MPLS   C. Villamizar, Ed.
Internet-Draft   OCCNC
Intended status: Informational   K. Kompella
Expires: August 16, 2014   Juniper Networks
                                                              S. Amante
                                                              Apple Inc.
                                                               A. Malis
                                                                 Huawei
                                                          C. Pignataro
                                                                  Cisco
                                                      February 12, 2014

MPLS Forwarding Compliance and Performance Requirements
draft-ietf-mpls-forwarding-07

Abstract

   This document provides guidelines for implementers regarding MPLS
   forwarding and a basis for evaluations of forwarding implementations.
   Guidelines cover many aspects of MPLS forwarding. Topics are
   highlighted where implementers might otherwise overlook practical
   requirements which are unstated or under emphasized or are optional
   for conformance to RFCs but are often considered mandatory by
   providers.

Copyright Notice

Table of Contents

1. Introduction and Document Scope

   The initial purpose of this document was to address concerns raised
   on the MPLS WG mailing list about shortcomings in implementations of
   MPLS forwarding. Documenting existing misconceptions and potential
   pitfalls might potentially avoid repeating past mistakes. The
   document has grown to address a broad set of forwarding requirements.

   The focus of this document is MPLS forwarding, base pseudowire
   forwarding, and MPLS Operations, Administration, and Maintenance
   (OAM). The use of pseudowire control word, and sequence number are
   discussed. Specific pseudowire Attachment Circuit (AC) and Native
   Service Processing (NSP) are out of scope. Specific pseudowire
   applications, such as various forms of Virtual Private Network (VPN),
   are out of scope.

MPLS support for multipath techniques is considered essential by many
service providers and is useful for other high capacity networks. In
order to obtain sufficient entropy from MPLS traffic service
providers and others find it essential for the MPLS implementation to
interpret the MPLS payload as IPv4 or IPv6 based on the contents of
the first nibble of payload. The use of IP addresses, the IP
protocol field, and UDP and TCP port number fields in multipath load
balancing are considered within scope. The use of any other IP
protocol fields, such as tunneling protocols carried within IP, are
out of scope.

Implementation details are a local matter and are out of scope. Most
interfaces today operate at 1 Gb/s or greater. It is assumed that
all forwarding operations are implemented in specialized forwarding
hardware rather than on a general purpose processor. This is often
referred to as "fast path" and "slow path" processing. Some
recommendations are made regarding implementing control or management
plane functionality in specialized hardware or with limited
assistance from specialized hardware. This advise is based on
expected control or management protocol loads and on the need for
denial of service (DoS) protection.

1.1. Abbreviations

The following abbreviations are used.

AC   Attachment Circuit ([RFC3985])

ACH   Associated Channel Header (pseudowires)

ACK   Acknowledgement (TCP flag and type of TCP packet)

AIS   Alarm Indication Signal (MPLS-TP OAM)

ATM   Asynchronous Transfer Mode (legacy switched circuits)

BFD   Bidirectional Forwarding Detection

BGP   Border Gateway Protocol

CC-CV Connectivity Check and Connectivity Verification

CE   Customer Edge (LDP, RSVP-TE, other protocols)

CPU   Central Processing Unit (computer or microprocessor)

CT   Class Type ([RFC4124])

CW   Control Word ([RFC4385])

DCCP Datagram Congestion Control Protocol

DDoS Distributed Denial of Service

DM   Delay Measurement (MPLS-TP OAM)

DSCP Differentiated Services Code Point ([RFC2474])

DWDM Dense Wave Division Multiplexing

DoS   Denial of Service

E-LSP EXP-Inferred-PSC LSP ([RFC3270])

EBGP External BGP

ECMP Equal Cost Multi-Path

ECN   Explicit Congestion Notification ([RFC3168] and [RFC5129])

EL   Entropy Label ([RFC6790])

ELI   Entropy Label Indicator ([RFC6790])

EXP   Experimental (field in MPLS renamed to TC in [RFC5462])

FEC   Forwarding Equivalence Classes (LDP), also Forward Error
      Correction in other context

FR   Frame Relay (legacy switched circuits)

FRR   Fast Reroute ([RFC4090])

G-ACh Generic Associated Channel ([RFC5586])

GAL   Generic Associated Channel Label ([RFC5586])

GFP   Generic Framing Protocol (used in OTN)

GMPLS Generalized MPLS ([RFC3471])

GTSM Generalized TTL Security Mechanism ([RFC5082])

Gb/s Gigabits per second (billion bits per second)

IANA Internet Assigned Numbers Authority

   ILM   Incoming Label Map ([RFC3031])

   IP   Internet Protocol

   IPVPN Internet Protocol VPN

   IPv4   Internet Protocol version 4

   IPv6   Internet Protocol version 6

   L-LSP Label-Only-Inferred-PSC LSP ([RFC3270])

   L2VPN Layer 2 VPN

   LDP   Label Distribution Protocol ([RFC5036])

   LER   Label Edge Router ([RFC3031])

   LM   Loss Measurement (MPLS-TP OAM)

   LSP   Label Switched Path ([RFC3031])

   LSR   Label Switching Router ([RFC3031])

   MP2MP Multipoint to Multipoint

   MPLS   MultiProtocol Label Switching ([RFC3031])

   MPLS-TP MPLS Transport Profile ([RFC5317])

   Mb/s   Megabits per second (million bits per second)

   NSP   Native Service Processing ([RFC3985])

   NTP   Network Time Protocol

   OAM   Operations, Administration, and Maintenance ([RFC6291])

   OOB   Out-of-band (not carried within a data channel)

   OTN   Optical Transport Network

   P   Provider router (LDP, RSVP-TE, other protocols)

   P2MP   Point to Multi-Point

   PE   Provider Edge router (LDP, RSVP-TE, other protocols)

PHB   Per-Hop-Behavior ([RFC2475])

PHP   Penultimate Hop Popping ([RFC3443])

POS   Packet over SONET

PSC   This abbreviation has multiple interpretations.

      1.   Packet Switch Capable ([RFC3471]

      2.   PHB Scheduling Class ([RFC3270])

      3.   Protection State Coordination ([RFC6378])

PTP   Precision Time Protocol

PW   Pseudowire

QoS   Quality of Service

RA   Router Alert ([RFC3032])

RDI   Remote Defect Indication (MPLS-TP OAM)

RSVP-TE RSVP Traffic Engineering

RTP   Real-Time Transport Protocol

SCTP   Stream Control Transmission Protocol

SDH   Synchronous Data Hierarchy (European SONET, a form of TDM)

SONET Synchronous Optical Network (US SDH, a form of TDM)

T-LDP Targeted LDP (LDP sessions over more than one hop)

TC   Traffic Class ([RFC5462])

TCP   Transmission Control Protocol

TDM   Time-Division Multiplexing (legacy encapsulations)

TOS   Type of Service (see [RFC2474])

TTL   Time-to-live (a field in IP and MPLS headers)

UDP   User Datagram Protocol

   UHP   Ultimate Hop Popping (opposite of PHP)

   VCCV Virtual Circuit Connectivity Verification ([RFC5085])

   VLAN Virtual Local Area Network (Ethernet)

   VOQ   Virtual Output Queuing (switch fabric design)

   VPN   Virtual Private Network

   WG   Working Group

1.2. Use of Requirements Language

   This document is informational. The upper case [RFC2119] key words
   are not used in this document, except in the following cases.

   1. RFC 2119 keywords are used where requirements stated in this
       document are called for in referenced RFCs. In most cases the
       RFC containing the requirement is cited within the statement
       using an RFC 2119 keyword.

   2. RFC 2119 keywords are used where explicitly noted that the
       keywords indicate that operator experiences indicate a
       requirement, but there are no existing RFC requirements.

   Advice provided by this document may be ignored by implementations.
   Similarly, implementations not claiming conformance to specific RFCs
   may ignore the requirements of those RFCs. In both cases,
   implementers should consider the risk of doing so.

1.3. Apparent Misconceptions

   In early generations of forwarding silicon (which might now be behind
   us), there apparently were some misconceptions about MPLS. The
   following statements provide clarifications.

   1. There are practical reasons to have more than one or two labels
       in an MPLS label stack. Under some circumstances the label stack
       can become quite deep. See Section 2.1.

   2. The label stack MUST be considered to be arbitrarily deep.
       Section 3.27.4. "Hierarchy: LSP Tunnels within LSPs" of RFC3031
       states "The label stack mechanism allows LSP tunneling to nest to
       any depth." [RFC3031] If a bottom of the label stack cannot be
       found, but sufficient number of labels exist to forward, an LSR
       MUST forward the packet. An LSR MUST NOT assume the packet is
       malformed unless the end of packet is found before bottom of
       stack. See Section 2.1.

   3. In networks where deep label stacks are encountered, they are not
       rare. Full packet rate performance is required regardless of
       label stack depth, except where multiple pop operations are
       required. See Section 2.1.

   4. Research has shown that long bursts of short packets with 40 byte
       or 44 byte IP payload sizes in these bursts are quite common.
       This is due to TCP ACK compression [ACK-compression]. The
       following two sub-bullets constitutes advice that reflects very
       common non-negotiable requirements of providers. Implementers
       may ignore this advice but should consider the risk of doing so.

       a. A forwarding engine SHOULD, if practical, be able to sustain
           an arbitrarily long sequence of small packets arriving at
           full interface rate.

       b. If indefinite full packet rate for small packets is not
           practical, a forwarding engine MUST be able to buffer a long
           sequence of small packets inbound to the on-chip decision
           engine and sustain full interface rate for some reasonable
           average packet rate. Absent this small on-chip buffering,
           QoS agnostic packet drops can occur.

       See Section 2.3.

   5. The implementations and system designs MUST support pseudowire
       control word (CW) if MPLS-TP is supported or if ACH [RFC5586] is
       being used on a pseudowire. The implementation and system design
       SHOULD support pseudowire CW even if MPLS-TP and ACH [RFC5586]
       are not used, using instead CW and VCCV Type 1 [RFC5085] to allow
       the use of multipath in the underlying network topology without
       impacting the PW traffic. [RFC7079] does note that there are
       still some deployments where the CW is not always used. It also
       notes that many service providers do enable the CW. See
       Section 2.4.1 for more discussion on why deployments SHOULD
       enable the pseudowire CW.

   The following statements provide clarification regarding more recent
   requirements that are often missed.

1.   The implementer and system designer SHOULD support adding a
     pseudowire Flow Label [RFC6391]. Deployments MAY enable this
     feature for appropriate pseudowire types. See Section 2.4.3.

2.   The implementer and system designer SHOULD support adding an MPLS
     entropy label [RFC6790]. Deployments MAY enable this feature.
     See Section 2.4.4.

1.4. Target Audience

This document is intended for multiple audiences: implementer
(implementing MPLS forwarding in silicon or in software); systems
designer (putting together a MPLS forwarding systems); deployer
(running an MPLS network). These guidelines are intended to serve
the following purposes:

1.   Explain what to do and what not to do when a deep label stack is
     encountered. (audience: implementer)

2.   Highlight pitfalls to look for when implementing an MPLS
     forwarding chip. (audience: implementer)

3.   Provide a checklist of features and performance specifications to
     request. (audience: systems designer, deployer)

4.   Provide a set of tests to perform. (audience: systems designer,
     deployer).

The implementer, systems designer, and deployer have a transitive
supplier customer relationship. It is in the best interest of the
supplier to review their product against their customer's checklist
and secondary customer's checklist if applicable.

This document identifies and explains many details and potential pit-
falls of MPLS forwarding. It is likely that the identified set of
potential pit-falls will later prove to be an incomplete set.

2. Forwarding Issues

A brief review of forwarding issues is provided in the subsections
that follow. This section provides some background on why some of
these requirements exist. The questions to ask of suppliers is
covered in Section 3. Some guidelines for testing are provided in
Section 4.

2.1. Forwarding Basics

Basic MPLS architecture and MPLS encapsulation, and therefore packet
forwarding are defined in [RFC3031] and [RFC3032]. RFC3031 and
RFC3032 are somewhat LDP centric. RSVP-TE supports traffic
engineering (TE) and fast reroute, features that LDP lacks. The base
document for RSVP-TE based MPLS is [RFC3209].

A few RFCs update RFC3032. Those with impact on forwarding include
the following.

1. TTL processing is clarified in [RFC3443].

2. The use of MPLS Explicit NULL is modified in [RFC4182].

3. Differentiated Services is supported by [RFC3270] and [RFC4124].
    The "EXP" field is renamed to "Traffic Class" in [RFC5462],
    removing any misconception that it was available for
    experimentation or could be ignored.

4. ECN is supported by [RFC5129].

5. The MPLS G-ACh and GAL are defined in [RFC5586].

6. [RFC5332] redefines the two data link layer codepoints for MPLS
    packets.

Tunneling encapsulations carrying MPLS, such as MPLS in IP [RFC4023],
MPLS in GRE [RFC4023], MPLS in L2TPv3 [RFC4817], or MPLS in UDP
[I-D.ietf-mpls-in-udp], are out of scope.

Other RFCs have implications to MPLS Forwarding and do not update
RFC3032 or RFC3209, including:

1. The pseudowire (PW) Associated Channel Header (ACH), defined by
    [RFC5085], later generalized by the MPLS G-ACh [RFC5586].

2. The entropy label indicator (ELI) and entropy label (EL) are
    defined by [RFC6790].

A few RFCs update RFC3209. Those that are listed as updating RFC3209
generally impact only RSVP-TE signaling. Forwarding is modified by
major extension built upon RFC3209.

RFCs which impact forwarding are discussed in the following
subsections.

2.1.1. MPLS Special Purpose Labels

[RFC3032] specifies that label values 0-15 are special purpose labels
with special meanings. [I-D.ietf-mpls-special-purpose-labels]
renamed these from the term "reserved labels" used in [RFC3032] to
"special purpose labels". Three values of NULL label are defined
(two of which are later updated by [RFC4182]) and a router-alert
label is defined. The original intent was that special purpose
labels, except the NULL labels, could be sent to the routing engine
CPU rather than be processed in forwarding hardware. Hardware
support is required by new RFCs such as those defining entropy label
and OAM processed as a result of receiving a GAL. For new special
purpose labels, some accommodation is needed for LSR that will send
the labels to a general purpose CPU or other highly programmable
hardware. For example, ELI will only be sent to LSR which have
signaled support for [RFC6790] and high OAM packet rate must be
negotiated among endpoints.

[RFC3429] reserves a label for ITU-T Y.1711, however Y.1711 does not
work with multipath and its use is strongly discouraged.

The current list of special purpose labels can be found on the
"Multiprotocol Label Switching Architecture (MPLS) Label Values"
registry reachable at IANA's pages at [1].

[I-D.ietf-mpls-special-purpose-labels] introduces an IANA "Extended
Special Purpose MPLS Label Values" registry and makes use of the
"extension" label, label 15, to indicate that the next label is an
extended special purpose label and requires special handling. The
range of only 16 values for special purpose labels allows a table to
be used. The range of extended special purpose labels with 20 bits
available for use may have to be handled in some other way in the
unlikely event that in the future the range of currently reserved
values 256-1048575 are used. If only the standards action range,
16-239, and the experimental range, 240-255, are used, then a table
of 256 entries can be used.

Unknown special purpose labels and unknown extended special purpose
labels are handled the same. When an unknown special purpose label
is encountered or a special purpose label not directly handled in
forwarding hardware is encountered, the packet should be sent to a
general purpose CPU by default. If this capability is supported,
there must be an option to either drop or rate limit such packets on
a per special purpose label value basis.

2.1.2. **MPLS Differentiated Services**

   [RFC2474] deprecates the IP Type of Service (TOS) and IP Precedence
   (Prec) fields and replaces them with the Differentiated Services
   Field more commonly known as the Differentiated Services Code Point
   (DSCP) field. [RFC2475] defines the Differentiated Services
   architecture, which in other forums, is often called a Quality of
   Service (QoS) architecture.

   MPLS uses the Traffic Class (TC) field to support Differentiated
   Services [RFC5462]. There are two primary documents describing how
   DSCP is mapped into TC.

   1. [RFC3270] defines E-LSP and L-LSP. E-LSP use a static mapping of
       DSCP into TC. L-LSP uses a per LSP mapping of DSCP into TC, with
       one PHB Scheduling Class (PSC) per L-LSP. Each PSC can use
       multiple Per-Hop Behavior (PHB) values. For example, the Assured
       Forwarding service defines three PSC, each with three PHB
       [RFC2597].

   2. [RFC4124] defines assignment of a class-type (CT) to an LSP,
       where a per CT static mapping of TC to PHB is used. [RFC4124]
       provides a means to support up to eight E-LSP-like mappings of
       DSCP to TC.

   To meet Differentiated Services requirements specified in [RFC3270],
   the following forwarding requirements must be met. An ingress LER
   MUST be able to select an LSP and then apply a per LSP map of DSCP
   into TC. A midpoint LSR MUST be able to apply a per LSP map of TC to
   PHB. The number of mappings supported will be far less than the
   number of LSP supported.

   To meet Differentiated Services requirements specified in [RFC4124],
   the following forwarding requirements must be met. An ingress LER
   MUST be able to select an LSP and then apply a per LSP map of DSCP
   into TC. A midpoint LSR MUST be able to apply a per LSP map to CT
   map and then use Class Type (CT) to map TC to PHB. Since there are
   only eight allowed values of CT, only eight maps of TC to PHB need to
   be supported. The LSP label can be used directly to find the TC to
   PHB mapping, as is needed to support [RFC3270] L-LSP.

   While support for [RFC4124] and not [RFC3270] would allow support for
   only eight mappings of TC to PHB, it is common to support both and
   simply state a limit on the number of unique TC to PHB mappings which
   can be supported.

2.1.3. **Time Synchronization**

PTP or NTP may be carried over MPLS [I-D.ietf-tictoc-1588overmpls].
Generally NTP will be carried within IP with IP carried in MPLS
[RFC5905]. Both PTP and NTP benefit from accurate time stamping of
incoming packets and the ability to insert accurate time stamps in
outgoing packets. PTP correction which occurs when forwarding
requires updating a timestamp compensation field based on the
difference between packet arrival at an LSR and packet transmit time
at that same LSR.

Since the label stack depth may vary, hardware should allow a
timestamp to be placed in an outgoing packet at any specified byte
position. It may be necessary to modify layer-2 checksums or frame
check sequences after insertion. PTP and NTP timestamp formats
differ slightly. If NTP or PTP is carried over UDP/IP or UDP/IP/
MPLS, the UDP checksum will also have to be updated.

Accurate time synchronization in addition to being generally useful
is required for MPLS-TP delay measurement (DM) OAM. See
Section 2.6.4.

2.1.4. Uses of Multiple Label Stack Entries

MPLS deployments in the early part of the prior decade (circa 2000)
tended to support either LDP or RSVP-TE. LDP was favored by some for
its ability to scale to a very large number of PE devices at the edge
of the network, without adding deployment complexity. RSVP-TE was
favored, generally in the network core, where traffic engineering and
/or fast reroute were considered important.

Both LDP and RSVP-TE are used simultaneously within major Service
Provider networks using a technique known as "LDP over RSVP-TE
Tunneling". This technique allows service providers to carry LDP
tunnels inside RSVP-TE tunnels. This makes it possible to take
advantage of the Traffic Engineering and Fast Re-Route on more
expensive Inter-City and Inter-Continental transport paths. The
ingress RSVP-TE PEs places many LDP tunnels on a single RSVP-TE LSP
and carries it to the egress RSVP-TE PE. The LDP PEs are situated
further from the core, for example within a metro network. LDP over
RSVP-TE tunneling requires a minimum of two MPLS labels: one each for
LDP and RSVP-TE.

The use of MPLS FRR [RFC4090] might add one more label to MPLS
traffic, but only when FRR protection is in use (active). If LDP
over RSVP-TE is in use, and FRR protection is in use, then at least
three MPLS labels are present on the label stack on the links through
which the Bypass LSP traverses. FRR is covered in Section 2.1.7.

LDP L2VPN, LDP IPVPN, BGP L2VPN, and BGP IPVPN added support for VPN
services that are deployed by the vast majority of service providers.
These VPN services added yet another label, bringing the label stack
depth (when FRR is active) to four.

Pseudowires and VPN are discussed in further detail in Section 2.1.8
and Section 2.1.9.

MPLS hierarchy as described in [RFC4206] and updated by [RFC7074] can
in principle add at least one additional label. MPLS hierarchy is
discussed in Section 2.1.6.

Other features such as Entropy Label (discussed in Section 2.4.4) and
Flow Label (discussed in Section 2.4.3) can add additional labels to
the label stack.

Although theoretical scenarios can easily result in eight or more
labels, such cases are rare if they occur at all today. For the
purpose of forwarding, only the top label needs to be examined if PHP
is used, a few more if UHP is used (see Section 2.5). For deep label
stacks, quite a few labels may have to be examined for the purpose of
load balancing across parallel links (see Section 2.4), however this
depth can be bounded by a provider through use of Entropy Label.

2.1.5. MPLS Link Bundling

MPLS Link Bundling was the first RFC to address the need for multiple
parallel links between nodes [RFC4201]. MPLS Link Bundling is
notable in that it tried not to change MPLS forwarding, except in
specifying the "All-Ones" component link. MPLS Link Bundling is
seldom if ever deployed. Instead multipath techniques described in
Section 2.4 are used.

2.1.6. MPLS Hierarchy

MPLS hierarchy is defined in [RFC4206] and updated by [RFC7074].
Although RFC4206 is considered part of GMPLS, the Packet Switching
Capable (PSC) portion of the MPLS hierarchy are applicable to MPLS
and may be supported in an otherwise GMPLS free implementation. The
MPLS PSC hierarchy remains the most likely means of providing further
scaling in an RSVP-TE MPLS network, particularly where the network is
designed to provide RSVP-TE connectivity to the edges. This is the
case for envisioned MPLS-TP networks. The use of the MPLS PSC
hierarchy can add at least one additional label to a label stack,
though it is likely that only one layer of PSC will be used in the
near future.

2.1.7. MPLS Fast Reroute (FRR)

   Fast reroute is defined by [RFC4090]. Two significantly different
   methods are defined in RFC4090, the "One-to-One Backup" method which
   uses the "Detour LSP" and the " Facility Backup" which uses a "bypass
   tunnel". These are commonly referred to as the detour and bypass
   methods respectively.

   The detour method makes use of a presignaled LSP. Hardware
   assistance is needed for detour FRR only if necessary to accomplish
   local repair of a large number of LSP within the 10s of milliseconds
   target. For each affected LSP a swap operation must be reprogrammed
   or otherwise switched over. The use of detour FRR doubles the number
   of LSP terminating at any given hop and will increase the number of
   LSP within a network by a factor dependent on the average detour path
   length.

   The bypass method makes use of a tunnel that is unused when no fault
   exists but may carry many LSP when a local repair is required. There
   is no presignaling indicating which working LSP will be diverted into
   any specific bypass LSP. The merge LSR (egress LSR of the bypass
   LSP) MUST use platform label space (as defined in [RFC3031]) so that
   an LSP working path on any given interface can be backed up using a
   bypass LSP terminating on any other interface. Hardware assistance
   is needed if necessary to accomplish local repair of a large number
   of LSP within the 10s of milliseconds target. For each affected LSP
   a swap operation must be reprogrammed or otherwise switched over with
   an additional push of the bypass LSP label. The use of platform
   label space impacts the size of the LSR ILM for LSR with a very large
   number of interfaces.

2.1.8. Pseudowire Encapsulation

   The pseudowire (PW) architecture is defined in [RFC3985]. A
   pseudowire, when carried over MPLS, adds one or more additional label
   entries to the MPLS label stack. A PW Control Word is defined in
   [RFC4385] with motivation for defining the control word in [RFC4928].
   The PW Associated Channel defined in [RFC4385] is used for OAM in
   [RFC5085]. The PW Flow Label is defined in [RFC6391] and is
   discussed further in this document in Section 2.4.3.

   There are numerous pseudowire encapsulations, supporting emulation of
   services such as Frame Relay, ATM, Ethernet, TDM, and SONET/SDH over
   packet switched networks (PSNs) using IP or MPLS.

   The pseudowire encapsulation is out of scope for this document.
   Pseudowire impact on MPLS forwarding at midpoint LSR is within scope.
   The impact on ingress MPLS push and egress MPLS UHP pop are within

scope. While pseudowire encapsulation is out of scope, some advice
is given on sequence number support.

[2.1.8.1](#). Pseudowire Sequence Number

Pseudowire (PW) sequence number support is most important for PW
payload types with a high expectation of lossless and/or in-order
delivery. Identifying lost PW packets and the exact amount of lost
payload is critical for PW services which maintain bit timing, such
as Time Division Multiplexing (TDM) services since these services
MUST compensate lost payload on a bit-for-bit basis.

With PW services which maintain bit timing, packets that have been
received out of order also MUST be identified and MAY be either re-
ordered or dropped. Resequencing requires, in addition to sequence
numbering, a "reorder buffer" in the egress PE, and ability to
reorder is limited by the depth of this buffer. The down side of
maintaining a large reorder buffer is added end-to-end service delay.

For PW services which maintain bit timing or any other service where
jitter must be bounded, a jitter buffer is always necessary. The
jitter buffer is needed regardless of whether reordering is done. In
order to be effective, a reorder buffer must often be larger than a
jitter buffer needs to be creating a tradeoff between reducing loss
and minimizing delay.

PW services which are not timing critical bit streams in nature are
cell oriented or frame oriented. Though resequencing support may be
beneficial to PW cell and frame oriented payloads such as ATM, FR and
Ethernet, this support is desirable but not required. Requirements
to handle out of order packets at all vary among services and
deployments. For example for Ethernet PW, occasional (very rare)
reordering is usually acceptable. If the Ethernet PW is carrying
MPLS-TP, then this reordering may be acceptable.

Reducing jitter is best done by an end-system, given that the
tradeoff of loss vs delay varies among services. For example with
interactive real time services low delay is preferred, while with
non-interactive (one way) real time services low loss is preferred.
The same end-site may be receiving both types of traffic. Regardless
of this, bounded jitter is sometimes a requirement for specific
deployments.

Packet reordering should be rare except in a small number of
circumstances, most of which are due to network design or equipment
design errors:

   1. The most common case is where reordering is rare, occurring only
       when a network or equipment fault forces traffic on a new path
       with different delay. The packet loss that accompanies a network
       or equipment fault is generally more disruptive than any
       reordering which may occur.

   2. A path change can be caused by reasons other than a network or
       equipment fault, such as administrative routing change. This may
       result in packet reordering but generally without any packet
       loss.

   3. If the edge is not using pseudowire control word (CW) and the
       core is using multipath, reordering will be far more common. If
       this is occurring, using CW on the edge will solve the problem.
       Without CW, resequencing is not possible since the sequence
       number is contained in the CW.

   4. Another avoidable case is where some core equipment has multipath
       and for some reason insists on periodically installing a new
       random number as the multipath hash seed. If supporting MPLS-TP,
       equipment MUST provide a means to disable periodic hash reseeding
       and deployments MUST disable periodic hash reseeding. Operator
       experience dictates that even if not supporting MPLS-TP,
       equipment SHOULD provide a means to disable periodic hash
       reseeding and deployments SHOULD disable periodic hash reseeding.

   In provider networks which use multipath techniques and which may
   occasionally rebalance traffic or which may change PW paths
   occasionally for other reasons, reordering may be far more common
   than loss. Where reordering is more common than loss, resequencing
   packets is beneficial, rather than dropping packets at egress when
   out of order arrival occurs. Resequencing is most important for PW
   payload types with a high expectation of lossless delivery since in
   such cases out of order delivery within the network results in PW
   loss.

2.1.9. Layer-2 and Layer-3 VPN

   Layer-2 VPN [RFC4664] and Layer-3 VPN [RFC4110] add one or more label
   entry to the MPLS label stack. VPN encapsulations are out of scope
   for this document. Its impact on forwarding at midpoint LSR are
   within scope.

Any of these services may be used on an MPLS entropy label enabled
ingress and egress (see Section 2.4.4 for discussion of entropy
label) which would add an additional two labels to the MPLS label
stack. The need to provide a useful entropy label value impacts the
requirements of the VPN ingress LER but is out of scope for this
document.

2.2. MPLS Multicast

MPLS Multicast encapsulation is clarified in [RFC5332]. MPLS
Multicast may be signaled using RSVP-TE [RFC4875] or LDP [RFC6388].

[RFC4875] defines a root initiated RSVP-TE LSP setup rather than leaf
initiated join used in IP multicast. [RFC6388] defines a leaf
initiated LDP setup. Both [RFC4875] and [RFC6388] define point to
multipoint (P2MP) LSP setup. [RFC6388] also defined multipoint to
multipoint (MP2MP) LSP setup.

The P2MP LSP have a single source. An LSR may be a leaf node, an
intermediate node, or a "bud" node. A bud serves as both a leaf and
intermediate. At a leaf an MPLS pop is performed. The payload may
be a IP Multicast packet that requires further replication. At an
intermediate node a MPLS swap operation is performed. The bud
requires that both a pop operation and a swap operation be performed
for the same incoming packet.

One strategy to support P2MP functionality is to pop at the LSR
interface serving as ingress to the P2MP traffic and then optionally
push labels at each LSR interface serving as egress to the P2MP
traffic at that same LSR. A given LSR egress chip may support
multiple egress interfaces, each of which requires a copy, but each
with a different set of added labels and layer-2 encapsulation. Some
physical interfaces may have multiple sub-interfaces (such as
Ethernet VLAN or channelized interfaces) each requiring a copy.

If packet replication is performed at LSR ingress, then the ingress
interface performance may suffer. If the packet replication is
performed within a LSR switching fabric and at LSR egress, congestion
of egress interfaces cannot make use of backpressure to ingress
interfaces using techniques such as virtual output queuing (VOQ). If
buffering is primarily supported at egress, then the need for
backpressure is minimized. There may be no good solution for high
volumes of multicast traffic if VOQ is used.

Careful consideration should be given to the performance
characteristics of high fanout multicast for equipment that is
intended to be used in such a role.

   MP2MP LSP differ in that any branch may provide an input, including a
   leaf. Packets must be replicated onto all other branches. This
   forwarding is often implemented as multiple P2MP forwarding trees,
   one for each potential input interface at a given LSR.

2.3. Packet Rates

   While average packet size of Internet traffic may be large, long
   sequences of small packets have both been predicted in theory and
   observed in practice. Traffic compression and TCP ACK compression
   can conspire to create long sequences of packets of 40-44 bytes in
   payload length. If carried over Ethernet, the 64 byte minimum
   payload applies, yielding a packet rate of approximately 150 Mpps
   (million packets per second) for the duration of the burst on a
   nominal 100 Gb/s link. The peak rate for other encapsulations can be
   as high as 250 Mpps (for example IP or MPLS encapsulated using GFP
   over OTN ODU4).

   It is possible that the packet rates achieved by a specific
   implementation is acceptable for a minimum payload size, such as 64
   byte (64B) payload for Ethernet, but the achieved rate declines to an
   unacceptable level for other packet sizes, such as 65B payload.
   There are other packet rates of interest besides TCP ACK. For
   example, a TCP ACK carried over an Ethernet PW over MPLS over
   Ethernet may occupy 82B or 82B plus an increment of 4B if additional
   MPLS labels are present.

   A graph of packet rate vs. packet size often displays a sawtooth.
   The sawtooth is commonly due to a memory bottleneck and memory
   widths, sometimes internal cache, but often a very wide external
   buffer memory interface. In some cases it may be due to a fabric
   transfer width. A fine packing, rounding up to the nearest 8B or 16B
   will result in a fine sawtooth with small degradation for 65B, and
   even less for 82B packets. A course packing, rounding up to 64B can
   yield a sharper drop in performance for 65B packets, or perhaps more
   important, a larger drop for 82B packets.

   The loss of some TCP ACK packets are not the primary concern when
   such a burst occurs. When a burst occurs, any other packets,
   regardless of packet length and packet QoS are dropped once on-chip
   input buffers prior to the decision engine are exceeded. Buffers in
   front of the packet decision engine are often very small or non-
   existent (less than one packet of buffer) causing significant QoS
   agnostic packet drop.

   Internet service providers and content providers at one time
   specified full rate forwarding with 40 byte payload packets as a
   requirement. Today, this requirement often can be waived if the

provider can be convinced that when long sequence of short packets
occur no packets will be dropped.

Many equipment suppliers have pointed out that the extra cost in
designing hardware capable of processing the minimum size packets at
full line rate is significant for very high speed interfaces. If
hardware is not capable of processing the minimum size packets at
full line rate, then that hardware MUST be capable of handling large
burst of small packets, a condition which is often observed. This
level of performance is necessary to meet Differentiated Services
[RFC2475] requirements for without it, packets are lost prior to
inspection of the IP DSCP field [RFC2474] or MPLS TC field [RFC5462].

With adequate on-chip buffers before the packet decision engine, an
LSR can absorb a long sequence of short packets. Even if the output
is slowed to the point where light congestion occurs, the packets,
having cleared the decision process, can make use of larger VOQ or
output side buffers and be dealt with according to configured QoS
treatment, rather than dropped completely at random.

These on-chip buffers need not contribute significant delay since
they are only used when the packet decision engine is unable to keep
up, not in response to congestion, plus these buffers are quite
small. For example, an on-chip buffer capable of handling 4K packets
of 64 bytes in length, or 256KB, corresponds to 2 msec on a 10 Mb/s
link and 0.2 usec on a 100 Gb/s link. If the packet decision engine
is capable of handling packets at 90% of the full rate for small
packets, then the maximum added delay is 0.2 msec and 20 nsec
respectively, and this delay only applies if a 4K burst of short
packets occurs. When no burst of short packets was being processed,
no delay is added.

Packet rate requirements apply regardless of which network tier
equipment is deployed in. Whether deployed in the network core or
near the network edges, one of the two conditions MUST be met if
Differentiated Services requirements are to be met:

1. Packets must be processed at full line rate with minimum sized
    packets. -OR-

2. Packets must be processed at a rate well under generally accepted
    average packet sizes, with sufficient buffering prior to the
    packet decision engine to accommodate long bursts of small
    packets.

2.4. MPLS Multipath Techniques

In any large provider, service providers and content providers, hash
based multipath techniques are used in the core and in the edge. In
many of these providers hash based multipath is also used in the
larger metro networks.

The Differentiated Services requirements for good reasons dictate
that packets within a common microflow SHOULD NOT be reordered
[RFC2474]. Service providers generally impose stronger requirements,
commonly requiring that packets within a microflow MUST NOT be
reordered except in rare circumstances such as load balancing across
multiple links or path change for load balancing or path change for
other reason.

The most common multipath techniques are ECMP applied at the IP
forwarding level, Ethernet LAG with inspection of the IP payload, and
multipath on links carrying both IP and MPLS, where the IP header is
inspected below the MPLS label stack. In most core networks, the
vast majority of traffic is MPLS encapsulated.

In order to support an adequately balanced load distribution across
multiple links, IP header information must be used. Common practice
today is to reinspect the IP headers at each LSR and use the label
stack and IP header information in a hash performed at each LSR.
Further details are provided in Section 2.4.5.

The use of this technique is so ubiquitous in provider networks that
lack of support for multipath makes any product unsuitable for use in
large core networks. This will continue to be the case in the near
future, even as deployment of MPLS entropy label begins to relax the
core LSR multipath performance requirements given the existing
deployed base of edge equipment without the ability to add an entropy
label.

A generation of edge equipment supporting the ability to add an MPLS
entropy label is needed before the performance requirements for core
LSR can be relaxed. However, it is likely that two generations of
deployment in the future will allow core LSR to support full packet
rate only when a relatively small number of MPLS labels need to be
inspected before hashing. For now, don't count on it.

Common practice today is to reinspect the packet at each LSR and use
information from the packet combined plus a hash seed that is
selected by each LSR. Where flow labels or entropy labels are used,
a hash seed must be used when creating these labels.

2.4.1. Pseudowire Control Word

Within the core of a network some form of multipath is almost certain
to be used. Multipath techniques deployed today are likely to be
looking beneath the label stack for an opportunity to hash on IP
addresses.

A pseudowire encapsulated at a network edge must have a means to
prevent reordering within the core if the pseudowire will be crossing
a network core, or any part of a network topology where multipath is
used (see [RFC4385] and [RFC4928]).

Not supporting the ability to encapsulate a pseudowire with a control
word may lock a product out from consideration. A pseudowire
capability without control word support might be sufficient for
applications that are strictly both intra-metro and low bandwidth.
However a provider with other applications will very likely not
tolerate having equipment which can only support a subset of their
pseudowire needs.

2.4.2. Large Microflows

Where multipath makes use of a simple hash and simple load balance
such as modulo or other fixed allocation (see Section 2.4) the
presence of large microflows that each consumes 10% of the capacity
of a component link of a potentially congested composite link, one
such microflow can upset the traffic balance and more than one can in
effect reduce the effective capacity of the entire composite link by
more than 10%.

When even a very small number of large microflows are present, there
is a significant probability that more than one of these large
microflows could fall on the same component link. If the traffic
contribution from large microflows is small, the probability for
three or more large microflows on the same component link drops
significantly. Therefore in a network where a significant number of
parallel 10 Gb/s links exists, even a 1 Gb/s pseudowire or other
large microflow that could not otherwise be subdivided into smaller
flows should carry a flow label or entropy label if possible.

Active management of the hash space to better accommodate large
microflows has been implemented and deployed in the past, however
such techniques are out of scope for this document.

2.4.3. Pseudowire Flow Label

Unlike a pseudowire control word, a pseudowire flow label [RFC6391],
is required only for relatively large capacity pseudowires. There
are many cases where a pseudowire flow label makes sense. Any
service such as a VPN which carries IP traffic within a pseudowire
can make use of a pseudowire flow label.

Any pseudowire carried over MPLS which makes use of the pseudowire
control word and does not carry a flow label is in effect a single
microflow (in [RFC2475] terms) and may result in the types of
problems described in Section 2.4.2.

2.4.4. MPLS Entropy Label

The MPLS entropy label simplifies flow group identification [RFC6790]
at midpoint LSRs. Prior to the MPLS entropy label midpoint LSRs
needed to inspect the entire label stack and often the IP headers to
provide an adequate distribution of traffic when using multipath
techniques (see Section 2.4.5). With the use of MPLS entropy label,
a hash can be performed closer to network edges, placed in the label
stack, and used by midpoint LSRs without fully reinspecting the label
stack and inspecting the payload.

The MPLS entropy label is capable of avoiding full label stack and
payload inspection within the core where performance levels are most
difficult to achieve (see Section 2.3). The label stack inspection
can be terminated as soon as the first entropy label is encountered,
which is generally after a small number of labels are inspected.

In order to provide these benefits in the core, LSR closer to the
edge must be capable of adding an entropy label. This support may
not be required in the access tier, the tier closest to the customer,
but is likely to be required in the edge or the border to the network
core. LSR peering with external networks will also need to be able
to add an entropy label on incoming traffic.

2.4.5. Fields Used for Multipath Load Balance

The most common multipath techniques are based on a hash over a set
of fields. Regardless of whether a hash is used or some other method
is used, the there is a limited set of fields which can safely be
used for multipath.

2.4.5.1. MPLS Fields in Multipath

If the "outer" or "first" layer of encapsulation is MPLS, then label
stack entries are used in the hash. Within a finite amount of time
(and for small packets arriving at high speed that time can be quite
limited) only a finite number of label entries can be inspected.

Pipelined or parallel architectures improve this, but the limit is
still finite.

The following guidelines are provided for use of MPLS fields in
multipath load balancing.

1. Only the 20 bit label field SHOULD be used. The TTL field SHOULD
    NOT be used. The S bit MUST NOT be used. The TC field (formerly
    EXP) MUST NOT be used. See text following this list for reasons.

2. If an ELI label is found, then if the LSR supports entropy label,
    the EL label field in the next label entry (the EL) SHOULD be
    used and label entries below that label SHOULD NOT be used and
    the MPLS payload SHOULD NOT be used. See below this list for
    reasons.

3. Special purpose labels (label values 0-15) MUST NOT be used.
    Extended special purpose labels (any label following label 15)
    MUST NOT be used. In particular, GAL and RA MUST NOT be used so
    that OAM traffic follows the same path as payload packets with
    the same label stack.

4. If a new special purpose label or extended special purpose label
    is defined which requires special load balance processing, then,
    as is the case for the ELI label, a special action may be needed
    rather than skipping the special purpose label or extended
    special purpose label.

5. The most entropy is generally found in the label stack entries
    near the bottom of the label stack (innermost label, closest to
    S=1 bit). If the entire label stack cannot be used (or entire
    stack up to an EL), then it is better to use as many labels as
    possible closest to the bottom of stack.

6. If no ELI is encountered, and the first nibble of payload
    contains a 4 (IPv4) or 6 (IPv6), an implementation SHOULD support
    the ability to interpret the payload as IPv4 or IPv6 and extract
    and use appropriate fields from the IP headers. This feature is
    considered a non-negotiable requirement by many service
    providers. If supported, there MUST be a way to disable it (if,
    for example, PW without CW are used). This ability to disable
    this feature is considered a non-negotiable requirement by many
    service providers. Therefore an implementation has a very strong
    incentive to support both options.

7. A label which is popped at egress (UHP pop) SHOULD NOT be used.
    A label which is popped at the penultimate hop (PHP pop) SHOULD
    be used.

Apparently some chips have made use of the TC (formerly EXP) bits as
a source of entropy. This is very harmful since it will reorder
Assured Forwarding (AF) traffic [RFC2597] when a subset does not
conform to the configured rates and is remarked but not dropped at a
prior LSR. Traffic which uses MPLS ECN [RFC5129] can also be
reordered if TC is used for entropy. Therefore, as stated in the
guidelines above, the TC field (formerly EXP) MUST NOT be used in
multipath load balancing as it violates Differentiated Services
Ordered Aggregate (OA) requirements in these two instances.

Use of the MPLS label entry S bit would result in putting OAM traffic
on a different path if the addition of a GAL at the bottom of stack
removed the S bit from the prior label.

If an ELI label is found, then if the LSR supports entropy label, the
EL label field in the next label entry (the EL) SHOULD be used and
the search for additional entropy within the packet SHOULD be
terminated. Failure to terminate the search will impact client MPLS-
TP LSP carried within server MPLS LSP. A network operator has the
option to use administrative attributes as a means to identify LSR
which do not terminate the entropy search at the first EL.
Administrative attributes are defined in [RFC3209]. Some
configuration is required to support this.

If the label removed by a PHP pop is not used, then for any PW for
which CW is used, there is no basis for multipath load split. In
some networks it is infeasible to put all PW traffic on one component
link. Any PW which does not use CW will be improperly split
regardless of whether the label removed by a PHP pop is used.
Therefore the PHP pop label SHOULD be used as recommended above.

2.4.5.2. IP Fields in Multipath

Inspecting the IP payload provides the most entropy in provider
networks. The practice of looking past the bottom of stack label for
an IP payload is well accepted and documented in [RFC4928] and in
other RFCs.

Where IP is mentioned in the document, both IPv4 and IPv6 apply. All
LSRs MUST fully support IPv6.

When information in the IP header is used, the following guidelines
apply:

1. Both the IP source address and IP destination address SHOULD be
    used. There MAY be an option to reverse the order of these
    addresses, improving the ability to provide symmetric paths in
    some cases. Many service providers require that both addresses
    be used.

2. Implementations SHOULD allow inspection of the IP protocol field
    and use of the UDP or TCP port numbers. For many service
    providers this feature is considered mandatory, particularly for
    enterprise, data center, or edge equipment. If this feature is
    provided, it SHOULD be possible to disable use of TCP and UDP
    ports. Many service providers consider it a non-negotiable
    requirement that use of UDP and TCP ports can be disabled.
    Therefore there is a strong incentive for implementations to
    provide both options.

3. Equipment suppliers MUST NOT make assumptions that because the IP
    version field is equal to 4 (an IPv4 packet) that the IP protocol
    will either be TCP (IP protocol 6) or UDP (IP protocol 17) and
    blindly fetch the data at the offset where the TCP or UDP ports
    would be found. With IPv6, TCP and UDP port numbers are not at
    fixed offsets. With IPv4 packets carrying IP options, TCP and
    UDP port numbers are not at fixed offsets.

4. The IPv6 header flow field SHOULD be used. This is the explicit
    purpose of the IPv6 flow field, however observed flow fields
    rarely contains a non-zero value. Some uses of the flow field
    have been defined such as [RFC6438]. In the absence of MPLS
    encapsulation, the IPv6 flow field can serve a role equivalent to
    entropy label.

5. Support for other protocols that share a common Layer-4 header
    such as RTP [RFC3550], UDP-Lite [RFC3828], SCTP [RFC4960] and
    DCCP [RFC4340] SHOULD be provided, particularly for edge or
    access equipment where additional entropy may be needed.
    Equipment SHOULD also use RTP, UDP-lite, SCTP and DCCP headers
    when creating an entropy label.

6. The following IP header fields should not or must not be used:

    a. Similar to avoiding TC in MPLS, the IP DSCP, and ECN bits
        MUST NOT be used.

    b. The IPv4 TTL or IPv6 Hop Count SHOULD NOT be used.

    c. Note that the IP TOS field was deprecated ([RFC0791] was
        updated by [RFC2474]). No part of the IP DSCP field can be
        used (formerly IP PREC and IP TOS bits).

7. Some IP encapsulations support tunneling, such as IP-in-IP, GRE,
    L2TPv3, and IPSEC. These provide a greater source of entropy
    which some provider networks carrying large amounts of tunneled
    traffic may need, for example as used in [RFC5640] for GRE and
    L2TPv3. The use of tunneling header information is out of scope
    for this document.

This document makes the following recommendations. These
recommendations are not required to claim compliance to any existing
RFC therefore implementers are free to ignore them, but due to
service provider requirements should consider the risk of doing so.
The use of IP addresses MUST be supported and TCP and UDP ports
(conditional on the protocol field and properly located) MUST be
supported. The ability to disable use of UDP and TCP ports MUST be
available. Though potentially very useful in some networks, it is
uncommon to support using payloads of tunneling protocols carried
over IP. Though the use of tunneling protocol header information is
out of scope for this document, it is not discouraged.

2.4.5.3. Fields Used in Flow Label

The ingress to a pseudowire (PW) can extract information from the
payload being encapsulated to create a flow label. [RFC6391]
references IP carried in Ethernet as an example. The Native Service
Processing (NSP) function defined in [RFC3985] differs with
pseudowire type. It is in the NSP function where information for a
specific type of PW can be extracted for use in a flow label. Which
fields to use for any given PW NSP is out of scope for this document.

2.4.5.4. Fields Used in Entropy Label

An entropy label is added at the ingress to an LSP. The payload
being encapsulated is most often MPLS, a PW, or IP. The payload type
is identified by the layer-2 encapsulation (Ethernet, GFP, POS, etc).

If the payload is MPLS, then the information used to create an
entropy label is the same information used for local load balancing
(see Section 2.4.5.1). This information MUST be extracted for use in
generating an entropy label even if the LSR local egress interface is
not a multipath.

Of the non-MPLS payload types, only payloads that are forwarded are
of interest. For example, ARP is not forwarded and CNLP (used only
for ISIS) is not forwarded.

The non-MPLS payload type of greatest interest are IPv4 and IPv6.
The guidelines in Section 2.4.5.2 apply to fields used to create and
entropy label.

   The IP tunneling protocols mentioned in Section 2.4.5.2 may be more
   applicable to generation of an entropy label at edge or access where
   deep packet inspection is practical due to lower interface speeds
   than in the core where deep packet inspection may be impractical.

2.5. MPLS-TP and UHP

   MPLS-TP introduces forwarding demands that will be extremely
   difficult to meet in a core network. Most troublesome is the
   requirement for Ultimate Hop Popping (UHP, the opposite of
   Penultimate Hop Popping or PHP). Using UHP opens the possibility of
   one or more MPLS pop operation plus an MPLS swap operation for each
   packet. The potential for multiple lookups and multiple counter
   instances per packet exists.

   As networks grow and tunneling of LDP LSPs into RSVP-TE LSPs is used,
   and/or RSVP-TE hierarchy is used, the requirement to perform one or
   two or more MPLS pop operations plus a MPLS swap operation (and
   possibly a push or two) increases. If MPLS-TP LM (link monitoring)
   OAM is enabled at each layer, then a packet and byte count MUST be
   maintained for each pop and swap operation so as to offer OAM for
   each layer.

2.6. Local Delivery of Packets

   There are a number of situations in which packets are destined to a
   local address or where a return packet must be generated. There is a
   need to mitigate the potential for outage as a result of either
   attacks on network infrastructure, or in some cases unintentional
   misconfiguration resulting in processor overload. Some hardware
   assistance is needed for all traffic destined to the general purpose
   CPU that is used in MPLS control protocol processing or network
   management protocol processing and in most cases to other general
   purpose CPUs residing on an LSR. This is due to the ease of
   overwhelming such a processor with traffic arriving on LSR high speed
   interfaces, whether the traffic is malicious or not.

   Denial of service (DoS) protection is an area requiring hardware
   support that is often overlooked or inadequately considered.
   Hardware assist is also needed for OAM, particularly the more
   demanding MPLS-TP OAM.

2.6.1. DoS Protection

   Modern equipment supports a number of control plane and management
   plane protocols. Generally no single means of protecting network
   equipment from denial of service (DoS) attacks is sufficient,
   particularly for high speed interfaces. This problem is not specific

to MPLS, but is a topic that cannot be ignored when implementing or
evaluating MPLS implementations.

Two types of protections are often cited as primary means of
protecting against attacks of all kinds.

Isolated Control/Management Traffic
    Control and Management traffic can be carried out-of-band (OOB),
    meaning not intermixed with payload. For MPLS, use of G-ACh and
    GAL to carry control and management traffic provides a means of
    isolation from potentially malicious payload. Used alone, the
    compromise of a single node, including a small computer at a
    network operations center, could compromise an entire network.
    Implementations which send all G-ACh/GAL traffic directly to a
    routing engine CPU are subject to DoS attack as a result of such
    a compromise.

Cryptographic Authentication
    Cryptographic authentication can very effectively prevent
    malicious injection of control or management traffic.
    Cryptographic authentication can in some circumstances be subject
    to DoS attack by overwhelming the capacity of the decryption with
    a high volume of malicious traffic. For very low speed
    interfaces, cryptographic authentication can be performed by the
    general purpose CPU used as a routing engine. For all other
    cases, cryptographic hardware may be needed. For very high speed
    interfaces, even cryptographic hardware can be overwhelmed.

Some control and management protocols are often carried with payload
traffic. This is commonly the case with BGP, T-LDP, and SNMP. It is
often the case with RSVP-TE. Even when carried over G-ACh/GAL
additional measures can reduce the potential for a minor breach to be
leveraged to a full network attack.

Some of the additional protections are supported by hardware packet
filtering.

GTSM
    [RFC5082] defines a mechanism that uses the IPv4 TTL or IPv6 Hop
    Limit fields to insure control traffic that can only originate
    from an immediate neighbor is not forged and originating from a
    distant source. GTSM can be applied to many control protocols
    which are routable, for example LDP [RFC6720].

IP Filtering
    At the very minimum, packet filtering plus classification and use
    of multiple queues supporting rate limiting is needed for traffic
    that could potentially be sent to a general purpose CPU used as a

routing engine. The first level of filtering only allows
connections to be initiated from specific IP prefixes to specific
destination ports and then preferably passes traffic directly to
a cryptographic engine and/or rate limits. The second level of
filtering passes connected traffic, such as TCP connections
having received at least one authenticated SYN or having been
locally initiated. The second level of filtering only passes
traffic to specific address and port pairs to be checked for
cryptographic authentication.

The cryptographic authentication is generally the last resort in DoS
attack mitigation. If a packet must be first sent to a general
purpose CPU, then sent to a cryptographic engine, a DoS attack is
possible on high speed interfaces. Only where hardware can identify
a signature and the portion of packet covered by the signature is
cryptographic authentication highly beneficial in protecting against
DoS attacks.

For chips supporting multiple 100 Gb/s interfaces, only a very large
number of parallel cryptographic engines can provide the processing
capacity to handle a large scale DoS or distributed DoS (DDoS)
attack. For many forwarding chips this much processing power
requires significant chip real estate and power, and therefore
reduces system space and power density. For this reason,
cryptographic authentication is not considered a viable first line of
defense.

For some networks the first line of defense is some means of
supporting OOB control and management traffic. In the past this OOB
channel might make use of overhead bits in SONET or OTN or a
dedicated DWDM wavelength. G-ACh and GAL provide an alternative OOB
mechanism which is independent of underlying layers. In other
networks, including most IP/MPLS networks, perimeter filtering serves
a similar purpose, though less effective without extreme vigilance.

A second line of defense is filtering, including GTSM. For protocols
such as EBGP, GTSM and other filtering is often the first line of
defense. Cryptographic authentication is usually the last line of
defense and insufficient by itself to mitigate DoS or DDoS attacks.

2.6.2. MPLS OAM

[RFC4377] defines requirements for MPLS OAM that predate MPLS-TP.
[RFC4379] defines what is commonly referred to as LSP Ping and LSP
Traceroute. [RFC4379] is updated by [RFC6424] supporting MPLS
tunnels and stitched LSP and P2MP LSP. [RFC4379] is updated by
[RFC6425] supporting P2MP LSP. [RFC4379] is updated by [RFC6426] to
support MPLS-TP connectivity verification (CV) and route tracing.

[RFC4950] extends the ICMP format to support TTL expiration that may
occur when using IP traceroute within an MPLS tunnel. The ICMP
message generation can be implemented in forwarding hardware, but if
sent to a general purpose CPU must be rate limited to avoid a
potential denial or service (DoS) attack.

[RFC5880] defines Bidirectional Forwarding Detection (BFD), a
protocol intended to detect faults in the bidirectional path between
two forwarding engines. [RFC5884] and [RFC5885] define BFD for MPLS.
BFD can provide failure detection on any kind of path between
systems, including direct physical links, virtual circuits, tunnels,
MPLS Label Switched Paths (LSPs), multihop routed paths, and
unidirectional links as long as there is some return path.

The processing requirements for BFD are less than for LSP Ping,
making BFD somewhat better suited for relatively high rate proactive
monitoring. BFD does not verify that the data plane matches the
control plane, where LSP Ping does. LSP Ping is somewhat better
suited for on-demand monitoring including relatively low rate
periodic verification of data plane and as a diagnostic tool.

Hardware assistance is often provided for BFD response where BFD
setup or parameter change is not involved and may be necessary for
relatively high rate proactive monitoring. If both BFD and LSP Ping
are recognized in filtering prior to passing traffic to a general
purpose CPU, appropriate DoS protection can be applied (see
Section 2.6.1). Failure to recognize BFD and LSP Ping and at least
rate limit creates the potential for misconfiguration to cause
outages rather than cause errors in the misconfigured OAM.

2.6.3. Pseudowire OAM

Pseudowire OAM makes use of the control channel provided by Virtual
Circuit Connectivity Verification (VCCV) [RFC5085]. VCCV makes use
of the Pseudowire Control Word. BFD support over VCCV is defined by
[RFC5885]. [RFC5885] is updated by [RFC6478] in support of static
pseudowires. [RFC4379] is updated by [RFC6829] supporting LSP Ping
for Pseudowire FEC advertised over IPv6.

G-ACh/GAL (defined in [RFC5586]) is the preferred MPLS-TP OAM control
channel and applies to any MPLS-TP end points, including Pseudowire.
See Section 2.6.4 for an overview of MPLS-TP OAM.

2.6.4. MPLS-TP OAM

[RFC6669] summarizes the MPLS-TP OAM toolset, the set of protocols
supporting the MPLS-TP OAM requirements specified in [RFC5860] and
supported by the MPLS-TP OAM framework defined in [RFC6371].

The MPLS-TP OAM toolset includes:

CC-CV
    [RFC6428] defines BFD extensions to support proactive
    Connectivity Check and Connectivity Verification (CC-CV)
    applications. [RFC6426] provides LSP ping extensions that are
    used to implement on-demand connectivity verification.

RDI
    Remote Defect Indication (RDI) is triggered by failure of
    proactive CC-CV, which is BFD based. For fast RDI initiation,
    RDI SHOULD be initiated and handled by hardware if BFD is handled
    in forwarding hardware. [RFC6428] provides an extension for BFD
    that includes the RDI indication in the BFD format and a
    specification of how this indication is to be used.

Route Tracing
    [RFC6426] specifies that the LSP ping enhancements for MPLS-TP
    on-demand connectivity verification include information on the
    use of LSP ping for route tracing of an MPLS-TP path.

Alarm Reporting
    [RFC6427] describes the details of a new protocol supporting
    Alarm Indication Signal (AIS), Link Down Indication, and fault
    management. Failure to support this functionality in forwarding
    hardware can potentially result in failure to meet protection
    recovery time requirements and is therefore strongly recommended.

Lock Instruct
    Lock instruct is initiated on-demand and therefore need not be
    implemented in forwarding hardware. [RFC6435] defines a lock
    instruct protocol.

Lock Reporting
    [RFC6427] covers lock reporting. Lock reporting need not be
    implemented in forwarding hardware.

Diagnostic
    [RFC6435] defines protocol support for loopback. Loopback
    initiation is on-demand and therefore need not be implemented in
    forwarding hardware. Loopback of packet traffic SHOULD be
    implemented in forwarding hardware on high speed interfaces.

Packet Loss and Delay Measurement
    [RFC6374] and [RFC6375] define a protocol and profile for packet
    loss measurement (LM) and delay measurement (DM). LM requires a
    very accurate capture and insertion of packet and byte counters
    when a packet is transmitted and capture of packet and byte
    counters when a packet is received. This capture and insertion
    MUST be implemented in forwarding hardware for LM OAM if high
    accuracy is needed. DM requires very accurate capture and
    insertion of a timestamp on transmission and capture of timestamp
    when a packet is received. This timestamp capture and insertion
    MUST be implemented in forwarding hardware for DM OAM if high
    accuracy is needed.

See Section 2.6.2 for discussion of hardware support necessary for
BFD and LSP Ping.

CC-CV and alarm reporting is tied to protection and therefore SHOULD
be supported in forwarding hardware in order to provide protection
for a large number of affected LSP within target response intervals.
Since CC-CV is supported by BFD, for MPLS-TP providing hardware
assistance for BFD processing helps insure that protection recovery
time requirements can be met even for faults affecting a large number
of LSP.

MPLS-TP Protection State Coordination (PSC) is defined by [RFC6378]
and updated by [I-D.ietf-mpls-psc-updates], correcting some errors in
[RFC6378].

2.6.5. MPLS OAM and Layer-2 OAM Interworking

[RFC6670] provides the reasons for selecting a single MPLS-TP OAM
solution and examines the consequences were ITU-T to develop a second
OAM solution that is based on Ethernet encodings and mechanisms.

[RFC6310] and [RFC7023] specifies the mapping of defect states
between many types of hardware Attachment Circuits (ACs) and
associated Pseudowires (PWs). This functionality SHOULD be supported
in forwarding hardware.

It is beneficial if an MPLS OAM implementation can interwork with the
underlying server layer and provide a means to interwork with a
client layer. For example, [RFC6427] specifies an inter-layer
propagation of AIS and LDI from MPLS server layer to client MPLS
layers. Where the server layer is a Layer-2, such as Ethernet, PPP
over SONET/SDH, or GFP over OTN, interwork among layers is also
beneficial. For high speed interfaces, supporting this interworking
in forwarding hardware helps insure that protection based on this
interworking can meet recovery time requirements even for faults
affecting a large number of LSP.

2.6.6. Extent of OAM Support by Hardware

Where certain requirements must be met, such as relatively high CC-CV
rates and a large number of interfaces, or strict protection recovery
time requirements and a moderate number of affected LSP, some OAM
functionality must be supported by forwarding hardware. In other
cases, such as highly accurate LM and DM OAM or strict protection
recovery time requirements with a large number of affected LSP, OAM
functionality must be entirely implemented in forwarding hardware.

Where possible, implementation in forwarding hardware should be in
programmable hardware such that if standards are later changed or
extended these changes are likely to be accommodated with hardware
reprogramming rather than replacement.

For some functionality there is a strong case for an implementation
in dedicated forwarding hardware. Examples include packet and byte
counters needed for LM OAM as well as needed for management
protocols. Similarly the capture and insertion of packet and byte
counts or timestamps needed for transmitted LM or DM or time
synchronization packets MUST be implemented in forwarding hardware if
high accuracy is required.

For some functions there is a strong case to provide limited support
in forwarding hardware but may make use of an external general
purpose processor if performance criteria can be met. For example
origination of RDI triggered by CC-CV, response to RDI, and
Protection State Coordination (PSC) functionality may be supported by
hardware, but expansion to a large number of client LSP and
transmission of AIS or RDI to the client LSP may occur in a general
purpose processor. Some forwarding hardware supports one or more on-
chip general purpose processors which may be well suited for such a
role. [I-D.ietf-mpls-psc-updates], being a very recent document that
affects a protection state machine that requires hardware support,
underscores the importance of having a degree of programmability in
forwarding hardware.

The customer (system supplier or provider) should not dictate design, but should independently validate target functionality and performance. However, it is not uncommon for service providers and system implementers to insist on reviewing design details (under NDA) due to past experiences with suppliers and to reject suppliers who are unwilling to provide details.

2.7. Number and Size of Flows

Service provider networks may carry up to hundreds of millions of flows on 10 Gb/s links. Most flows are very short lived, many under a second. A subset of the flows are low capacity and somewhat long lived. When Internet traffic dominates capacity a very small subset of flows are high capacity and/or very long lived.

Two types of limitations with regard to number and size of flows have been observed.

1. Some hardware cannot handle some high capacity flows because of internal paths which are limited, such as per packet backplane paths or paths internal or external to chips such as buffer memory paths. Such designs can handle aggregates of smaller flows. Some hardware with acknowledged limitations has been successfully deployed but may be increasingly problematic if the capacity of large microflows in deployed networks continues to grow.

2. Some hardware approaches cannot handle a large number of flows, or a large number of large flows due to attempting to count per flow, rather than deal with aggregates of flows. Hash techniques scale with regard to number of flows due to a fixed hash size with many flows falling into the same hash bucket. Techniques that identify individual flows have been implemented but have never successfully deployed for Internet traffic.

3. Questions for Suppliers

The following questions should be asked of a supplier. These questions are grouped into broad categories. The questions themselves are intended to be an open ended question to the supplier. The tests in Section 4 are intended to verify whether the supplier disclosed any compliance or performance limitations completely and accurately.

3.1. Basic Compliance

   Q#1   Can the implementation forward packets with an arbitrarily
         large stack depth? What limitations exist, and under what

circumstances do further limitations come into play (such as high
packet rate or specific features enabled or specific types of
packet processing)? See Section 2.1.

Q#2   Is the entire set of basic MPLS functionality described in
      Section 2.1 supported?

Q#3   Are the set of MPLS special purpose labels handled correctly
      and with adequate performance? Are extended special purpose
      labels handled correctly and with adequate performance? See
      Section 2.1.1.

Q#4   Are mappings of label value and TC to PHB handled correctly,
      including RFC3270 L-LSP mappings and RFC4124 CT mappings to PHB?
      See Section 2.1.2.

Q#5   Is time synchronization adequately supported in forwarding
      hardware?

      a. Are both PTP and NTP formats supported?

      b. Is the accuracy of timestamp insertion and incoming stamping
          sufficient?

      See Section 2.1.3.

Q#6   Is link bundling supported?

      a. Can LSP be pinned to specific components?

      b. Is the "all-ones" component link supported?

      See Section 2.1.5.

Q#7   Is MPLS hierarchy supported?

      a. Are both PHP and UHP supported? What limitations exist on
          the number of pop operations with UHP?

      b. Are the pipe, short-pipe, and uniform models supported? Are
          TTL and TC values updated correctly at egress where
          applicable?

      See Section 2.1.6 regarding MPLS hierarchy. See [RFC3443]
      regarding PHP, UHP, and pipe, short-pipe, and uniform models.

Q#8   Are pseudowire sequence numbers handled correctly? See
      Section 2.1.8.1.

Q#9   Is VPN LER functionality handled correctly and without
      performance issues? See Section 2.1.9.

Q#10 Is MPLS multicast (P2MP and MP2MP) handled correctly?

   a. Are packets dropped on uncongested outputs if some outputs
       are congested?

   b. Is performance limited in high fanout situations?

   See Section 2.2.

3.2. Basic Performance

Q#11 Can very small packets be forwarded at full line rate on all
     interfaces indefinitely? What limitations exist, and under what
     circumstances do further limitations come into play (such as
     specific features enabled or specific types of packet
     processing)?

Q#12 Customers must decide whether to relax the prior requirement and
     to what extent. If the answer to the prior question indicates
     that limitations exist, then:

   a. What is the smallest packet size where full line rate
       forwarding can be supported?

   b. What is the longest burst of full rate small packets that can
       be supported?

   Specify circumstances (such as specific features enabled or
   specific types of packet processing) often impact these rates and
   burst sizes.

Q#13 How many pop operations can be supported along with a swap
     operation at full line rate while maintaining per LSP packet and
     byte counts for each pop and swap? This requirement is
     particularly relevant for MPLS-TP.

Q#14 How many label push operations can be supported. While this
     limitation is rarely an issue, it applies to both PHP and UHP,
     unlike the pop limit which applies to UHP.

Q#15 For a worst case where all packets arrive on one LSP, what is
     the counter overflow time? Are any means provided to avoid
     polling all counters at short intervals? This applies to both
     MPLS and MPLS-TP.

3.3. Multipath Capabilities and Performance

   Multipath capabilities and performance do not apply to MPLS-TP but
   apply to MPLS and apply if MPLS-TP is carried in MPLS.

   Q#16 How are large microflows accommodated? Is there active
        management of the hash space mapping to output ports? See
        Section 2.4.2.

   Q#17 How many MPLS labels can be included in a hash based on the MPLS
        label stack?

   Q#18 Is packet rate performance decreased beyond some number of
        labels?

   Q#19 Can the IP header and payload information below the MPLS stack
        be used in the hash? If so, which IP fields, payload types and
        payload fields are supported?

   Q#20 At what maximum MPLS label stack depth can Bottom of Stack and
        an IP header appear without impacting packet rate performance?

   Q#21 Are special purpose labels excluded from the label stack hash?
        Are extended purpose labels excluded from the label stack hash?
        See Section 2.4.5.1.

   Q#22 How is multipath performance affected by high capacity flows or
        an extremely large number of flows, or by very short lived flows?
        See Section 2.7.

3.4. Pseudowire Capabilities and Performance

   Q#23 Is the pseudowire control word supported?

   Q#24 What is the maximum rate of pseudowire encapsulation and
        decapsulation? Apply the same questions as in Base Performance
        for any packet based pseudowire such as IP VPN or Ethernet.

   Q#25 Does inclusion of a pseudowire control word impact performance?

   Q#26 Are flow labels supported?

   Q#27 If so, what fields are hashed on for the flow label for
        different types of pseudowires?

   Q#28 Does inclusion of a flow label impact performance?

3.5. **Entropy Label Support and Performance**

   Q#29 Can an entropy label be added when acting as in ingress LER and
        can it be removed when acting as an egress LER?

   Q#30 If so, what fields are hashed on for the entropy label?

   Q#31 Does adding or removing an entropy label impact packet rate
        performance?

   Q#32 Can an entropy label be detected in the label stack, used in the
        hash, and properly terminate the search for further information
        to hash on?

   Q#33 Does using an entropy label have any negative impact on
        performance? It should have no impact or a positive impact.

3.6. **DoS Protection**

   Q#34 For each control and management plane protocol in use, what
        measures are taken to provide DoS attack hardening?

   Q#35 Have DoS attack tests been performed?

   Q#36 Can compromise of an internal computer on a management subnet be
        leveraged for any form of attack including DoS attack?

3.7. **OAM Capabilities and Performance**

   Q#37 What OAM proactive and on-demand mechanisms are supported?

   Q#38 What performance limits exist under high proactive monitoring
        rates?

   Q#39 Can excessively high proactive monitoring rates impact control
        plane performance or cause control plane instability?

   Q#40 Ask the prior questions for each of the following.

        a. MPLS OAM

        b. Pseudowire OAM

        c. MPLS-TP OAM

        d. Layer-2 OAM Interworking

        See Section 2.6.2.

4. **Forwarding Compliance and Performance Testing**

   Packet rate performance of equipment supporting a large number of 10
   Gb/s or 100 Gb/s links is not possible using desktop computers or
   workstations. The use of high end workstations as a source of test
   traffic was barely viable 20 years ago, but is no longer at all
   viable. Though custom microcode has been used on specialized router
   forwarding cards to serve the purpose of generating test traffic and
   measuring it, for the most part performance testing will require
   specialized test equipment. There are multiple sources of suitable
   equipment.

   The set of tests listed here do not correspond one-to-one to the set
   of questions in Section 3. The same categorization is used and these
   tests largely serve to validate answers provided to the prior
   questions, and can also provide answers where a supplier is unwilling
   to disclose compliance or performance.

   Performance testing is the domain of the IETF Benchmark Methodology
   Working Group (BMWG). Below are brief descriptions of conformance
   and performance tests. Some very basic tests are specified in
   [RFC5695] which partially cover only the basic performance test T#3.

   The following tests should be performed by the systems designer, or
   deployer, or performed by the supplier on their behalf if it is not
   practical for the potential customer to perform the tests directly.
   These tests are grouped into broad categories.

   The tests in Section 4.1 should be repeated under various conditions
   to retest basic performance when critical capabilities are enabled.
   Complete repetition of the performance tests enabling each capability
   and combinations of capabilities would be very time intensive,
   therefore a reduced set of performance tests can be used to gauge the
   impact of enabling specific capabilities.

4.1. **Basic Compliance**

   T#1 Test forwarding at a high rate for packets with varying number
        of label entries. While packets with more than a dozen label
        entries are unlikely to be used in any practical scenario today,
        it is useful to know if limitations exists.

   T#2 For each of the questions listed under "Basic Compliance" in
        Section 3, verify the claimed compliance. For any functionality
        considered critical to a deployment, where applicable performance
        using each capability under load should be verified in addition
        to basic compliance.

4.2. Basic Performance

   T#3 Test packet forwarding at full line rate with small packets.
        See [RFC5695]. The most likely case to fail is the smallest
        packet size. Also test with packet sizes in four byte increments
        ranging from payload sizes or 40 to 128 bytes.

   T#4 If the prior tests did not succeed for all packet sizes, then
        perform the following tests.

        a. Increase the packet size by 4 bytes until a size is found
            that can be forwarded at full rate.

        b. Inject bursts of consecutive small packets into a stream of
            larger packets. Allow some time for recovery between bursts.
            Increase the number of packets in the burst until packets are
            dropped.

   T#5 Send test traffic where a swap operation is required. Also set
        up multiple LSP carried over other LSP where the device under
        test (DUT) is the egress of these LSP. Create test packets such
        that the swap operation is performed after pop operations,
        increasing the number of pop operations until forwarding of small
        packets at full line rate can no longer be supported. Also check
        to see how many pop operations can be supported before the full
        set of counters can no longer be maintained. This requirement is
        particularly relevant for MPLS-TP.

   T#6 Send all traffic on one LSP and see if the counters become
        inaccurate. Often counters on silicon are much smaller than the
        64 bit packet and byte counters in various IETF MIBs. System
        developers should consider what counter polling rate is necessary
        to maintain accurate counters and whether those polling rates are
        practical. Relevant MIBs for MPLS are discussed in [RFC4221] and
        [RFC6639].

4.3. Multipath Capabilities and Performance

   Multipath capabilities do not apply to MPLS-TP but apply to MPLS and
   apply if MPLS-TP is carried in MPLS.

   T#7 Send traffic at a rate well exceeding the capacity of a single
        multipath component link, and where entropy exists only below the
        top of stack. If only the top label is used this test will fail
        immediately.

T#8 Move the labels with entropy down in the stack until either the
     full forwarding rate can no longer be supported or most or all
     packets try to use the same component link.

T#9 Repeat the two tests above with the entropy contained in IP
     headers or IP payload fields below the label stack rather than in
     the label stack. Test with the set of IP headers or IP payload
     fields considered relevant to the deployment or to the target
     market.

T#10 Determine whether traffic that contains a pseudowire control
     word is interpreted as IP traffic. Information in the payload
     MUST NOT be used in the load balancing if the first nibble of the
     packet is not 4 or 6 (IPv4 or IPv6).

T#11 Determine whether special purpose labels and extended special
     purpose labels are excluded from the label stack hash. They MUST
     be excluded.

T#12 Perform testing in the presence of combinations of:

     a. Very large microflows.

     b. Relatively short lived high capacity flows.

     c. Extremely large numbers of flows.

     d. Very short lived small flows.

4.4. Pseudowire Capabilities and Performance

T#13 Ensure that pseudowire can be set up with a pseudowire label and
     pseudowire control word added at ingress and the pseudowire label
     and pseudowire control word removed at egress.

T#14 For pseudowire that contains variable length payload packets,
     repeat performance tests listed under "Basic Performance" for
     pseudowire ingress and egress functions.

T#15 Repeat pseudowire performance tests with and without a
     pseudowire control word.

T#16 Determine whether pseudowire can be set up with a pseudowire
     label, flow label, and pseudowire control word added at ingress
     and the pseudowire label, flow label, and pseudowire control word
     removed at egress.

T#17 Determine which payload fields are used to create the flow label
     and whether the set of fields and algorithm provide sufficient
     entropy for load balancing.

T#18 Repeat pseudowire performance tests with flow labels included.

4.5. Entropy Label Support and Performance

T#19 Determine whether entropy labels can be added at ingress and
     removed at egress.

T#20 Determine which fields are used to create an entropy label.
     Labels further down in the stack, including entropy labels
     further down and IP headers or IP payload fields where applicable
     should be used. Determine whether the set of fields and
     algorithm provide sufficient entropy for load balancing.

T#21 Repeat performance tests under "Basic Performance" when entropy
     labels are used, where ingress or egress is the device under test
     (DUT).

T#22 Determine whether an ELI is detected when acting as a midpoint
     LSR and whether the search for further information on which to
     base the load balancing is used. Information below the entropy
     label SHOULD NOT be used.

T#23 Ensure that the entropy label indicator and entropy label (ELI
     and EL) are removed from the label stack during UHP and PHP
     operations.

T#24 Insure that operations on the TC field when adding and removing
     entropy label are correctly carried out. If TC is changed during
     a swap operation, the ability to transfer that change MUST be
     provided. The ability to suppress the transfer of TC MUST also
     be provided. See "pipe", "short pipe", and "uniform" models in
     [RFC3443].

T#25 Repeat performance tests for a midpoint LSR with entropy labels
     found at various label stack depths.

4.6. DoS Protection

T#26 Actively attack LSR under high protocol churn load and determine
     control plane performance impact or successful DoS under test
     conditions. Specifically test for the following.

a. TCP SYN attack against control plane and management plane
    protocols using TCP, including CLI access (typically SSH
    protected login), NETCONF, etc.

b. High traffic volume attack against control plane and
    management plane protocols not using TCP.

c. Attacks which can be performed from a compromised management
    subnet computer, but not one with authentication keys.

d. Attacks which can be performed from a compromised peer within
    the control plane (internal domain and external domain).
    Assume that per peering keys and per router ID keys rather
    than network wide keys are in use.

See [Section 2.6.1](#).

[4.7](#). OAM Capabilities and Performance

T#27 Determine maximum sustainable rates of BFD traffic. If BFD
     requires CPU intervention, determine both maximum rates and CPU
     loading when multiple interfaces are active.

T#28 Verify LSP Ping and LSP Traceroute capability.

T#29 Determine maximum rates of MPLS-TP CC-CV traffic. If CC-CV
     requires CPU intervention, determine both maximum rates and CPU
     loading when multiple interfaces are active.

T#30 Determine MPLS-TP DM precision.

T#31 Determine MPLS-TP LM accuracy.

T#32 Verify MPLS-TP AIS/RDI and Protection State Coordination (PSC)
     functionality, protection speed, and AIS/RDI notification speed
     when a large number of Management Entities (ME) must be notified
     with AIS/RDI.

[5](#). Acknowledgements

Numerous very useful comments have been received in private email.
Some of these contributions are acknowledged here, approximately in
chronologic order.

Paul Doolan provided a brief review resulting in a number of
clarifications, most notably regarding on-chip vs. system buffering,
100 Gb/s link speed assumptions in the 150 Mpps figure, and handling
of large microflows. Pablo Frank reminded us of the sawtooth effect

in PPS vs. packet size graphs, prompting the addition of a few
paragraphs on this. Comments from Lou Berger at IETF-85 prompted the
addition of Section 2.7.

Valuable comments were received on the BMWG mailing list. Jay
Karthik pointed out testing methodology hints that after discussion
were deemed out of scope and were removed but may benefit later work
in BMWG.

Nabil Bitar pointed out the need to cover QoS (Differentiated
Services), MPLS multicast (P2MP and MP2MP), and MPLS-TP OAM. Nabil
also provided a number of clarifications to the questions and tests
in Section 3 and Section 4.

Mark Szczesniak provided a thorough review and a number of useful
comments and suggestions that improved the document.

Gregory Mirsky and Thomas Beckhaus provided useful comments during
the MPLS RT review.

Tal Mizrahi provided comments that prompted clarifications regarding
timestamp processing, local delivery of packets, and the need for
hardware assistance in processing OAM traffic.

Alexander (Sasha) Vainshtein pointed out errors in Section 2.1.8.1
and suggested new text which after lengthy discussion resulted in
restating the summarization of requirements from PWE3 RFCs and more
clearly stating the benefits and drawbacks of packet resequencing
based on PW sequence number.

Loa Anderson provided useful comments and corrections prior to WGLC.
Adrian Farrel provided useful comments and corrections prior as part
of the AD review.

Discussion with Steve Kent during SecDir review resulted in expansion
of Section 7, briefly summarizing security considerations related to
forwarding in normative references. Tom Petch pointed out some
editorial errors in private email. Al Morton during OpsDir review
prompted clarification in the target audience section, suggested more
clear wording in places, and found numerous editorial errors.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

This document reviews forwarding behavior specified elsewhere and
points out compliance and performance requirements. As such it
introduces no new security requirements or concerns.

Discussion of hardware support and other equipment hardening against
DoS attack can be found in Section 2.6.1. Section 3.6 provides a
list of question regarding DoS to be asked of suppliers. Section 4.6
suggests types of testing that can provide some assurance of the
effectiveness of supplier DoS hardening claims.

Knowledge of potential performance shortcomings may serve to help new
implementations avoid pitfalls. It is unlikely that such knowledge
could be the basis of new denial of service as these pitfalls are
already widely known in the service provider community and among
leading equipment suppliers. In practice extreme data and packet
rate are needed to affect existing equipment and to affect networks
that may be still vulnerable due to failure to implement adequate
protection. The extreme data and packet rates make this type of
denial of service unlikely and make undetectable denial of service of
this type impossible.

The set of normative references each contain security considerations.
A brief summarization of MPLS security considerations applicable to
forwarding follows:

1.   MPLS encapsulation does not support an authentication extension.
     This is reflected in the security section of [RFC3032].
     Documents which clarify MPLS header fields such as TTL
     [RFC3443], the explicit null label [RFC4182], renaming EXP to TC
     [RFC5462], ECN for MPLS [RFC5129], and MPLS Ethernet
     encapsulation [RFC5332] make no changes to security
     considerations in [RFC3032].

2.   Some cited RFCs are related to Diffserv forwarding. [RFC3270]
     refers to MPLS and Diffserv security. [RFC2474] mentions theft
     of service and denial of service due to mismarking. [RFC2474]
     mentions IPsec interaction, but with MPLS, not being carried by
     IP, this type of interaction in [RFC2474] is not relevant.

3.   [RFC3209] is cited here due only to make-before-break forwarding
     requirements. This is related to resource sharing and the theft
     of service and denial of service concerns in [RFC2474] apply.

4.   [RFC4090] defines FRR which provides protection but does not add
     security concerns. RFC4201 defines link bundling but raises no
     additional security concerns.

5.   Various OAM control channels are defined in [RFC4385] (PW CW),
     [RFC5085] (VCCV), [RFC5586] (G-Ach and GAL). These documents
     describe potential abuse of these OAM control channels.

6.   [RFC4950] defines ICMP extensions when MPLS TTL expires and
     payload is IP. This provides MPLS header information which is
     of no use to an IP attacker, but sending this information can be
     suppressed through configuration.

7.   GTSM [RFC5082] provides a means to improve protection against
     high traffic volume spoofing as a form of DoS attack.

8.   BFD [RFC5880] [RFC5884] [RFC5885] provides a form of OAM used in
     MPLS and MPLS-TP. The security considerations related to the
     OAM control channel are relevant. The BFD payload supports
     authentication unlike the MPLS encapsulation or MPLS or PW
     control channel encapsulation is carried in. Where an IP return
     OAM path is used IPsec is suggested as a means of securing the
     return path.

9.   Other forms of OAM are supported by [RFC6374] [RFC6375] (Loss
     and Delay Measurement), [RFC6428] (Connectivity Check/
     Verification based on BFD), and [RFC6427] (Fault Management).
     The security considerations related to the OAM control channel
     are relevant. IP return paths, where used, can be secured with
     IPsec.

10. Linear protection is defined by [RFC6378] and updated by
     [I-D.ietf-mpls-psc-updates]. Security concerns related to MPLS
     encapsulation and OAM control channels apply. Security concerns
     reiterate [RFC5920] as applied to protection switching.

11. The PW Flow Label [RFC6391] and MPLS Entropy Label [RFC6790]
     affect multipath load balancing. Security concerns reiterate
     [RFC5920]. Security impacts would be limited to load
     distribution.

MPLS security including data plane security is discussed in greater
detail in [RFC5920] (MPLS/GMPLS Security Framework). The MPLS-TP
security framework [RFC6941] build upon this, focusing largely on the
MPLS-TP OAM additions and OAM channels with some attention given to
using network management in place of control plane setup. In both
security framework documents MPLS is assumed to run within a "trusted
zone", defined as being where a single service provider (SP) has
total operational control over that part of the network.

If control plane security and management plane security are
sufficiently robust, compromise of a single network element may

result in chaos in the data plane anywhere in the network through
denial of service attacks, but not a Byzantine security failure in
which other network elements are fully compromised.

MPLS security, or lack of, can affect whether traffic can be
misrouted and lost, or intercepted, or intercepted and reinserted (a
man-in-the-middle attack) or spoofed. End user applications,
including control plane and management plane protocols used by the
SP, are expected to make use of appropriate end-to-end authentication
and where appropriate end-to-end encryption.

8. References

8.1. Normative References

[I-D.ietf-mpls-psc-updates]
          Osborne, E., "Updates to PSC", draft-ietf-mpls-psc-
          updates-01 (work in progress), January 2014.

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

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y.,
          Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack
          Encoding", RFC 3032, January 2001.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
          and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
          Tunnels", RFC 3209, December 2001.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen,
          P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-
          Protocol Label Switching (MPLS) Support of Differentiated
          Services", RFC 3270, May 2002.

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

[RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute
          Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May
          2005.

[RFC4182] Rosen, E., "Removing a Restriction on the use of MPLS
          Explicit NULL", RFC 4182, September 2005.

[RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling
          in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.

   [RFC4385]   Bryant, S., Swallow, G., Martini, L., and D. McPherson,
               "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for
               Use over an MPLS PSN", RFC 4385, February 2006.

   [RFC4950]   Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "ICMP
               Extensions for Multiprotocol Label Switching", RFC 4950,
               August 2007.

   [RFC5082]   Gill, V., Heasley, J., Meyer, D., Savola, P., and C.
               Pignataro, "The Generalized TTL Security Mechanism
               (GTSM)", RFC 5082, October 2007.

   [RFC5085]   Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit
               Connectivity Verification (VCCV): A Control Channel for
               Pseudowires", RFC 5085, December 2007.

   [RFC5129]   Davie, B., Briscoe, B., and J. Tay, "Explicit Congestion
               Marking in MPLS", RFC 5129, January 2008.

   [RFC5332]   Eckert, T., Rosen, E., Aggarwal, R., and Y. Rekhter, "MPLS
               Multicast Encapsulations", RFC 5332, August 2008.

   [RFC5586]   Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic
               Associated Channel", RFC 5586, June 2009.

   [RFC5880]   Katz, D. and D. Ward, "Bidirectional Forwarding Detection
               (BFD)", RFC 5880, June 2010.

   [RFC5884]   Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow,
               "Bidirectional Forwarding Detection (BFD) for MPLS Label
               Switched Paths (LSPs)", RFC 5884, June 2010.

   [RFC5885]   Nadeau, T. and C. Pignataro, "Bidirectional Forwarding
               Detection (BFD) for the Pseudowire Virtual Circuit
               Connectivity Verification (VCCV)", RFC 5885, June 2010.

   [RFC6374]   Frost, D. and S. Bryant, "Packet Loss and Delay
               Measurement for MPLS Networks", RFC 6374, September 2011.

   [RFC6375]   Frost, D. and S. Bryant, "A Packet Loss and Delay
               Measurement Profile for MPLS-Based Transport Networks",
               RFC 6375, September 2011.

   [RFC6378]   Weingarten, Y., Bryant, S., Osborne, E., Sprecher, N., and
               A. Fulignoli, "MPLS Transport Profile (MPLS-TP) Linear
               Protection", RFC 6378, October 2011.

   [RFC6391]   Bryant, S., Filsfils, C., Drafz, U., Kompella, V., Regan,
               J., and S. Amante, "Flow-Aware Transport of Pseudowires
               over an MPLS Packet Switched Network", RFC 6391, November
               2011.

   [RFC6427]   Swallow, G., Fulignoli, A., Vigoureux, M., Boutros, S.,
               and D. Ward, "MPLS Fault Management Operations,
               Administration, and Maintenance (OAM)", RFC 6427, November
               2011.

   [RFC6428]   Allan, D., Swallow Ed. , G., and J. Drake Ed. , "Proactive
               Connectivity Verification, Continuity Check, and Remote
               Defect Indication for the MPLS Transport Profile", RFC
               6428, November 2011.

   [RFC6790]   Kompella, K., Drake, J., Amante, S., Henderickx, W., and
               L. Yong, "The Use of Entropy Labels in MPLS Forwarding",
               RFC 6790, November 2012.

8.2. Informative References

   [ACK-compression]
               , , , "Observations and Dynamics of a Congestion Control
               Algorithm: The Effects of Two-Way Traffic", Proc. ACM
               SIGCOMM, ACM Computer Communications Review (CCR) Vol 21,
               No 4, 1991, pp.133-147., 1991.

   [I-D.ietf-mpls-in-udp]
               Building, K., Sheth, N., Yong, L., Pignataro, C., and F.
               Yongbing, "Encapsulating MPLS in UDP", draft-ietf-mpls-in-
               udp-05 (work in progress), January 2014.

   [I-D.ietf-mpls-special-purpose-labels]
               Kompella, K., Andersson, L., and A. Farrel, "Allocating
               and Retiring Special Purpose MPLS Labels", draft-ietf-
               mpls-special-purpose-labels-03 (work in progress), July
               2013.

   [I-D.ietf-tictoc-1588overmpls]
               Davari, S., Oren, A., Bhatia, M., Roberts, P., and L.
               Montini, "Transporting Timing messages over MPLS
               Networks", draft-ietf-tictoc-1588overmpls-05 (work in
               progress), June 2013.

   [RFC0791]   Postel, J., "Internet Protocol", STD 5, RFC 791, September
               1981.

   [RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black,
              "Definition of the Differentiated Services Field (DS
              Field) in the IPv4 and IPv6 Headers", RFC 2474, December
              1998.

   [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z.,
              and W. Weiss, "An Architecture for Differentiated
              Services", RFC 2475, December 1998.

   [RFC2597] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski,
              "Assured Forwarding PHB Group", RFC 2597, June 1999.

   [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
              Label Switching Architecture", RFC 3031, January 2001.

   [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition
              of Explicit Congestion Notification (ECN) to IP", RFC
              3168, September 2001.

   [RFC3429] Ohta, H., "Assignment of the 'OAM Alert Label' for
              Multiprotocol Label Switching Architecture (MPLS)
              Operation and Maintenance (OAM) Functions", RFC 3429,
              November 2002.

   [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching
              (GMPLS) Signaling Functional Description", RFC 3471,
              January 2003.

   [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.
              Jacobson, "RTP: A Transport Protocol for Real-Time
              Applications", STD 64, RFC 3550, July 2003.

   [RFC3828] Larzon, L-A., Degermark, M., Pink, S., Jonsson, L-E., and
              G. Fairhurst, "The Lightweight User Datagram Protocol
              (UDP-Lite)", RFC 3828, July 2004.

   [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-
              Edge (PWE3) Architecture", RFC 3985, March 2005.

   [RFC4023] Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating
              MPLS in IP or Generic Routing Encapsulation (GRE)", RFC
              4023, March 2005.

   [RFC4110] Callon, R. and M. Suzuki, "A Framework for Layer 3
              Provider-Provisioned Virtual Private Networks (PPVPNs)",
              RFC 4110, July 2005.

   [RFC4124] Le Faucheur, F., "Protocol Extensions for Support of
              Diffserv-aware MPLS Traffic Engineering", RFC 4124, June
              2005.

   [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP)
              Hierarchy with Generalized Multi-Protocol Label Switching
              (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

   [RFC4221] Nadeau, T., Srinivasan, C., and A. Farrel, "Multiprotocol
              Label Switching (MPLS) Management Overview", RFC 4221,
              November 2005.

   [RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram
              Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

   [RFC4377] Nadeau, T., Morrow, M., Swallow, G., Allan, D., and S.
              Matsushima, "Operations and Management (OAM) Requirements
              for Multi-Protocol Label Switched (MPLS) Networks", RFC
              4377, February 2006.

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

   [RFC4664] Andersson, L. and E. Rosen, "Framework for Layer 2 Virtual
              Private Networks (L2VPNs)", RFC 4664, September 2006.

   [RFC4817] Townsley, M., Pignataro, C., Wainner, S., Seely, T., and
              J. Young, "Encapsulation of MPLS over Layer 2 Tunneling
              Protocol Version 3", RFC 4817, March 2007.

   [RFC4875] Aggarwal, R., Papadimitriou, D., and S. Yasukawa,
              "Extensions to Resource Reservation Protocol - Traffic
              Engineering (RSVP-TE) for Point-to-Multipoint TE Label
              Switched Paths (LSPs)", RFC 4875, May 2007.

   [RFC4928] Swallow, G., Bryant, S., and L. Andersson, "Avoiding Equal
              Cost Multipath Treatment in MPLS Networks", BCP 128, RFC
              4928, June 2007.

   [RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC
              4960, September 2007.

   [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP
              Specification", RFC 5036, October 2007.

   [RFC5317] Bryant, S. and L. Andersson, "Joint Working Team (JWT)
              Report on MPLS Architectural Considerations for a
              Transport Profile", RFC 5317, February 2009.

   [RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching
              (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic
              Class" Field", RFC 5462, February 2009.

   [RFC5640] Filsfils, C., Mohapatra, P., and C. Pignataro, "Load-
              Balancing for Mesh Softwires", RFC 5640, August 2009.

   [RFC5695] Akhter, A., Asati, R., and C. Pignataro, "MPLS Forwarding
              Benchmarking Methodology for IP Flows", RFC 5695, November
              2009.

   [RFC5860] Vigoureux, M., Ward, D., and M. Betts, "Requirements for
              Operations, Administration, and Maintenance (OAM) in MPLS
              Transport Networks", RFC 5860, May 2010.

   [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network
              Time Protocol Version 4: Protocol and Algorithms
              Specification", RFC 5905, June 2010.

   [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS
              Networks", RFC 5920, July 2010.

   [RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu,
              D., and S. Mansfield, "Guidelines for the Use of the "OAM"
              Acronym in the IETF", BCP 161, RFC 6291, June 2011.

   [RFC6310] Aissaoui, M., Busschbach, P., Martini, L., Morrow, M.,
              Nadeau, T., and Y(J). Stein, "Pseudowire (PW) Operations,
              Administration, and Maintenance (OAM) Message Mapping",
              RFC 6310, July 2011.

   [RFC6371] Busi, I. and D. Allan, "Operations, Administration, and
              Maintenance Framework for MPLS-Based Transport Networks",
              RFC 6371, September 2011.

   [RFC6388] Wijnands, IJ., Minei, I., Kompella, K., and B. Thomas,
              "Label Distribution Protocol Extensions for Point-to-
              Multipoint and Multipoint-to-Multipoint Label Switched
              Paths", RFC 6388, November 2011.

   [RFC6424] Bahadur, N., Kompella, K., and G. Swallow, "Mechanism for
              Performing Label Switched Path Ping (LSP Ping) over MPLS
              Tunnels", RFC 6424, November 2011.

   [RFC6425] Saxena, S., Swallow, G., Ali, Z., Farrel, A., Yasukawa,
              S., and T. Nadeau, "Detecting Data-Plane Failures in
              Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC
              6425, November 2011.

   [RFC6426] Gray, E., Bahadur, N., Boutros, S., and R. Aggarwal, "MPLS
              On-Demand Connectivity Verification and Route Tracing",
              RFC 6426, November 2011.

   [RFC6435] Boutros, S., Sivabalan, S., Aggarwal, R., Vigoureux, M.,
              and X. Dai, "MPLS Transport Profile Lock Instruct and
              Loopback Functions", RFC 6435, November 2011.

   [RFC6438] Carpenter, B. and S. Amante, "Using the IPv6 Flow Label
              for Equal Cost Multipath Routing and Link Aggregation in
              Tunnels", RFC 6438, November 2011.

   [RFC6478] Martini, L., Swallow, G., Heron, G., and M. Bocci,
              "Pseudowire Status for Static Pseudowires", RFC 6478, May
              2012.

   [RFC6639] King, D. and M. Venkatesan, "Multiprotocol Label Switching
              Transport Profile (MPLS-TP) MIB-Based Management
              Overview", RFC 6639, June 2012.

   [RFC6669] Sprecher, N. and L. Fang, "An Overview of the Operations,
              Administration, and Maintenance (OAM) Toolset for MPLS-
              Based Transport Networks", RFC 6669, July 2012.

   [RFC6670] Sprecher, N. and KY. Hong, "The Reasons for Selecting a
              Single Solution for MPLS Transport Profile (MPLS-TP)
              Operations, Administration, and Maintenance (OAM)", RFC
              6670, July 2012.

   [RFC6720] Pignataro, C. and R. Asati, "The Generalized TTL Security
              Mechanism (GTSM) for the Label Distribution Protocol
              (LDP)", RFC 6720, August 2012.

   [RFC6829] Chen, M., Pan, P., Pignataro, C., and R. Asati, "Label
              Switched Path (LSP) Ping for Pseudowire Forwarding
              Equivalence Classes (FECs) Advertised over IPv6", RFC
              6829, January 2013.

   [RFC6941] Fang, L., Niven-Jenkins, B., Mansfield, S., and R.
              Graveman, "MPLS Transport Profile (MPLS-TP) Security
              Framework", RFC 6941, April 2013.

   [RFC7023] Mohan, D., Bitar, N., Sajassi, A., DeLord, S., Niger, P.,
              and R. Qiu, "MPLS and Ethernet Operations, Administration,
              and Maintenance (OAM) Interworking", RFC 7023, October
              2013.

   [RFC7074] Berger, L. and J. Meuric, "Revised Definition of the GMPLS
              Switching Capability and Type Fields", RFC 7074, November
              2013.

   [RFC7079] Del Regno, N. and A. Malis, "The Pseudowire (PW) and
              Virtual Circuit Connectivity Verification (VCCV)
              Implementation Survey Results", RFC 7079, November 2013.

Appendix A. Organization of References Section

   The References section is split into Normative and Informative
   subsections. References that directly specify forwarding
   encapsulations or behaviors are listed as normative. References
   which describe signaling only, though normative with respect to
   signaling, are listed as informative. They are informative with
   respect to MPLS forwarding.

Authors' Addresses

   Curtis Villamizar (editor)
   Outer Cape Cod Network Consulting, LLC

   Email: curtis@occnc.com

   Kireeti Kompella
   Juniper Networks

   Email: kireeti@juniper.net

   Shane Amante
   Apple Inc.
   1 Infinite Loop
   Cupertino, California 95014

   Email: samante@apple.com

   Andrew Malis
   Huawei Technologies

   Email: agmalis@gmail.com

   Carlos Pignataro
   Cisco Systems
   7200-12 Kit Creek Road
   Research Triangle Park, NC 27709
   US

   Email: cpignata@cisco.com