Network Working Group   T. Nadeau
Request for Comments: 4377   M. Morrow
Category: Informational   G. Swallow
                                                    Cisco Systems, Inc.
                                                              D. Allan
                                                       Nortel Networks
                                                         S. Matsushima
                                                         Japan Telecom
                                                         February 2006

              Operations and Management (OAM) Requirements
           for Multi-Protocol Label Switched (MPLS) Networks

Status of This Memo

   This memo provides information for the Internet community. It does
   not specify an Internet standard of any kind. Distribution of this
   memo is unlimited.

Copyright Notice

Abstract

   This document specifies Operations and Management (OAM) requirements
   for Multi-Protocol Label Switching (MPLS), as well as for
   applications of MPLS, such as pseudo-wire voice and virtual private
   network services. These requirements have been gathered from network
   operators who have extensive experience deploying MPLS networks.

Table of Contents

1. Introduction

   This document describes requirements for user and data plane
   Operations and Management (OAM) for Multi-Protocol Label Switching
   (MPLS). These requirements have been gathered from network operators
   who have extensive experience deploying MPLS networks. This document
   specifies OAM requirements for MPLS, as well as for applications of
   MPLS.

   Currently, there are no specific mechanisms proposed to address these
   requirements. The goal of this document is to identify a commonly
   applicable set of requirements for MPLS OAM at this time.
   Specifically, a set of requirements that apply to the most common set
   of MPLS networks deployed by service provider organizations at the
   time this document was written. These requirements can then be used
   as a base for network management tool development and to guide the
   evolution of currently specified tools, as well as the specification
   of OAM functions that are intrinsic to protocols used in MPLS
   networks.

2. Document Conventions

2.1. Terminology

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in RFC 2119 [RFC2119].

   Queuing/buffering Latency: The delay caused by packet queuing (value
                              is variable since it is dependent on the
                              packet arrival rate, the packet length,
                              and the link throughput).

   Probe-based-detection: Active measurement tool that can measure
                              the consistency of an LSP [RFC4379].

   Defect: Any error condition that prevents a Label
                              Switched Path (LSP) from functioning
                              correctly. For example, loss of an
                              Interior Gateway Protocol (IGP) path will
                              most likely result in an LSP not being
                              able to deliver traffic to its
                              destination. Another example is the
                              interruption of the path for a TE tunnel.
                              These may be due to physical circuit
                              failures or failure of switching nodes to
                              operate as expected.

	Multi-vendor/multi-provider network operation typically requires agreed upon definitions of defects (when it is broken and when it is not) such that both recovery procedures and service level specification impact can be specified.
Head-end Label Switching Router (LSR):	The beginning of an LSP. A head-end LSR is also referred to as an ingress LSR.
Tail-end Label Switching Router (LSR):	The end of an LSP. A tail-end LSR is also referred to as an egress LSR.
Propagation Latency:	The delay added by the propagation of the packet through the link (fixed value that depends on the distance of the link and the propagation speed).
Transmission Latency:	The delay added by the transmission of the packet over the link, i.e., the time it takes to put the packet over the media (value that depends on the link throughput and packet length).
Processing Latency:	The delay added by all the operations related to the switching of labeled packets (value is node implementation specific and may be considered fixed and constant for a given type of equipment).
Node Latency:	The delay added by the network element resulting from of the sum of the transmission, processing, and queuing/buffering latency.
One-hop Delay:	The fixed delay experienced by a packet to reach the next hop resulting from the of the propagation latency, the transmission latency, and the processing latency.
Minimum Path Latency:	The sum of the one-hop delays experienced by the packet when traveling from the ingress to the egress LSR.

      Variable Path Latency: The variation in the sum of the delays
                                  experienced by packets transiting the
                                  path, otherwise know as jitter.

2.2. Acronyms

   ASBR: Autonomous System Border Router

   CE: Customer Edge

   PE: Provider Edge

   SP: Service Provider

   ECMP: Equal-Cost Multi-path

   LSP: Label Switched Path

   LSP Ping: Label Switched Path Ping

   LSR: Label Switching Router

   OAM: Operations and Management

   RSVP: Resource reSerVation Protocol

   LDP: Label Distribution Protocol

   DoS: Denial of Service

3. Motivations

   This document was created to provide requirements that could be used
   to create consistent and useful OAM functionality that meets
   operational requirements of those service providers (SPs) who have
   deployed or are deploying MPLS.

4. Requirements

   The following sections enumerate the OAM requirements gathered from
   service providers who have deployed MPLS and services based on MPLS
   networks. Each requirement is specified in detail to clarify its
   applicability. Although the requirements specified herein are
   defined by the IETF, they have been made consistent with requirements
   gathered by other standards bodies such as the ITU [Y1710].

4.1. Detection of Label Switched Path Defects

   The ability to detect defects in a broken LSP MUST not require manual
   hop-by-hop troubleshooting of each LSR used to switch traffic for
   that LSP. For example, it is not desirable to manually visit each
   LSR along the data plane path transited by an LSP; instead, this
   function MUST be automated and able to be performed at some operator
   specified frequency from the origination point of that LSP. This
   implies solutions that are interoperable to allow for such automatic
   operation.

   Furthermore, the automation of path liveliness is desired in cases
   where large numbers of LSPs might be tested. For example, automated
   ingress LSR to egress LSR testing functionality is desired for some
   LSPs. The goal is to detect LSP path defects before customers do,
   which requires detection and correction of LSP defects in a manner
   that is both predictable and within the constraints of the service
   level agreement under which the service is being offered. Simply
   put, the sum of the time it takes an OAM tool to detect a defect and
   the time needed for an operational support system to react to this
   defect, by possibly correcting it or notifying the customer, must
   fall within the bounds of the service level agreement in question.

   Synchronization of detection time bounds by tools used to detect
   broken LSPs is required. Failure to specify defect detection time
   bounds may result in an ambiguity in test results. If the time to
   detect broken LSPs is known, then automated responses can be
   specified with respect and regard to resiliency and service level
   specification reporting. Further, if synchronization of detection
   time bounds is possible, an operational framework can be established
   to guide the design and specification of MPLS applications.

   Although an ICMP-based ping [RFC792] can be sent through an LSP as an
   IP payload, the use of this tool to verify the defect-free operation
   of an LSP has the potential of returning erroneous results (both
   positive and negative) for a number of reasons. For example, in some
   cases, because the ICMP traffic is based on legally addressable IP
   addressing, it is possible for ICMP messages that are originally
   transmitted inside of an LSP to "fall out of the LSP" at some point
   along the path. In these cases, since ICMP packets are routable, a
   falsely positive response may be returned. In other cases, where the
   data plane of a specific LSP needs to be tested, it is difficult to
   guarantee that traffic based on an ICMP ping header is parsed and
   hashed to the same equal-cost multi-paths (ECMP) as the data traffic.

   Any detection mechanisms that depend on receiving the status via a
   return path SHOULD provide multiple return options with the
   expectation that one of them will not be impacted by the original

defect. An example of a case where a false negative might occur
would be a mechanism that requires a functional MPLS return path.
Since MPLS LSPs are unidirectional, it is possible that although the
forward LSP, which is the LSP under test, might be functioning, the
response from the destination LSR might be lost, thus giving the
source LSR the false impression that the forward LSP is defective.
However, if an alternate return path could be specified -- say IP for
example -- then the source could specify this as the return path to
the destination, and in this case, would receive a response
indicating that the return LSP is defective.

The OAM packet MUST follow the customer data path exactly in order to
reflect path liveliness used by customer data. Particular cases of
interest are forwarding mechanisms, such as ECMP scenarios within the
operator's network, whereby flows are load-shared across parallel
paths (i.e., equal IGP cost). Where the customer traffic may be
spread over multiple paths, the ability to detect failures on any of
the path permutations is required. Where the spreading mechanism is
payload specific, payloads need to have forwarding that is common
with the traffic under test. Satisfying these requirements
introduces complexity into ensuring that ECMP connectivity
permutations are exercised and that defect detection occurs in a
reasonable amount of time.

4.2. Diagnosis of a Broken Label Switched Path

The ability to diagnose a broken LSP and to isolate the failed
component (i.e., link or node) in the path is required. For example,
note that specifying recovery actions for mis-branching defects in an
LDP network is a particularly difficult case. Diagnosis of defects
and isolation of the failed component is best accomplished via a path
trace function that can return the entire list of LSRs and links used
by a certain LSP (or at least the set of LSRs/links up to the
location of the defect). The tracing capability SHOULD include the
ability to trace recursive paths, such as when nested LSPs are used.
This path trace function MUST also be capable of diagnosing LSP mis-
merging by permitting comparison of expected vs. actual forwarding
behavior at any LSR in the path. The path trace capability SHOULD be
capable of being executed from the head-end Label Switching Router
(LSR) and may permit downstream path components to be traced from an
intermediate mid-point LSR. Additionally, the path trace function
MUST have the ability to support ECMP scenarios described in Section
4.1.

4.3. Path Characterization

   The path characterization function is the ability to reveal details
   of LSR forwarding operations. These details can then be compared
   during subsequent testing relevant to OAM functionality. This
   includes but is not limited to:

      - consistent use of pipe or uniform time to live (TTL) models by
         an LSR [RFC3443].

      - sufficient details that allow the test origin to exercise all
         path permutations related to load spreading (e.g., ECMP).

      - stack operations performed by the LSR, such as pushes, pops,
         and TTL propagation at penultimate hop LSRs.

4.4. Service Level Agreement Measurement

   Mechanisms are required to measure the diverse aspects of Service
   Level Agreements, which include:

      - latency – amount of time required for traffic to transit the
         network

      - packet loss

      - jitter – measurement of latency variation

      - defect free forwarding – the service is considered to be
         available, or the service is unavailable and other aspects of
         performance measurement do not have meaning.

   Such measurements can be made independently of the user traffic or
   via a hybrid of user traffic measurement and OAM probing.

   At least one mechanism is required to measure the number of OAM
   packets. In addition, the ability to measure the quantitative
   aspects of LSPs, such as jitter, delay, latency, and loss, MUST be
   available in order to determine whether the traffic for a specific
   LSP is traveling within the operator-specified tolerances.

   Any method considered SHOULD be capable of measuring the latency of
   an LSP with minimal impact on network resources. See Section 2.1 for
   definitions of the various quantitative aspects of LSPs.

4.5. Frequency of OAM Execution

   The operator MUST have the flexibility to configure OAM parameters to
   meet their specific operational requirements.

   This includes the frequency of the execution of any OAM functions.
   The ability to synchronize OAM operations is required to permit a
   consistent measurement of service level agreements. To elaborate,
   there are defect conditions, such as mis-branching or misdirection of
   traffic, for which probe-based detection mechanisms that incur
   significant mismatches in their detection frequency may result in
   flapping. This can be addressed either by synchronizing the rate or
   having the probes self-identify their probe rate. For example, when
   the probing mechanisms are bootstrapping, they might negotiate and
   ultimately agree on a probing rate, therefore providing a consistent
   probing frequency and avoiding the aforementioned problems.

   One observation would be that wide-spread deployment of MPLS, common
   implementation of monitoring tools, and the need for inter-carrier
   synchronization of defect and service level specification handling
   will drive specification of OAM parameters to commonly agreed on
   values. Such values will have to be harmonized with the surrounding
   technologies (e.g., SONET/SDH, ATM) to be useful. This will become
   particularly important as networks scale and mis-configuration can
   result in churn, alarm flapping, etc.

4.6. Alarm Suppression, Aggregation, and Layer Coordination

   Network elements MUST provide alarm suppression functionality that
   prevents the generation of a superfluous generation of alarms by
   simply discarding them (or not generating them in the first place),
   or by aggregating them together, thereby greatly reducing the number
   of notifications emitted. When viewed in conjunction with the
   requirement in Section 4.7 below, this typically requires fault
   notification to the LSP egress that may have specific time
   constraints if the application using the LSP independently implements
   path continuity testing (for example, ATM I.610 Continuity check
   (CC)[I610]). These constraints apply to LSPs that are monitored.
   The nature of MPLS applications allows for the possibility of having
   multiple MPLS applications attempt to respond to defects
   simultaneously, e.g., layer-3 MPLS VPNs that utilize Traffic
   Engineered tunnels where a failure occurs on the LSP carrying the
   Traffic Engineered tunnel. This failure would affect the VPN traffic
   that uses the tunnel's LSP. Mechanisms are required to coordinate
   network responses to defects.

4.7. Support for OAM Inter-working for Fault Notification

   An LSR supporting the inter-working of one or more networking
   technologies over MPLS MUST be able to translate an MPLS defect into
   the native technology's error condition. For example, errors
   occurring over an MPLS transport LSP that supports an emulated ATM VC
   MUST translate errors into native ATM OAM Alarm Indication Signal
   (AIS) cells at the termination points of the LSP. The mechanism
   SHOULD consider possible bounded detection time parameters, e.g., a
   "hold off" function before reacting to synchronize with the OAM
   functions.

   One goal would be alarm suppression by the upper layer using the LSP.
   As observed in Section 4.5, this requires that MPLS perform detection
   in a bounded timeframe in order to initiate alarm suppression prior
   to the upper layer independently detecting the defect.

4.8. Error Detection and Recovery

   Recovery from a fault by a network element can be facilitated by MPLS
   OAM procedures. These procedures will detect a broader range of
   defects than that of simple link and node failures. Since MPLS LSPs
   may span multiple routing areas and service provider domains, fault
   recovery and error detection should be possible in these
   configurations as well as in the more simplified single-area/domain
   configurations.

   Recovery from faults SHOULD be automatic. It is a requirement that
   faults SHOULD be detected (and possibly corrected) by the network
   operator prior to customers of the service in question detecting
   them.

4.9. Standard Management Interfaces

   The wide-spread deployment of MPLS requires common information
   modeling of management and control of OAM functionality. Evidence of
   this is reflected in the standard IETF MPLS-related MIB modules
   (e.g., [RFC3813][RFC3812][RFC3814]) for fault, statistics, and
   configuration management. These standard interfaces provide
   operators with common programmatic interface access to Operations and
   Management functions and their statuses. However, gaps in coverage
   of MIB modules to OAM and other features exist; therefore, MIB
   modules corresponding to new protocol functions or network tools are
   required.

4.10. Detection of Denial of Service Attacks

   The ability to detect denial of service (DoS) attacks against the
   data or control planes MUST be part of any security management
   related to MPLS OAM tools or techniques.

4.11. Per-LSP Accounting Requirements

   In an MPLS network, service providers can measure traffic from an LSR
   to the egress of the network using some MPLS related MIBs, for
   example. This means that it is reasonable to know how much traffic
   is traveling from location to location (i.e., a traffic matrix) by
   analyzing the flow of traffic. Therefore, traffic accounting in an
   MPLS network can be summarized as the following three items:

      (1) Collecting information to design network

         For the purpose of optimized network design, a service
         provider may offer the traffic information. Optimizing
         network design needs this information.

      (2) Providing a Service Level Specification

         Providers and their customers MAY need to verify high-level
         service level specifications, either to continuously optimize
         their networks, or to offer guaranteed bandwidth services.
         Therefore, traffic accounting to monitor MPLS applications is
         required.

      (3) Inter-AS environment

         Service providers that offer inter-AS services require
         accounting of those services.

   These three motivations need to satisfy the following:

      - In (1) and (2), collection of information on a per-LSP
         basis is a minimum level of granularity for collecting
         accounting information at both of ingress and egress of an
         LSP.

      - In (3), SP's ASBR carry out interconnection functions as an
         intermediate LSR. Therefore, identifying a pair of ingress
         and egress LSRs using each LSP is needed to determine the
         cost of the service that a customer is using.

4.11.1. Requirements

   Accounting on a per-LSP basis encompasses the following set of
   functions:

      (1) At an ingress LSR, accounting of traffic through LSPs that
          begin at each egress in question.

      (2) At an intermediate LSR, accounting of traffic through LSPs for
          each pair of ingress to egress.

      (3) At egress LSR, accounting of traffic through LSPs for each
          ingress.

      (4) All LSRs containing LSPs that are being measured need to have
          a common identifier to distinguish each LSP. The identifier
          MUST be unique to each LSP, and its mapping to LSP SHOULD be
          provided whether from manual or automatic configuration.

   In the case of non-merged LSPs, this can be achieved by simply
   reading traffic counters for the label stack associated with the
   LSP at any LSR along its path. However, in order to measure
   merged LSPs, an LSR MUST have a means to distinguish the source of
   each flow so as to disambiguate the statistics.

4.11.2. Location of Accounting

   It is not realistic for LSRs to perform the described operations on
   all LSPs that exist in a network. At a minimum, per-LSP based
   accounting SHOULD be performed on the edges of the network -- at the
   edges of both LSPs and the MPLS domain.

5. Security Considerations

   Provisions to any of the network mechanisms designed to satisfy the
   requirements described herein are required to prevent their
   unauthorized use. Likewise, these network mechanisms MUST provide a
   means by which an operator can prevent denial of service attacks if
   those network mechanisms are used in such an attack.

   LSP mis-merging has security implications beyond that of simply being
   a network defect. LSP mis-merging can happen due to a number of
   potential sources of failure, some of which (due to MPLS label
   stacking) are new to MPLS.

   The performance of diagnostic functions and path characterization
   involve extracting a significant amount of information about network
   construction that the network operator MAY consider private.

6. References

6.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997.

6.2. Informative References

   [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol
              Label Switched (MPLS) Data Plane Failures", RFC 4379,
              February 2006.

   [RFC3812] Srinivasan, C., Viswanathan, A., and T. Nadeau,
              "Multiprotocol Label Switching (MPLS) Traffic Engineering
              (TE) Management Information Base (MIB)", RFC 3812, June
              2004.

   [RFC3813] Srinivasan, C., Viswanathan, A., and T. Nadeau,
              "Multiprotocol Label Switching (MPLS) Label Switching
              Router (LSR) Management Information Base (MIB)", RFC 3813,
              June 2004.

   [RFC3814] Nadeau, T., Srinivasan, C., and A. Viswanathan,
              "Multiprotocol Label Switching (MPLS) Forwarding
              Equivalence Class To Next Hop Label Forwarding Entry
              (FEC-To-NHLFE) Management Information Base (MIB)", RFC
              3814, June 2004.

   [Y1710]   ITU-T Recommendation Y.1710, "Requirements for OAM
              Functionality In MPLS Networks"

   [I610]   ITU-T Recommendation I.610, "B-ISDN operations and
              maintenance principles and functions", February 1999

   [RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an
              IANA Considerations Section in RFCs", BCP 26, RFC 2434,
              October 1998.

   [RFC792]   Postel, J., "Internet Control Message Protocol", STD 5, RFC
              792, September 1981.

   [RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing in
             Multi-Protocol Label Switching (MPLS) Networks", RFC 3443,
             January 2003.

7. Acknowledgements

Authors' Addresses

   Comments should be made directly to the MPLS mailing list
   at mpls@lists.ietf.org.

   Thomas D. Nadeau
   Cisco Systems, Inc.
   300 Beaver Brook Road
   Boxboro, MA 01719

   Phone: +1-978-936-1470
   EMail: tnadeau@cisco.com

   Monique Jeanne Morrow
   Cisco Systems, Inc.
   Glatt-Com, 2nd Floor
   CH-8301
   Switzerland

   Phone: (0)1 878-9412
   EMail: mmorrow@cisco.com

   George Swallow
   Cisco Systems, Inc.
   300 Beaver Brook Road
   Boxboro, MA 01719

   Phone: +1-978-936-1398
   EMail: swallow@cisco.com

   David Allan
   Nortel Networks
   3500 Carling Ave.
   Ottawa, Ontario, CANADA

   Phone: 1-613-763-6362
   EMail: dallan@nortel.com

   Satoru Matsushima
   Japan Telecom
   1-9-1, Higashi-Shinbashi, Minato-ku
   Tokyo, 105-7316 Japan

   Phone: +81-3-6889-1092
   EMail: satoru@ft.solteria.net