Operation, Administration, and Maintenance in MPLS Networks

Dirceu Cavendish, NEC Laboratories America
Hiroshi Ohta, NTT Network Service Systems Laboratories
Hari Rakotoranto, Cisco Systems Inc.

ABSTRACT

The boundaries between packet and circuit networks have long disappeared, with many traditional circuit-switched applications such as voice and video now being carried over packet-switched IP/MPLS or Ethernet networks. However, this transition has happened so fast that many OAM functions supported by circuit-switched networks such as SONET/SDH are still unmatched in packet networks. In order to match the quality sustained by circuit-switched networks, OAM functions must also be developed for such packet networks. A number of recent efforts have started to address OAM functions for IP/MPLS and Ethernet packet technologies. Service providers and carriers alike are the driving force behind the work, as there is general recognition that to generate sustained revenues, services must be efficiently managed. In this article we discuss issues in providing OAM features and capabilities for MPLS-based packet networks.

INTRODUCTION

Packet networks have evolved considerably since the early days of the ARPANET. Various technologies have appeared; some have established themselves, others have faded away. The surge of packet-based service providers has underlined the importance of operations, administration, and maintenance (OAM) functions in packet networks as a necessary step to establish viable business models for new services. Mirrored on OAM functions of circuit-switched (time-division multiplexing, TDM) networks, OAM functions are gradually making their way into various packet technologies such as asynchronous transfer mode (ATM), IP, and Ethernet. Obviously, each packet technology may require specific OAM functions, so new ones are being developed as well.

ATM has long been a packet technology embraced by carriers to provide revenue generating services due to its advanced OAM features relative to existing IP and Ethernet technologies. Recently, however, an effort to equip IP technology with traffic engineering and service level agreement (SLA) supportive functions that allow revenue generating services to be deployed has emerged. Multiprotocol label switching (MPLS) is part of this effort, as well as OAM protocols.

This article addresses OAM functions and protocols needed by service providers to offer viable services built on MPLS-based packet networks. In what follows, we first examine currently available IP-based OAM functions. We then discuss recent OAM proposals for MPLS networks, as well as issues in MPLS OAM yet to be addressed, mainly based on the activities within International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Study Group 13 (SG13) and the Internet Engineering Task Force (IETF).

MPLS OAM REQUIREMENTS

Early IP OAM functions were somewhat limited [1], with ping and traceroute the most used. IP ping allows a host to probe reachability and round-trip delay of a certain IP address via a series of (variable-packet-length) Internet Control Message Protocol (ICMP) echo messages [2]. IP traceroute allows a host to probe the most likely path taken by IP packets to reach a given IP address.

With the introduction of real-time and quality of service (QoS)-sensitive traffic, such as voice and video over IP, the need to control the switching, routing, and delivering of IP packets across the Internet became apparent. MPLS [3] has been regarded as the main technology to bring control into an otherwise unpredictable IP packet network. MPLS labels are used to pin down paths in the Internet over which IP streams must travel. The pinned path is referred to as a label switched path (LSP).

An MPLS label contains four fields [4]:

- **Label** (20 bits): Carries the actual value of the MPLS label.
- **Exp** (3 bits): Experimental bits. It can be used to control the queuing and discarding behaviors of the packet at each label switched router (LSR).

0163-6804/04/$20.00 © 2004 IEEE

	ITU-T SG13	IETF
Requirements	Y.1710 [5]	draft-ietf-mpls-oam-requirements-02.txt [5]
Detection	Y.1711 [6] Y.1713 [7]	Bidirectional forwarding detection [8]
Diagnostic		MPLS ping/trace [9], VCCV, OAM state mapping,[1] LSR Self Test [10]
Fault management		MIBs (LSR, LDP, TE, FTN, RSVP …)
Recovery	Y.1720 [11]	Fast reroute [12] MPLS high availability: graceful restart and fault tolerance for LDP, BGP [13][2]
Performance	Y.1561 [15]	Netflow
Security		

[1] Virtual Circuit Connection Verification and OAM State mapping are not addressed in this manuscript, for space considerations. The interested reader is referred to draft-ietf-pwe3pvccv-02.txt and draft-nadeau-pwe3-oam-msg-map-04.txt IETF drafts.
[2] These topics are not addressed in this article for space considerations. The interested reader is referred to [13] and its references.

Table 1. *Ongoing MPLS OAM efforts within the ITU-T and IETF.*

- **Stack** (1 bit): Supports label stacking by indicating the bottom label stack entry.
- **TTL** (8 bits): Supports loop prevention at each label hierarchy level.

The label field carries the local label used to identify to which LSP the packet belongs. The Exp field enables features such as differentiated packet queuing and switching for traffic within a single LSP (e.g., using the same label). The stack field allows embedding an inner LSP in an outer LSP, or hierarchical LSP tunneling. Of particular interest to us is the label time to live (TTL) field. This field is decremented at each LSP hop, with packets discarded when it reaches zero, thus preventing looping packets from surviving indefinitely.

MPLS OAM Requirements and Standardization Efforts

OAM has become paramount to the deployment of new network technologies such as MPLS as network operators recognize the importance of OAM in their business plans. Considerable effort has been made recently to realize MPLS OAM functions to reflect operator requirements, particularly within standardization bodies such as ITU-T SG13 and the IETF. Since this work began in earnest in 2000, several standards documents have been produced, with new and enhanced OAM functions currently under study. Table 1 summarizes this work.

MPLS OAM requirements [5, 14] can be summarized as follows:

- **Separation between control plane and data plane OAM**: OAM packets should follow the data path.
- **Detection of failed LSPs**: Detection of equal cost multiple path defects; detection of defects independent of customer traffic activity and before customer complaints.
- **Defect detection and recovery**: Loss of LSP connectivity; degradation of LSP service with subsequent LSP reroute; swapped LSP defect; detection of LSP traffic replication into another LSP; looping.
- **Defect localization on a failed LSP**: Localization of defects on a tunneled LSP scenario. (Note: The tracing capability should apply to both "native" and hierarchical LSPs).
- **LSP characterization and hierarchy**: Since LSPs can nest, management of nested LSPs is needed.
- **LSP defect notification**: Alarm suppression in multiple-layer network scenarios; interworking with other defect notification technologies, such as those in ATM and SONET/SDH, at the endpoints of an LSP.
- **SLA measurement**: In particular, service availability, traffic latency and jitter, and traffic loss.
- **Recovery**: Capability to recover from failure automatically is necessary for some services.
- **Detection of denial of service attacks**.

Moreover, to scale to large networks, OAM functions should be simple and easy to configure, backward compatible with legacy LSRs, and perform under degraded network and link conditions. Finally, they should operate at various administrative domains, such as customer and provider domains. In the following sections we discuss recent initiatives to fulfill some of the OAM requirements mentioned above, and also introduce issues yet to be addressed.

MPLS Fault Detection

This section deals with fault detection issues and protocols on MPLS networks, including detection of a failed LSP, as well as LSP recovery.

Separation between Control and Data Planes

Separation between control and data planes is directly related to the format and handling of an OAM packet. The format of a label stack entry, which is a header for MPLS packets, consists of a label value (20 bits), EXP bits (3 bits), S bit (1 bit), and TTL (8 bits)[4]. To distinguish an OAM packet from a data packet, several methods can be used. ITU-T Y.1711specifies the use of a label stack with two entries. The top label has the same value as user packets (transport label), ensuring that in most cases OAM packets can be routed on the same path as user packets. The second label has a special reserved value of 14 to be distinguished from user packets [16]. However, the introduction of the second label might not be fully compatible with existing load balancing algorithms and therefore may not correctly handle networks where equal cost multipath (ECMP) is in use. Furthermore, in networks where penultimate hop popping (PHP) optimization is in use, further limitation applies, as outlined in Table 2 (see also [16] for details). Another option to separate OAM packets from data packets is to use as the MPLS payload a specific IP packet destined to a well-known port, as in MPLS Ping. In this case the transport label is still the same as for data packets. This

method, however, assumes that the destination address is a nonroutable address to make sure that the IP packet will be processed by the route processor of the "egress" router. This case allows the support of PHP and limited ECMP.

Other methods to separate control and data planes may include a hybrid mode, where an OAM packet might be identified by a specific label (router alert) or using a specific field in the packet that makes recognition easier. Virtual Circuit Connection Verification relies on this technique.

Detection of a Failed Label Switched Path

LSP failures require the testing of specific packet flows, besides network failure detection/localization mechanisms, because in many instances packet flows may get interrupted without network (link/node) failure. This may be due to routing/forwarding table problems, a broken label binding, network congestion, or other causes. Connectivity verification, as well as ping/traceroute types of OAM functions are appropriate for this type of fault detection. The implementation of these protocols may differ, depending on the particular packet technology. Regardless of the technology, however, it is important that the OAM packets used travel the same path as regular data packets.

LSP Defect Scenarios

There have been two efforts in the standard world regarding LSP defect detection and recovery. ITU-T Recommendation Y.1711 has standardized connectivity verification (CV) function, whereas bidirectional forwarding detection (BFD) function is under study within IETF. Figure 1 shows various kinds of plausible defect scenarios. A' and B' are the intended receiver for A and B, respectively. Defect scenarios are:

- Simple loss of connection
- Misconnection
- Swapped connection
- Mismerging
- Loop/unintended replication

Connectivity Verification (ITU-T Y.1711)

The basic idea of CV function is to send test packets (CV packets) periodically (one per second) from the ingress LSR to the egress LSR with the identity of the ingress LSR and originating LSP. The egress LSR analyzes the identification information of the received CV packets to detect defects. Figure 2 shows the CV packet format. It has two label stack entries. The top label is the same as that in user packets. The second label identifies the OAM packet, via a reserved value of 14 [16]. The EXP bits in the top label and the OAM label are both set to the highest priority value used across the entire network. The S bit of the OAM label is set to 1 because it is the bottom label. Moreover, to ensure that CV packets will not go beyond the egress LSR, the MPLS network is treated as one-hop. The function type field indicates the OAM type: a code point 01H is assigned for CV packets. The next two fields, LSR ID and LSP ID, form the LSP trail termination source identifier (TTSI). The TTSI is used to identify the ingress LSR (via the IP address of its output port) and originating LSP (via the LSP ID). The BIP 16 field is used for error detection/correction.

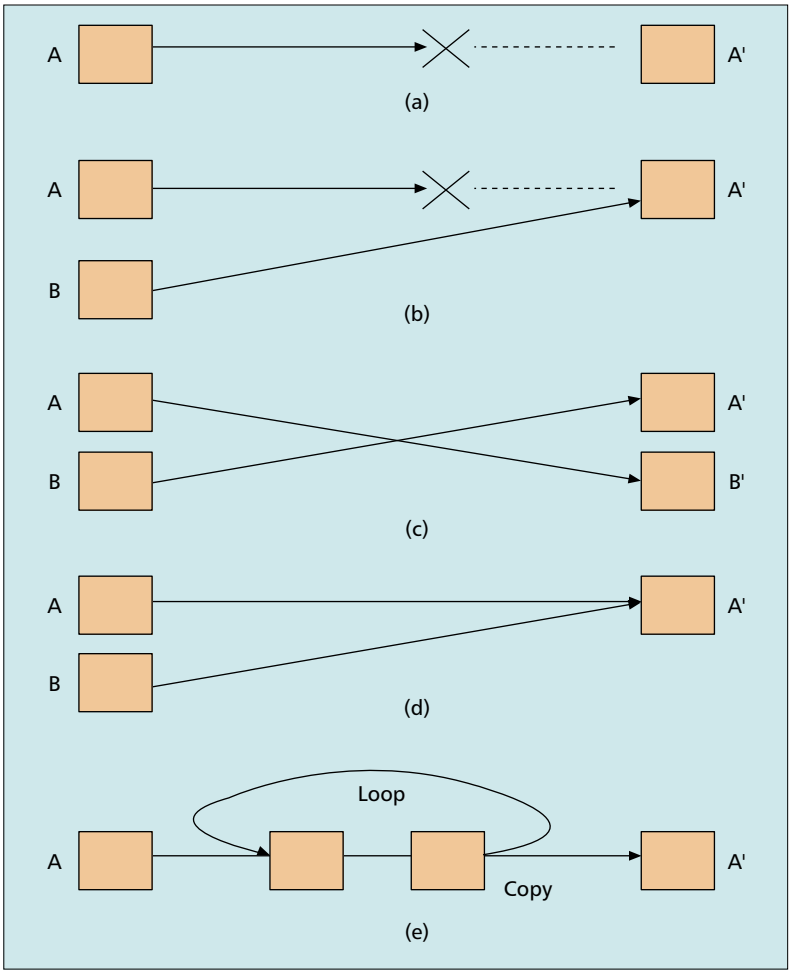

Figure 1. *LSP defect scenarios: a) simple loss of connection; b) misconnection; c) swapped connection; d)mismerging; e) loop/unintended replication.*

The ingress LSR sends CV packets with its TTSI value at every second. The egress LSR receives CV packets and analyzes their TTSI values to determine whether they are expected values or not. If the egress LSR receives an expected TTSI, it declares no defect. If it does not receive any CV packet for more than 3 s, it declares a loss of connectivity defect (dLOCV) (Fig. 1a). The egress LSR waits before declaring dLOCV to avoid misdetection of dLOCV due to an OAM packet loss. If there are two consecutive OAM packet losses, the egress LSR may mistakenly declare dLOCV, but this happens very rarely in normal operating conditions. In addition, as an egress LSR waits for 3 s, the expected delay variation, which is normally much less than a second, is acceptable. If an egress LRS receives an unexpected TTSI, it declares a TTSI mismatch defect (dTTSI-mismatch) (Fig. 1b and c). If it receives both unexpected and expected TTSI values, it declares a TTSI mismerge defect (dTTSI-miss-merge) (Fig. 1d). If it receives more than five CV packets with expected TTSI values within 3 s, it declares excessive reception defect (dExcess) (Fig. 1e).

	Y.1711	MPLS ping/trace
Hardware	"Probably" need new hardware (scalability, TTSI handling, …)	Uses existing hardware
Applicability	• Mainly a detection mechanism • Point-to-point connection-oriented LSP (e.g., RSVP-TE) • Detect LSP mismerge, misbranching	• Mainly a troubleshooting tool • LDP, RSVP-TE, any other MPLS signaling • Detect FEC consistency (i.e. LSP misbranching label to FEC mapping problem) • Troubleshooting: hop-by-hop tracing and problem localization
ECMP friendly	No, use reserved label, might affect load balancing if ECMP occurs at the intermediate LSRs of a given LSP	Yes Basic ECMP algorithm that will explore all 127/8 address range; draft suggests per-hop ECMP exploration
PHP friendly	There is a limitation depending on the type of PHP [16]	Yes
Scalability/Frequency	Packet injection frequency every 1 s	Frequency as per operator request
	Requires management of TTSI	Use native information
	BIP 16 calculation	IP cyclic redundancy check
	Head-end and tail-end need to keep track of LSP state machine	No state machine, echo reply contains code that is interpreted by operator and/or management platform
Service level agreement	No	No

Table 2. *Y.1711 vs MPLS ping/trace functionality.*

In addition to CV, fast failure detection (FFD) function was introduced in the revised version of ITU-T Recommendation Y.1711 in order to realize faster protection switching. The function of FFD is the same as CV. However, FFD sends OAM packets with shorter intervals, allowing faster defect detection.

Bidirectional Forwarding Detection (IETF)

Bidirectional forwarding detection is intended to be a low-overhead short-duration failure detection mechanism in the forwarding path between two adjacent network elements [17]. A verification packet can be used at any of the protocol layers, which makes BFD a very versatile tool. BFD can provide failure detection on any kind of path and media, such as physical links, virtual circuits, and an MPLS LSP between two pairs of network elements. The protocol is based on a three-way handshake, which is applied when establishing or tearing down BFD sessions, ensuring at any given time that the participants in the session are aware of state changes.

A session is operational when two-way communication is established between the pairs' system. Applied to MPLS networks [8], a BFD session can be established over each forwarding equivalency class (FEC) associated with an MPLS LSP. Session establishment is done via LSP ping protocol. Parameters such as timers and detection failure rate are negotiated initially and can also be changed on the fly (by the operator) should a need arise (e.g., to speed up or reduce the detection failure rate). BFD can be used in the following modes:

- Connectivity verification in asynchronous mode: BFD packets are sent continuously

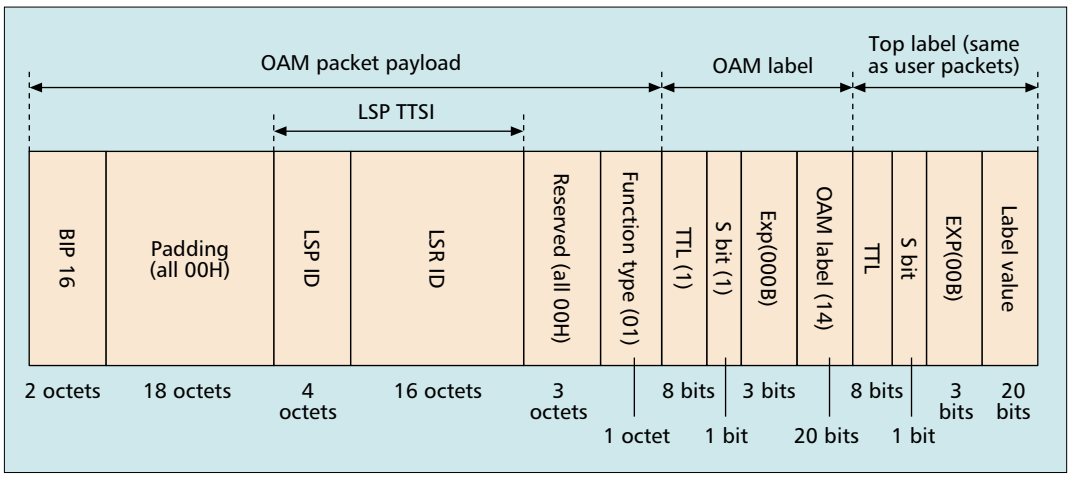

Figure 2. *Connectivity verification packet format.*

from one system to the remote peer. Connectivity is declared lost upon missing a consecutive number of packets. The threshold number is configurable and should take into account several parameters, such as link characteristics. On demand verification is also possible to assess link/media status.

- Loopback test (echo function) : BFD packets are sent to the remote peer, which in turn transmits them back to the originating node. Typically, echo mode provides forwarding plane testing. One can also perform connectivity verification with the echo function.

As for any detection protocol, timers and frequency of the probe packets need to be correctly chosen. The biggest advantage of BFD is the possibility to fine tune those parameters to satisfy specific detection criteria for either mode.

While CV and BFD are defect detection functions, MPLS LSP ping and MPLS LSP traceroute, to be introduced shortly, are diagnostic functions. These two types of functions complement each other: CV and BFD detect defects first, while ping and traceroute localize where the defects occurred.

MPLS Fault Localization and Management

In this section we address protocols that facilitate defect localization. Ping, traceroute protocols, the LSR test procedure, as well as LSP hierarchy and its relation to defect notification are addressed.

Defect Localization on a Failed Label Switched Path

MPLS LSP Ping — Reference [9] is a recent proposal for MPLS LSP ping and traceroute functions. MPLS echo request and reply, defined as UDP packets, are used to implement these functions (Fig. 3). The header of an MPLS echo request carries a return code field, which allows the far end router (LSR at the end of the LSP) to indicate whether a given prefix is reachable via the LSP at which the MPLS echo message was sent. This prefix can be of various types, such as IPv4, IPv6, Resource Reservation Protocol (RSVP), a layer 2/3 virtual private network (L2/L3 VPN), and so on. Optionally, a downstream mapping type length value (TLV) is carried by the echo request as a way to check consistency of labels within the LSP. An echo reply has the same packet format as the request. Figure 3 illustrates a MPLS echo packet format.

An LSR wishing to ping a given prefix over a particular LSP originates an MPLS echo UDP packet with IP TTL = 1 (the packet does not leave the MPLS domain) and label TTL=255 (the packet should be able to reach the last LSR of a LSP), and an FEC TLV containing the prefix to be pinged. An intermediate LSR, upon receiving the MPLS echo request, checks whether it is a transit or egress LSR for that LSP (error conditions are also covered). The LSR also checks whether the label used in its incoming interface has been advertised for the FEC

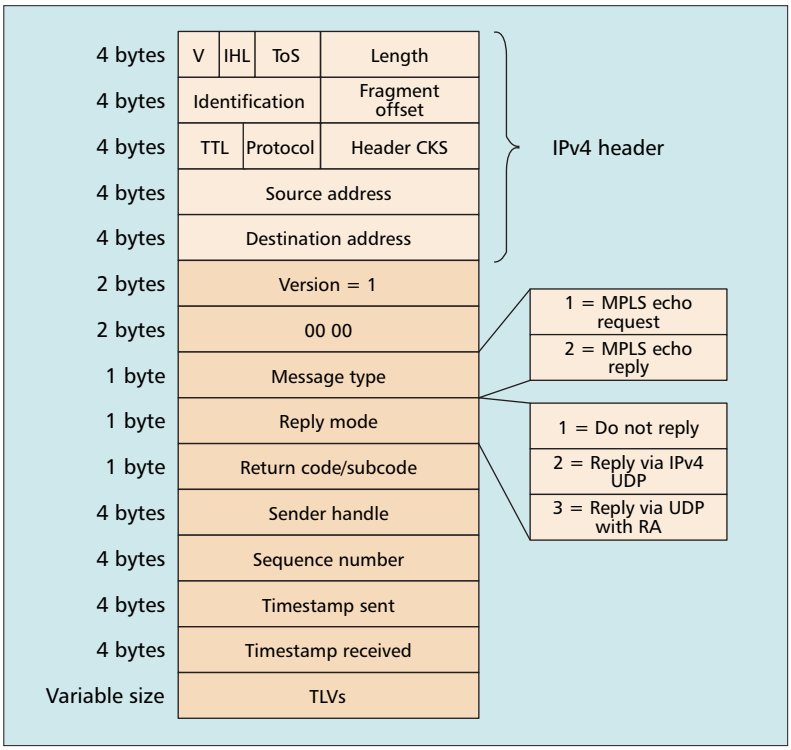

Figure 3. *MPLS echo request/reply packet format.*

corresponding to the prefix carried in the echo request. Furthermore, if a downstream mapping TLV is present in the request, the LSR in question checks if its router ID or one of its interface addresses matches one of the addresses in the downstream mapping TLV of the request message. An MPLS echo reply UDP packet is generated with the results of these checks. The echo reply has its TTL set to 255, and a return code, reflecting the status of the checks, destination IP address, and UDP port, which are copied from the source address/port of the request packet. If a downstream mapping TLV was contained in the request packet, the LSR issuing the reply computes its own downstream LSR address and label, and includes those in the echo reply packet. Timestamps are used as an aid to latency estimation of future OAM protocols.

MPLS LSP Traceroute — An LSR wishing to traceroute a given LSP issues a series of MPLS echo requests with increasing label TTL values. Initially, it issues an echo request with label TTL = 1 and an FEC TLV with the prefix of its next neighbor in the LSP of interest. The next neighbor LSR receives the request, and processes it in the same manner as described under MPLS ping, issuing an MPLS echo reply. Upon receipt of the echo reply, the LSR that has initiated the traceroute copies the downstream mapping TLV received in the reply onto its next echo request, increments the label TTL value to reach the next LSR in the LSP, and sends its next echo request. A series of echo replies are gathered as LSP traceroute information. The protocol is similar to IP traceroute, except that LSP routes are not expected to change.

Often, LSPs for a given FEC may have multi-

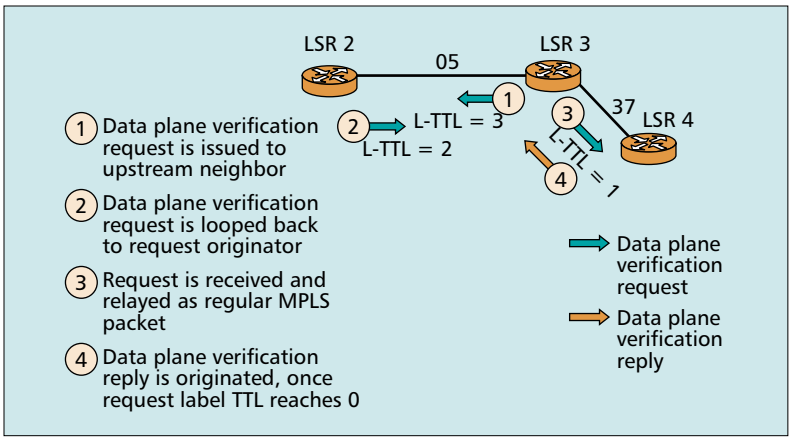

Figure 4. *Data plane LSP verification protocol.*

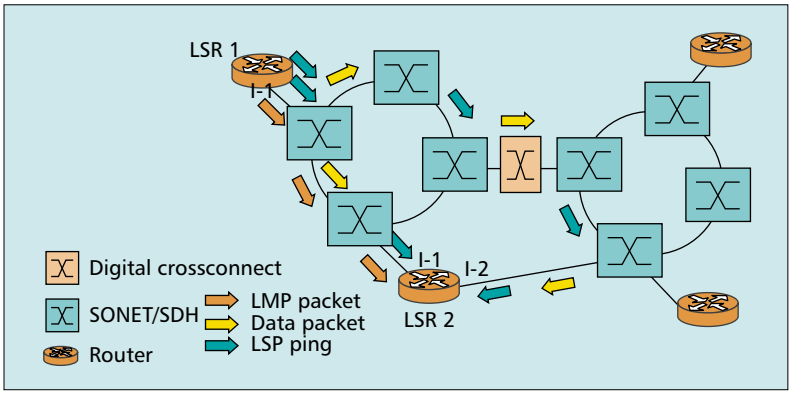

Figure 5. *LSP OAM and LMP.*

ple next hops at transit LSRs. LSPs may have backup paths, detour paths, and other alternative paths to take in the case of a failure in the primary LSP. Altough it is desirable that MPLS echo requests exercise all possible paths, this may not be practical, as the algorithms that a given LSP uses to load balance packets over alternative paths may not be publicly available. To achieve some degree of coverage on alternate paths, the MPLS ping/trace mechanism may use 127/8 addresses as the destination addresses of the MPLS echo request packet. This destination address might affect load balancing if the LSR uses it in the IP header as the decision for load balancing. Furthermore, in the case of traceroute, each transit LSR will provide information about how each of its downstream routers can be reached. The ingress can then send MPLS echo requests that exercise these paths. Table 2 summarizes ITU and IETF MPLS ping and traceroute characteristics.

LSP Self-Test — MPLS traceroute exposes the whole set of LSP labels of a given LSP. However, it may be useful for an LSR to test only its local label biding, perhaps as part of an LSP fault localization procedure for a defective LSP. An LSP self-test proposal [10] addresses this issue. Figure 4 illustrates the self-test protocol.

The idea is for an LSR willing to test its LSP label biding (LSR 3) to issue an MPLS data verification request to an upstream neighbor (LSR2) to send a labeled packet back to itself over a certain incoming interface with an expected label. The packet request has its label TTL value set to expire at the next downstream neighbor from the LSR initiating the protocol (LSR3), in this case LSR 4. The MPLS packet sent by LSR2 is processed as any other labeled packet by LSR 3 and forwarded over to LSR 4. Upon TTL expiration, LSR 4 sends a reply to LSR 3, completing the test. The request and reply messages are special LSP ping messages, optimized for fast processing (e.g., no timestamps are supported; see Fig. 3).

MPLS Link Management Protocol — Another fault management initiative is the Link Management Protocol [18]. In essence, LMP runs between a pair of generic nodes that have established a control channel adjacency between two IP interfaces for management purposes. Once the adjacency is established, a hello protocol is used to monitor control link connectivity. The protocol includes the relaying of faults detected over multiple parallel data links and fault localization features. That is, once a failure is detected on a given direction of a link, channel status messages are sent on the reverse direction to indicate to an upstream or downstream node the location of the faulty link. These messages may also be relayed to adjacencies that are chained together in both upstream/downstream directions so that other nodes may locate the failure. This is analogous to Alarm Indication Signal and Remote Defect Indication messages in ATM technology.

Notice that the LMP is intended to operate over control channels that may run over diverse

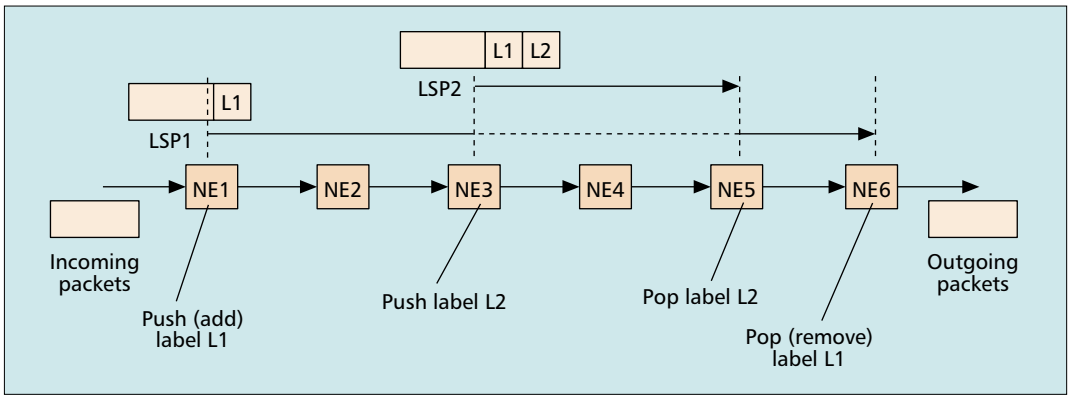

Figure 6. *Maintenance entity for MPLS OAM.*

data links. That is, its purpose is to manage control capabilities of routers, not data forwarding capabilities, as is the case for LSP fault management protocols. A healthy control channel is typically used to exchange IP routing information, signaling, and management information between IP interfaces connected via a variety of transport technologies, such as Ethernet, ATM, and synchronous optical network/synchronous digital hierarchy (SONET/SDH.) Moreover, tests on data links can be coordinated by LMP messages running over control channels. Figure 5 illustrates the distinction between data and control planes with regard to LMP and LSP OAM protocols.

In the figure LSR1 and LSR2 establish an LMP control adjacency between LSR1 interface I-1 and LSR2 interface I-1. They use this control interface to coordinate tests to be executed in two distinct LSPs , one connecting I-1 of LSR1 to I-1 of LSR2, and another connecting I-1 of LSR1 with I-2 of LSR2. LSP ping messages are used to test data link connectivity on both LSPs.

LSP Characterization and Hierarchy

A maintenance entity is an object to which OAM functions are applied. In ITU-T Recommendation Y.1711, the maintenance entity is an LSP. OAM functions can be applied to each level of LSPs when LSPs are nested. Unlike ATM, which defines both segment and end-to-end OAM functions, only end-to-end OAM functions are defined in Y.1711. However, there are cases where supervision of a specific span that is not an end-to-end LSP is necessary because the span to be supervised does not always coincide with the entire LSP. In that case, the span in question can be supervised by defining a new LSP over that span and accommodating LSPs that need to be supervised. Figure 6 shows an example. LSP1 is defined from NE1 to NE6. If one needs to supervise from NE3 to NE5, one can define an LSP2 using label stacking. NE3 pushes (adds) a label L2 on top of L1 and NE5 pops (removes) label L2. With this operation, LSP2 is defined from NE3 to NE5, accommodating LSP1. The span from NE3 to NE5 can be supervised using OAM functions for LSP2.

LSP Defect Notification

ITU Recommendation Y.1711 [6] specifies forward defect indication (FDI) and backward defect indication (BDI) as defect information transfer functions, in the forward and backward directions, respectively. The purposes of these functions is to announce the existence of a defect and suppress unnecessary alarms. When the defect detection is announced, nodes downstream and upstream take necessary action (e.g., stop usage-based billing). When a transmission line that accommodates several LSPs fails, dLOCV is detected at the egress of each LSP. Since these dLOCVs are secondary defects caused by a defect on the transmission line, they should be suppressed if the root cause is properly handled. FDI packets are generated and transmitted downstream every second by the LSR that detects a server (lower) layer defect, which may include defects in lower-level LSPs when LSPs are nested using label stacking. When an egress LSR detects a dLOCV and this LSP accommodates another higher-level LSP, the egress LSR generates FDI in that higher level LSP. Figure 7 illustrates how alarm suppression works. When the link between E and F breaks, LSPs A-H, B-J, and C-K are interrupted. Then H, J ,and K detect dLOCV. LSR F detects the link defect and reports it to the OSS. Then F generates FDI packets toward H, J, and K. When H, J and K receive FDI packets, they do not report dLOCV to the operations support system (OSS) even though they detect dLOCV, because they understand that this defect has already been detected and reported upstream.

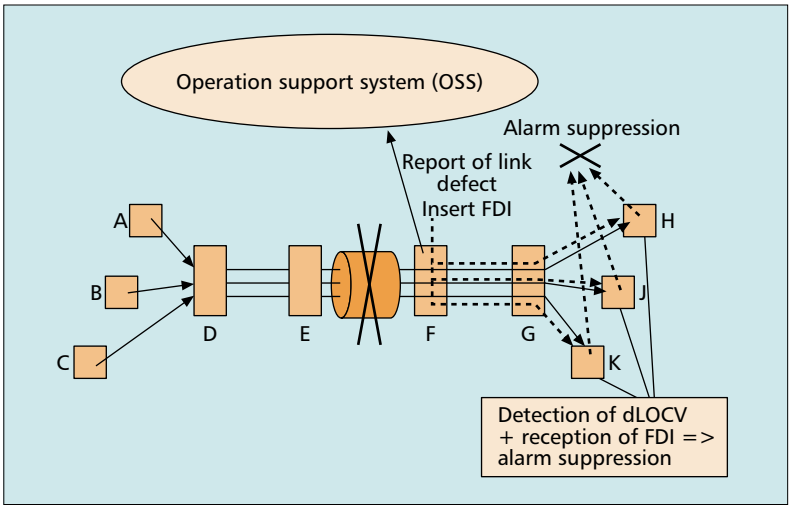

Figure 7. *Alarm suppression mechanism.*

BDI is designed for indicating the existence and detection of a defect upstream. When an egress LSR receives an FDI packet, it generates a BDI packet and transmits it in the backward direction if an LSP in this direction exists. Since LSPs are unidirectional, if there is no LSP in the backward direction BDI would be either not transmitted or transmitted by other methods.

MPLS Service Management

In this section we discuss issues related to the management of an MPLS domain. Specifically, we address MPLS tools to support SLAs, including LSP packet loss, throughput, and protection mechanisms, and discuss OAM domains.

Service Level Agreement Measurement

Traditionally, IP services have been limited to datagram applications with loose or no QoS guarantees. For these applications, connectivity was the main concern, so IP ping and traceroute functions were quite adequate. This is no longer the case, as real-time applications such as IP telephony and videoconferencing have made their way into the Internet. In the near future, stringent QoS requirements must be met by MPLS networks if such point-to-point applications are to become widespread and generate revenue. As noted in [14], LSP performance monitoring involves the measurement of LSPs' packet service properties such as service availability, packet delay, jitter, and loss. Service availability within an SLA should be calculated based on some of these measurers, depending

For a given UNI, the incoming traffic can be delivered to all other UNIs, to some other UNIs, or to any of the other UNIs, depending on the service provided. QoS may need to be supported, which may entail the definition of packet loss and delay guarantees.

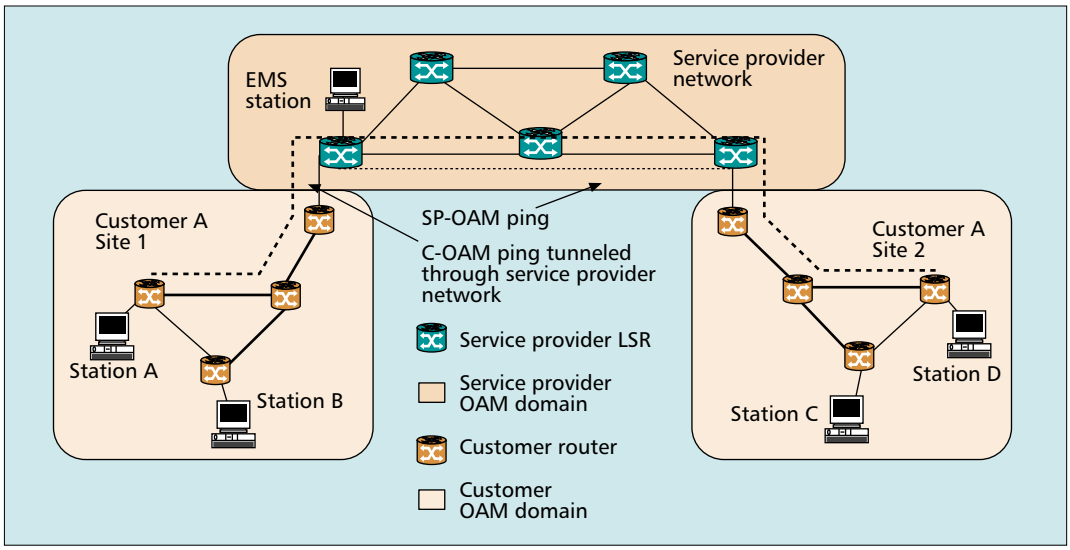

Figure 8. *Customer and service provider OAM domains.*

on the type of SLA. Morever, SLAs are likely to contain QoS guarantees, such as packet delay, jitter, and loss. Therefore, delay, jitter, and loss measurements of a packet flow are indispensable. Furthermore, measurement of the throughput may also be necessary.

MPLS Protection and Recovery

Protection and recovery can be classified as either local or global. Global protection applies to an LSP end-to-end, while local protection applies to a failed link or node. Protection restoration time is a direct function of where the fault is processed. Local protection generally provides a shorter convergence time because the fault message is processed by the local node (link protection) or an immediately adjacent node (node protection). Convergence times of the order of 50 ms are achievable. Path protection requires a longer convergence time as the fault message must propagate back to the head-end, which initiates the protection mechanism.

The simplest method is 1+1 and 1:1 protection switching architectures specified by ITU-T Recommendation Y.1720 [11]. In 1+1 architecture, traffic is copied and sent both to working and protection LSPs simultaneously. In 1:1 architecture, traffic is sent to either a working or protection LSP. As such, the protection LSP can be used to convey other traffic (so-called extra traffic) when not being used to carry working traffic. The efficiency of protection bandwidth may be further improved via mesh protection.

Based on the signaling protocol used, several resilience options are available. For example, an LDP-based network might rely on fast IGP convergence and high-availability enhancements to LDP.An RSVP with traffic engineering (RSVP-TE)-based network allows the use of multiple explicit paths when signaling an LSP. Upon primary path failure, an optimization procedure may signal the other paths, enabling traffic to be switched to the new working LSP. Moreover, RSVP-TE has been enhanced to provide local repair of LSP tunnels via establishment of a bypass LSP [12] that allows a single backup tunnel to protect several primary LSPs (*n*:*m* protection model).

Point-to-Point and Multipoint Service Management

Telecommunication service providers have long been providing point-to-point services, with the OAM framework already developed geared toward these services. However, the management of multipoint services such as VPNs and distributed storage systems that are gaining popularity is not well understood as of this writing.

Typically, user-network interfaces (UNIs) are used to characterize customer/provider points of attachment. In the multipoint case, N (> 2) UNIs are needed, with each UNI having an incoming traffic profile (traffic injected into the service provider network) and an outgoing traffic profile (traffic exiting the provider network). For a given UNI, the incoming traffic can be delivered to all other UNIs, to some other UNIs, or to any of the other UNIs, depending on the service provided. QoS may need to be supported (e.g., broadcast video applications), which may entail the definition of packet loss and delay guarantees. Such guarantees may also need to be made on a per-UNI (or even pair-of-interfaces) basis, as interfaces of diverse link speeds may be involved.

Once appropriate SLAs are defined, the challenge is the reservation of resources to achieve the performance objectives stated in the SLAs, and the traffic engineering of multipoint services. Finally, protection requirements need to be addressed for multipoint services as well, since the per-path protection approach used for point-to-point services may not scale for multipoint services, given $N(N-1)$ paths are involved in an N-node multipoint service, especially if QoS requirements are involved.

OAM Domains

OAM functions operate on network resources under control of a given entity: a network operator, a carrier, or an enterprise. Often, multiple

entities must interact to provide services across a wide geographical area, bringing OAM domains into play. An OAM domain encompasses network resources that are visible, controllable, and manageable by a given operator. It may involve multiple transport technologies, such as SONET/SDH, ATM, IP, and Ethernet. Two basic domains are the customer network and the provider network, illustrated in Fig. 8.

Other OAM Considerations

The list of OAM functions so far discussed is by no means exhaustive. Other MPLS OAM functions are likely to appear in the future, as new IP/MPLS services and applications emerge.

One area requiring attention is the profusion of worms and denial of service (DoS) attacks in the Internet today. MPLS OAM must support detection and neutralization of DoS attacks [19, 20]. DoS attack detection is based on screening data packet streams via a packet selection process using filtering or sampling [19]. In filtering a mask is used to select packets with a certain property, such as an IP source/destination address, a TCP port, an FEC, or an LSP label. With sampling, packets are selected based on a random process and thus may not have any resemblance to each other. Once a DoS attack is detected, mechanisms can be used to neutralize it and track its origin [20].

Conclusion

With the increase of data services, an efficient OAM platform is the key to a successful business model of a service provider/carrier. We have discussed issues in the design of OAM protocols for MPLS packet technologies, as well as recent proposals.

Large-scale MPLS deployment must be accompanied by appropriate OAM tools to efficiently manage packet networks, since service providers and carriers are willing to adopt new technologies only if they come with appropriate management tools.

Acknowledgments

The authors are indebted to the many members of the IETF and ITU-T SG13, Q 3/13 dealing with MPLS OAM for fruitful technical discussions, and also acknowledge the comments of the reviewers.

References

[1] G. Kessler *et al.*, "A Primer On Internet and TCP/IP Tools and Utilities," IETF RFC 2151, June 1997.
[2] J. Postel, "Echo Protocol," RFC 862, May 1983.
[3] E. Rosen *et al.*, "Multiprotocol Label Switching Architecture," IETF RFC 3031, Jan. 2001.
[4] E. Rosen *et al.*, "MPLS Label Stack Encoding," IETF RFC 3032, Jan. 2001.
[5] ITU-T Rec. Y.1710, "Requirements for OAM Functionality in MPLS Networks," Nov. 2002.
[6] ITU-T Rec. Y.1711, "Operation & Maintenance Mechanism for MPLS Networks," Feb. 2004.
[7] ITU-T Rec. Y.1713, "Misbranching Detection for MPLS Networks," Mar. 2004.
[8] R. Aggarwal and K. Kompella, "BFD for MPLS LSPs," IETF Internet draft draft-raggarwa-mpls-bfd-00.txt, Oct. 2003.
[9] K. Kompella *et al.*, "Detecting MPLS Data Plane Failures," IETF Internet draft draft-ietf-mpls-lsp-ping-04.txt, Oct. 2003.
[10] G. Swallow, K. Kompella, D. Tappan, "Label Switching Router Self-Test," IETF Internet draft draft-ietf-mpls-lsr-self-test-01.txt, Oct. 2003.
[11] ITU-T Rec. Y.1720, "Protection Switching for MPLS Networks," Sept. 2003.
[12] Ping Pan *et al.*, "Fast Reroute Extension to RSVP-TE for LSP tunnel" IETF Internet draft ,draft-ietf-mpls-rsvp-lsp-fastreroute-03.txt, Dec. 2003.
[13] V. Sharma *et al.*, "Framework for Multi-Protocol Label Switching (MPLS)-Based Recovery," IETF RFC 3469, Feb. 2003.
[14] ITU-T Draft Rec. Y.1561, "Performance and Availability Parameters for MPLS Networks," TD45(PLEN), Geneva, Switzerland, Feb. 2004.
[15] T. D. Nadeau *et al.*, "OAM Requirements for MPLS Networks," IETF Internet draft draft-ietf-mpls-oam-requirements-01.txt, June 2003.
[16] H. Ohta, "Assignment of the "OAM Alert Label" for Multiprotocol Label Switching Architecture (MPLS) Operation and Maintenance (OAM) Functions," IETF RFC 3429, Nov. 2002.
[17] D. Katz, and D. Ward, "Bidirectional Forwarding Detection," IETF Internet Draft draft-katz-ward-bfd-01.txt, Aug. 2003.
[18] J. Lang, "Link Management Protocol (LMP)," IETF Internet draft draft-ietf-ccamp-lmp-10.txt, Oct. 2003.
[19] T. Zseby *et al.*, "Sampling and Filtering Techniques for IP Packet Selection," IETF Internet draft draft-ietf-psamp-sample-tech-02.txt, June 2003.
[20] A. Belenky and N. Ansari, "On IP Traceback," *IEEE Commun. Mag.*, July 2003, pp. 142–53.

Biographies

DIRCEU CAVENDISH (dirceu@sv.nec-labs.com) received his Bachelor's degree in electronics from Federal University of Pernambuco, Brazil, in 1986. He spent five years as a development engineer at the Business Communications Division of Philips. He received his M.S. in computer science from Kyushu Institute of Technology, Japan, in 1994, and his Ph.D. from the Computer Science Department of the University of California at Los Angeles in 1998, working with congestion control and routing in QoS-supportive high-speed networks. Since 1998 he has been with NEC Laboratories America, Princeton, New Jersey, and most recently Cupertino, California. His current research interests include network management, optical transport technologies, and Web services.

HIROSHI OHTA [M] (ohta.hiroshi@lab.ntt.co.jp) received his B.S. degree in electrical engineering, and his M.S. and Dr. Eng. degrees in electronics engineering from Kyoto University, Japan, in 1985, 1987, and 2000, respectively. He joined Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Kanagawa, Japan in 1987. From 1987 to 1999 he was engaged in research and development of ATM-based transport systems, in particular optical subscriber loops, cell loss analysis/recovery, OAM functions, and protection switching, as well as development of an ATM crossconnect system. Since 2000 he has been engaged in development of services for corporate users such as IP-VPN and metro Ethernet services, and development of services for consumers such as content delivery services. Since 1992 he has actively participated in standardization meetings of ITU-T SG13, IEEE 802 LAN/MAN Standards Committee, and the IETF. He is a rapporteur for Question 3/13 (OAM and network management) of ITU-T SG13. He is currently a senior research engineer at NTT Network Service Systems Laboratories. He is a member of the IEICE Japan.

HARI RAKOTORANTO (hrakotor@cisco.com) received his Master of Science in applied mathematics from INSA, France, and McGill University, Canada. He then joined Novell working on UNIX kernel design, routing protocols, and integration between Netware and UNIX services. He spent four years at Bay Networks focusing on network management, layer 2/3 switches, and routing protocols. Since 2000 he has been with Cisco Systems Internet Technologies Division as a technical marketing engineer working closely with service providers in the field of MPLS (VPN, traffic engineering, QoS, etc.) on technology deployment and overseeing future directions. He is an active member of ITU-T SG13 focusing mainly on MPLS, MPLS OAM, and MPLS management. He is currently product manager for GMPLS.

Large-scale MPLS deployment must be accompanied by appropriate OAM tools to efficiently manage packet networks, since service providers and carriers are willing to adopt new technologies only if they come with appropriate management tools.