TECHNOLOGY WHITE PAPER

Ethernet Link and Service OAM Tools

Author:
Michael Ritter,
ADVA Optical Networking

Ethernet succeeded in being the ubiquitous bandwidth bearer for enterprise and carrier networks and this scale of demand has driven unbeatable levels of standardization and commoditization for the technology. The factor-of-10 leaps in bandwidth, from 10Mbit/s to 10Gbit/s and soon to 100Gbit/s, have kept Ethernet ahead of the ever-growing demand for bandwidth. Preservation of frame structure across different media and bandwidths has ensured seamless compatibility with earlier generations as technology evolves.

Nonetheless, Ethernet was not originally conceived for carrier use, so it would be surprising if it perfectly satisfied all carrier requirements immediately. The ongoing transition to a technology satisfying carrier requirements is summarized under the name "Carrier Ethernet". The most important missing feature set in traditional Ethernet equipment is what carriers call "OAM" – Operations, Administration and Maintenance. Essentially, this is the capability of the network to automatically report faults and the tools to isolate the failure point and repair faults quickly when they occur.

Carrier networks differ from enterprise LANs in scale, geographical reach and services offered. The first two points may be obvious, but the third is also important. In contrast to enterprise networks, which generally present a shared and un-metered resource for any-to-any connectivity, carriers sell secure connectivity between specific customer sites or between customer sites and various services providers. Carrier resources are backed by strict Service Level Agreements (SLAs) and terminate at clearly identified demarcation points on the customer premises. All of the logical and physical connections, starting at the Customer Premise Equipment (CPE), must be actively monitored. Furthermore, carriers sell services to a large and changing customer base. The vast geographical scale of most carrier networks requires a high degree of automation in monitoring, fault finding and connection administration because it is prohibitively expensive and slow to send technicians to customer sites.

One of the biggest reasons that Ethernet is attractive as a next-generation access technology is that it is a flexible, high-bandwidth bearer for multiple services. However, in order to offer services that can be abstracted from the underlying network links, it is necessary to monitor both the physical network links and the logical connections that they support. These logical connections are end-to-end connections which may cross many physical Ethernet links in the network. They may also be multipoint, rather than point-to-point connections. In the context of Carrier Ethernet, these connections are called Ethernet Virtual Circuits (EVCs). There is no concept of EVC in the connectionless world of enterprise Ethernet networks, but Virtual Local Area Networks (VLANs) provide some similar functions.

OAM standards should cover both the link and end-to-end aspects of carrier networks and services, and standards bodies such as IEEE and ITU are already well advanced in creating those standards.

OAM standards should cover both the link and end-to-end aspects of carrier networks and services, and standards bodies such as IEEE and ITU are already well advanced in creating those standards. Furthermore, the individual standards bodies cooperate closely to develop OAM standards. The emergence of standards-based OAM for native Ethernet networks is a crucial step in spreading carrier adoption of this technology, which plays a dominant role in next-generation metro and access networks.

IEEE 802.3ah Link Fault Management

The IEEE 802.3ah standard, commonly known as Ethernet in the First Mile (EFM), relates to Ethernet link OAM and defines three different media for EFM connections – fiber, copper and passive optical net-work. It also defines a simple OAM protocol, referred to as EFMOAM, for testing the Ethernet link between neighboring Ethernet nodes using these links. Although this could theoretically be used on any Ethernet link, EFMOAM is targeted for use on the last- or first-mile, typically between Provider Edge (PE) equipment and CPEs deployed by the carrier for service termination.

EFMOAM is designed to provide a basic level of visibility into a native Ethernet access network, while allowing the CPE devices to remain relatively simple and cheap.

Carriers usually have good management access to their own PE equipment, including backup routes. However, it is prohibitively expensive to provide this out-of-band management to CPEs. In fact, doing so would require a second access line to the customer site, which negates the economic advantage of providing services through a single access link. Furthermore, the PE equipment is relatively expensive, aggregating many customer connections into the network, so it usually has extensive on-board management capabilities. In summary, carriers have good visibility between PEs into their networks but poor visibility outside into the first mile. EFMOAM is designed to provide a basic level of visibility into a native Ethernet access network, while allowing the CPE devices to remain relatively simple and cheap.

The IEEE 802.3ah OAM functions are generally initiated by the PE device to check the status of the CPE without requiring expensive on-site visits. The standard defines these functions:

- Discovery
- Link Performance Monitoring
- Remote Failure Indication
- Remote Loopback
- MIB Variable Retrieval
- Vendor-Specific Enhancements

EFMOAM defines Protocol Data Units (OAMPDUs) for communication. These are standard Ethernet frames with a specific reserved multicast destination address and Ethernet type, which allows the receiving station to identify them as OAM traffic and treat them differently than customer data traffic. EFMOAM is a so-called "slow protocol," which transmits fewer than 10 packets per second, to avoid impacting throughput for customer data traffic. Furthermore, OAMPDUs should always be intercepted and not forwarded by the receiver, thus keeping the scope of the protocol inherently link-local.

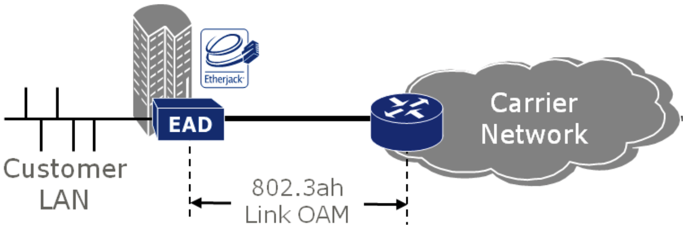

Figure 1: EFMOAM network scenario

Discovery occurs when EFM stations connect, allowing both ends to discover the existence, the identity and the OAM capabilities of its neighbor. For instance, neighbors may be in active or passive mode. This allows a master and slave relationship between PE device and CPE. Active mode stations, for example, do not respond to loopback requests from passive mode neighbors. An OAMPDU must be sent at least once every second and received every five seconds. If this time expires, the discovery process restarts.

Link Performance Monitoring sends events to its neighbor when one of these events is detected on the link:

- Errored symbols per second exceed a threshold within a specified period
- Errored frames per second exceed a threshold within a specified period
- Errored frames per N frames exceed a threshold
- Errored frame seconds within N seconds exceed a threshold

Remote Failure Indication allows a device to indicate failure conditions to its neighbor. Because Ethernet does not provide any framing around the data frames, faults are not always easy to identify. These faults can be indicated:

- Link Fault
- Dying Gasp
- Critical Event

Link Fault Indication is sent once per second when the receiver detects a loss of signal from its neighbor. Note that conventional Ethernet equipment typically shuts down the port, including the transmit path back to the neighbor, when it detects a loss of signal. This simplifies the failure event for higher-layer protocols. For EFMOAM, it is desirable that the return transmit path remains open for link fault OAMPDUs only. This provides superior fault localization. Dying Gasp indicates that an unrecoverable condition, such as a power failure, has occurred, while Critical Event is used for any other critical event indication.

Remote Loopback allows an EFMOAM device to place its EFM neighbor into a loopback state for test purposes during installation or fault finding. Subsequent frames sent on this link are returned to the loopback-invoking device until the loopback is disabled. These returned frames can in turn be used by the loopback-invoking device to test throughput, latency and jitter on the link, all of which may relate to the carrier's SLA. Note, however, that customer frames, and thus the offered services, are interrupted for the duration of a loopback.

MIB Variable Retrieval allows an EFMOAM device to get management information about Ethernet variables, such as traffic statistics, from its neighbor. The Management Information Base (MIB) describes variables and parameters of the network element that are accessible by Simple Network Management Protocol (SNMP).

Finally, the standard allows for the definition of organization-specific extensions within EFMOAM, referred to as Vendor Specific Enhancements. These allow vendors to implement practically any future functions they may need by organizing specific OAMPDUs. ADVA Optical Networking uses this to implement many enhancements over and above these listed functions, including full configuration and management of the remote device by using EFMOAM. Optionally, SNMP can also be used for this purpose.

IEEE 802.1ag Connectivity Fault Management (CFM)

While IEEE 802.3ah defines link-level OAM between CPE and PE for the first-mile, another standard from IEEE, termed IEEE 802.1ag, defines OAM not only for physical links but for the logical connections, which constitute the end-to-end service offered. These logical connections, or EVCs, generally traverse multiple hops in Carrier Ethernet networks. The purpose of IEEE 802.1ag is Connectivity Fault Management (CFM) for Carrier Ethernet services.

One obvious question is whether EFMOAM and CFM need to interwork as the goal of end-to-end OAM. Certainly, the interworking of these standards in PE devices could be useful, as the CPE offers no IEEE 802.1ag support. However, IEEE 802.1ag is often incorporated into demarcation devices, especially for high-value services, in which case end-to-end CFM does not require such interworking.

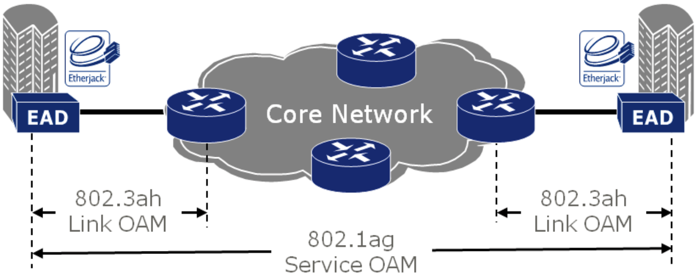

Figure 2: Comparing the scope of IEEE 802.3ah and 802.1ag

IEEE 802.1ag is designed to work for EVCs offered across multiple levels and concatenations of carrier networks.

The scope of IEEE 802.1ag is much greater than that of 802.3ah, and therefore the framework it defines is necessarily more complex. IEEE 802.1ag is designed to work for EVCs offered across multiple levels and concatenations of carrier networks. Thus, IEEE 802.1ag provides a framework for CFM support at every ingress and egress port of every hop along an EVC path, in which the intervening nodes are separated for OAM purposes into their respective management domains. This makes it possible not only to detect connectivity failures or performance problems for an end-to-end service, but also to isolate the exact location of the service degradation.

A quick introduction to CFM terminology would be helpful at this point. IEEE 802.1ag defines the concepts of both Maintenance Domain – that is, a network domain for which faults in connectivity need to be managed – and Domain Service Access Points (DSAPs) at the edges of this domain. A service or Service Instance creates a Maintenance Association between various DSAPs, and these paths consist of Maintenance End Points (MEPs) and the intervening Ethernet hops or ports, referred to as Maintenance Intermediate Points (MIPs). Both the ingress and the egress ports of an Ethernet switching or transport device may be MIPs. The decision of where to activate MIP functionality is ultimately up to the network operator.

This concept is hierarchical, allowing provider OAM domains to be built on services from other domains, as shown in Figure 3 below. Faults identified on a lower layer are alarmed only to the next-higher level, so that appropriate actions such as traffic rerouting can be performed, while the details of the problem remain and can be resolved at the layer where the incident occurred. This ensures that the fault is regulated within the appropriate domain, while preventing a mass broadcast of alarms throughout all layers of the network. IEEE 802.1ag defines a hierarchy of eight levels in which the highest, level 7, always represents the whole connection path from the customer's point of view and the lowest, level 0, always represents the Ethernet section – that is, the physical links, or network hops, on the EVC path. The six intermediate domains leave ample room for nested carrier's carrier scenarios.

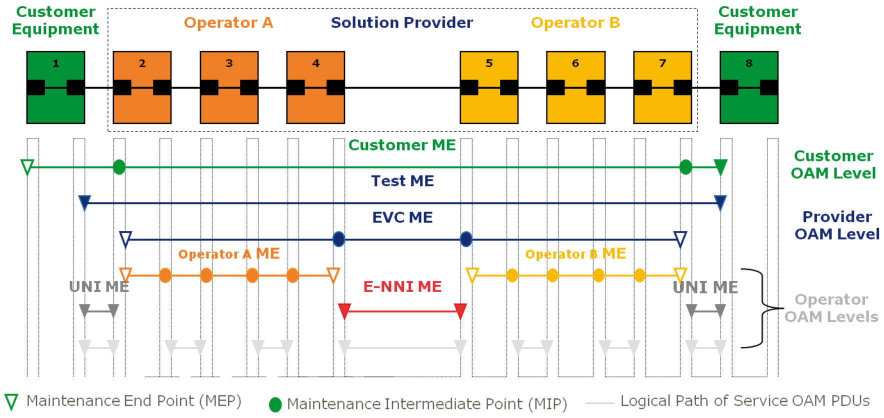

Figure 3: IEEE 802.1ag Connectivity Fault Management levels

For example, a customer's end-to-end connection on level 7 may be supplied by a solution provider's EVC on level 5, which in turn uses Ethernet connectivity from two network operators on level 3, whose ETH sections on level 0 reside in their respective transport networks. The CFM domain of the solution provider will be above the domains of the network operators in the CFM hierarchy and will see the underlying carrier services as a single hop, respectively. Each can use IEEE 802.1ag functions to isolate problems within their own domain and then refer these down the hierarchy for closer attention. Note that both network operators in this example use levels 3 and 0 but have separate CFM domains. Domains may be interconnected but do not intersect.

Thus, in this example the customer sees three CFM hops at level 7 – local MEP-MIP, MIP-MIP and remote MIP-MEP. The two MEP-MIP hops can be used to verify the customer's connections to the demarcation or PE device. The MIP-MIP hop in the middle is the EVC between the demarcation or PE devices. Remember, the exact position of the MIP function is up to the service provider. If CFM shows this EVC to be degraded or broken, the problem can be referred to the solution provider. The solution provider can isolate the problem to one of the two network operators or to his own links between these. The affected network operator can in turn isolate the problem down to hops in the overlay packet data network or further still in the underlying transport network. To distinguish between different levels in the hierarchy, CFM Protocol Data Units (PDUs) have a reserved multicast address and a domain-level field for each of the eight levels. CFM PDUs sent by the solution provider on level 5 will be ignored by the levels below, namely level 3 and 0, although they must obviously traverse these links. To distinguish between different OAM domains and service instances, CFM PDUs contain a globally unique Maintenance Association IDentifier (MAID).

Based on the hierarchical concept described, IEEE 802.1ag defines the following OAM tools:

- Connectivity Check
- Loopback
- Link Trace

These tools make use of CFM-specific Ethernet frames with unicast and reserved multicast destination addresses. In all cases, the CFM frames follow the same path as the EVC they are monitoring. This has the additional advantage that no explicit interworking with Spanning Tree Protocol (STP) or other dynamic network changes is required for compatibility. The CFM frames follow the same alternate path as normal data packets in the event of STP-driven or other path changes.

Connectivity Check (CC) messages are multicast frames sent at regular and configurable intervals, down to as little as 3.3 milliseconds, and are used for fault

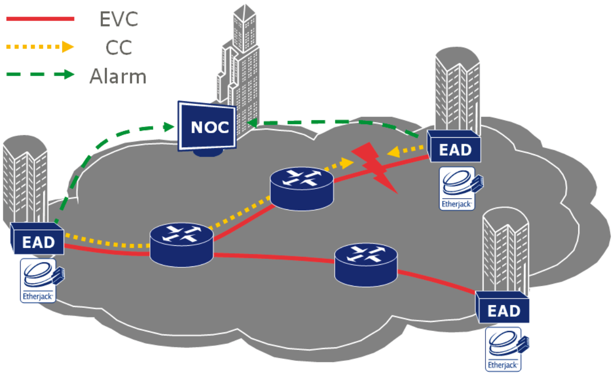

Figure 4: Fault detection using Connectivity Check Messages

detection. These are received by all MIPs and MEPs for the given Maintenance Association. If the EVC service provided is multipoint, then the CC messages will be received by multiple MEPs. The receiving MEPs do not reply to CC messages but are rather configured to react – for instance, inform the NMS via an alarm – when they are not received and also when the wrong CC frames are received. Service faults or interruptions are therefore detected. The MEPs retain a database of other MEPs from which they expect to receive CC messages. In this way, CC can detect other problems, such as mis-configurations, in addition to loss of connectivity.

The *Loopback* function is used for verifying service faults and can be compared to IP ping but at the MAC layer. It allows an MEP to issue a LoopBack Message (LBM) addressed to the specific unicast MAC address of a given MEP or MIP. The MAC addresses for all intervening MIPs are assumed to have been learned earlier via Link Trace or other means. If the MEP or MIP is reachable, it will send a LoopBack Reply (LBR) message to the originating MEP. Performance measurement tools could make use of loopback messages for analyzing link characteristics such as throughput, latency and jitter but this is not defined within the scope of IEEE 802.1ag.

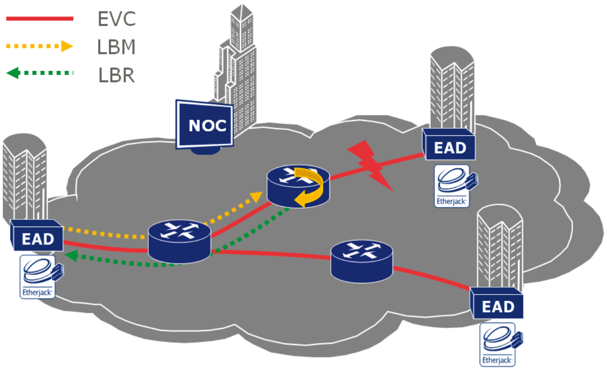

Figure 5: Fault verification using Loopback messages

Link Trace can be compared to IP traceroute but at the MAC layer and is used for fault location and isolation purposes. A MEP can send link trace to any other MIP or MEP, in fact to any unicast destination MAC address. If sent to a remote MEP, all MIPs on the path to that MEP and the target MEP itself, will send replies to the originating MEP, allowing this to learn their MAC addresses and their position on the path. The link trace message uses a multicast address. The target unicast MAC address is contained within the message. However, the message is not multicasted to all ports. Rather, each MIP on the path to the MEP replies and then forwards the packet via the appropriate path. This process continues until the target is reached or until a MIP does not know how to reach the target MAC address. The returned Link Trace messages and those not returned uniquely identify the segment or node where the fault originates. While IP traceroute uses multiple ping packets, each with increasing TTL values to get "TTL-expired" replies from each hop, IEEE 802.1ag defines a specific link trace packet that needs only to be sent once.

In addition to enhancements for fault indication and diagnostics, the mechanisms defined in ITU-T Y.1731 enable the service provider to measure frame delay, delay variation and frame loss SLA parameters.

ITU-T Y.1731 Ethernet Service OAM

The ITU-T Y.1731 builds on IEEE 802.1ag to add in performance monitoring features on and end-to-end service basis. In addition to enhancements for fault indication and diagnostics, the mechanisms defined in ITU-T Y.1731 enable the service provider to measure frame delay, delay variation and frame loss SLA parameters. For fault management, the following tools are supported based on standards originally defined in the ATM protocol:

- Ethernet Alarm Indication Signal
- Ethernet Remote Defect Indication

The *Alarm Indication Signal* (AIS) is a multicast message which is propagated toward the downstream service endpoint by MEPs detecting a connectivity failure to affected MEPs at the next higher level. It serves two purposes: to inform higher layers of the failure and also to suppress the fault alarms that would normally be caused in the higher layers. MEPs at each of the higher layers can pass AIS upwards to the next higher layer so that MEPs at all affected levels receive the alarm information. To ensure that a failure state is maintained, AIS messages are sent periodically until the service is restored.

Remote Defect Indication (RDI) ensures that, upon loss of the receiving signal, the downstream node signals the fault back upstream. It is the upstream twin of the AIS downstream signal.

In order to report the actual quality of a service, ITU-T Y.1731 specifies techniques for measuring both one-way and round-trip frame delay, delay variation and frame loss.

- Frame Delay Measurements
- Frame Delay Variation Measurements
- Frame Loss Measurements

Frame Delay measurement tests the travel time between two MEPs across the network for delivered frames. One-way delay measurement requires that the service endpoints have synchronized reference clocks, while round-trip measurements do not. One-way delay is measured by sending a time stamped

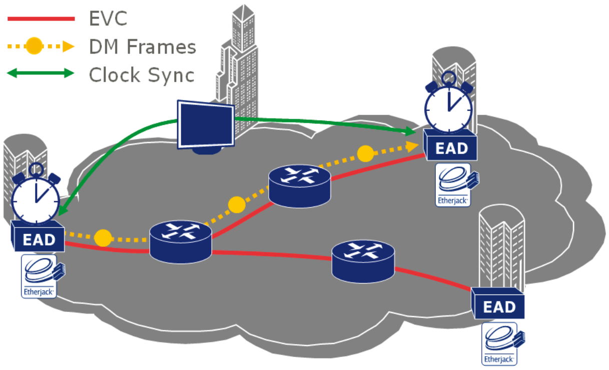

Figure 6: Delay measurement based on ITU-T Y.1731

Delay Measurement (DM) test frame through the network to the far-end MEP that compares the original timestamp to its current reference clock to calculate the delay. Round-trip delay is measured by sending a DM message frame, whose time stamp is returned to the originating node in a DM reply frame.

Frame Delay Variation is the difference in the delay of two subsequently received DM packets belonging to the same EVC and is calculated over a predefined measurement period.

The *Frame Loss* ratio is determined by analyzing the counters for sent and received CC messages at the service end points and measuring the number of lost or discarded frames out of all frames that should have been delivered within a specific time interval. ITU-T Y.1731 defines both uni-directional and bi-directional frame loss ratio measurement, where the bi-directional measurement involves the exchange of Loss Measurement (LM) messages and LM replies.

The *Ethernet Test Signal* function completes the ITU-T Y.1731 standard and is used to perform one-way, in-service and out-of-service diagnostic tests including throughput measurement. It enables the accurate reflection of service performance under real-world conditions in addition to intrusive tests performed at service setup or during early-hour maintenance windows.

RFC 2544 Ethernet Service Testing

The Internet Engineering Task Force (IETF) standard RFC 2544 is the de facto methodology that outlines the tests required to measure and prove performance criteria for Carrier Ethernet networks. The method is widely used as a turn-up solution in order to test the quality of a connection prior to enabling customer traffic. The standard provides an out-of-service benchmarking methodology to evaluate the performance of services and network devices, among which are:

> *The standard provides an out-of-service benchmarking methodology to evaluate the performance of services and network devices ...*

- Throughput Test
- Back-to-Back Test
- Frame Loss Test
- Delay Test

Each of the tests validates a specific part of an SLA. The methodology defines the frame size, test duration and number of test iterations. Once completed, these tests will provide performance metrics of the Ethernet network under test. In order to ensure that an Ethernet network is capable of supporting a variety of services, such as VoIP, video and data, the RFC 2544 test suite supports seven pre-defined frame sizes (64, 128, 256, 512, 1024, 1280 and 1518 bytes) to simulate various traffic conditions. Small frame sizes increase the number of frames transmitted, thereby stressing the network device since it must forward a large number of frames.

The *Throughput Test* is the most widely used method of RFC 2544 and defines the maximum number of frames per second that can be transmitted without any error. This test is performed to measure the rate-limiting capability of Ethernet devices as found in Carrier Ethernet networks. The methodology involves starting at a maximum frame rate and then comparing the number of transmitted and received frames. Should frame loss occur, the transmission rate is divided by two and the test is restarted. If during this trial there is no frame loss, then the transmission rate is increased by half of the difference from the previous trial. This methodology is known as the half/doubling method. This trial-and-error methodology is repeated until a rate is established at which there is no frame loss found. The throughput test must be performed for each frame size. Although the test time during which frames are transmitted can be short, it must be at least 60 seconds for the final validation.

The *Back-to-Back Test*, also known as the "burst test", assesses the buffering capability of a network node. It measures the maximum number of frames received at full line rate before a frame is lost. In Carrier Ethernet networks, this measurement is quite useful as it validates the Excess Information Rate (EIR) as defined in many SLAs. As demonstrated in Figure 7, a burst of back-to-back frames is transmitted across the network with minimum inter-frame gap. Should a frame be dropped, the burst length is shortened. Should it be received without any errors, the burst length will be increased. The trial length must be at least two seconds long and the measurement should be repeated at least 50 times, with the average of the recorded values being reported for each frame size.

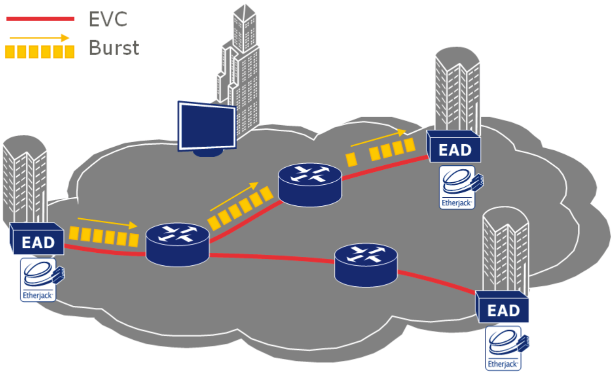

Figure 7: Back-to-Back Test example

The *Frame Loss Test* measures the network's response in overload conditions, a critical indicator of the network's ability to support real-time applications in which a large amount of lost frames will rapidly degrade service quality. As there is no retransmission in real-time applications, these services might rapidly become unusable if frame loss is not controlled. The test instrument sends traffic at maximum line rate and then measures if the network dropped any frames. If so, the values are recorded and the test will restart at a slower rate. This test is repeated until there is no frame loss for three consecutive iterations.

The *Delay Test* measures the time required for a frame to travel from the originating device through the network to the destination device. This test can also be configured to measure the round-trip time. When the delay time varies from frame to frame, it causes issues with real-time services. For example, delay variation in VoIP applications would degrade the voice quality and create pops or clicks on the line. Long delay can also degrade Ethernet service quality.

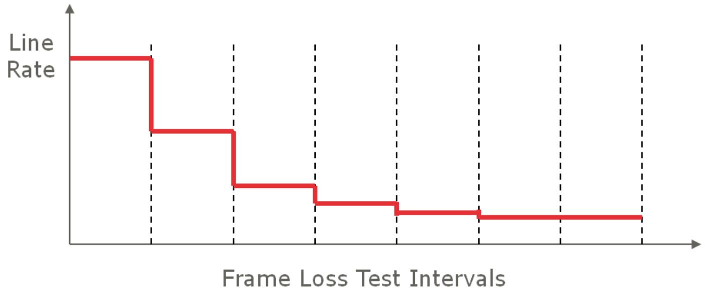

Figure 8: Frame Loss Test scenario

The test procedure begins by measuring and benchmarking the throughput for each frame size to ensure the frames are transmitted without being discarded. This fills all device buffers, therefore measuring delay in the worst conditions. The second step is for the test device to send traffic for 120 seconds. At mid-point in the transmission, a frame must be tagged with a time-stamp and when it is received back at the test device, the delay is measured. The transmission should continue for the rest of the time period.

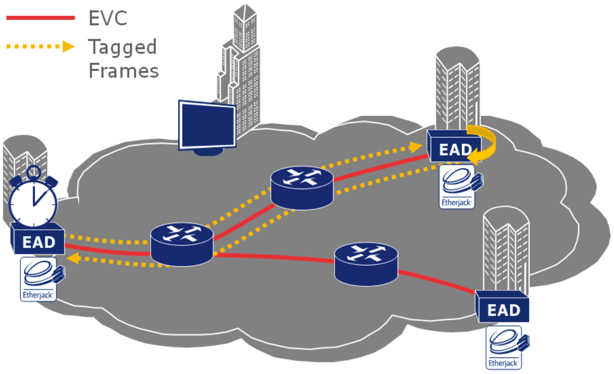

Figure 9: Delay Test setup

Carrier-Class Ethernet Service Demarcation

ADVA Optical Networking's patent-pending Etherjack™ demarcation technology allows a carrier to provide an intelligent Ethernet service demarcation point compliant with the latest OAM standards ...

As service providers scale their Ethernet services, the application of Ethernet link and service OAM tools has become a critical issue to delivering profitable Ethernet services. With the adoption of Ethernet OAM standards in the access network and for customer premises equipment, Ethernet service providers can dramatically reduce operational expenses and service outages, enabling delivery of a true carrier-class Ethernet service. A rich implementation of standardized OAM functions will enable Ethernet services to scale and Carrier Ethernet to become the one single converged access network of the future.

ADVA Optical Networking's patent-pending Etherjack® demarcation technology allows a carrier to provide an intelligent Ethernet service demarcation point compliant with the latest OAM standards such as IEEE

802.3ah, IEEE 802.1ag, ITU-T Y.1731 and RFC 2544. Featuring the full Etherjack functionality, all members of the FSP 150 family of Ethernet demarcation devices are capable of acting as Maintenance End Points (MEPs) or Intermediate Points (MIPs) in any topology including multi-vendor, multi-carrier and multi-technology networks.

Combined with an MEF-certified User Network Interface (UNI), the FSP 150CC family of Ethernet access products enables carriers to deliver Ethernet services that can be remotely monitored and managed with a minimal number of truck rolls. It provides the service intelligence necessary to encourage enterprise data users to make the switch from Frame Relay, Private Line and ATM services to a carrier-class Ethernet service.

About ADVA Optical Networking

ADVA Optical Networking is a global provider of intelligent telecommunications infrastructure solutions. With software-automated Optical+Ethernet transmission technology, the Company builds the foundation for high-speed, next-generation networks. The Company's FSP product family adds scalability and intelligence to customers' networks while removing complexity and cost. Thanks to reliable performance for more than 15 years, the Company has become a trusted partner for more than 250 carriers and 10,000 enterprises across the globe.

Products

FSP 3000

ADVA Optical Networking's scalable optical transport solution is a modular WDM system specifically designed to maximize the bandwidth and service flexibility of access, metro and core networks. The unique optical layer design supports WDM-PON, CWDM and DWDM technology, including 100Gbit/s line speeds with colorless, directionless and contentionless ROADMs. RAYcontrol™, our integrated, industry-leading multi-layer GMPLS control plane, guarantees operational simplicity, even in complex meshed-network topologies. Thanks to OTN, Ethernet and low-latency aggregation, the FSP 3000 represents a highly versatile and cost-effective solution for packet optical transport.

FSP 150

ADVA Optical Networking's family of intelligent Ethernet access products provides devices for Carrier Ethernet service demarcation, extension and aggregation. It supports delivery of intelligent Ethernet services both in-region and out-of-region. Incorporating an MEF-certified UNI and the latest OAM and advanced Etherjack™ demarcation capabilities, the FSP 150 products enable delivery of SLA-based services with full end-to-end assurance. Its comprehensive Syncjack™ technology for timing distribution, monitoring and timing service assurance opens new revenue opportunities from the delivery of synchronization services.

ADVA Optical Networking
North America, Inc.
5755 Peachtree Industrial Blvd.
Norcross, Georgia 30092
USA

ADVA Optical Networking SE
Campus Martinsried
Fraunhoferstrasse 9a
82152 Martinsried/Munich
Germany

ADVA Optical Networking
Singapore Pte. Ltd.
25 International Business Park
#05-106 German Centre
Singapore 609916

For more information visit us at www.advaoptical.com

ADVA™
Optical Networking

Version 07/2012