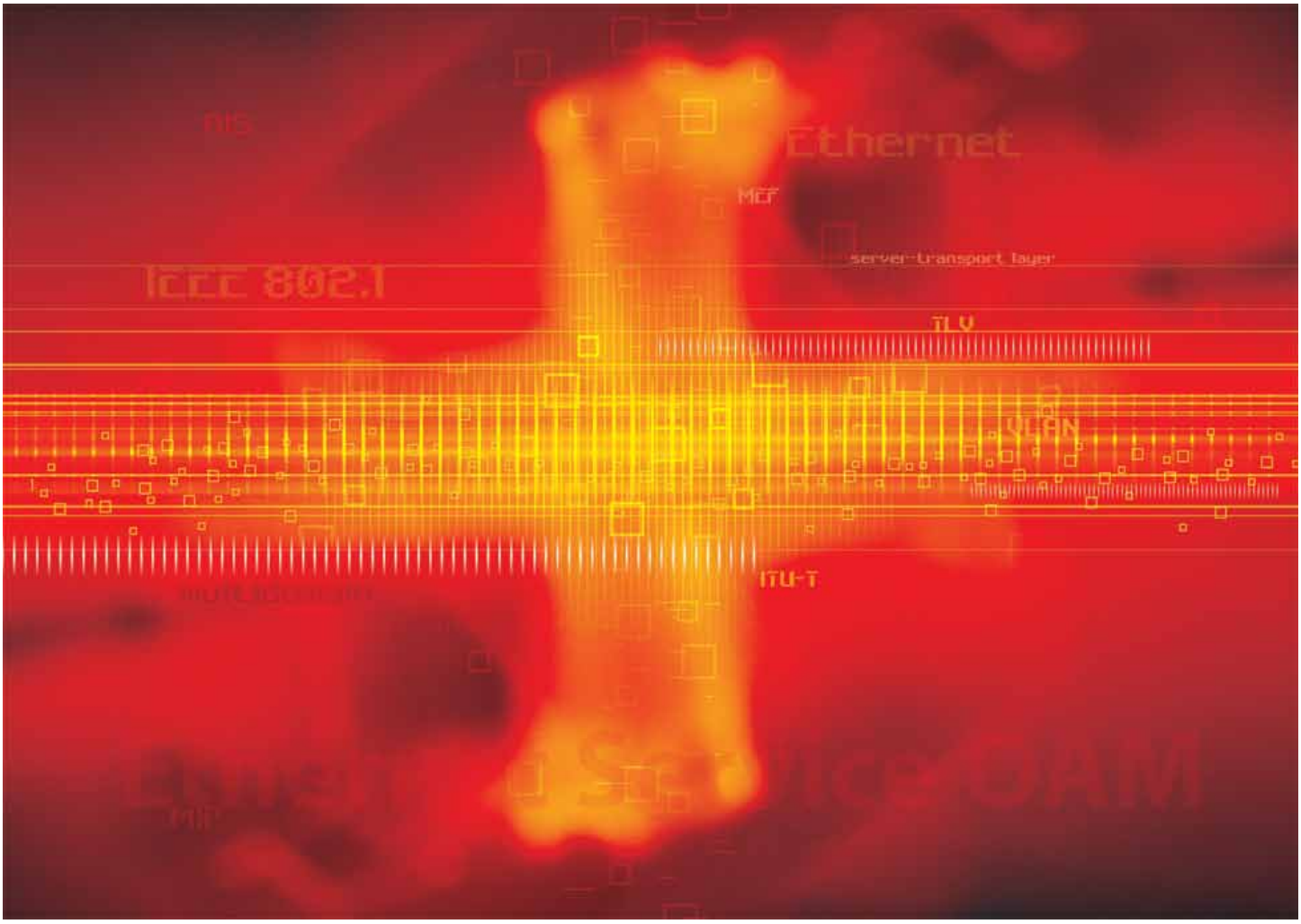
FUJITSU

Ethernet Service OAM: Overview, Applications, Deployment, and Issues



Introduction

As the number of Ethernet service instances grows, service providers will require a robust set of management tools to maintain Ethernet service networks. The positive news is that IEEE 802.1, ITU SG 13, and the MEF have been working in close cooperation to develop complimentary standards for multi-domain Ethernet Service OAM. Ethernet Service OAM encompasses Fault Management and Performance Management capabilities that are incorporated in NEs that support Ethernet services. In a layered network model, these capabilities are active at the Ethernet Service Layer.

The goals of this paper are to provide an overview of relevant standards, summarize the multi-domain model and key capabilities of Ethernet Service OAM, define the major protocol aspects, and provide some examples of how service providers may use this toolkit of OAM functionality. The paper will address how the requisite technology can be integrated into network elements, such as 802.1AD provider edge bridges, and will summarize relevant operational issues and open standards issues.

Ethernet Service OAM – Relevant Standards

The following forums are standardizing the relevant Ethernet Service OAM technology:

- The IEEE 802.1 committee is developing IEEE 802.1ag – Connectivity Fault Management
- The ITU-T SG 13 Q5 WG is developing recommendation Y.1731 – OAM Functions and Mechanisms for Ethernet Based Networks
- The MEF is developing MEF Service OAM Requirements & Framework – Phase 1 Technical Specification

ITU-T SG 13 recently approved Y.1731, however 802.1ag and the MEF standard are still works in progress. This paper will provide a snapshot of where this technology is at today. Open issues and topics not yet resolved will be highlighted. The three committees are working in close cooperation to ensure that the family of Ethernet Service OAM standards is compatible and complimentary. There are differences in the terminology used by IEEE 802.1, ITU-T, and MEF, and this paper will highlight and clarify these differences.

IEEE 802.1 and ITU SG 13 Q5 WG are working together to define a common frame format and protocol elements. IEEE 802.1 is defining the protocol element encoding and OpCodes for a specific set of functionality under CFM. IEEE 802.1 is also defining the detailed implementation of CFM in an 802.1 standard VLAN bridge. ITU-T SG 13 and MEF standards will reference the base CFM functionality and add extensions for additional OAM functionality such as Performance Management and Discovery. ITU-T Y.1731 has been developed to be consistent with the ITU G.8010 Ethernet Layer Network Architecture, and G.8021 Ethernet Equipment recommendations. There are also plans in ITU-T to extend the Ethernet Service OAM protocol to support Ethernet Protection Switching.

The frame format and base functionality is generally agreed across 802.1ag and Y.1731. It is expected that IEEE 802.1ag will be an approved IEEE 802 standard by the end of 2006. The MEF recently agreed to combine previous Fault Management and Performance Management drafts into a common specification. The MEF standard will build on agreements already reached by IEEE 802.1 and ITU-T SG 13. The MEF standard should be also completed by the end of 2006. It is important to note that the standard Ethernet Service OAM protocol is extensible so that phase 2 standards or equipment vendors can add additional functionality.

FUJITSU NETWORK COMMUNICATIONS INC.
2801 Telecom Parkway, Richardson, Texas 75082-3515
Telephone: (972) 690-6000
(800) 777-FAST (U.S.)
us.fujitsu.com/telecom

Ethernet Service OAM – Technology Overview

Multi-domain Network Model

Recognizing the fact that Ethernet networks often encompass multiple administrative domains, IEEE 802.1, ITU-T SG 13, and the MEF have adopted a common multi-domain network model. This model is illustrated in Figure 1.

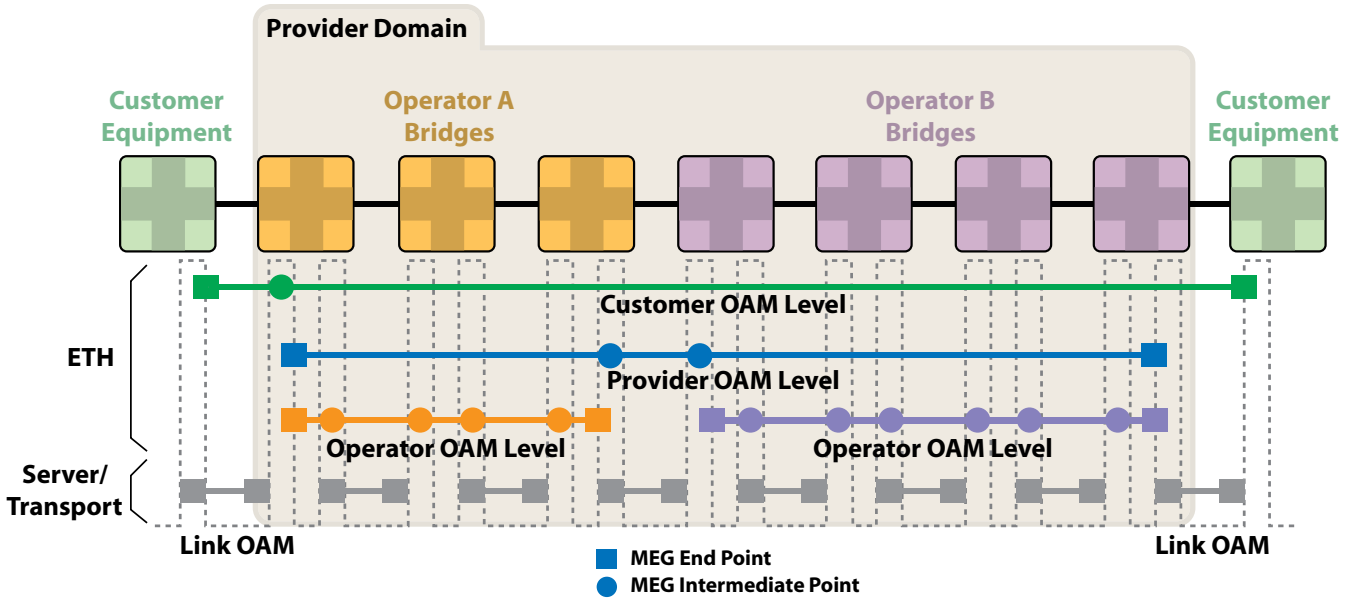


Figure 1: Multi-domain Ethernet Service OAM

The service network is partitioned into customer, provider, and operator maintenance levels. Providers have end-to-end service responsibility. Operators provide service transport across a subnetwork. The ETH Layer consists of customer service Ethernet frames that may include both customer VLAN tags and provider VLAN tags.

The Server/Transport Layer consists of underlying packet transport links. These links may be single hop Ethernet links or multihop MPLS pseudowire or SONET/SDH paths. For pseudowires or SONET/SDH, the Server/Transport layer may consist of multiple sub-layers. Appropriate layer OAM mechanisms are used at these sub-layers. For Ethernet links, IEEE 802.3ah link OAM can be used. For MPLS pseudowires, ITU-T SG 13 has developed recommendation Y.1711, and IETF is developing appropriate layer OAM protocols such as VCCV and MPLS LSP Ping. The focus of this paper is on Ethernet Service Layer OAM, but it will address layer interactions in the course of generating AIS.

ITU-T and the MEF use the following terminology to describe the entities that are managed and the management functional components. An ME is an entity that requires management. A MEG includes a set of MEs that satisfy the following conditions:

- MEs in an MEG exist in the same administrative domain and have the same ME level.
- MEs in an MEG belong to the same service provider VLAN (S-VLAN). In ITU-T terminology, this is a point-to-point or multipoint Ethernet connection.

For a point-to-point Ethernet connection/S-VLAN, an MEG contains a single ME. For a multipoint Ethernet connection, a MEG contains n*(n-1)/2 MEs, where n is the number of Ethernet connection end points. An MEP is a maintenance functional entity that is implemented at the ends of a ME. It generates and receives OAM frames. A ME represents a relationship between two MEPs. An MIP is a maintenance functional entity that is located at intermediate points along the end-to-end path where Ethernet frames are bridged to a set of transport links. It reacts and responds to OAM frames.

Figure 1 illustrates MEP and MIP locations. MEPs are implemented at administrative domain boundaries. Figure 1 also illustrates that for a given S-VAN, an NE port may implement multiple MEPs and MIPs contingent on the number of domain levels. MEP functions may also be used for the server layer packet transport links.

For IEEE 802.1ag, the terms Maintenance Entity, Maintenance Level, and Maintenance Domain have the same meaning as in ITU-T Y.1731. IEEE 802.1ag uses the term MA in the same context as Y.1731 MEG. For IEEE 802.1ag, MEPs and MIPs are the short form of MA End Points and Intermediate Points. They are functionally equivalent to ITU-T MEPs and MIPs.

Fault Management Functions – Overview

Ethernet Service OAM encompasses the following Fault Management functions:

1) **Continuity Check Messages (CCM)** – MEPs periodically exchange Continuity Check OAM messages to detect loss of continuity or incorrect network connections. A CCM is multicast to each MEP in a MA/MEG at each administrative level. CC Messages can also be used to perform two way dual-ended Frame Loss measurements. A Flags field is incorporated in CC Messages. This field includes a bit for Remote Defect Indication (RDI) and an indication of the period at which CC Messages are transmitted.

2) **Loopback Message (LBM)** – MEPs send loopback messages to verify connectivity with another MEP or MIP for a specific MA/MEG. Loopback is a ping-like request/reply function. A MEP sends a loopback request message to another MEP or MIP, which generates a subsequent LRM. LBMs/LRMs are used to verify bidirectional connectivity. They are typically initiated by operator command. However, an MEP can be provisioned to send LBMs periodically. For IEEE 802.1ag, loopback is a unicast OAM message. Y.1731 allows both unicast and multicast loopback.

Loopback can also be used as an out of service diagnostic test. For this applications, which only applies to unicast loopback frames, the loopback OAM PDU also includes a Test Pattern TLV parameter)

3) **Link Trace Message (LTM)** – MEPs multicast LTMs on a particular MA/MEG to identify adjacency relationships with remote MEPs and MIPs at the same administrative level. LTM can also be used for fault isolation. The message body of an LTM includes a destination MAC address of a target MEP that terminates the linktrace. When a MIP or MEP receives an LTM, it generates a unicast LTR to the initiating MEP. It also forwards the LTM to the target MEP destination MAC address. An LTM effectively traces the path to the target MEP.

4) **Alarm Indication Signal (AIS)** – When an MEP detects a connectivity failure at level N, it will multicast AIS in the direction away from the detected failure at the next most superior level. It will multicast AIS on each S-VLAN affected by the failure. AIS serves two purposes: 1) alarm suppression so that an NMS does not receive an excessive number of redundant alarms for a particular fault; 2) Informs clients that a transport path and/or a service instance has failed. For point-to-point S-VLANs/Ethernet connections, there is only one remote peer MEP that cannot be reached. However, for multipoint S-LANs/Ethernet connections, a client layer MEP, upon receiving an AIS, cannot determine which of its remote peers have lost connectivity. It is recommended that for multipoint the client layer MEP should suppress alarms for all peer MEPs.

Use of AIS is not recommended for environments that utilize the spanning tree protocol, which provides an independent restoration capability. Due to the spanning tree and multipoint limitation associated with AIS, IEEE 802.1 committee has chosen not to support AIS in 802.1ag.

5) **Remote Defect Indication (RDI)** – When a downstream MEP detects a defect condition, such as receive signal failure or AIS, it will send an RDI in the opposite upstream direction to its peer MEP or MEPs. This informs the upstream MEPs that there has been a downstream failure. RDI is subject to the same multipoint issue as AIS. A MEP that receives an RDI cannot determine what subset of peer MEPs have experienced a defect. For Y.1711, RDI is encoded as a bit in the Flags field in CC messages. IEEE 802.1ag does not support RDI.

6) **Locked Signal Function (LCK)** – Locked Signal is transmitted by a MEP to communicate intentional administrative or diagnostic actions at that MEP. LCK is sent to all associated client layer (layer N+1) MEPs. It is used for client layer alarm suppression, and enables client MEPs to differentiate between the defect conditions and intentional administrative/diagnostic actions at the server layer MEP. This capability is only supported in Y.1731.

7) **Test Signal (TST)** – Y.1731 has defined a Test Signal message. IEEE 802.1ag CFM does not include this capability. A MEP sends an OAM message that includes test data, which can be used to test throughput, measure bit errors, or detect frames delivered out of sequence. The test data can be configured to be a pseudorandom sequence or a test pattern. The test function is one way only, and can be accomplished either in service or out of service.

8) **Maintenance Communications Channel (MCC)** – MCC provides a maintenance communications channel between a pair of MEPs. This channel can be used to perform remote maintenance, such as requesting remote maintenance from a peer MEP. Specific applications and the protocol for the MCC are not defined. The PDU includes an OUI, which encodes an organization specific use of the MCC. MCC capability is only supported in Y.1731.

9) **Vendor Specific and Experimental OAM (VSM and VSR/EXM and EXR)** – Y.1731 has reserved two of its OpCode points for Vendor Specific and Experimental OAM frames. Vendor Specific allows for vendor value add Ethernet OAM extensions. Experimental allows for functionality, which can be used within an administrative domain on a temporary basis. Both of these OAM frame types include an OUI field to identify a specific vendor or administration.

Performance Management Functions – Overview

To date, only Y.1731 has defined the ability to measure performance parameters. The MEF plans to address performance management requirements that are not addressed by ITU-T or IEEE 802.1ag. For Phase 1 Ethernet Service OAM, the measurement of performance management parameters is limited to point-to-point MA/MEG. Y.1731 uses the same performance parameter definitions as used in the MEF 10 Standard, "Ethernet Service Attributes Phase 1."

The following performance parameters are measured by appropriate OAM messages:

1) **Frame Loss Ratio (FLR)** – FLR is defined as a ratio, expressed as a percentage, of the number of service frames not delivered divided by the total number of service frames during time interval T, where the number of service frames not delivered is the difference between the number of service frames sent to an ingress UNI and the number of service frames received at an egress UNI.

Two types of FLR measurement are possible, Dual-ended LM and Single-ended LM. Dual-ended LM is accomplished by exchanging CCM OAM frames that include appropriate counts of frames transmitted and frames received. These counts do not include OAM frames at the MEPs ME Level. Dual-ended LM enables the proactive measurement of both Near End and Far End FLR at each end of a MEG.

Single-ended LM is accomplished by the on-demand exchange of LMM and LMR OAM frames. These frames include appropriate counts of frames transmitted and received. Single-ended LM only provides Near End and Far End FLR at the end that initiated the LM Request.

2) **Frame Delay (FD)** – FD is specified as round trip delay for a frame, where FD is defined as the time elapsed since the start of transmission of the first bit of the frame by a source node until the reception of the last bit of the loop backed frame by the same source node, when the loopback is performed at the frame's destination node.

3) **Frame Delay Variation (FDV)** – FDV is a measure of the variations in the FD between a pair of service frames, where the service frames belong to the same CoS instance on a point-to-point ETH connection.

There are two types of FD measurements, One-way and Two-way. One-way FD is measured by MEPs periodically sending 1DM frames, which include appropriate Transmit Time Stamps. FD is calculated at the receiving MEP by taking the difference between the Transmit Time Stamp and a Receive Time Stamp, which is created when the 1DM frame is received. One-way DM requires synchronized clocks between the two MEPs.

Two-way DM measures round trip delay and does not require synchronized clocks. It is accomplished by MEPs exchanging DMM and DMR frames. Each of these DM OAM frames includes Transmit Time Stamps. Y.1731 allows an option for inclusion of additional time stamps such as a Receive Time Stamp and a return Transmit Time Stamp. These additional time stamps compensate for DMR processing time.

FDV is calculated exactly as the difference between two consecutive Two-way FD measurements.

Ethernet Service OAM Frame Format and Protocol Elements

Figure 2 illustrates the common Ethernet Service OAM frame format. Each specific OAM message type will add additional fields to the common PDU format.

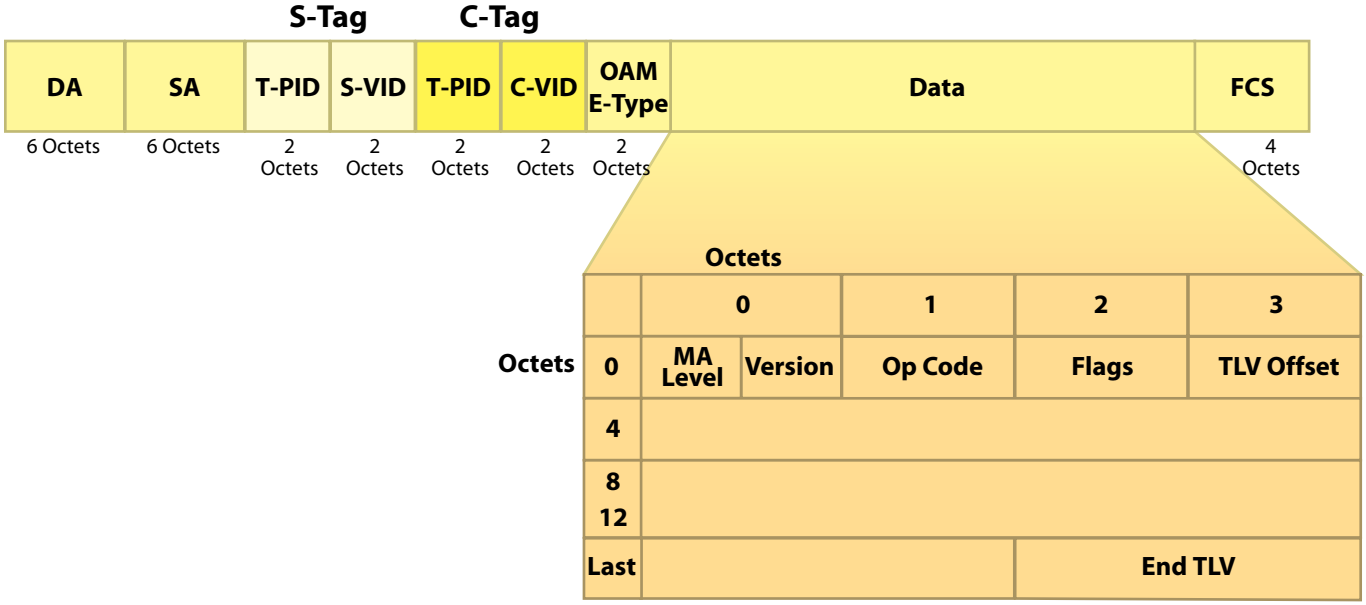


Figure 2: Common Ethernet Service OAM Frame Format

As with all Ethernet frames, the destination and source MAC address (DA/SA), is preceded by a seven octet preamble and a one octet start of frame delimiter. The frame may or may not include a customer VLAN tag. It may also include some form of service provider Ethernet transport connection delineation tag, such as an 802.1ad Provider Bridge S-Tag. A new OAM Ethertype will be assigned for this application by IEEE 802 committee. IEEE 802.1ag also supports the LLC/SNAP encoded frame format, which includes a LLC header in addition to the OAM Ethertype.

The Version field facilitates the development of future extensions to the initial Ethernet Service OAM protocol. The MA level corresponds to the administrative domains shown in Figure 1. A numerically higher MA level corresponds to domains with the greatest physical reach. MA levels 5 through 7 are reserved for customer domains, MA levels 3 and 4 are reserved for service provider domains, and MA levels 0 through 2 are reserved for operator domains.

OpCodes indicate the OAM message type, Continuity Check, Loopback, etc. OpCodes 0 - 31 are assigned by IEEE 802.1. Continuity Check, Loopback, and Link Trace use OpCodes in this range, and the OpCode assignments are the same for Y.1711 and 802.1ag. OpCodes 32 - 63 are assigned by ITU SG 13. Performance Management functions, which are only supported in Y.1731, fall into this range. OpCodes 64 - 255 are assigned by IEEE 802.1.

Flags is an 8-bit field. Use of the bits in this field is dependent on the OAM PDU type. TLV Offset is a 1-octet field, which contains the offset to the first TLV in an OAM PDU relative to the TLV Offset field.

TLVs are type, length, and value, encoded parameters that are included in the message body. For each TLV, there is a one octet type field, followed by a two octet length field and an N octet value field. The End TLV has a Type field equal to 0 and the Length and Value fields are not used.

Each OAM message type adds additional Information Elements.

1) CCM – MEG ID, MEP ID and Transmission Period. MEG ID is 48 octets, globally unique, and identifies the network operator that is responsible for the MEG. The MEP ID is a two octet field whose 13 least significant bits identify the MEP within the MEG. Transmission Period is encoded in the 3 least significant bits of the Flags field, and can be in the range of 3.3 ms to 10 minutes.

RDI is encoded as 1 bit in the flags field. When CCM is used to support Dual-ended Loss Measurement, the PDU includes the following Information Elements

- TxFCf - is a 4-octet field which carries the value of the counter of in-profile data frames transmitted by the MEP towards its peer MEP, at the time of CCM frame transmission.
- RxFCb - is a 4-octet field which carries the value of the counter of data frames received by the MEP from its peer MEP, at the time of receiving the last CCM frame from that peer MEP.
- TxFCb - is a 4-octet field which carries the value of the TxFCf field in the last CCM frame received by the MEP from its peer MEP.

2) LBM/LBR – Transaction ID / Sequence Number is mandatory and Data / Test Pattern TLV is optional. Transaction ID/Sequence Number is a 4-octet field that contains the transaction ID/sequence number for the LBM. The length and contents of the Data / Test Pattern TLV are determined by the transmitting MEP.

3) LTM – Transaction ID, TTL, Origin MAC address and Target MAC address. TTL is the number of hops remaining in the linktrace. It is decremented by one by each MEP or MIP along the path. If the TTL is 0 on input to an MEP/MIP, it is discarded. If TTL is 0 on output from an MEP/MIP, it is not forwarded. The Origin MAC address is the source MAC address of the MEP that originated the LTM. The Target MAC address is the MAC address of the MEP at the end of the path, which is being traced. The LTM will not be forwarded beyond this point. The flags field contains one bit known as "Hardware Only". If this bit is set, then only MAC addresses learned in a bridge's active forwarding tables (and not information saved in software) is to be used to forward LTM.

4) LTR – Transaction ID and TTL and two TLVs that are reserved for 802.1ag. There is also one octet called Relay Action, which is also reserved but not supported by Y.1731. It is supported by 802.1ag. Relay Action includes 4 flag bits which report how a data frame sent to the target LTM address would pass through the MAC relay entity to the egress bridge port (e.g. port blocked, target address in forwarding data base, etc.)

The reserved TLVs are Reply Ingress and Reply Egress. The Reply Ingress TLV is only returned if the LTM is received by a bridge port that includes a MEP / MIP at the correct level. The Reply Egress TLV is only returned if the egress bridge port, onto which the LTM is relayed, includes a MEP / MIP at the correct level. The Reply Ingress/Egress TLVs include port MAC address and a set of flags, which indicated how a data frame to the target address would be forwarded.

5) AIS – Transmission Period encoded the same as a CCM period

6) LCK - Transmission Period encoded the same as a CCM period

7) TST – Sequence Number and Test Data/Pattern

8) MCC – OUI, SubOpCode, and MCC Data. OUI is an Organizationally Unique Identifier of the organization defining the format of MCC Data and the values of the SubOpCode

9) LMM – TxFCf

10) LMR – TxFCf, TxFCb, RxFCf. These counters are defined as
- TxFCf is a 4-octet field, which carries the value of the TxFCf field in the last LMM PDU received by the MEP from its peer MEP.
- TxFCb is a 4-octet field, which carries the value of the counter of in-profile data frames transmitted by the MEP towards its peer MEP at the time of LMR frame transmission.
- RxFCf is a 4-octet field, which carries the value of the counter of data frames received by the MEP from its peer MEP, at the time of receiving last LMM frame from that peer MEP.

11) 1DM – TxTimeStampf - an 8-octet field that contains the timestamp of 1DM transmission. The format of TxTimeStampf is equal to the Time Representation format in IEEE 1588-2002

12) DMM – TxTimeStampf

13) DMR – TxTimeStampf, RxTimeStampf, TxTimeStampb. These time stamps are defined as follows:
- TxTimeStampf is an 8-octet field that contains the copy of TxTimeStampf field in received DMM.
- RxTimeStampf is an optional 8-octet field that contains the timestamp of DMM reception. The format of RxTimeStampf is equal to the Time Representation format in IEEE 1588-2002.
- TxTimeStampb is an optional 8-octet field that contains the timestamp of DMR transmission. The format of TxTimeStampb is equal to the Time Representation format in IEEE 1588-2002.

14) EXM/EXR – OUI, SubOpCode, and EXM Data. OUI identifies the organization using this OpCode.

15) VSM/VSR – OUI, SubOpCode, and Vendor Specific Data. OUI identifies the vendor that is using this OpCode.

OAM Addressing and Message Processing

Addressing

IEEE 802.1ag supports two addressing modes: the Bridge Port Model and the Master Port Model. In the Bridge Port Model, MEPs and MIPs assume the same MAC address as the bridge port. In this model, MEPs and MIPs are implemented in a shim layer in each bridge port. This model supports the most management functionality and flexibility. For the Master Port Model, MEPs are implemented in a logical bridge master port, which could be a control CPU. All MEPs use the same master port MAC address. These Master Port MEPs have some functional limitations such as an ambiguity in the identification of a MEP to which a loopback reply is destined. The principal reason for this model is that the NE is a legacy device, which cannot support port-based service OAM. With respect to MIPs and the Master Port Model there are two options: 1) do not implement MIPs, which adds more functional limitations (e.g. no linktrace); and 2) implement MIPs in a port-based shim layer and they can use the Master Port MAC address. The issue with the second approach is that it partially defeats the purpose of the Master Port Model. Clearly the Port Addressing Model is preferred. This model will be followed for the remainder of this section.

The source MAC address for all OAM frames is always a unicast MAC address. The destination MAC address may be either a unicast or a multicast address dependent on the message type and application. Two types of multicast MAC addresses have been assigned for Ethernet OAM, Multicast DA Class 1, which address all MEPs in a MEG and Multicast DA Class 2, which address all MEPs and MIPs in a MEG. Multicast DA Type 2 is only used for Link Trace Messages. Multicast DA Class 1 is used for all other applications that require a multicast DA.

IEEE 802.1ag allows an additional addressing option to grandfather legacy equipment. This option, which is not recommended for new equipment, allows the multicast DA to encode the MEG level.

Continuity Check Messages

CCMs are multicast to all MIPs and MEPs associated with a given MA/MEG. Use of a multicast DA allows for discovery of remote MEP MAC addresses and the detection of network misconnections. Use of a unicast MAC DA is also allowed if the detection of misconnections is not required. Every MEP transmits a CCM on its associated Ethernet connection at its configured transmission rate.

The CCM transmission interval can range from 3.3 msec to 10 minutes. If the port associated with an MEP experiences a fault condition, the MEP will encode RDI in the flags field.

MEPs are configured (MEG ID and MEP ID) with a list of all the other MEPs in their maintenance level.. Every active MEP maintains a CCM database. As an MEP receives CCMs, it catalogues them in the database indexed by MEP ID. If no CCM frames from a peer MEP are received within the interval equal to 3.5 times the receiving MEP's CCM transmission period, loss of continuity with peer MEP is detected. In addition to loss of continuity, exchange of CCMs between MEPs in a MEG allow for the detection of the following additional defects:

- If a CCM frame with a MEG Level lower than the receiving MEP's MEG Level is received, Unexpected MEG Level is detected.
- If a CCM frame with same MEG Level but with a MEG ID different than the receiving MEP's own MEG ID is received, Mismerge is detected.
- If a CCM frame with the same MEG Level and a correct MEG ID but with an incorrect MEP ID, including receiving MEP's own MEP ID, is received, Unexpected MEP is detected.
- If a CCM frame is received with a correct MEG Level, a correct MEG ID, a correct MEP ID, but with a period field value different than the receiving MEP's own CCM transmission period, Unexpected Period is detected.

Loopback Message/Loopback Reply

There are two Loopback applications, verification of bidirectional connectivity with peer MEP or MIP and a bidirectional diagnostic test between a pair of MEPs. Both of these applications are on demand and are not a continuous OAM exchange like CCM.

For the connectivity test LBMs can be transmitted with either a unicast or a multicast DA. This address can be learned from the CCM database. The unicast DA can address either a MEP or a MIP. The multicast DA is only used to address MEPs. The LBM includes a Transaction ID/Sequence Number, which is retained by the transmitting MEP for at least five seconds. After Unicast LBM frame transmission, a MEP expects to receive a Unicast LBR frame, with the same Transaction ID / Sequence Number within 5 seconds.

When an LBM is received by a remote MEP/MIP, that matches its address, a LBR will be generated. Every field in the LBM is copied to the LBR with the exception that: 1) the source and destination MAC addresses are swapped and 2) the OpCode field is changed from LBM to LBR. The Transaction ID/Sequence Number and Data TLV fields are returned to the originating MEP unchanged. These fields are verified by the originating MEP. For multipoint Loopback, each MEP returns a LBR after a randomized delay.

A Loopback diagnostic text is only for point-to-point applications between MEPs, and uses unicast destination MAC addresses. The LBM includes a Test Pattern and the LBR returns the same Test Pattern.

Linktrace Message/Linktrace Reply
On an on demand basis, a MEP will multicast LTM on its associated Ethernet connection. The multicast destination MAC address is a Class 2 multicast address. The Transaction ID, TTL, Origin MAC address and Target MAC address are encoded in the LTM PDU. The target MAC address is the address of the MEP at the end of the Ethernet connection, which is being traced. It can be learned from CCM. The Origin MAC address is the address of the MEP that initiates the linktrace. The transaction ID/Sequence number and target MAC address are retained for at least five seconds after the LTM is transmitted. This is for comparison with the linktrace reply.

When the LTM is received by a bridge port, its TTL and target MAC address fields are checked on the path being traced. If TTL > 0 and if a data frame with the target MAC Address would pass through an MEP or MIP on a ingress or egress port and not be filtered, then an LTR will be sent to the originating MEP. In addition, the LTM will be forwarded out the same egress port that would be used for a data frame with the target MAC address if the TTL is > 1. The TTL will be decremented by 1. The LTR is sent after a random delay with a unicast destination MAC address to the MEP, which originated the linktrace. This destination MAC address matches the Origin MAC address field in the LTM Origin field and the Transaction ID/Sequence Number is copied from the LTM. For implementations, which support 802.1ag, a Relay Action field and Reply Ingress and Reply Egress TLVs are added. The MEP that originated the LTM checks returned LTR for a correct Transaction ID/Sequence Number.

AIS

Upon detecting a defect condition the MEP can immediately start transmitting periodic frames with ETH-AIS information at a configured client MEG Level. AIS is generally transmitted with a Class 1 multicast DA. A unicast DA is allowed for point-to-point applications. The Transmission Period can be in the range of 3.3 msec to 10 minutes. A Transmission Period of 1 second is recommended. A MEP continues to transmit periodic frames with ETH-AIS information until the defect condition is removed. Upon receiving a frame with ETH-AIS information from its server (sub) layer, a client (sub) layer MEP detects AIS condition and suppresses alarms associated with all its peer MEPs. Following detection of AIS defect condition, if no AIS frames are received within an interval of 3.5 times the AIS transmission period, the MEP clears AIS defect condition.

Lock (LCK)
LCK functions like AIS with the exception that it is initiated when a MEP is configured for an intentional administrative/diagnostic action that disrupts the data traffic. The LCK Transmission Period is the same as the AIS Transmission Period.

Text (TST)
TST OAM messages are generally sent with a unicast DA. The use of a Class 1 multicast DA is also allowed for multipoint testing. TST is a one-way diagnostic function, which can be done either in service or out of service. If done in service, the repetition rate must not be disruptive of client layer traffic. If done out of service the affected MEPs will initiate LCK messages. The TST OAM PDU includes a Transaction ID / Sequence number and also typically includes pseudorandom test data, which is checked for bit errors by the receiving MEP.

Management Communications Channel (MCC)
MCC OAM PDUs are typically sent with a unicast DA. For cases where a dedicated point-to-point VLAN is used for the MCC, a Class 1 multicast DA may be used. The MCC message content and processing is organization specific and is not defined.

Experimental and Vendor Specific
Addressing and message processing is vendor and organization specific and is outside the scope of Y.1731

Loss Measurement
Generally Frame Loss Ratio (FLR) measurement is done on a point-to-point basis and uses a unicast DA. However Y.1731 allows the use of Class 1 multicast DA to support multipoint testing. For Dual-ended Loss Measurement, CCM OAM frames are used. These frames include Information Elements for TxFCf, RxFCb, and TxFCb as previously defined. FLR is calculated across pairs of consecutive frames which compensates for the lack of synchronization across the initial counter values. Near End and Far End FLR is calculated from the following equations:

Frame Loss (far-end) = |TxFCb[tc] – TxFCb[tp]| - |RxFCb[tc] – RxFCb[tp]|
Frame Loss (near-end) = |TxFCf[tc] – TxFCf[tp]| - |RxFCl[tc] – RxFCl[tp]|

tc = Counter values for the current CCM frame
tp = Counter values for the previous CCM frame

Delay Measurement
With respect to addressing delay measurement is the same as loss measurement. Delay measurement can be accomplished on either a one-way or round trip basis. One-way delay measurement, by a 1DM message, requires synchronized clocks between the transmitting and receiving MEPs to achieve an accurate measurement. The one-way delay is calculated by

Frame Delay = RxTimef – TxTimeStampf

RxTimef = time that the 1DM PDU was received
TxTimeStampf = time stamp at the time the 1DM PDU was sent

Two-way delay measurement avoids the clock synchronization issue, but could incur inaccuracy due to the DMM to DMR processing in the target MEP. Consequently Y.1711 allows for two options in the measurement of two-way delay. If the target MAP turn around delay is not considered significant then the round trip delay can be calculated by

Frame Delay = RxTimeb – TxTimeStampf

RxTimeb = time that the DMR PDU is received by the initiating MEP

A more accurate two-way delay measurement can be achieved if the target MEP turn around delay is subtracted out. In this cast the round trip delay can be calculated by

Frame Delay = (RxTimeb–TxTimeStampf)–(TxTimeStampb–RxTimeStampf)

TxTimeStampb = time that the DMR PDU is sent by the target MEP
RxTimeStampf = time that the DMM PDU is received by the target MEP

This second option requires that the DMR PDU include two additional time stamps, TxTimeStampb and RxTimeStampf

NE Implementation Aspects

Ethernet Service OAM MEP and MIP functions will typically be implemented in the context of a bridge function within an Ethernet Service transport NE. This NE may bridge customer Ethernet frames to an Ethernet, SONET/SDH, or MPLS pseudowire-based transport link. IEEE 802.1ag defines the implementation aspects of integrating MEP and MIP functions within the Ethernet bridge function.

An 802.1ag CFM Maintenance Point (MP), which can be either a MEP or MIP, is modeled as a shim sub-layer within a bridge port. The MP is bounded by two Service Access Points (SAP) that pass frames between it and adjacent entities. There is an external SAP (DSAP), which is on the outside of a maintenance domain and internal SAP (ISAP), which is internal to a maintenance domain. A MEP shim has directionality in that it sources OAM PDU in one direction. It is therefore modeled as a triangle is associated standards. A bridge port can support MEPs, which either source frames toward the bridge relay function (inward-facing MEP), or source frames toward the bridge port (outward-facing MEP).

A MIP is located within the interior of a maintenance domain, and it can source frames in either direction. It is consequently modeled as two back-to-back MIP Half Functions with an interior ISAP. A circle is used within applicable standards to represent a MIP.

A single service instance may require a bridge port to support MAs at multiple MA/MEG levels. In addition, a bridge port may support MAs for multiple services instances. VLAN tags are used to distinguish the OAM flow for each set of MAs for each service instance. The VID is used as the means by which data for the various MAs are multiplexed. Each MA shim admits and emits only OAM frames for its VIDs. In order to support the multiplexing of multiple MAs by VID, 802.1ag has defined a new functional entity that resides within a bridge port; know as an "ISS-EISS Multiplex Entity". An aggregate set of multiple MAs with multiple VIDs at a single EISS (Enhanced Internal Sub-layer Service) interface is demultiplexed to a set of separate ISS (Internal Sub-layer Service entity) interfaces each supporting one MA. The ISS augments the Ethernet MAC layer in support of the bridge relay function. The EISS is supported by tagging and detagging functions that use the ISS. The MA shim entity resides between two ISS interfaces.

Figure 3 provides a 802.1 bridge port model which illustrates the concepts discussed above. DSAPs are marked with "D" and ISAPs are marked with "IS." MS in Figure 3 indicates the MAC Service Interface across which pass Ethernet frames that terminate and originate on the bridge.

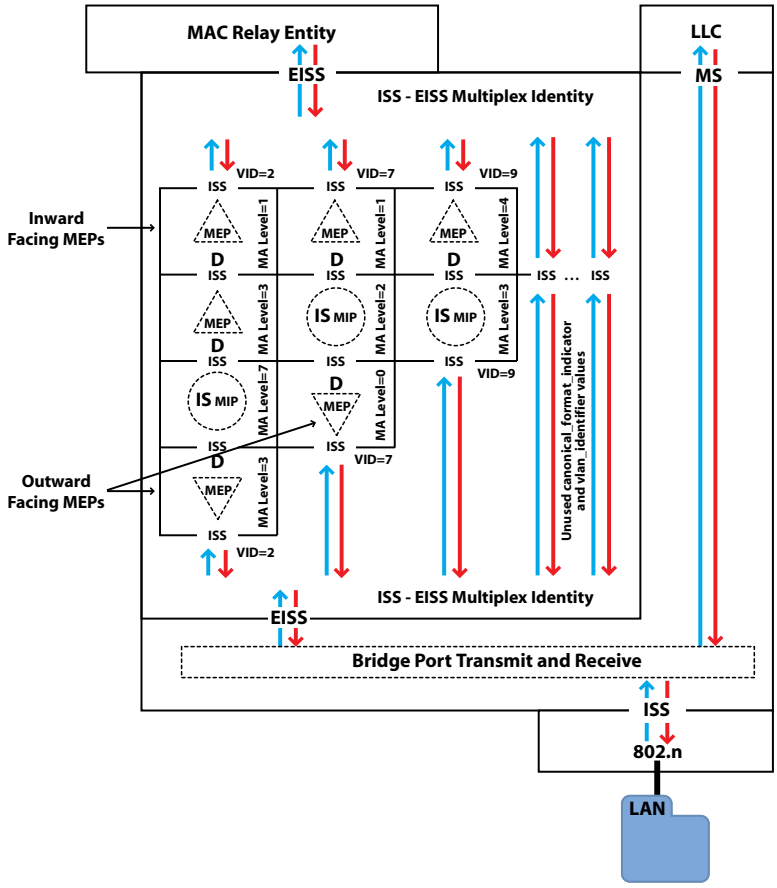


Figure 3: CFM Bridge Model

Figure 4 shows a complete network model with inward-facing MEPs, outward-facing MEPs, MEP AIS Generators, MIPs, and maintenance levels (shown as the number inside the triangle or circle).

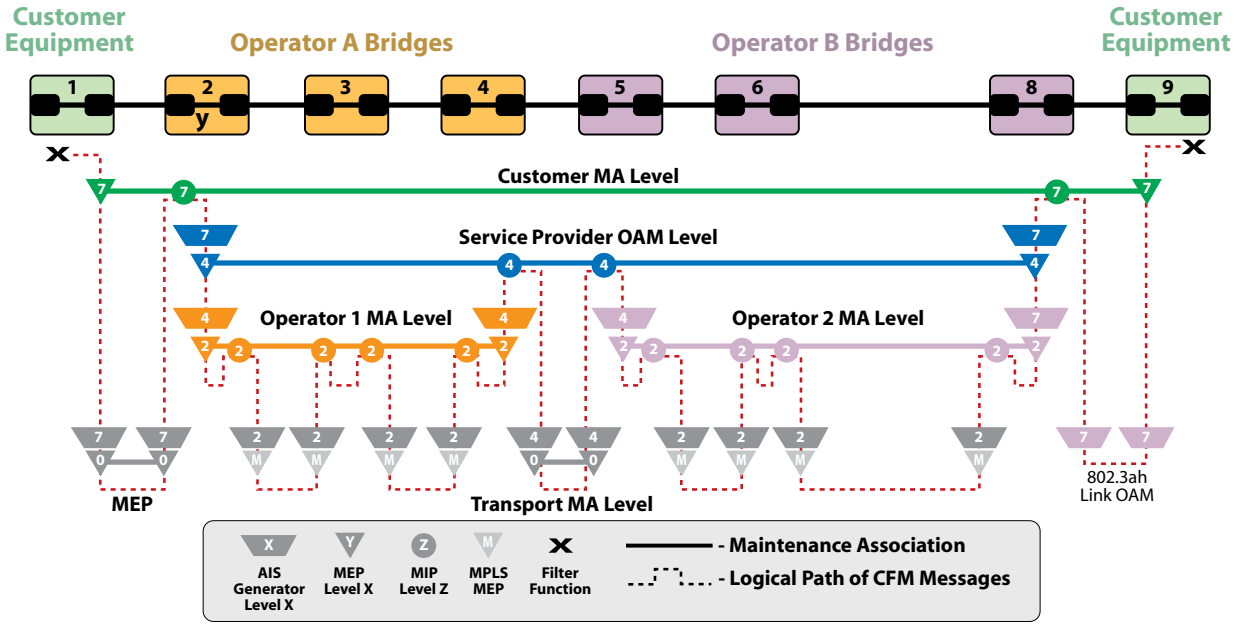


Figure 4: MEPs, MIPs, and ME Levels

In addition to generating and replying to OAM messages, an important NE requirement is filtering OAM messages between administrative domains. The MEP at an edge of a domain must filter by maintenance level. MEPs filter OAM messages at their own level and at higher levels. Referring to Figure 4, the MEP at level 4 (service provider level) in Bridge 2 port Y must filter level 4 OAM frames to prevent forwarding these frames to the customer network. Bridge 2 port will pass level 7 customer OAM frames.

Figure 4 illustrates the MEP AIS Generation Function. IEEE 802.1 has decided not to support this function in phase 1 802.1ag. However it is supported in Y.731, and is expected to be included in ITU Recommendation G.8021 on Ethernet Equipment. Upon fault detection, AIS is generated in the upstream direction at the maintenance level of the client layer.

Figure 4 also shows OAM flows between domains and Transport Layer OAM. Inter-domain OAM requires coordination between provides, operators, and customers for issues such as MA Level assignments. At the interface between Operator A and the Customer, Operator A is implementing 802.1ag / Y.1731 level 0 MEPs. At the interface between Operator B and the Customer, Operator B and the Customer have decided to use 802.3ah Link OAM.

The Transport Layer can be a MPLS pseudowire path, a SONET path, or an Ethernet Link. At this layer it is necessary to implement an Ethernet OAM AIS Generation Function for the Ethernet client layer at the appropriate MA Level. This is illustrated in Figure 4 for both MPLS and Ethernet Transport links.

Service Provider Applications and Operational Issues

Within the realm of fault management, Ethernet OAM can support fault detection, fault verification, fault isolation, fault notification, and fault recovery. In the realm of performance management, Ethernet OAM provides the tools to measure frame loss, delay, and delay variation, and service availability.

For fault detection, Ethernet OAM provides a means to detect both hard and soft failures such as miss-configuration or software failure. Due to the fact that CCMs are multicast, if an MEP receives a CCM with a MEP ID that is not within its configured MA/MEG, a miss-configuration or cross connect error is likely. A customer's EVC may include an unauthorized site and an appropriate alarm will be generated. A good feature of Ethernet OAM is that if a service instance is taken out of service, then in order to avoid triggering false failure detection, the associated MEP indicates its upcoming out-of-service status by issuing a LOCK message to other member NEs for each MA/MEG.

The principal operational issue for Ethernet OAM is scalability. CCM can be sent as fast as every 3.3 ms. There can be 4,094 VLANs per port and up to eight maintenance levels. This yields a worst case CCM transmission rate of 9.8 million CCMs per second. Also as previously noted supporting an optional MIP CCM database may present some scalability issues.

An operational issue related to Ethernet OAM is MEP and MIP provisioning and discovery. An MEP must be provisioned with information about its peer MEPs. This information can be potentially discovered. MEPs can proactively discover other MEPs by CCM messages. ITU-T has defined a multicast loopback, which can be used to discover other MEPs on an on-demand basis. MIPs can be discovered by using linktrace. Another administrative issue is negotiation, agreement, and provisioning of ME Levels across customer, provider, and operator. An associated issue with MIPs and multiple administrative levels is this question: will service providers support customer MIP functions within their network?

The Test Message can be used to detect service frames delivered out of order or excessive service frame error rates. The current Ethernet OAM standards do not support an out of service payload loopback function. The reason for this is that miss-provisioning may cause permanently looping test frames. However there is some belief that such a test could help diagnose data pattern sensitive errors.

Fault verification is accomplished by using loopback messages. The principal operational issue is MEP knowledge of remote MEP/MIP addresses. Fault isolation can be addressed by using the linktrace message. The main operational issue for linktrace is Ethernet MAC address learning and aging. When there is a network fault, the MAC address of a target node can age out in several minutes (e.g. typically five minutes). Solutions are to launch linktrace within the age out time or to maintain a separate target MEP database at intermediate MIPs. However, this requires a MIP CCM database.

Fault notification and alarm suppression is accomplished by using SNMP notifications and AIS/RDI. AIS can provide both alarm suppression and upstream notification. RDI provides downstream notification. The main issues with AIS are multipoint service instances, and the potential interaction with Ethernet STP loop prevention and recovery. An STP based network reconfiguration may result in AIS interruption or redirection. The issues with RDI are multipoint service instances and bidirectional faults, which would block RDI downstream transmission.

With respect to fault recovery, ITU-T SG 15 is in the process of developing a recommendation for Ethernet protection switching, which uses fast CCM frames for detection, and a Y.1731 protocol format message for coordination. A Y.1731 OpCode has been reserved for protection switching messages .

With respect to measuring performance parameters, the principal issues are the complexity of measuring one-way delay by Ethernet Service OAM frames, and measuring performance parameters for multipoint service instances.

Summary and Conclusions

This paper provides an overview of Ethernet Service OAM including relevant standards, protocol message processing and NE implementation aspects, applications and operation issues. The positive news is that there is excellent coordination across IEEE 802.1, ITU-T, and MEF in the generation of a robust set of OAM standards. Phase 1 of this work should be completed in 2006. When completed, this set of standards will provide an extremely useful toolkit of fault management and performance management functions. A key aspect of this management toolkit is that it supports multiple domain Ethernet networks.

As with any emerging technology, there are still some open issues and administrative and operational challenges in the deployment of Ethernet Service OAM. These issues are summarized in the previous section. Possible solutions are suggested and additional creative solutions will emerge as implementation experience grows.

References

[1] IEEE 802.1ag, Virtual Bridged Local Area Networks – Amendment 5: Connectivity Fault Management, Draft 5.2, December, 2005, IEEE 802.1 Committee

[2] Draft Recommendation Y.1731 – OAM Functions and Mechanisms for Ethernet based Networks, January 2006 Draft, ITU-T SG 13 WP 4/Q5

[3] MEF Service OAM Requirements & Framework – Phase 1 Technical Specification, Draft Version 3.2, January 2006, Metro Ethernet Forum

[4] Draft Recommendation G.8021 – Characteristics of Ethernet Transport Network Equipment Functional Blocks, December 2005 Draft, ITU-T SG 15 WP3/Q9

Acronym	Descriptor
AIS	Alarm Indication Signals
CCM	Continuity Check Messages
CFM	Connectivity Fault Management
CoS	Class of Service
DM	Delay Measurement
DMM	Delay Measurement Message
DMR	Delay Measurement Response
EISS	Enhanced Internal Sublayer Service
FD	Frame Delay
FDV	Frame Delay Variation
FLR	Frame Loss Ratio
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
ITU	International Telecommunications Union
LBM	Loopback Message
LCK	Locked Signal Function
LLC	Logical Link Control
LM	Loss Measurement
LMM	Loss Measurement Request
LMR	Loss Measurement Reply
LRM	Loopback Reply Message
LSP	Local Service Provider
LTM	Linktrace Message
LTM	Linktrace Message
LTR	Linktrace Replay Message
MA	Maintenance Association

Acronym	Descriptor
MAC	Medium Access Control
MCC	Maintenance Communications Channel
ME	Maintenance Entity
MEF	Metro Ethernet Forum
MEG	Maintenance Entity Group
MEP	MEG End Point
MIP	MEG Intermediate Point
MPLS	Multiprotocol Label Switching
ms	Millisecond
NE	Network Element
NMS	Network Management System
OAM	Operations, Administration and Maintenance
OUI	Organizational Unique Identifier
PDU	Protocol Data Unit
RDI	Remote Defect Indication
SDH	Synchronous Digital Hierarchy
SNAP	Subnetwork Access Protocol
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical Network
STP	Spanning Tree Protocol
TLV	Time, Length, Value
TST	Test Signal
UNI	User Network Interface
VCCV	Virtual Circuit Connectivity Verification
VID	VLAN ID
VLAN	Virtual Local Area Network

FUJITSU NETWORK COMMUNICATIONS INC.
2801 Telecom Parkway, Richardson, Texas 75082-3515
Telephone: (972) 690-6000
(800) 777-FAST (U.S.)
us.fujitsu.com/telecom