          MPLS Forwarding Benchmarking Methodology for IP Flows

Abstract

   This document describes a methodology specific to the benchmarking
   of Multiprotocol Label Switching (MPLS) forwarding devices, limited
   to the most common MPLS packet forwarding scenarios and delay
   measurements for each, considering IP flows. It builds upon the
   tenets set forth in RFC 2544, RFC 1242, and other IETF Benchmarking
   Methodology Working Group (BMWG) efforts. This document seeks to
   extend these efforts to the MPLS paradigm.

Status of This Memo

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

Copyright Notice

not be created outside the IETF Standards Process, except to format
it for publication as an RFC or to translate it into languages other
than English.

Table of Contents

1. Introduction

   Over the past several years, there has been an increase in the use of
   MPLS as a forwarding architecture in new and existing network
   designs. MPLS, defined in [RFC3031], is a foundation technology and
   the basis for many advanced technologies such as Layer 3 MPLS VPNs,
   Layer 2 MPLS VPNs, and MPLS Traffic Engineering.

However, there is no standard method defined to compare and contrast
the foundational MPLS packet forwarding capabilities of network
devices. This document proposes a methodology using common criteria
(such as throughput, latency, frame-loss rate, system recovery,
reset, etc.) to evaluate MPLS forwarding of any implementation.

2. Document Scope

The benchmarking methodology principles outlined in RFC 2544
[RFC2544] are independent of forwarding techniques; however, they
don't fully address MPLS benchmarking. The workload on network
forwarding device resources that MPLS forwarding places is different
from that of IP forwarding; therefore, MPLS forwarding benchmarking
specifics are desired.

The purpose of this document is to describe a methodology specific to
the benchmarking of MPLS forwarding devices. The methods described
are limited in scope to the most common MPLS packet forwarding
scenarios and corresponding performance measurements in a laboratory
setting. It builds upon the tenets set forth in RFC 2544 [RFC2544],
RFC 1242 [RFC1242], and other IETF Benchmarking Methodology Working
Group (BMWG) efforts. In other words, this document is not a
replacement for, but a complement to, RFC 2544.

This document focuses on the MPLS label stack [RFC3032] that has only
one entry, as it is the fundamental of MPLS forwarding. It is
expected that future documents may cover the benchmarking of MPLS
applications such as Layer 3 VPN (L3VPN) [RFC4364], Layer 2 VPN
(L2VPN) [RFC4664], Fast ReRoute [RFC4090], etc., which require more
than one entry in the MPLS label stack.

Moreover, to address the majority of current deployments' needs, this
document focuses on having IP packets as the MPLS payload. In other
words, label distribution for IP Forwarding Equivalence Class (FEC)
[RFC3031] is prescribed (see Section 4.1.3) by this document. It is
expected that future documents may focus on having non-IP packets as
the MPLS payload.

Note that the presence of an MPLS label stack does not require the
length of MPLS payload (which is an IP packet, per this document) to
be changed; hence, the effective maximum size of a frame can increase
by Z octets (where Z = 4 x number of label stack entries), as
observed in current deployments. This document focuses on
benchmarking such a scenario.

3. Key Words To Reflect Requirements

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in BCP 14, RFC 2119
   [RFC2119]. RFC 2119 defines the use of these key words to help make
   the intent of Standards Track documents as clear as possible. While
   this document uses these keywords, this document is not a Standards
   Track document.

4. Test Methodology

   The set of methodologies described in this document will use the
   topology described in this section. An effort has been made to
   exclude superfluous equipment needs such that each test can be
   carried out with a minimal number of devices. Figure 1 illustrates
   the sample topology in which the Device Under Test (DUT) is connected
   to the test ports on the test tool in accord with Figure 1 of RFC
   2544.

```
                        +-----------------+
        +---------+     |                 |     +---------+
        | Test    |     |                 |     | Test    |
        | Port A1 +-----+ DA1         DB1 +-----+ Port B1 |
        +---------+     |                 |     +---------+
        +---------+     |      DUT        |     +---------+
        | Test    |     |                 |     | Test    |
        | Port A2 +-----+ DA2         DB2 +-----+ Port B2 |
        +---------+     |                 |     +---------+
           ...          | ...         ... |        ...
        +---------+     |                 |     +---------+
        | Test    |     |                 |     | Test    |
        | Port Ap +-----+ DAp         DBp +-----+ Port Bp |
        +---------+     +-----------------+     +---------+
```

            Figure 1: Topology for MPLS Forwarding Benchmarking
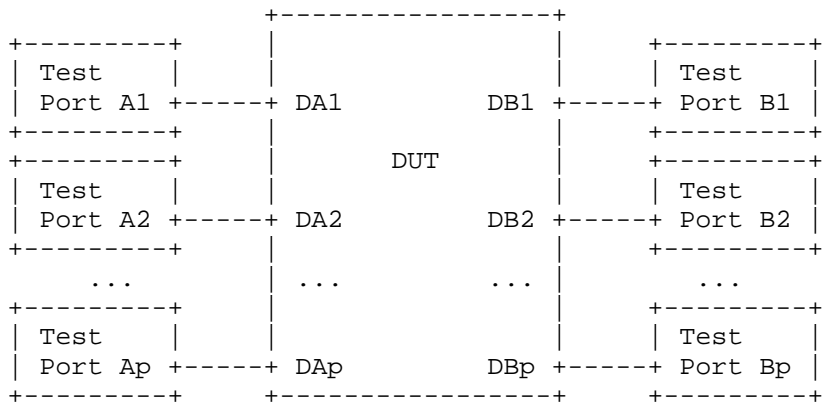
   A represents a Tx-side Module of the test tool, whereas B represents
   an Rx-side Module of the same test tool. Of course, the suffixed
   numbers (1, 2, ..., p) represent ports on a Module.

   Similarly, DA represents an Rx-side Module of the DUT, whereas DB
   represents a Tx-side Module. The suffixed numbers (1, 2, ..., p)
   represent ports on a Module.

   p = the number of {DA, DB} pair ports on the DUT. It is determined
   by the maximum unidirectional forwarding throughput of the DUT and
   the load capacity of the port media (e.g., interface) connecting the
   DUT to the test tool.

   For example, if the DUT's maximum forwarding throughput is 100 frames
   per second (fps) and the load capacity of the port media (e.g.,
   interface) is 50 fps, then p >= 2 is needed to sufficiently test the
   maximum frame forwarding.

   The exact throughput is a measured quantity obtained through testing.
   Throughput may vary depending on the number of ports used and other
   factors. The number of ports (p) used SHOULD be reported. Please
   see "Test Setup" in Section 6. Following Figure 1 from Section 6 of
   RFC 2544 is recommended.

4.1. Test Considerations

   This methodology assumes a full-duplex, uniform medium topology. The
   medium used MUST be reported in each test result. Issues regarding
   mixed transmission media, speed mismatches, media header differences,
   etc., are not under consideration. Traffic affecting features such
   as Flow control, Quality of Service (QoS), Graceful Restart, etc.
   MUST be disabled, unless explicitly requested in the test case.
   Additionally, any non-essential traffic MUST also be avoided.

4.1.1. Abbreviations Used

   The terms used in this document remain consistent with those defined
   in "Benchmarking Terminology for Network Interconnect Devices" RFC
   1242 [RFC1242]. This terminology SHOULD be consulted before using or
   applying the recommendations of this document.

   Please refer to Figure 1 for a topology view of the network. The
   following abbreviations are used in this document:

   M := Module on a device (i.e., Line-Card or Slot; could be A or B)

   p := Port number (i.e., port on the Module; could be 1, 2, etc.)

   RN := Remote Network (i.e., network that is reachable via a port of a
   module; could be B1RN1 or B2RN5 to mean the first network reachable
   via port 1 of module B, e.g., B1, or the fifth network reachable via
   port 2 of module B, etc.). RN is considered to be the IP Prefix FEC
   from the MPLS perspective.

4.1.2. IGP Support

   It is RECOMMENDED that all of the ports (A1, DA1, DB1, and A2) on the
   DUT and test tool support a dynamic Interior Gateway Protocol (IGP)
   for routing such as IS-IS, OSPF, RIP, etc. Furthermore, there are
   testing considerations in this document that the device be able to
   provide a stable control plane during heavy forwarding workloads. In
   particular, the procedures defined in Section 11.3 of RFC 2544 must
   be followed. This is to ensure that control plane instability during
   load conditions is not the contributing factor towards frame
   forwarding performance.

   The route distribution method (OSPF, IS-IS, Enhanced Interior Gateway
   Routing Protocol (EIGRP), RIP, Static, etc.), if used, MUST be
   reported. Furthermore, if any specific configuration is used to
   maintain control plane stability during the test (i.e., Control Plane
   Protection, Control Plane Rate Limiting, etc.), then it MUST also be
   reported.

4.1.3. Label Distribution Support

   The DUT and test tool must support at least one protocol for
   exchanging MPLS label/FEC bindings for Prefix Forwarding Equivalence
   Class (FEC) [RFC3031]. The DUT and test tool MUST be capable of
   learning and advertising MPLS label/FEC bindings via the chosen
   protocol(s) and use them during packet forwarding all the time
   (including when the label/FEC bindings change). The most commonly
   used protocols are Label Distribution Protocol (LDP) [RFC5036],
   Resource Reservation Protocol-Traffic Engineering (RSVP-TE)
   [RFC3209], and Border Gateway Protocol (BGP) [RFC3107].

   All of the ports (A1, DA1, DB1, B1, etc.) either on the DUT or the
   test tool used in the testing SHOULD support LDP, RSVP-TE, and BGP
   for IPv4 or IPv6 Prefix Forwarding Equivalence Classes (FECs).

   Static labels SHOULD NOT be used to establish the MPLS label switched
   paths (LSPs), unless specified explicitly by the test case.

   This is because the use of a static label is quite uncommon in the
   production networks.

   The IPv4 and IPv6 Explicit NULL labels (label values 0 and 2) are
   sometimes used to identify the payload of an MPLS packet on an LSP
   [RFC3032]. Explicit NULL labels are not used in the tests described
   in this document because the tests are limited to the use of no more
   than one non-reserved MPLS label in the label stack of all packets
   to, from, or through the DUT.

4.1.4. Frame Formats

   This section explains the frame formats for IP and MPLS packets
   (Section 4.1.4.1), the usage of IP as the mandatory Layer 3 protocol
   and as the MPLS packet payload (Section 4.1.4.2), change in frame
   format during forwarding (Section 4.1.4.3), and recommended frame
   formats for the MPLS benchmarking (Section 4.1.4.4).

4.1.4.1. Frame Format for IP versus MPLS

   A test frame carrying an IP packet is illustrated in Figure 2 below.
   Note that RFC 2544 [RFC2544] prescribes using such a frame as the
   test frame over the chosen Layer 2 media.

```
        +---------+--------------+-----------------------+
        | Layer 2 | Layer 3 = IP | Layer 4 = UDP         |
        +---------+--------------+-----------------------+
```
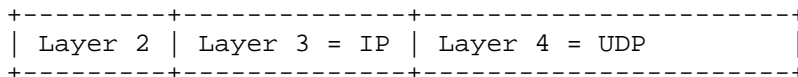
                Figure 2: Frame Format for IP Packets

   Unlike a test frame carrying an IP packet, a test frame carrying an
   MPLS packet contains an "MPLS label stack" [RFC3032] immediately
   after the Layer 2 header (and before the IP header, if any) as
   illustrated in Figure 3 below.

```
        +---------+-------+--------------+-----------------------+
        | Layer 2 | MPLS  | Layer 3 = IP | Layer 4 = UDP         |
        +---------+-------+--------------+-----------------------+
```
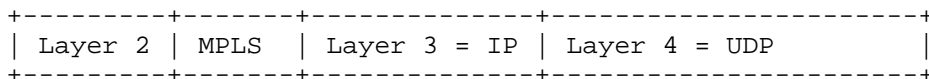
                Figure 3: Frame Format for MPLS Packets

   The MPLS label stack is represented as a sequence of "label stack
   entries", where each label stack entry is 4 octets, as illustrated in
   Figure 1 of [RFC3032]. This document requires exactly one entry in
   the MPLS label stack in an MPLS packet.

   MPLS label values used in any test case MUST be outside the reserved
   label value (0-15) unless stated otherwise.

4.1.4.2. MPLS Packet Payload

   This document prescribes using an IP packet as the MPLS payload (as
   illustrated in Figure 3 above). Generically speaking, this document
   mandates the test frame to include IP (either IPv4 or IPv6) as the
   Layer 3 protocol, in accord with Section 8 of [RFC2544] and
   independent of the MPLS label stack presence, for three reasons:

1. This enables using IP Prefix Forwarding Equivalence Class (FEC)
   [RFC3031], which is a must for every MPLS network.

2. This provides the foundation or baseline for the benchmarking of
   various other MPLS applications such as L3VPN, L2VPN, TE-FRR, etc.

3. This enables leveraging RFC 2544 [RFC2544], which prescribes using
   IP packets with UDP data as the test frames. (Note that [RFC5180]
   also uses this prescription for IPv6 benchmarking).

While the usage of non-IP payloads is possible, it requires an MPLS
application, e.g., L2VPN, whose benchmarking may be covered in
separate BMWG documents (MPLS L2VPN Benchmarking, for example) in the
future. This is also explained in Section 2.

4.1.4.3. Change in Frame Format Due to MPLS Push and Pop

A frame carrying an IP or MPLS packet may go through any of the three
MPLS forwarding operations: label push (or LSP Ingress), label swap,
and label pop (or LSP Egress), as defined in [RFC3031]. It is
important to understand the change of the frame format from IP to
MPLS or vice versa depending on the forwarding operation.

In a label push (or LSP Ingress) operation, the DUT receives a frame
containing an IP packet and forwards a frame containing an MPLS
packet if the corresponding forwarding lookup for the IP destination
points to a label push operation.

In a label swap operation, the DUT receives a frame containing an
MPLS packet and forwards a frame containing an MPLS packet if the
corresponding forwarding lookup for the label value points to a label
swap operation.

In a label pop (or LSP Egress) operation, the DUT receives a frame
containing an MPLS packet and forwards a frame containing an IP
packet if the corresponding forwarding lookup for the label value
points to a label pop operation.

4.1.4.4. Frame Formats to Be Used for Benchmarking

This document prescribes using two test frame formats to
appropriately test the forwarding operations: (1) Frame format for IP
and (2) Frame format for MPLS. Both formats are explained in Section
4.1.4.1. Additionally, the format of the test frame may change
depending on the forwarding operation.

1. For test cases involving the label push operation, the test tool
   must use the frame format for IP packets (Figure 2) to send the
   test frames to the DUT, and must use the frame format for MPLS
   packets (Figure 3) to receive the test frames from the DUT.

2. For test cases involving the label swap operation, the test tool
   must use the frame format for MPLS packets (Figure 3) to send the
   test frames to the DUT, and must use the frame format for MPLS
   packets (Figure 3) to receive the test frames from the DUT.

3. For test cases involving the label pop operation, the test tool
   must use the frame format for MPLS packets (Figure 3) to send the
   test frames to the DUT, and must use the frame format for IP
   packets (Figure 2) to receive the test frames from the DUT.

4.1.5. Frame Sizes

   Two types of port media are commonly deployed: Ethernet and POS
   (Packet Over Synchronous Optical Network). This section identifies
   the frame sizes that SHOULD be used for each media type, if supported
   by the DUT; Section 4.1.5.1 covers Ethernet and Section 4.1.5.2
   covers POS.

   First, it is important to note the possible increase in frame size
   due to the presence of an MPLS label stack in the frame (as explained
   in Section 4.1.4.3).

   As observed in the current deployments, presence of an MPLS label
   stack in a Layer 2 frame is assumed to be transparent to Layer3=IP,
   which continues to follow the conventional maximum frame payload size
   [RFC3032] (1500 octets for Ethernet, say). This means that the
   effective maximum frame payload size [RFC3032] of the resulting Layer
   2 frame is Z octets more than the conventional maximum frame payload
   size, where Z = 4 x number of entries in the label stack.

   Hence, to ensure successful delivery of Layer 2 frames carrying MPLS
   packets and realistic benchmarking, it is RECOMMENDED to set the
   media MTU value to the effective maximum frame payload size
   [RFC3032], which equals Z octets + conventional maximum frame payload
   size. It is expected that such a change in the media MTU value only
   impacts the effective Maximum Frame Payload Size for MPLS packets,
   but not for IP packets.

   Note that this document requires exactly a single entry in the MPLS
   label stack in an MPLS packet. In other words, the depth of the
   label stack is set to one, e.g., Z = 4 x 1 = 4 octets. Furthermore,
   in accord with Sections 9 and 9.1 of RFC 2544, this document
   prescribes that each test case is run with different (Layer 2) frame

   sizes in different trials. Additionally, results MAY also be
   collected with multiple simultaneous frame sizes (sometimes referred
   to as an Interactive Multimodal Information Extraction (IMIX) to
   simulate real network traffic according to the frame size ordering
   and usage). There is no standard for mixtures of frame sizes, and
   the results are subject to wide interpretation (see Section 18 of RFC
   2544). When running trials using multiple simultaneous frame sizes,
   the DUT configuration MUST remain the same.

4.1.5.1. Frame Sizes To Be Used on Ethernet Media

   Ethernet media, in all its types, has become the most commonly
   deployed port media in MPLS networks. If any test case execution
   (such as the Label Push case) requires the test tool to send (or
   receive) a Layer 2 frame containing an IP packet, then the following
   frame sizes SHOULD be used for benchmarking over Ethernet media: 64,
   128, 256, 512, 1024, 1280, and 1518 octets. This is in-line with
   Sections 9 and 9.1 of RFC 2544. Figure 4 illustrates the header
   sizes for an untagged Ethernet frame containing an IP payload (per
   RFC 2544).

```
            <----------------64-1518B------------------------>
            <--18B---><-----------46-1500B------------------->
            +---------+---------+---------------------------+
            | Layer 2 | Layer 3 | Layer 4 (and higher)      |
            +---------+---------+---------------------------+
```
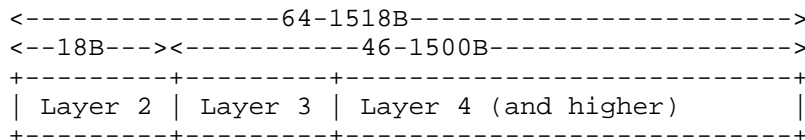
             Figure 4: Frame Size for Label Push Cases

   Note 1: The 64- and 1518-octet frame size represents the minimum
   and maximum length of an untagged Ethernet frame, as per IEEE
   802.3 [IEE8023]. A frame size commonly used in operational
   environments may range from 68 to 1522 octets, which are the
   minimum and maximum lengths of a single VLAN-tagged frame, as per
   IEEE 802.1D [IEE8021].

   Note 2: While jumbo frames are outside the scope of the 802.3 IEEE
   standard, tests SHOULD be executed with the frame sizes that are
   supported by the DUT. Examples of commonly used jumbo (Ethernet)
   frame sizes are: 4096, 8192, and 9216 octets.

   If any test case execution (such as Label Swap and Label Pop cases)
   requires the test tool to transmit (or receive) a Layer 2 frame
   containing an MPLS packet, then the untagged Layer 2 frame must

include an additional 4 octets for the MPLS header, resulting in the
following frame sizes to be used for benchmarking over Ethernet
media: 68, 132, 260, 516, 1028, 1284, and 1522 octets. Figure 5
illustrates the header sizes for an untagged Ethernet frame
containing an MPLS packet.

```
        <-----------------68-1522B----------------------------->
        <--18B---><--4B--><-----------46-1500B------------------>
        +---------+-------+---------+---------------------------+
        | Layer 2 | MPLS  | Layer 3 | Layer 4 (and higher)      |
        +---------+-------+---------+---------------------------+
```
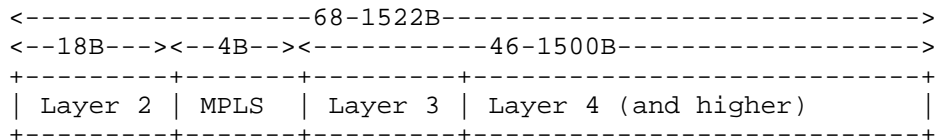
              Figure 5: Frame Size for Label Swap and Pop Cases

   Note: The Media MTU on the link between the test tool and the DUT
   must be changed, if needed, to accommodate the effective maximum
   frame payload size [RFC3032] resulting from adding an MPLS label
   stack to the IP packet. As clarified in Section 3.1 of RFC 3032,
   most Layer 2 media drivers are capable of sending and receiving
   Layer 2 frames with effective maximum frame payload size. Many
   vendors also allow the Media MTU to be changed for MPLS, without
   changing it for IP. The recommended link MTU value for MPLS is Z
   octets more than the conventional maximum frame payload size
   [RFC3032] (for example, the conventional maximum frame payload
   size for Ethernet is 1500 octets). This document prescribes Z=4
   octets. If a vendor DUT doesn't allow such an MTU change, then
   the benchmarking cannot be performed for the true maximum frame
   payload size [RFC3032] and this must be reported.

4.1.5.2. Frame Sizes to Be Used on POS Media

   Packet over SONET (POS) media are commonly used for edge uplinks and
   high-bandwidth core links. POS may use one of various encapsulations
   techniques (such as PPP, High-Level Data Link Control (HDLC), Frame
   Relay, etc.), resulting in the Layer 2 header (~4 octets) being less
   than that of the Ethernet media. The rest of the frame format
   (illustrated in Figures 2 and 3) remains pretty much unchanged.

   If the MPLS forwarding characterization of POS interfaces on the DUT
   is desired, then the following frame sizes SHOULD be used:

       Label Push test cases: 47, 64, 128, 256, 512, 1024,
                                       1280, 1518, 2048, and 4096 octets.

       Label Swap and Pop test cases: 51, 68, 132, 260, 516, 1028,
                                       1284, 1522, 2052, and 4100 octets.

4.1.6. Time-to-Live (TTL) or Hop Limit

   The TTL value in the frame header MUST be large enough to allow a TTL
   decrement to happen and still be forwarded through the DUT. The
   aforementioned TTL field may be referring to either the MPLS TTL,
   IPv4 TTL, or IPv6 Hop Limit depending on the exact forwarding
   scenario under evaluation.

   If TTL/Hop Limit decrement, as specified in [RFC3443], is a
   configurable option on the DUT, the setting SHOULD be reported.

4.1.7. Trial Duration

   Unless otherwise specified, the test portion of each trial SHOULD be
   no less than 30 seconds when static routing is in place, and no less
   than 200 seconds when a dynamic routing protocol and LDP (default LDP
   holddown timer is 180 seconds) are being used. If the holddown timer
   default value is changed, then it should be reported and the trial
   duration should still be 20 seconds more than the holddown timer
   value.

   The longer trial time used for dynamic routing protocols is to verify
   that the DUT is able to maintain a stable control plane when the
   data-forwarding plane is under stress.

4.1.8. Traffic Verification

   In all cases, sent traffic MUST be accounted for, whether it was
   received on the wrong port, the correct port, or not received at all.
   Specifically, traffic loss (also referred to as frame loss) is
   defined as the traffic (i.e., one or more frames) not received where
   expected (i.e., received on the incorrect port, or received with
   incorrect Layer 2 or above header information, etc.). In addition,
   the presence or absence of the MPLS label stack, every field value
   inside the label stack, if present, ethertype (0x8847 or 0x8848
   versus 0x0800 or 0x86DD), frame sequencing, and frame check sequence
   (FCS) MUST be verified in the received frame.

   Many test tools may, by default, only verify that they have received
   the embedded signature on the receive side. However, for MPLS header
   presence verification, some tests will require the MPLS header to be
   pushed while others will require a swap or pop. Hence, this document
   requires the test tool to verify the MPLS stack depth. An even
   greater level of verification would be to check if the correct label
   was pushed. However, some test tools are not capable of checking the
   received label value for correctness. Test tools SHOULD verify that
   the packets received carry the expected MPLS label.

4.1.9. Address Resolution and Dynamic Protocol State

   If a test setup utilizes any dynamic protocols for control plane
   signaling (e.g., ARP, PPP (including MPLSCP), OSPF, LDP, etc.), then
   all state for the protocols MUST be pre-established before the test
   case is executed (i.e., packet streams are started).

5. Reporting Format

   For each test case, it is RECOMMENDED that the following variables be
   reported in addition to the specific parameters requested by the test
   case:

Parameter	Units or Examples
Prefix Forwarding Equivalence Class (FEC)	IPv4, IPv6, Both
Label Distribution Protocol	LDP, RSVP-TE, BGP (or combinations)
MPLS Forwarding Operation	Push, Swap, Pop
IGP	ISIS, OSPF, EIGRP, RIP, static.
Throughput	Frames per second and bits per second
Port Media	GigE (Gigabit Ethernet), POS, ATM, etc.
Port Speed	1 gbps, 100 Mbps, etc.
Interface Encapsulation	Ethernet, Ethernet VLAN, PPP, HDLC, etc.
Frame Size (Section 4.1.5)	Octets
p (Number of {DA, DB} pair ports per Figure 1)	1,2, etc.

   The individual test cases may have additional reporting requirements
   that may refer to other RFCs.

6. MPLS Forwarding Benchmarking Tests

   MPLS is a different forwarding paradigm from IP. Unlike IP packets
   and IP forwarding, an MPLS packet may contain more than one MPLS
   header and may go through one of three forwarding operations: push
   (or LSP Ingress), swap, or pop (or LSP Egress), as defined in
   [RFC3031]. Such characteristics desire further granularity in MPLS
   forwarding benchmarking than those described in RFC 2544. Thus, the
   benchmarking may include, but is not limited to:

      1. Throughput

      2. Latency

      3. Frame-Loss Rate

      4. System Recovery

      5. Reset

      6. MPLS TC (previously known as EXP [RFC5462]) field Operations
         (including explicit-null cases)

      7. Negative Scenarios (TTL expiry, etc.)

      8. Multicast

   However, this document focuses only on the first five categories,
   inline with the spirit of RFC 2544. All the benchmarking test cases
   described in this document are expected to, at a minimum, follow the
   "Test Setup" and "Test Procedure" below:

   Test Setup

      Referring to Figure 1, a single port (p = 1) on both A and B
      Modules SHOULD be used. However, if the forwarding throughput of
      the DUT is more than that of the media rate of a single port, then
      additional ports on A and B Modules MUST be enabled as follows: if
      the DUT can be configured with the A and B ports so as to exceed
      the DUT's forwarding throughput without overloading any B ports,
      then those MUST be enabled; if, on the other hand, the DUT's
      forwarding throughput capacity is greater than what can be
      achieved enabling all ports, then all An and Bn ports MUST be
      enabled. In the case where more than one A and B port is enabled,
      the procedures described in Section 16 of RFC 2544 must be

followed to accommodate the multi-port scenario. The frame
formats transmitted and received must be in accord with Sections
4.1.4.3 and 4.1.4.4, and frame sizes must be in accord with
Section 4.1.5.

Note: The test tool must be configured not to advertise a prefix
or FEC to the DUT on more than one port. In other words, the DUT
must associate a FEC with one and only one DB port. The Equal
Cost Multi-Path (ECMP) behavior in MPLS networks uses heuristics
[RFC4928]; hence, the usage of ECMP is NOT permitted by this
document to ensure the deterministic forwarding behavior during
benchmarking.

Test Procedure

In accord with Section 26 of RFC 2544 [RFC2544], the traffic is
sent from test tool port(s) Ap to the DUT at a constant load for a
fixed-time interval, and is received from the DUT on test tool
port(s) Bp. As described in Section 4.1.4.3, the frame may
contain either an IP packet or an MPLS packet depending on the
test case need. Furthermore, the IP packet must be either an IPv4
or IPv6 packet, depending on whether the MPLS benchmarking is done
for IPv4 or IPv6.

If any frame loss is detected, then a new iteration is needed
where the offered load is decreased and the sender will transmit
again. An iterative search algorithm MUST be used to determine
the maximum offered frame rate with a zero frame loss.

This maximum offered frame rate that results in zero frame loss
through the DUT is defined as the Throughput in Section 3.17 of
[RFC1242] for that test case. Informally, this rate is referred
to as the No-Drop Rate (NDR).

Each iteration should involve varying the offered load of the
traffic, while keeping the other parameters (test duration, number
of ports, number of addresses, frame size, etc.) constant, until
the maximum rate at which none of the offered frames are dropped
is determined.

6.1. Throughput

This section contains the description of the tests that are related
to the characterization of a DUT's MPLS traffic forwarding.

6.1.1. Throughput for MPLS Label Push

   Objective

      To obtain the DUT's Throughput (as per RFC 2544) during label push
      or LSP Ingress forwarding operation (i.e., IP to MPLS).

   Test Setup

      In addition to the "Test Setup" described in Section 6, the test
      tool must advertise the IP prefix(es), i.e., RNx (using a routing
      protocol as per Section 4.1.2) and associated MPLS label-FEC
      binding(s) (using a label distribution protocol as per Section
      4.1.3) on its receive ports Bp to the DUT. The test tool may
      learn the IP prefix(es) on its transmit ports Ap from the DUT.

      MPLS and/or the label distribution protocol must be enabled only
      on the test tool receive ports Bp and DUT transmit ports DBp.

   Discussion

      The DUT's MPLS forwarding table (also referred to as Incoming
      Label Map (ILM) to Next Hop Label Forwarding Entry (NHLFE) mapping
      table per Section 3.11 of [RFC3031]) must contain a non-reserved
      MPLS label value as the outgoing label for each learned IP prefix
      corresponding to the label-FEC binding, resulting in the DUT
      performing the IP-to-MPLS forwarding operation. The test tool
      must receive MPLS packets on receive ports Bp (from the DUT) with
      the same label values that were advertised.

   Procedure

      Please see "Test Procedure" in Section 6. Additionally, the test
      tool MUST send the frames containing IP packets (with the IP
      destination belonging to the advertised IP prefix(es)) on transmit
      ports Ap, and expect to receive frames containing MPLS packets on
      receive ports Bp, as described in Section 4.1.4.4.

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

      Results for each test SHOULD be in the form of a table with a row
      for each of the tested frame sizes. Additional columns SHOULD
      include offered load and measured throughput.

6.1.2. Throughput for MPLS Label Swap

   Objective

      To obtain the DUT's Throughput (as per RFC 2544) during label
      swapping operation (i.e., MPLS-to-MPLS).

   Test Setup

      In addition to the setup described in Section 6, the test tool
      must advertise IP prefix(es) (using a routing protocol as per
      Section 4.1.2) and associated MPLS label-FEC bindings (using a
      label distribution protocol as per Section 4.1.3) on the receive
      ports Bp, and then learn the IP prefix(es) as well as the label-
      FEC binding(s) on the transmit ports Ap. The test tool must use
      the learned MPLS label values and learned IP prefix values in the
      frames transmitted on ports Ap to the DUT.

      MPLS and/or label distribution protocol must be enabled on the
      test tool ports Bp and Ap, and the DUT ports DBp and DAp.

   Discussion

      The DUT's MPLS forwarding table (also referred to as ILM to NHLFE
      mapping table per Section 3.11 of [RFC3031]) must contain non-
      reserved MPLS label values as the outgoing and incoming labels for
      the learned IP prefixes, resulting in an MPLS-to-MPLS forwarding
      operation, e.g., label swap. The test tool must receive MPLS
      packets on receive ports Bp (from the DUT) with the same label
      values that were advertised using the label distribution protocol.
      The received frames must contain the same number of MPLS headers
      as those of transmitted frames.

   Procedure

      Please see "Test Procedure" in Section 6. Additionally, the test
      tool must send frames containing MPLS packets (with the IP
      destination belonging to the advertised IP prefix(es)) on its
      transmit ports Ap, and expect to receive frames containing MPLS
      packets on its receive ports Bp, as described in Section 4.1.4.4.

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

      Results for each test SHOULD be in the form of a table with a row
      for each of the tested frame sizes.

6.1.3. Throughput for MPLS Label Pop (Unlabeled)

   Objective

      To obtain the DUT's Throughput (as per RFC 2544) during label pop
      or LSP Egress forwarding operation (i.e., MPLS-to-IP) using
      "Unlabeled" outgoing label.

   Test Setup

      In addition to the setup described in Section 6, the test tool
      must advertise the IP prefix(es) (using a routing protocol as per
      Section 4.1.2) without any MPLS label-FEC bindings on the receive
      ports Bp, and then learn the IP prefix(es) as well as the FEC-
      label binding(s) on the transmit ports Ap. The test tool must use
      the learned MPLS label values and learned IP prefix values in the
      frames transmitted on ports Ap.

      MPLS and/or label distribution protocol must be enabled only on
      the test tool port(s) Ap and DUT port(s) DAp.

   Discussion

      The DUT's MPLS forwarding table (also referred to as ILM to NHLFE
      mapping table per Section 3.11 of [RFC3031]) must contain an
      Unlabeled outgoing label (also known as untagged) for the learned
      IP prefix, resulting in an MPLS-to-IP forwarding operation.

   Procedure

      Please see "Test Procedure" in Section 6. Additionally, the test
      tool must send frames containing MPLS packets on its transmit
      ports Ap (with the IP destination belonging to the IP prefix(es)
      advertised on port Bp), and expect to receive frames containing IP
      packets on its receive ports Bp, as described in Section 4.1.4.4.

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

      Results for each test SHOULD be in the form of a table with a row
      for each of the tested frame sizes.

6.1.4. Throughput for MPLS Label Pop (Aggregate)

   Objective

      To obtain the DUT's Throughput (as per RFC 2544) during label pop
      or LSP Egress forwarding operation (i.e., MPLS-to-IP) using the
      "Aggregate" outgoing label [RFC3031].

   Test Setup

      In addition to the setup described in Section 6, the DUT must be
      provisioned such that it allocates an aggregate outgoing label
      (please see Section 3.20 in [RFC3031]) to an IP prefix, which
      aggregates a set of prefixes learned on ports DBp from the test
      tool. The prefix aggregation can be performed using BGP
      aggregation (Section 9.2.2.2 of [RFC4271]), IGP aggregation
      (Section 16.5 of [RFC2328]), etc.

      The DUT must advertise the aggregating IP prefix along with the
      associated MPLS label-FEC binding on ports DAp to the test tool.

      The test tool then must use the learned MPLS label values and
      learned IP prefix values in frames transmitted on ports Ap to the
      DUT. The test tool must receive frames containing IP packets on
      ports Bp from the DUT.

      MPLS and/or label distribution protocol must be enabled only on
      the test tool port(s) Ap and DUT port(s) DAp.

   Discussion

      The DUT's MPLS forwarding table (also referred to as ILM to NHLFE
      mapping table per Section 3.11 of [RFC3031]) must contain an
      aggregate outgoing label and IP forwarding table must contain a
      valid entry for the learned prefix(es), resulting in MPLS-to-IP
      forwarding operation (i.e., MPLS header removal followed by IP
      lookup).

   Procedure

      Please see "Test Procedure" in Section 6. Additionally, the test
      tool must send frames containing MPLS packets on its transmit
      ports Ap (with IP destination belonging to the IP prefix(es)
      advertised on port Bp), and expect to receive frames containing IP
      packets on its receive ports Bp, as described in Section 4.1.4.4.

Reporting Format

    The result should be reported as per Section 5 and RFC 2544.

    Results for each test SHOULD be in the form of a table with a row
    for each of the tested frame sizes.

6.1.5. Throughput for MPLS Label Pop (PHP)

   Objective

    To obtain the DUT's Throughput (as per RFC 2544) during label pop
    (i.e., MPLS-to-IP) or penultimate hop popping (PHP) using the
    "imp-null" outgoing label.

   Test Setup

    In addition to the setup described in Section 6, the test tool
    must be set up to advertise the IP prefix(es) (using a routing
    protocol as per Section 4.1.2) and associated MPLS label-FEC
    binding with a reserved MPLS label value = 3 (using a label
    distribution protocol as per Section 4.1.3) on its receive ports
    Bp. The test tool must learn the IP prefix(es) as well as the
    MPLS label-FEC bindings on its transmit ports Ap. The test tool
    then must use the learned MPLS label values and learned IP prefix
    values in the frames transmitted on ports Ap to the DUT. The test
    tool must receive frames containing IP packets on receive ports Bp
    (from the DUT).

    MPLS and/or label distribution protocol must be enabled on the
    test tool ports Bp and Ap, and DUT ports DBp and DAp.

   Discussion

    This test case characterizes Penultimate Hop Popping (PHP), which
    is described in RFC 3031.

    The DUT's MPLS forwarding table (also referred to as ILM to NHLFE
    mapping table per Section 3.11 of [RFC3031]) must contain a
    reserved MPLS label value = 3 (e.g., pop or imp-null) as the
    outgoing label for the learned prefix(es), resulting in MPLS-to-IP
    forwarding operation.

    This test case characterizes DUT's penultimate hop popping (PHP)
    functionality.

Procedure

    Please see "Test Procedure" in Section 6. Additionally, the test
    tool must send frames containing MPLS packets on its transmit
    ports Ap (with IP destination belonging to advertised IP
    prefix(es)), and expect to receive frames containing IP packets on
    its receive ports Bp, as described in Section 4.1.4.4.

Reporting Format

    The result should be reported as per Section 5 and RFC 2544.

    Results for each test SHOULD be in the form of a table with a row
    for each of the tested frame sizes.

6.2. Latency Measurement

    Latency measurement measures the time taken by the DUT to forward the
    MPLS packet during various MPLS switching paths such as IP-to-MPLS,
    MPLS-to-MPLS, or MPLS-to-IP involving an MPLS label stack.

    Objective

    To obtain the average latency induced by the DUT during MPLS
    packet forwarding for each of five forwarding operations.

    Test Setup

    Follow the "Test Setup" guidelines established for each of the
    five MPLS forwarding operations in Sections 6.1.1 (for IP-to-
    MPLS), 6.1.2 (for MPLS-to-MPLS), 6.1.3 (for MPLS-to-IP Unlabeled),
    6.1.4 (for MPLS-to-IP Aggregate), and 6.1.5 (for MPLS-to-IP PHP),
    one by one.

    Procedure

    Please refer to Section 26.2 in RFC 2544 in addition to following
    the associated procedure for each MPLS forwarding operation in
    accord with the test setup described earlier:

        IP-to-MPLS forwarding   (Push)   Section 6.1.1
        MPLS-to-MPLS forwarding   (Swap)   Section 6.1.2
        MPLS-to-IP forwarding   (Pop)   Section 6.1.3
        MPLS-to-IP forwarding   (Aggregate)   Section 6.1.4
        MPLS-to-IP forwarding   (PHP)   Section 6.1.5

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

6.3. Frame-Loss Rate (FLR) Measurement

   Frame-Loss Rate (FLR) measurement measures the percentage of MPLS
   frames that were not forwarded during various switching paths such as
   IP-to-MPLS (push), MPLS-to-IP (swap), or MPLS-IP (pop) by the DUT
   under overloaded state.

   Please refer to RFC 2544, Section 26.3, for more details.

   Objective

      To obtain the frame-loss rate, as defined in RFC 1242, for each of
      the three MPLS forwarding operations of a DUT, throughout the
      range of input data rates and frame sizes.

   Test Setup

      Follow the "Test Setup" guidelines established for each of the
      five MPLS forwarding operations in Sections 6.1.1 (for IP-to-
      MPLS), 6.1.2 (for MPLS-to-MPLS), 6.1.3 (for MPLS-to-IP Unlabeled),
      6.1.4 (for MPLS-to-IP Aggregate), and 6.1.5 (for MPLS-to-IP PHP),
      one by one.

   Procedure

      Please refer to Section 26.3 of RFC 2544 [RFC2544] and follow the
      associated procedure for each MPLS forwarding operation one-by-one
      in accord with the test setup described earlier:

         IP-to-MPLS forwarding   (Push)   Section 6.1.1
         MPLS-to-MPLS forwarding   (Swap)   Section 6.1.2
         MPLS-to-IP forwarding   (Pop)   Section 6.1.3
         MPLS-to-IP forwarding   (Aggregate)   Section 6.1.4
         MPLS-to-IP forwarding   (PHP)   Section 6.1.5

      A misdirected frame, that is, a frame received on the wrong Bn, is
      considered lost. If the test tool is capable of checking received
      MPLS label values, a frame with the wrong MPLS label is considered
      lost.

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

6.4. System Recovery

   Objective

      To characterize the speed at which a DUT recovers from an overload
      condition.

   Test Setup

      Follow the "Test Setup" guidelines established for each of the
      five MPLS forwarding operations in Sections 6.1.1 (for IP-to-
      MPLS), 6.1.2 (for MPLS-to-MPLS), 6.1.3 (for MPLS-to-IP Unlabeled),
      6.1.4 (for MPLS-to-IP Aggregate), and 6.1.5 (for MPLS-to-IP PHP),
      one by one.

   Procedure

      Please refer to Section 26.5 of RFC 2544 and follow the associated
      procedure for each MPLS forwarding operation in the referenced
      sections one-by-one in accord with the test setup described
      earlier:

         IP-to-MPLS forwarding   (Push)   Section 6.1.1
         MPLS-to-MPLS forwarding   (Swap)   Section 6.1.2
         MPLS-to-IP forwarding   (Pop)   Section 6.1.3
         MPLS-to-IP forwarding   (Aggregate) Section 6.1.4
         MPLS-to-IP forwarding   (PHP)   Section 6.1.5

   Reporting Format

      The result should be reported as per Section 5 and RFC 2544.

6.5. Reset

   The "reset" aspects of benchmarking are described in [RFC2544], but
   these procedures need to be clarified in order to ensure consistency.
   This document does not specify the reset procedures. These need to
   be addressed in a separate document and will more generally apply to
   IP and MPLS test cases.

   The text below describes the MPLS forwarding benchmarking-specific
   setup that will have to be used in conjunction with the procedures
   from the separate document to make this test case meaningful.

   Objective

      To characterize the speed at which a DUT recovers from a device or
      software reset.

Test Setup

   Follow the "Test Setup" guidelines established for each of the
   five MPLS forwarding operations in Sections 6.1.1 (for IP-to-
   MPLS), 6.1.2 (for MPLS-to-MPLS), 6.1.3 (for MPLS-to-IP Unlabeled),
   6.1.4 (for MPLS-to-IP Aggregate), and 6.1.5 (for MPLS-to-IP PHP),
   one by one.

   For this test case, the requirements of LDP and a routing protocol
   are removed and replaced by static configurations. For the IP-to-
   MPLS forwarding, static route configurations should be applied.
   For the MPLS-to-MPLS and MPLS-to-IP, static label configurations
   must be applied.

   For this test, all Graceful Restart features MUST be disabled.

Discussion

   This test case is intended to provide insight into how long an
   MPLS device could take to be fully operational after any of the
   reset events. It is quite likely that the time an IP/MPLS device
   takes to become fully operational after any of the reset events
   may be different from that of an IP-only device.

   Modern devices now have many more reset options that were not
   available when Section 26.6 of RFC 2544 was published. Moreover,
   different reset events on modern devices may produce different
   results, hence, needing clarity and consistency in reset
   procedures beyond what's specified in RFC 2544.

Procedure

   Please follow the procedure from the separate document for each
   MPLS forwarding operation one-by-one:

      IP-to-MPLS forwarding   (Push)   Section 6.1.1
      MPLS-to-MPLS forwarding   (Swap)   Section 6.1.2
      MPLS-to-IP forwarding   (Pop)   Section 6.1.3
      MPLS-to-IP forwarding   (Aggregate)   Section 6.1.4
      MPLS-to-IP forwarding   (PHP)   Section 6.1.5

Reporting Format

   The result should be reported as per Section 5 and as per the
   separate document.

7. Security Considerations

   Benchmarking activities, as described in this memo, are limited to
   technology characterization using controlled stimuli in a laboratory
   environment, with dedicated address space and the constraints
   specified in the sections above.

   The benchmarking network topology will be an independent test setup
   and MUST NOT be connected to devices that may forward the test
   traffic into a production network or misroute traffic to the test
   management network.

   Furthermore, benchmarking is performed on a "black-box" basis,
   relying solely on measurements observable external to the DUT/SUT
   (System Under Test).

   Special capabilities SHOULD NOT exist in the DUT/SUT specifically for
   benchmarking purposes. Any implications for network security arising
   from the DUT/SUT SHOULD be identical in the lab and in production
   networks.

   There are no specific security considerations within the scope of
   this document.

8. Acknowledgement

   The authors would like to thank Mo Khalid, who motivated us to write
   this document. We would like to thank Rodney Dunn, Chip Popoviciu,
   Jeff Byzek, Jay Karthik, Rajiv Papneja, Samir Vapiwala, Silvija
   Andrijic Dry, Scott Bradner, Al Morton, and Bill Cerveny for their
   careful review and suggestions.

   This document was originally prepared using 2-Word-v2.0.template.dot.

9. References

9.1. Normative References

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

   [RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for
             Network Interconnect Devices", RFC 2544, March 1999.

   [RFC1242] Bradner, S., "Benchmarking Terminology for Network
             Interconnection Devices", RFC 1242, July 1991.

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

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

   [RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in
              BGP-4", RFC 3107, May 2001.

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

9.2. Informative References

   [RFC5180] Popoviciu, C., Hamza, A., Van de Velde, G., and D.
              Dugatkin, "IPv6 Benchmarking Methodology for Network
              Interconnect Devices", RFC 5180, May 2008.

   [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.

   [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
              Networks (VPNs)", RFC 4364, February 2006.

   [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border
              Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

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

   [IEE8021] Mick Seaman (editor), "IEEE Std 802.1D-2004, MAC Bridges",
              Feb 2004.

   [IEE8023] LAN/MAN Standards Committee of the IEEE Computer Society,
              "IEEE Std 802.3as-2006, Part 3: Carrier Sense Multiple
              Access with Collision Detection (CSMA/CD) Access Method and
              Physical Layer Specifications, Amendment 3: Frame format
              extensions", Nov 2006.

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

   [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

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

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

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

Authors' Addresses

   Aamer Akhter
   Cisco Systems
   7025 Kit Creek Road
   RTP, NC 27709
   USA

   EMail: aakhter@cisco.com

   Rajiv Asati
   Cisco Systems
   7025 Kit Creek Road
   RTP, NC 27709
   USA

   EMail: rajiva@cisco.com

   Carlos Pignataro
   Cisco Systems
   7200-12 Kit Creek Road
   RTP, NC 27709
   USA

   EMail: cpignata@cisco.com