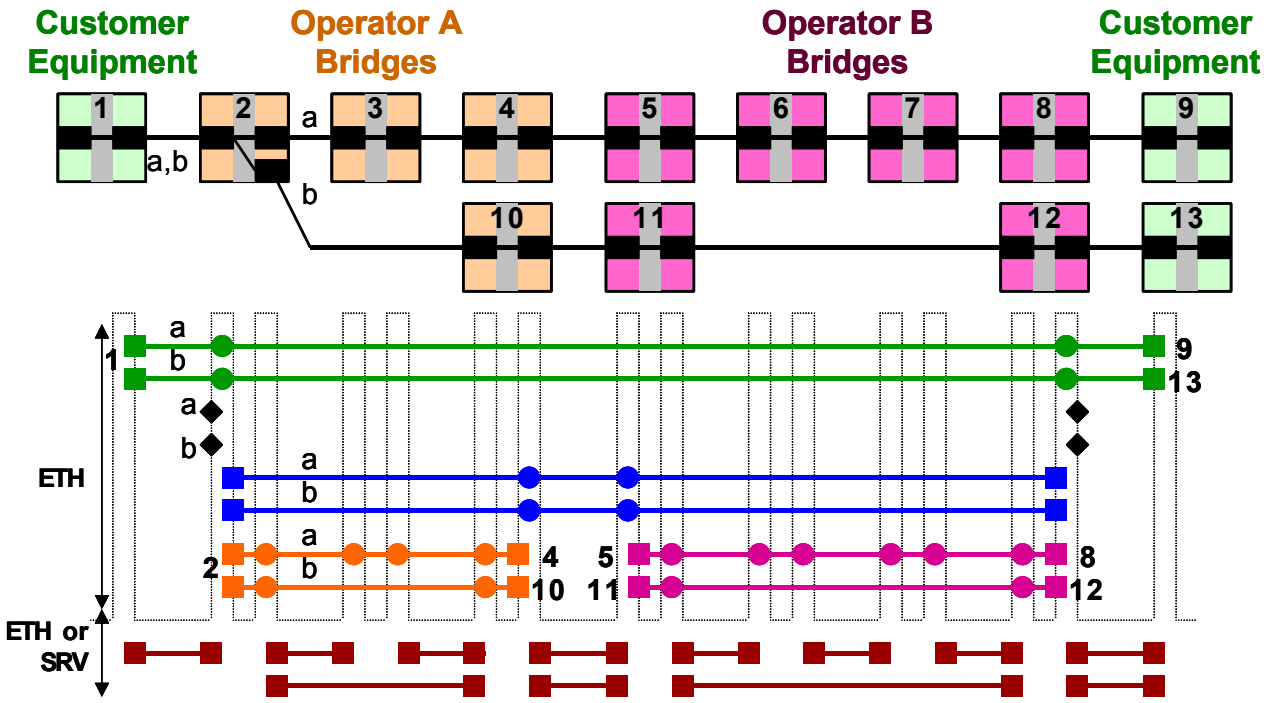


Figure II-10 – Illustrating the location of MEPs, MIPs and TCPs in the "baggy pants" model for the case of a two p2p connections with multiplexed access

- two p2p connections (a,b) with associated CM are depicted
- bridge 2 has two parallel sets of MEPs/MIP/TCP in the UNI facing port
- note: the figure depicts a single ME between CE1 and B2. This implies that this ME is a SRV ME. If it would have been an ME at ETH layer, then there should have been two ETH MEs, one for each p2p connection

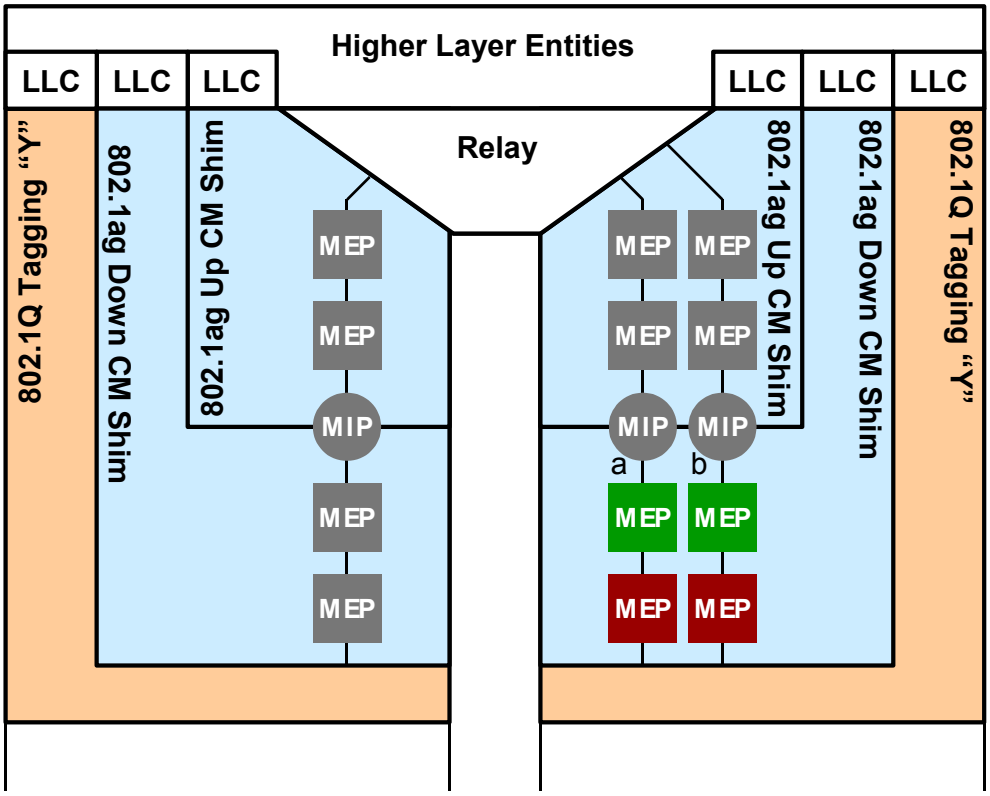


Figure II-11 – Illustrating the location of MEPs, MIPs and TCPs in the "baggy pants" model for the case of a two p2p connections with multiplexed access

- customer equipment number 1 has for the two ETH connections of figure II-10 two MEPs activated in the down shim to monitor the two UNI-C to UNI-C connections and the CE1 to B2 links

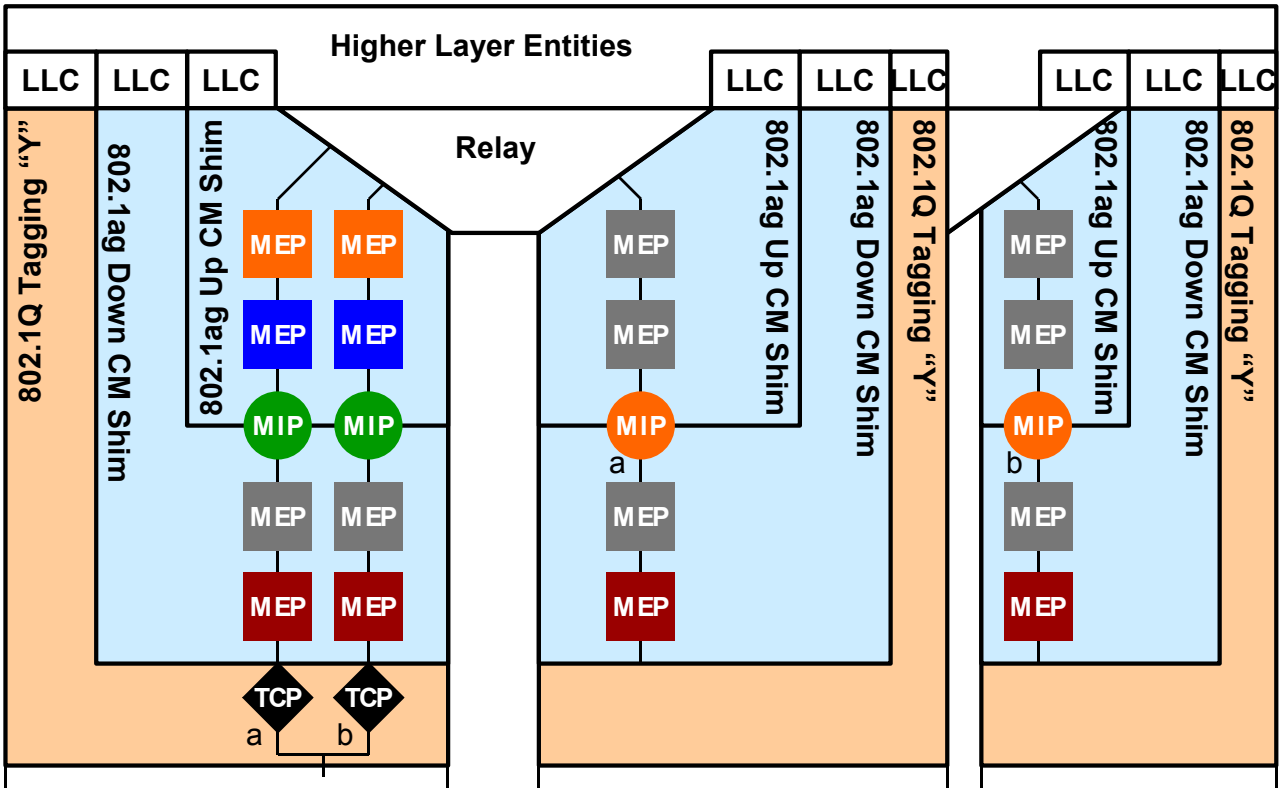


Figure II-12– Illustrating the location of MEPs, MIPs and TCPs in the "baggy pants" model for the case of a two p2p connections with multiplexed access

- operator A bridge number 2 has at its customer equipment facing port for each of the two p2p connections three MEPs active; one to terminate the link ME, one to terminate the service provider's ETH ME (blue) and one to terminate the operator A's ETH ME
- this port has also an active MIPfor each of the two p2p connections for use by the customer's UNI-C to UNI-C ETH MEs
- there are two interface ports facing the network, one for each of the two p2p connections

MEP, MIP, TCP for Dual Relay Model & Bundling MEP, MIP, TCP for Dual Relay Model & Bundling

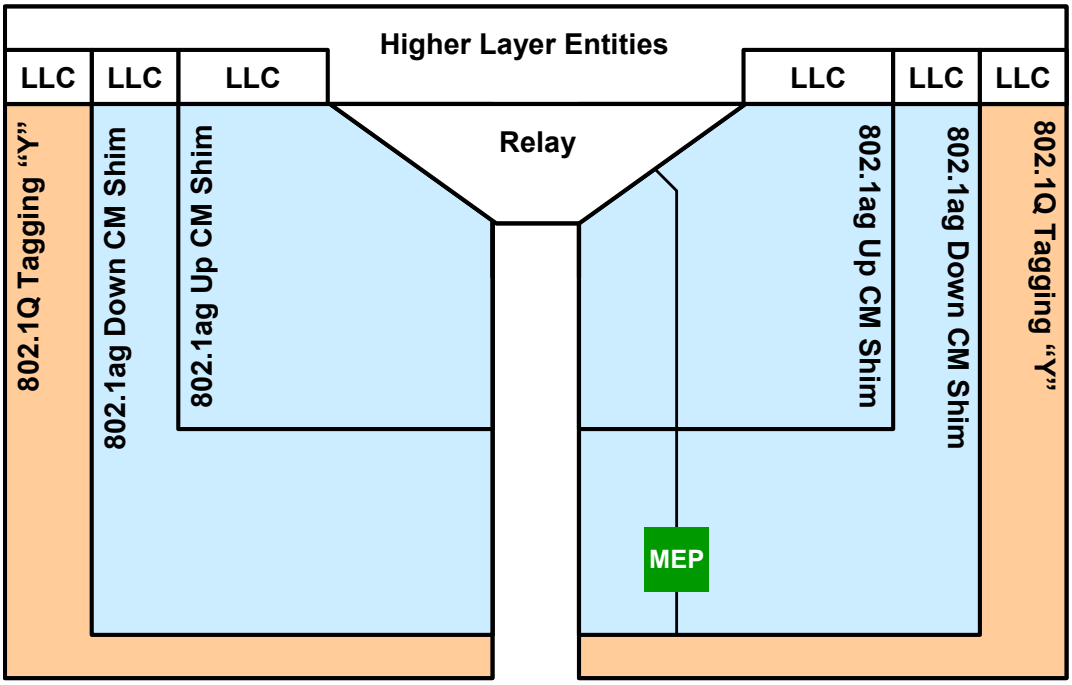


Figure II-13: Customer Bridge 1, example without ETH link ME

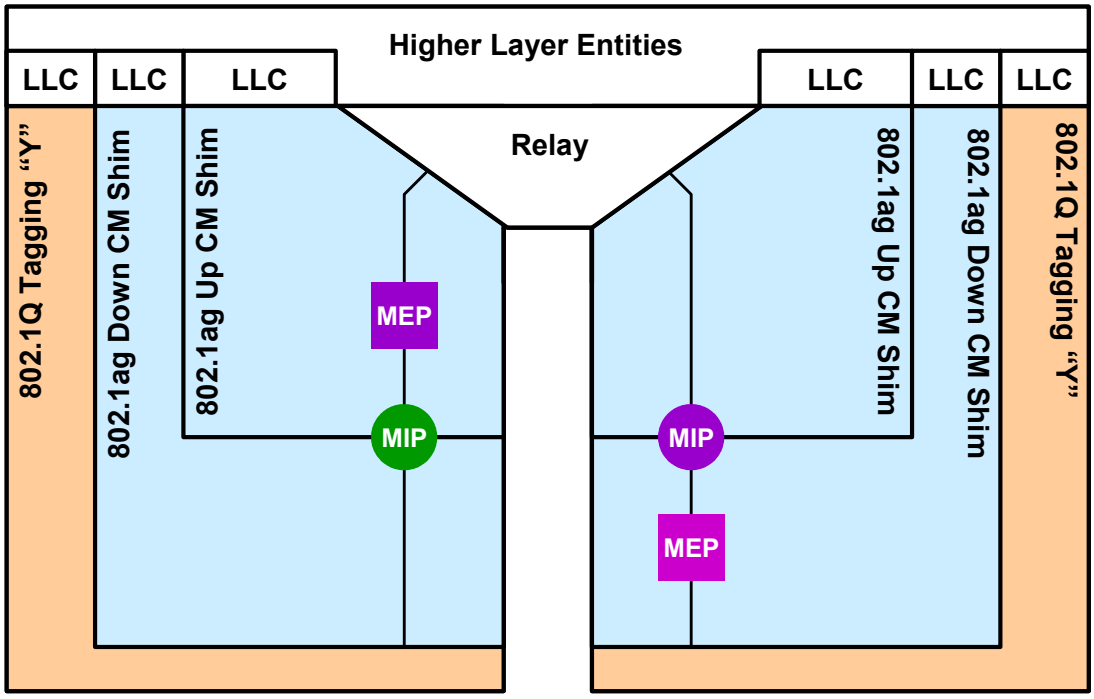


Figure II-14: Provider Bridge 2a
Example without ETH link ME

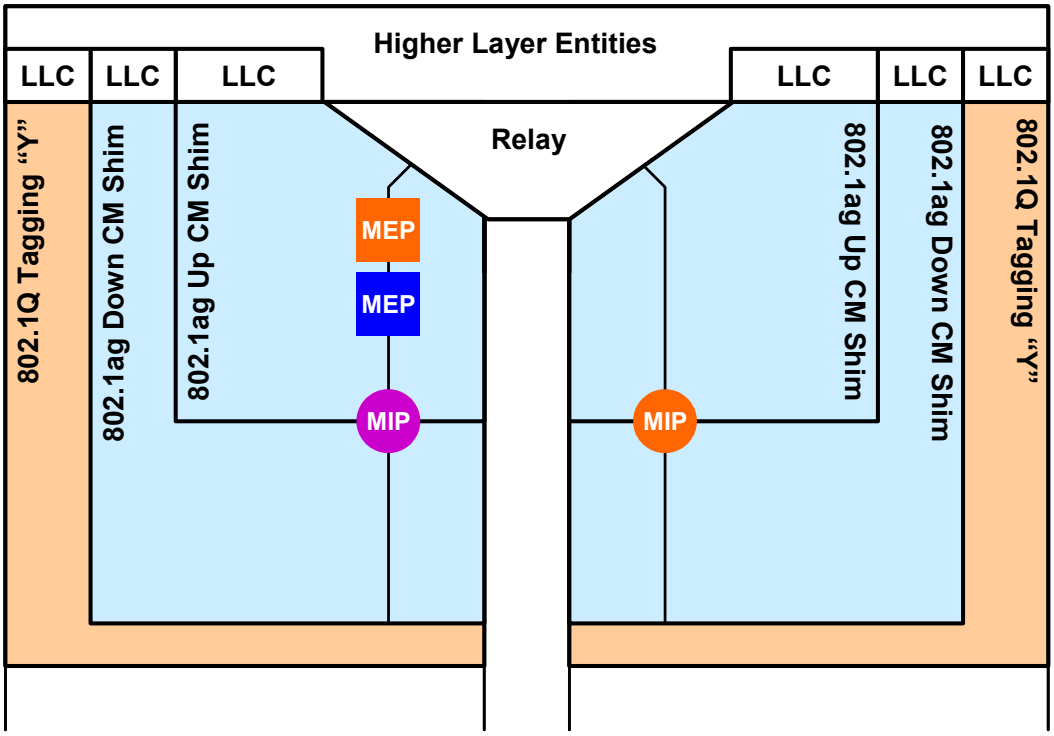


Figure II-15: Provider Bridge 2b, example without ETH link ME
Dual Relay Model with Single Relay as Provider Device

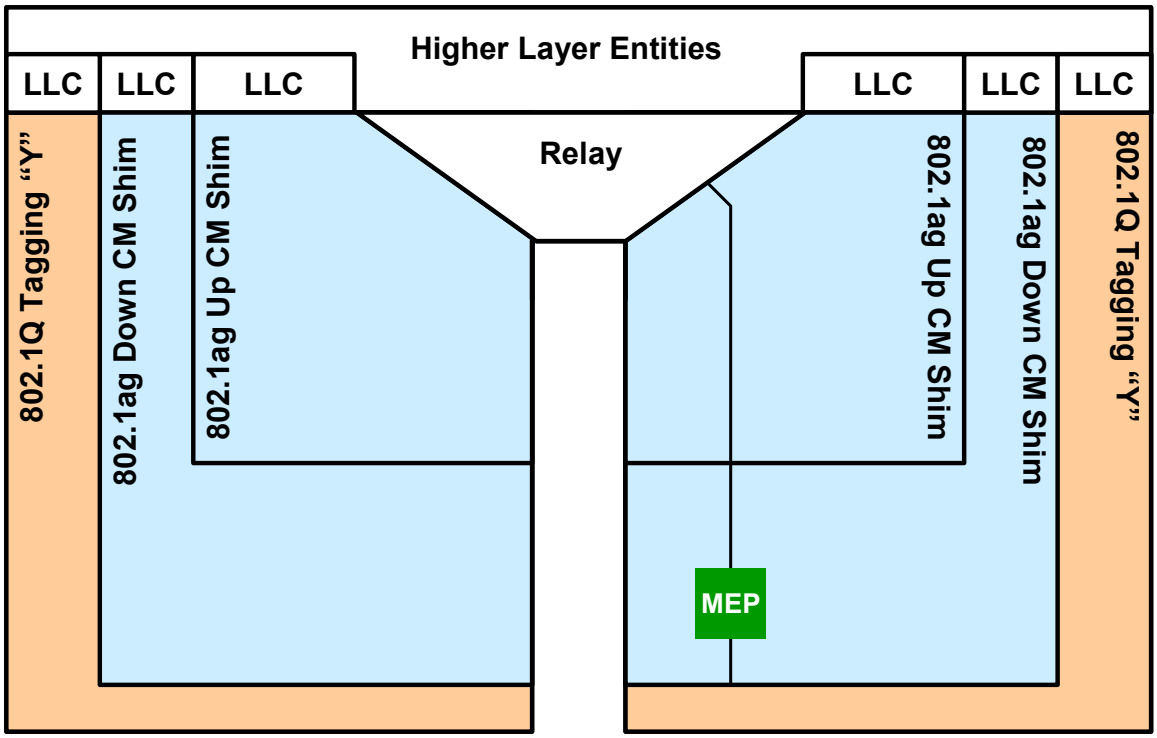


Figure II-16: Customer Bridge 1,
example without ETH link ME

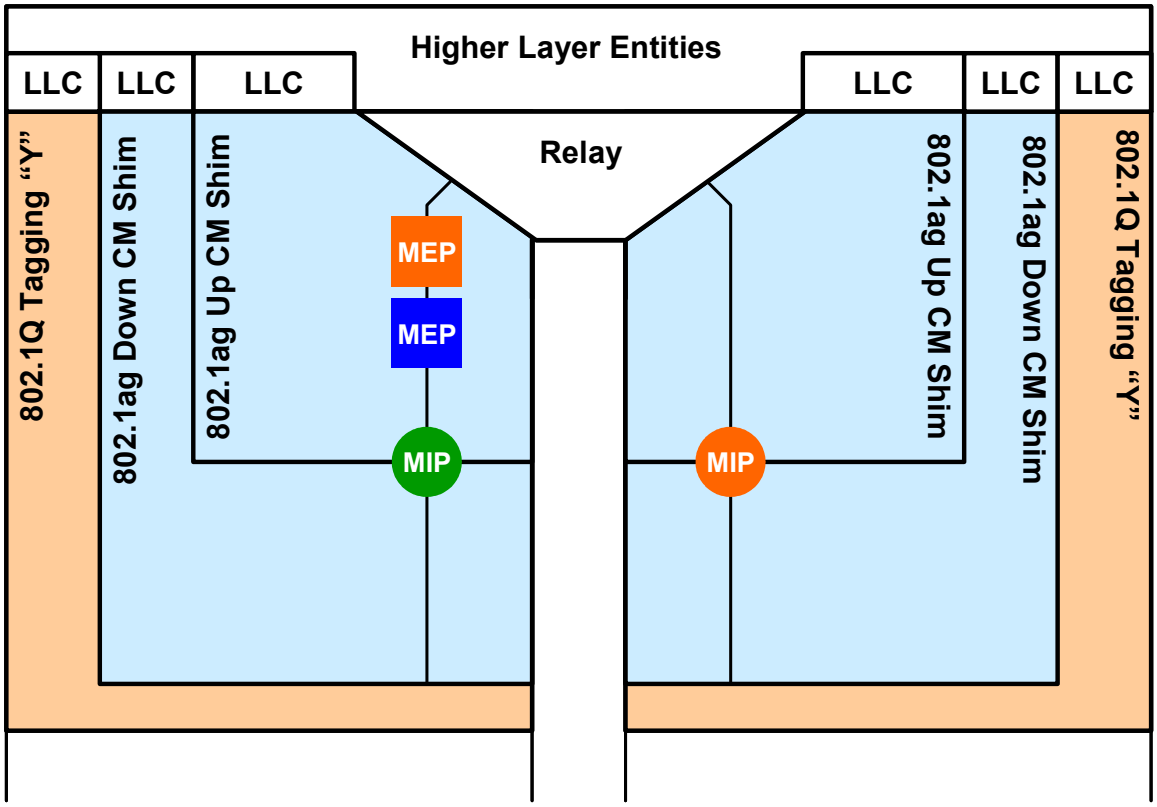


Figure II-17: Provider Bridge 2b, example without ETH link ME
Dual Relay Model with bundling for Single Integrated Provider Device

Appendix III

EDITOR'S NOTE: THE FIRST PART OF THIS APPENDIX (UP TO FIGURE III-7) WAS TAKEN FROM A WD BROUGHT IN NOVEMBER 2003, IT IS NECESSARY TO REVISE IT AND DECIDE WHICH PART OR IF ALL SHOULD BE DELETED OR REPLACED. IT NEEDS ALSO TO BE CONSOLIDATED WITH THE SECOND PART OF THIS APPENDIX (AFTER FIGURE III-7). NECESSARY INPUT FROM DINESH AND MAARTEN AS BOTH WERE THE CONTRIBUTORS

III-1 ETH alarm suppresion OAM considerations (ETH-AS considerations)

WD27 introduces a multipoint ETH connection example in Figures 3 and 4/WD27. WD28 illustrates the ETH-AS insertion points and the ETH maintenance entities present on the ETH links. WD28 also introduces three alternatives to identify the maintenance entity level. Two of these alternatives (MELI ID, STID) are being used in this contribution to analyse the ETH-AS behaviour.

Figure III-1 illustrates the maintenance entities present on some of the links in a multipoint ETH connection (see also WD28).

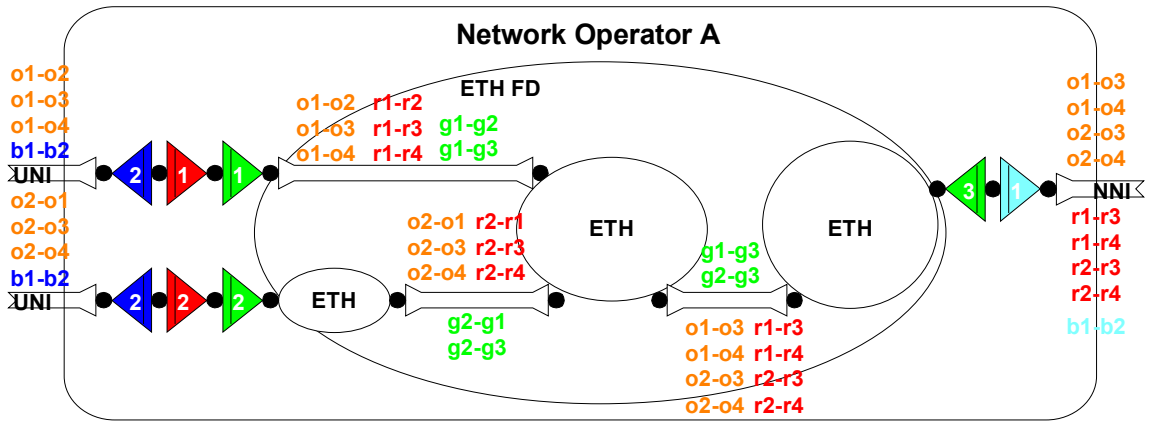


Figure III-1 – ETH maintenance entities on ETH links

III-2 ETH-AS when deploying MELI ID in ETH-CC

When deploying an ETH maintenance entity level instance ID (MELI ID) in ETH-CC OAM frames to identify the maintenance entity level the CC frame belongs to, this MELI ID information can be used at an ETH link end (and an ETH segment end) to learn the set of ETH maintenance entity levels passing through the ETH link and ETH segment. From the port identifier information present in the ETH-CC frames an ETH link end (and an ETH segment end) is able to learn the set of upstream ports that connect through the link or segment. Figure III-2 illustrates this learning at ETH link ends (Srv/ETH(-m)_A_Sk) and ETH segment ends (ETHS/ETH_A_Sk).

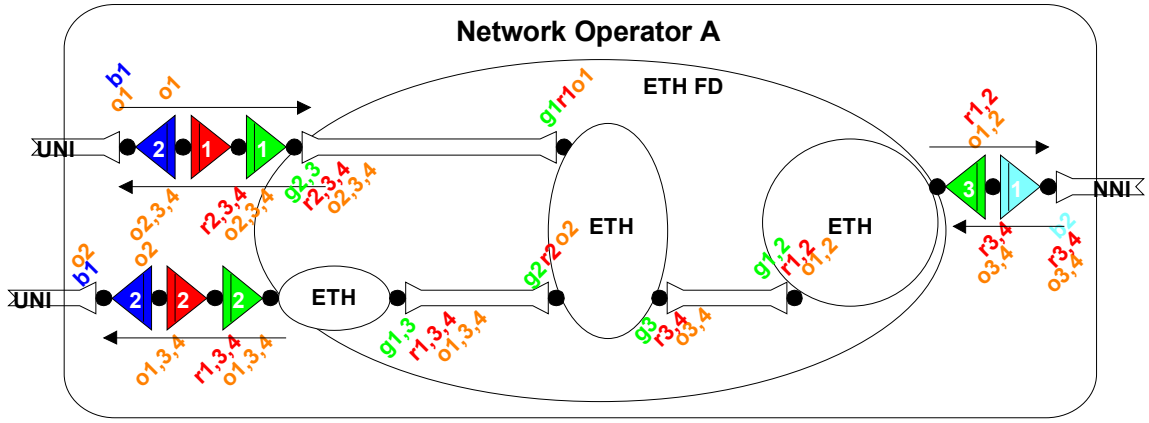


Figure III- 2 – ETH maintenance entity levels and upstream ports learned

Consider a fault occurring in an ETH link in one direction (Figure III-3), a set of maintenance entities (at multiple levels) is impacted. ETH-AS signal generation would in such case use the learned maintenance entity level and upstream port number information and generates ETH-AS frames per maintenance entity level instance, including the set of upstream port numbers.

For the case STP is present, an ETH link fault in a single direction will disable the use of the other direction of the link (for the traffic frames). At this point it is assumed that we will specify a kind of "Reverse Direction Link Down (RDLD)" maintenance signal[2] that runs between a Srv/ETH(-m)_A_So and a Srv/ETH(-m)_A_Sk function to inform the far end of the link that it is down. This signal should then result in ETH-AS signal generation at the far end of the link as well (aAIS = aSSF or dRDLD).

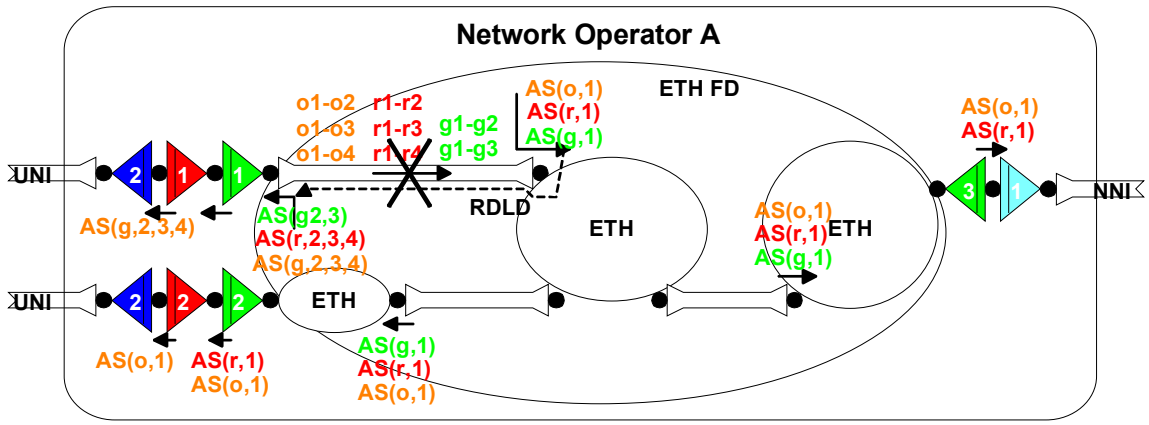


Figure III-3 – ETH-AS insertion example I

The different ETH-AS signals are forwarded[3] by the ETH flow domains and each ETHS_FT_Sk function extracts the ETH-AS signals of its maintenance entity level and processes the included

---

[2] As a first approximation (and perhaps already sufficient), RDI/BDI signals can be used as RDLD signal. The parameter controlling the port state MAC_Operational = CI_SSF or dRDLD. As a first and perhaps sufficient approximation MAC_Operational = AI_TSF or dRDI (from e.g. Sn_TT_Sk or Sn-X-L_TT_Sk).

[3] On a link fault, the port state changes as far as I understand… will this have any impact on the forwarding of these generated and inserted ETH-AS signals?

information (upstream port numbers that are disconnected due to fault). It will use this information to suppress the associated loss of continuity fault causes that will be detected as a consequence of the link fault.

The ETH-AS signals for other maintenance entity levels are simply passed through these ETHS_TT_Sk functions.

Figure III-4 present a second example with a bi-directional ETH link fault. Figure III-5 assumes an alternative link being available in the topology, which is initially blocked by spanning tree (or network management, or …). After ETH link fault is detected e.g. STP will restore the ETH connection by taking the black link part of the active topology. At the same time it will block traffic (including ETH-AS OAM) incoming to the ETH-FDs at the end of the failed link. A blocked port will have to flush their learned set of maintenance entity level instances and upstream port numbers.

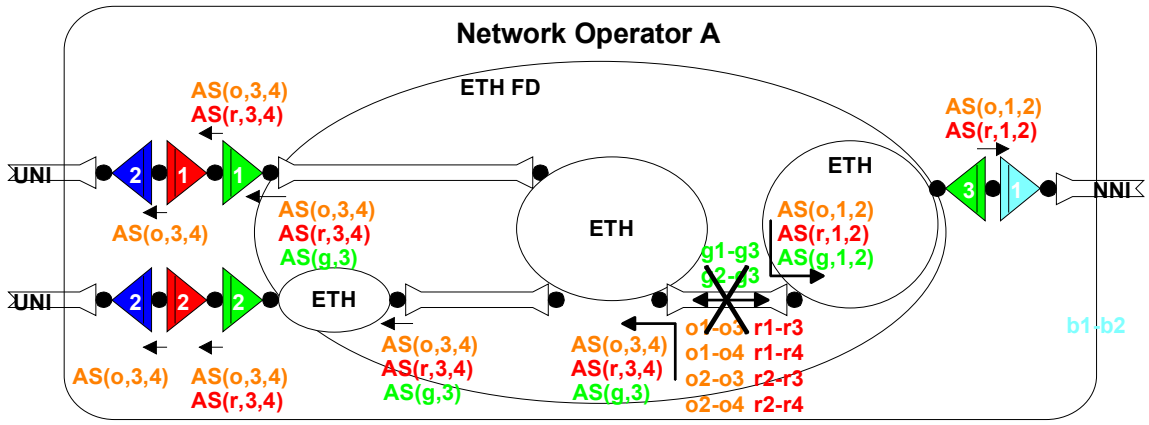


Figure III-4 – ETH-AS insertion example II

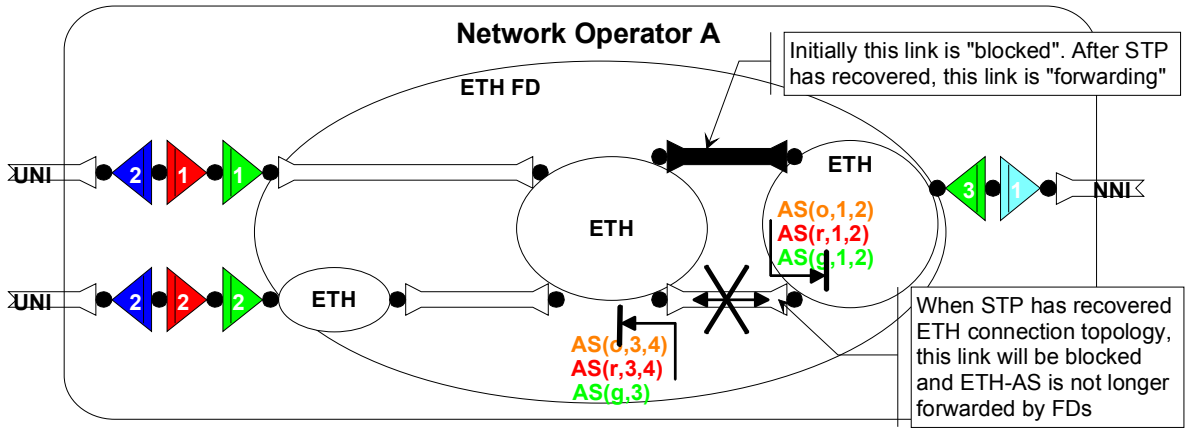


Figure III- 5 – ETH-AS insertion example II with restoration capability

ISSUE: what if the topology only can be partially recovered…

NOTE – if instead of bridges an MPLS (VPLS) network would be used that would run Y.1711 OAM, there would be a look alike, feel alike management behaviour; the ETH maintenance entities are now replaced by MPLS maintenance entities…

III-3 ETH-AS when deploying STID in ETH-CC

Figure III-6 illustrates the port identifiers of the maintenance entity at the top of the stack within a Srv/ETH adaptation sink function (link end) or ETHS/ETH adaptation sink function (segment end) in a multipoint ETH connection. Much less learning is required in this situation, and that is what is attractive… it also has a price…

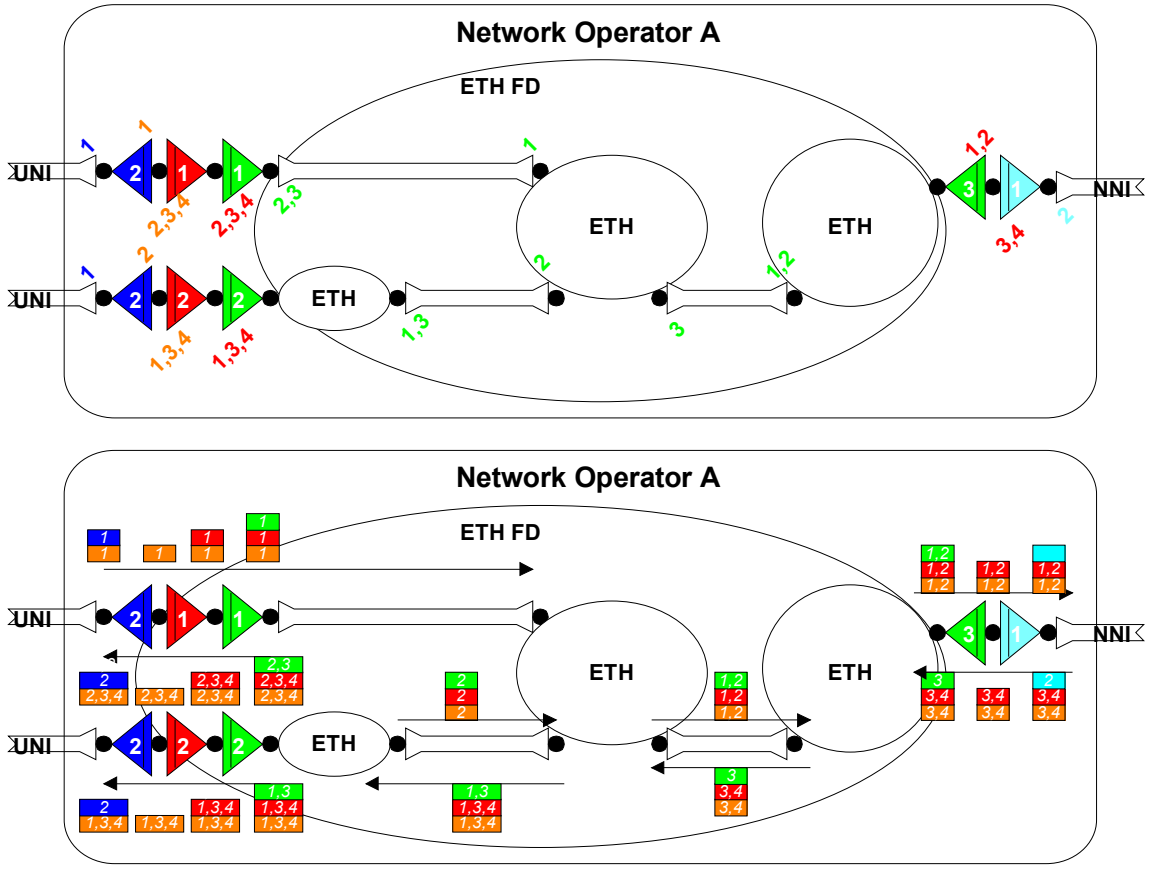


Figure III-6 – ETH maintenance entity port identifiers at the top of the stack (top) and full stack (bottom)

A link fault (Figure III-7) will now generate a single ETH-AS frame with upstream port numbers from the ETH link ends for the top level maintenance entity. Then at the first segment endpoints (green) these ETH-AS signals are extracted and processed. The signal fail status is forwarded to the adaptation sink function in the segment endpoint, where it has to trigger insertion of ETH-AS for the interrupted top level (red) maintenance entity. Unfortunately there is insufficient information at these points to generate ETH-AS frames with specific upstream port number list.

So, should we generate non-specific ETH-AS frames (then also at link ends)? The consequence is that it also will suppress the reporting of a true ETH layer continuity or connectivity fault located elsewhere in the ETH connection… should our ETH OAM be able to detect and report a dual fault condition?

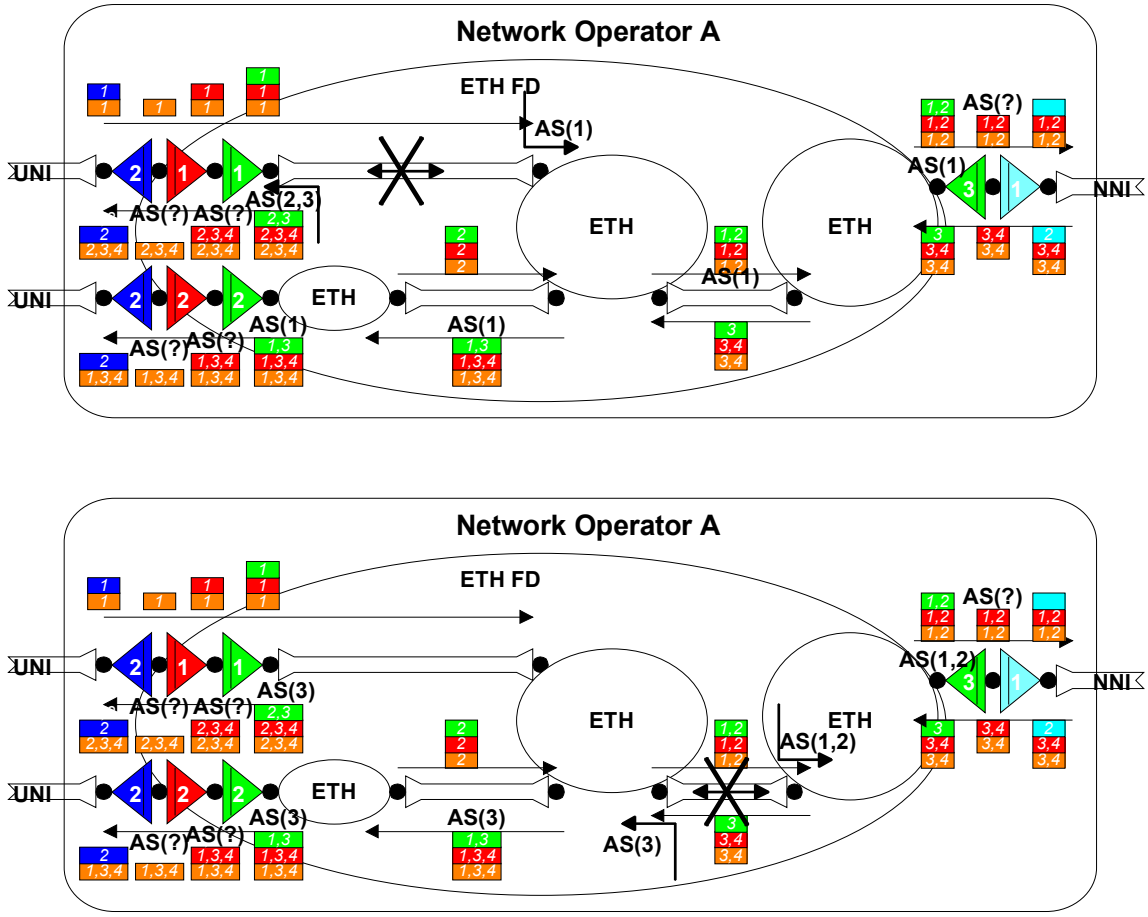

Network Operator A
ETH FD
AS(1)
UNI
UNI
AS(?) AS(?)
AS(2,3)
ETH
ETH
ETH
AS(1)
AS(1)
AS(?)
AS(?)
AS(?)
AS(1)
NNI
Network Operator A
ETH FD
UNI
UNI
AS(3)
AS(?) AS(?)
ETH
ETH
ETH
AS(1,2)
AS(1,2)
AS(?)
AS(?)
AS(3)
AS(3)
AS(3)
AS(?)
NNI

Figure III-7

EDITOR'S NOTE: THE FOLLOWING MATERIAL WAS ADDED FROM WD 12 TO JUNE 7-11 MEETING. IT SHOULD BE CONSOLIDATED WITH THE PREVIOUS MATERIAL: INPUT REQUESTED FROM DINESH AND MAARTEN

Figure III-8 represents a reference network where 3 bidirectional point-to-point services are assumed i.e. S12 (CE1-CE2), S13 (CE1-CE3), and S14 (CE1-CE4). Nodes PE1, PE2, PE3, and PE4 represent the provider edge nodes, while nodes P1, P2, and P3 represent the provider core nodes. The distinction between the core and edge provider nodes is simply that core nodes are not connected to any CE nodes, as per the reference network in Figure III-8.

Since redundancy is shown to exist in the network, links $P2_2$-$P3_3$ and $P1_3$-$P2_1$ may get blocked, either by Spanning Tree Protocol (STP) [3] or manual provisioning. The callouts in Figure 1 represent a view of Forwarding Information Base (FIB).

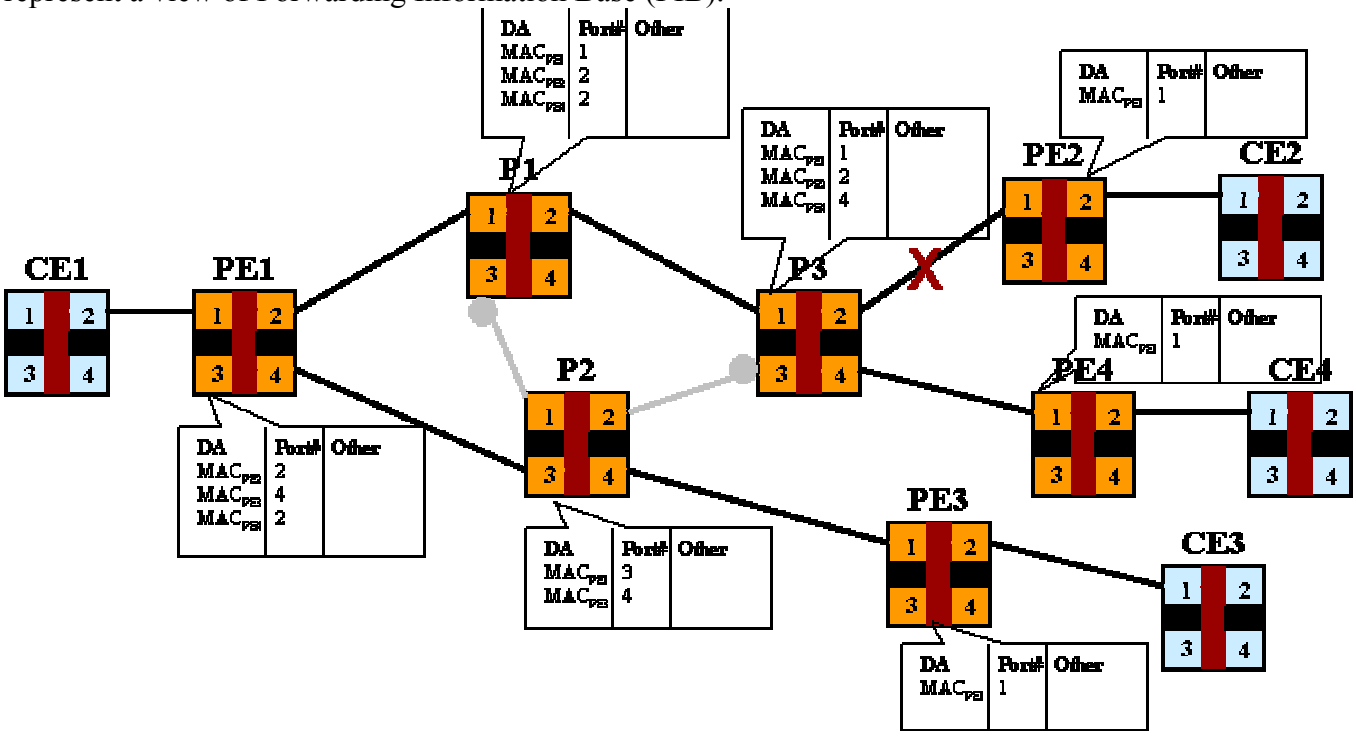


Figure III-8: Connectionless reference network for AIS with link failure scenario #1

A) Link Failure Scenario 1

When a link failure is considered, e.g. link $P3_2PE2_1$, service S12 is affected. Assuming that the link failure is detected on either end of the link, port $P3_2$ and port $PE2_1$ detect this failure. Now the possible options for node P3, if it supports AIS capability, are:

i. Send AIS across all other ports

ii. Send AIS selectively across selective ports

iii. Not send AIS at all

When considering option (i), sending AIS to all ports is not very useful, e.g. PE3 does not have any use for this AIS as the service instance S13 supported by PE3 is not effected by link $P3_2PE2_1$ failure

Option (ii) seems viable as the determination to forward AIS can be made on the basis of service instances e.g. P3 could determine that port $P3_2$ belongs to say VLAN 20, which is also associated with port $P3_1$ for the same point-to-point service instance S12. When sent out across port $P3_1$, the AIS is now received by node P1 across port $P1_2$. Since at P1, only other port associated with same service instance is port $P1_1$, AIS is forwarded to port $P1_1$. Such hop-by-hop forwarding of Ethernet AIS seems pragmatic.

However, one issue may arise when STP or its variants are used which result in flushing of FIBs due to Topology Change Notification (TCN) BPDUs. Under such circumstances hop-by-hop forwarding of AIS is not feasible, as the association of VLANs and corresponding ports on each node is lost due to TCN related flushing.

B) Link Failure Scenario 2

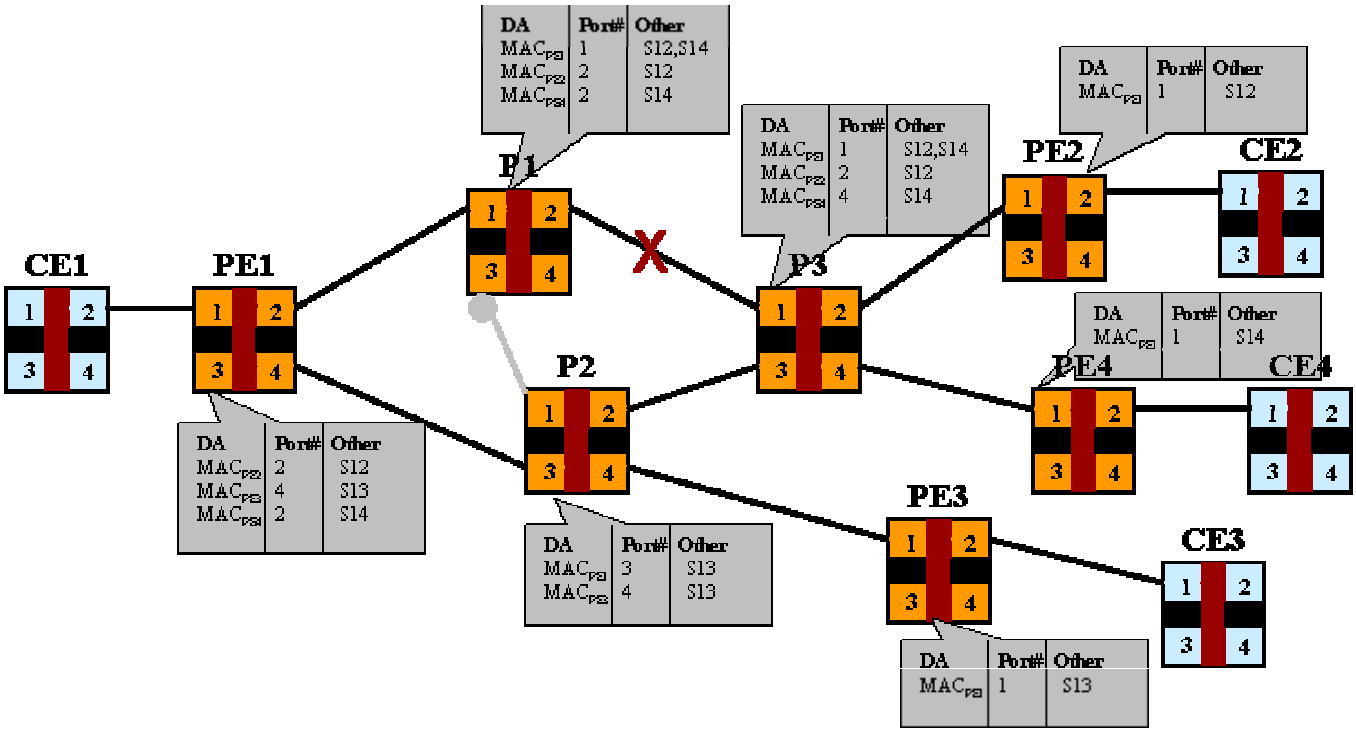


Figure III-9: Connectionless reference network for AIS with link failure scenario #2

When a link failure is considered, e.g. link $P1_2P3_1$, service S12 and S14 are initially affected since link $P2_2P3_3$ is initially blocked, either by STP or its variants or by manual provisioning. However, since this link failure is not a network isolating failure, e.g. link $P2_2P3_3$ is unblocked eventually, and no permanent loss of connectivity is experienced between PE1and PE2 or PE4.

Assuming that the link $P1_2P3_1$ failure is detected on either end of the link, port $P1_2$ and port $P3_1$ detect this failure. Now the possible options for node P3, if it supports AIS capability, are:

i. Send AIS across all other ports

ii. Send AIS selectively across selective ports

iii. Not send AIS at all

Similar to discussions in A), option (ii) is desirable when AIS functionality is supported and required.

However, one issue may arise when link $P2_2P3_3$ is unblocked and port $P3_3$ on node P3 now joins the same service instance as port $P3_1$. Following questions arise:

a) Whether node P3 should forward AIS along ports $P3_2$, $P3_3$, and $P3_4$ or not generate AIS at all i.e. option (iii)?

b) Under what circumstances does the node sending AIS stop sending AIS?

c) If node P3 does send the AIS, what does these AIS mean to node PE2 or PE4 or PE1 since the service is already restored?

d) If node P3 should not send AIS or should stop sending AIS after link $P2_2P3_3$ is unblocked, how does node P3 establish association between the failure and restoration events?

Similarly, when it is assumed from above discussion that node P3 does forward AIS along ports $P3_2$, $P3_3$, and $P3_4$, node P2 is likely to receive both AIS and service frames and other OAM frames (e.g. CC) for the same service instance across the same port. Question arises:

e) Whether node P2 should forward AIS along ports $P2_3$ or should ignore AIS and not forward it?

Further, if now another service instance S23 is created between nodes CE2 and CE3, ports $P3_3$ and $P2_2$ and $P2_4$ are now also associated with S23 service instance. This reflects the need for per service level AIS since otherwise AIS related to link $P1_2P3_1$ failure would get forwarded to node PE3 since port $P3_3$ is now associated with different service instances including S23 and port $P2_2$ is associated with different service instances including S23 as well.

C) Other Issues

Based on the above discussions, it is also important to consider following additional issues:

f) If AIS is required to be generated per service basis, given a single facility could carry thousands of services, the amount of AIS related traffic can be significant, especially around the time when the network has just experienced a fault condition!

g) The above situation is further problematic when the AIS is required to be forwarded along each higher level ME within the network operator, service provider and/or customer domains.

h) Is it always desirable to suppress service level alarms, if the facility level alarms have been detected, OR it is possible that service level alarms are still required independent of network level alarms since the OSS/NMS systems might be set up such.

APPENDIX IV

Some existing Management Objects(MOs) that can be used for the performance management mechanisms mentioned in Section 8 include:

- **IEEE 802.3-2002**
  - – aFramesTransmittedOK [5 – section 5.2.2.1.2]
  - – aFramesReceivedOK [5 - 5.2.2.1.5]

- **IEEE 802.1Q-2003**
  - –Frames Received [6 - 12.6.1.1.3]
  - –Frames Outbound [6 - 12.6.1.1.3]

- **RFC 3635 - Ethernet-like interface MIB (Obsoletes 2665)**
  - –IF-MIB
    - ifOutUCastPkts
    - ifOutMulticastPkts
    - ifOutBroadcastPkts
    - ifOutErrors
    - ifOutDiscards
    - ifInUCastPkts
    - ifInMulticastPkts
    - ifInBroadcastPkts
    - ifInErrors
    - ifInDiscards
  - –aFramesTransmittedOK = ifOutUCastPkts + ifOutMulticastPkts + ifOutBroadcastPkts – (ifOutErrors + ifOutDiscards)
  - –aFramesReceivedOK = ifInUCastPkts + ifInMulticastPkts + ifInBroadcastPkts + (ifInErrors + ifInDiscards)

- **RFC 2674 – VLAN Bridge MIB**
  - –dot1qPortVlanStatisticsTable
    - dot1qTpVlanPortInFrames
    - dot1qTpVlanPortOutFrames

Note: It may be noted that these managed objects values eventually wrap. This can lead to inaccurate results when such an event occurs. However, if the time interval of observation is small, the inaccuracy can be avoided. Averaging of the results over the period of observation can alleviate the in flight frames issue.

APPENDIX V

V-1 Frame Loss Calculations

For the frame loss calculation, the four cases below should be taken into account when counters with finite digits (bits) are used.

A) No wrapping around for both Transmit and Receive Counters

B) Only Transmit Counter wraps around

C) Only Receive Counter wraps around

D) Both Transmit and Receive Counters wrap around

For each case, the frame loss can be calculated as following.

A) No wrapping around for both Transmit and Receive Counters

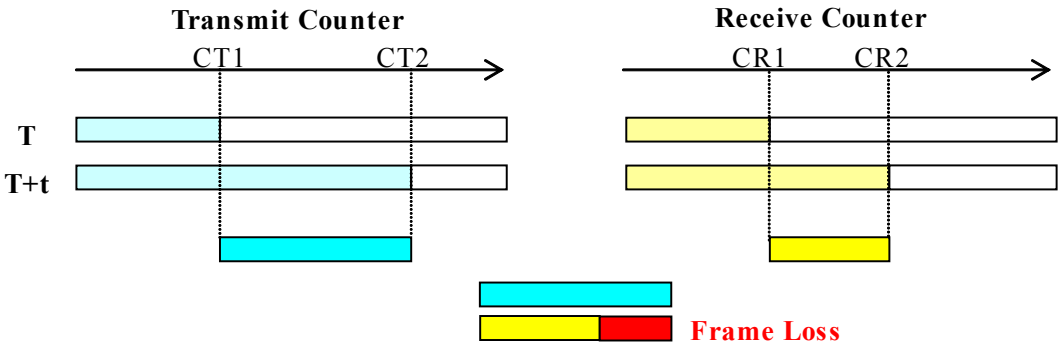


Figure V-1: A) No wrapping around

For this case, the frame loss can be calculated by the simple calculation.

Frame Loss = (CT2 – CT1) – (CR2 – CR1)

B) Only Transmit Counter wraps around

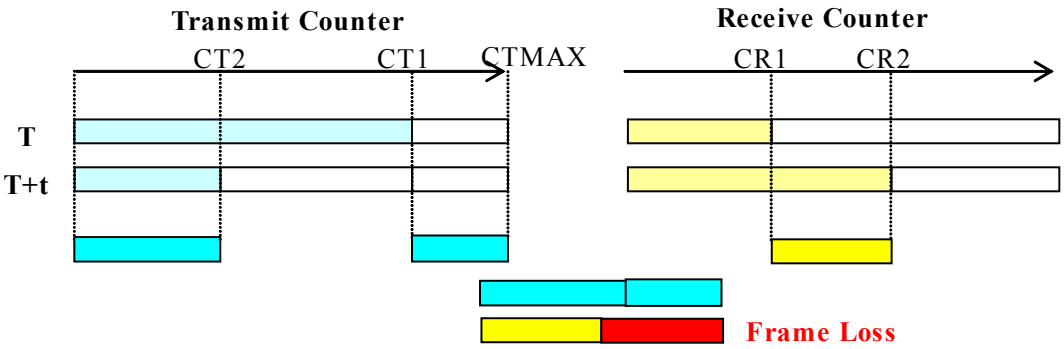


Figure V-2: B) Transmit Counter wraps around

In this case, it can be calculated by the following calculation as is described in the previous section

Frame Loss = ((CTMAX – CT1) + CT2+1) – (CR2 – CR1)

= (CT2 – CT1) – (CR2 – CR1) + (CTMAX+1)

C) Only Receive Counter wraps around

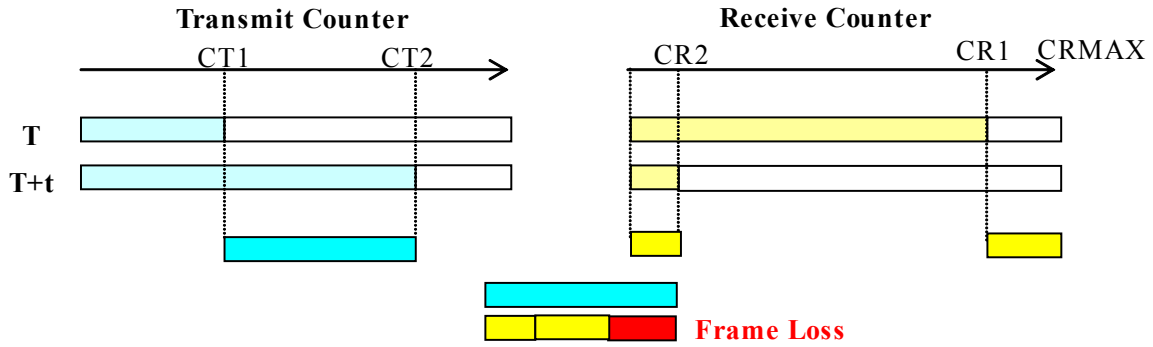


Figure V-3: C) Receive Counter wraps around

Frame Loss = (CT2 – CT1) – ((CRMAX – CR1) + CR2+1)

= (CT2 – CT1) – (CR2 – CR1) - (CRMAX+1)

D) Both Transmit and Receive Counters wrap around

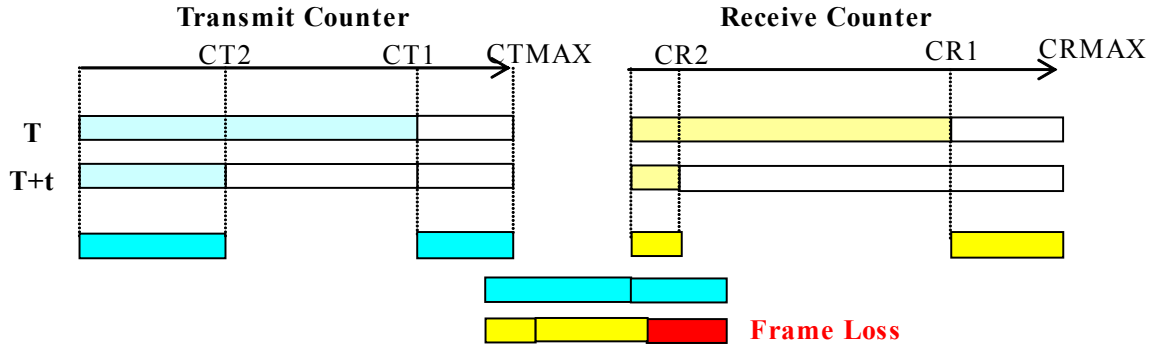


Figure V-4: D) Both Counters wrap around

Frame Loss = ((CTMAX – CT1)+CT2+1) – ((CRMAX – CR1) + CR2+1)

= (CT2 – CT1) – (CR2 – CR1) + (CTMAX+1) - (CRMAX+1)

V-1-1 Simplified calculation for Frame Loss

If the calculation is processed in unsigned value schema, the calculation formula for the frame loss can be greatly simplified by the following characteristics.

N+(MAX+1) ≡N mod(MAX+1)

N-(MAX+1) ≡ N mod(MAX+1)

Therefore each calculation formulas for frame loss which are described in the section 8.2.3 and 8.2.4 can be transformed as below.

A) **Frame Loss = (CT2 - CT1) - (CR2 - CR1)**

B) **Frame Loss** = (CT2 – CT1) – (CR2 – CR1) + CTMAX+1

= ((CT2 + (CTMAX+1)) – CT1) – (CR2 – CR1)

= (CT2 – CT1) – (CR2 – CR1)

C) **Frame Loss** = (CT2 – CT1) – (CR2 – CR1) – (CRMAX+1)

= (CT2 – CT1) – ((CR2 + CRMAX+1) – CR1)

= **(CT2 – CT1) – (CR2 – CR1)**

D) **Frame Loss** = (CT2 – CT1) – (CR2 – CR1) + (CTMAX+1) – (CRMAX+1)

= ((CT2 + (CTMAX+1)) – CT1) – ((CR2 + (CRMAX+1)) – CR1)

= **(CT2 – CT1) – (CR2 – CR1)**

As described above, the frame loss can be calculated by the single calculation formula for any case if it is calculated in unsigned value schema.

If wrapping around of counters happen more than twice, the counters for the wrapping around are required to calculate the frame loss correctly.