Network Working Group
Internet Draft
Intended status: Informational
Expires: March 2009

Aamer Akhter
Rajiv Asati
Cisco Systems

September 15, 2008

MPLS Forwarding Benchmarking Methodology

draft-ietf-bmwg-mpls-forwarding-meth-00.txt

Status of this Memo

Abstract

The purpose of this draft is to describe a methodology specific to the benchmarking of MPLS forwarding devices. The scope of this benchmarking will be limited to various types of packet-forwarding and delay measurements. It builds upon the tenets set forth in RFC2544 [RFC2544], RFC1242 [RFC1242] and other IETF Benchmarking Methodology Working Group (BMWG) efforts. This document seeks to extend these efforts to the MPLS paradigm.

Table of Contents

1. Introduction

Over the past several years MPLS networks have gained greater popularity. However, there is no standard method to compare and

contrast the varying implementations and their strong and weak points. This document proposes a methodology using common criteria for the comparison of various implementations of basic MPLS forwarding devices.

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

2. Document Scope

MPLS [RFC3031] is a foundation enabling technology for other more advanced technologies such as Layer 3 MPLS-VPNs, Layer 2 MPLS-VPNs, and MPLS Traffic Engineering. This document focuses on MPLS forwarding characterization.

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 topologies described in this section. An effort has been made to exclude superfluous equipment needs such that each test can be carried out with the minimum number of requirements.

Figure 1 illustrates the sample topology in which the DUT is connected to the test ports on the test tool.

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

Figure 1 Topology #1 for MPLS Forwarding Benchmarking

Where number of ports (p) is determined by the maximum unidirectional forwarding throughput of the DUT and the load capacity of the media between the Test Ports and DUT. For example, if the DUT's forwarding throughput is 100 frames per second (fps), and the media capacity is 50 fps, then p = 2.

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, QoS, Graceful Restart etc. MUST be disabled, unless explicitly requested by the test case. Additionally, any non-essential traffic MUST also be avoided.

4.1.1. IGP Support

It is highly RECOMMENDED that all of the interfaces (A1, DA1, DB1, A2..) on DUT and test tool support an IGP such as IS-IS, OSPF, EIGRP, RIP etc. Furthermore, there are testing considerations in this document that the device is able to provide a stable control-plane during heavy forwarding workloads. The route distribution method used (OSPF, IS-IS, EIGRP, RIP etc.) MUST be reported.

4.1.2. Label Distribution Support

The DUT and test tool must support at least one protocol for exchanging MPLS labels. The DUT and test tool MUST be capable of learning and advertising MPLS label bindings via the chosen protocol(s), and use them during packet forwarding all the time (includes when the label bindings change). The most commonly used protocol is Label Distribution Protocol (LDP) [RFC5036], RSVP-TE [RFC5151] and MP-BGP [RFC4364].

All of the interfaces connected to the DUT such as A1, DA1, DB1, A2 etc., SHOULD support Label Distribution Protocol (LDP), RSVP-TE, and MP-BGP for IPv4 or IPv6 FECs.

This draft discourages the use of static label to establish the MPLS label switched paths, since it is not commonly used in the production networks.

4.1.3. Frame Sizes

Each test SHOULD be run with different frame sizes in different trials. For better reference, the recommended sizes are 64, 128, 256, 512, 1024, 1280 and 1518 for IPv4. Recommended sizes for other media can be found in RFC 2544 and IPv6 Benchmarking [RFC5180]. Frame sizes MUST be based on the pre-MPLS shim version of the frame.

In addition to the individual frame size trials, an IMIX traffic run MAY also be included.

When running trials between different frame sizes, the DUT configuration MUST remain the same.

4.1.4. TTL or Hop Count

The MPLS TTL or IPv4 TTL or IPv6 Hop Count (depending on which portion of the packet the DUT is basing the forwarding behavior) MUST be large enough to traverse the DUT.

4.1.5. 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 holddown timer is 180 seconds) are being used.

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

4.1.5.1. Traffic Verification

In all cases the sent traffic MUST be accounted for, whether it was received on the wrong port, 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 incorrect port, or received with incorrect layer2 or above header information etc.). In addition, the MPLS header presence or non-presence of the packet MUST be verified, as well as checksum, frame sequencing and correct MPLS TTL decrementing.

The MPLS header presence will be determined by the test. Some tests will require the MPLS header to be imposed while others will require a swap or disposition. In general, many test tools will by default only verify that they have received the embedded signature on the receive side, but will not validate MPLS stack depth. An even greater level of verification would be to check if the correct label was imposed, but that is considered out of scope for these tests.

"In all cases the sent traffic MUST be accounted for, whether it was received on the wrong port, correct port or not received at all. In addition, the MPLS header...."

4.1.6. Address Resolution and Dynamic Protocol State

If the test or media is making use of a dynamic protocol (eg ARP, OSPF, LDP), all state for the protocols should be pre-established before the start of the trial.

4.1.7. Abbreviations Used

Please refer to Figure 1, "Port based Remote Network" for a topology view of the network. The following abbreviations are used in this document -

M := Module Side (could be A or B)

P := port number

RN := Remote Network (can also be thought of as a network that is reachable via) Mp.

Y := number of network. (ie the first network reachable via B1 would be called B1RN1 and the 5th network would be called B1RN5)

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	Unit
Internet Protocol	IPv4, IPv6, Dual-Stack
Label Distribution Protocol	LDP, RSVP-TE, BGP (or combinations)
MPLS Forwarding Operation	Imposition, Swap, Disposition
IGP	ISIS, OSPF, EIGRP, RIP, static, etc.
Throughput	Frames per second
Interface Type	GigE, POS, ATM etc
Interface Speed	1 gbps, 100 Mbps, etc
Interface Encapsulation	VLAN, PPP, HDLC

Packet Size	Bytes
Number of A and B interfaces (see Figure 1)	1A, 2B

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

6. MPLS Forwarding Benchmarking tests

MPLS is altogether a different forwarding paradigm from IP. Unlike IP packet and IP forwarding, MPLS packet is likely to contain more than one MPLS headers and may go through one of three forwarding operations - imposition, swap and disposition. Such characteristics desire further granularity in MPLS forwarding benchmarking than those of described in RFC2544. Thus the benchmarking includes, but not limited to:

1. Throughput

2. Latency

3. Frame Loss rate

4. System Recovery

5. Reset

6. MPLS EXP field Operations (including explicit-null cases)

7. Negative Scenarios (TTL expiry, etc)

8. Multicast

This document focuses on the first five categories. All the benchmarking test cases described in this document are expected to at a minimum follow the below 'Test Setup' and 'Test Procedure.'

Test Setup

It is recommended that a single A and B interface SHOULD be used. However, if the forwarding throughput of the DUT is more than that of the media rate of a single interface, then additional A and B interfaces MUST be enabled so as to exceed the DUT's forwarding throughput. In such case, the tool traffic should use BpRN1 and BpAN as the IP destinations in a weighted round robin fashion. The weighting ratio between BpRN1 and BpAN is a constant test parameter. A suggested ratio is 1:100 with BpAN:BpRN1. The traffic streams offered MUST conform to section 16 of RFC 2544.

Test Procedure

Send traffic from port Ap towards DUT at a constant load towards IP prefixes (BpRN1 addresses) advertised by the tool on the receive ports, for a fixed duration of time.

If any frame loss is detected, 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.

Each iteration will involve varying the offered load of the regular traffic, while keeping the other parameters (test duration, number of interfaces, 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 DUT's MPLS frame forwarding.

6.1.1. Throughput for MPLS Label Imposition

Objective

To obtain the maximum forwarding rate during label imposition (i.e. IP to MPLS) for a regular (IPv4 or IPv6) packet by the DUT.

Test Setup

In addition to setup described in section 6, the test tool should advertise the IP prefix(es) i.e. RNx(using a routing protocol as per section 1.1) and associated MPLS label (using a label distribution protocol as per section 1.2) on its receive ports Bp to DUT. The test tool may learn these IP prefixes on it's transmit ports Ap from DUT.

Discussion

The DUT's MPLS forwarding table must contain non-reserved MPLS label value as the outgoing label for the learned prefix, resulting in IP-to-MPLS forwarding operation. The testool must receive MPLS packets on receive ports Bp (from DUT) with the same label values that are advertised.

Procedure

Please see Test Procedure in section 6. Additionally, the test tool must send unlabeled IP packets on transmit ports Ap (with IP destination belonging to above IP prefix(es)), and expect to receive MPLS packets on receive ports Bp.

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

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. Throughout for MPLS Label Swap

Objective

To obtain the maximum label swap rate for a labeled packet (i.e. MPLS to MPLS) by the DUT.

Test Setup

In addition to setup described in section 6, the test tool must be set up to advertise IP prefix (using a routing protocol as per section 1.1) and associated MPLS label (using a label distribution protocol as per section 1.2) on the receive ports Bp, and learn the IP prefix(es) with the appropriate MPLS labels on the transmit ports Ap. The test tool then must use the learned MPLS label

values and learned IP prefix values in MPLS packets transmitted on ports Ap.

Discussion

The DUT's MPLS forwarding table must contain non-reserved MPLS label values as the outgoing and incoming labels for the learned prefix, resulting in MPLS-to-MPLS forwarding operation. The testool must receive MPLS packets on receive ports Bp (from DUT). The received MPLS packets must contain the same number of MPLS headers as those of transmitted MPLS Packets.

Procedure

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

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

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.

6.1.3. Throughout for MPLS Label Disposition

Objective

To obtain the maximum label disposition rate for MPLS packet (i.e. MPLS to IP) by the DUT, when DUT installs ''Untagged'' outgoing label.

Test Setup

In addition to 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 1.1) without any MPLS label on the receive ports Bp, and learn the IP prefix(es) with the appropriate MPLS labels on the transmit ports Ap. The test tool then must use the learned MPLS label values and learned IP prefix values in MPLS packets transmitted on ports Ap.

Discussion

The DUT's MPLS forwarding table must contain an untagged outgoing label for the learned prefix, resulting in MPLS-to-IP forwarding operation. The testool must receive IP packets on receive ports Bp (from DUT).

Procedure

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

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

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.

6.1.4. Throughput for MPLS Label Disposition (Aggregate)

Objective

To obtain the maximum label disposition rate for MPLS packet (i.e. MPLS to IP) by the DUT, when DUT installs ''Aggregate'' outgoing label.

Test Setup

In addition to setup described in section 6, the DUT should be provisioned such that it allocates an aggregate outgoing label to a prefix (where the prefix may be a 'BGP aggregated prefix' , 'BGP VPN connected prefix' or an IGP aggregation that results in an aggregate label, etc. and must include the addresses belonging to the DUT receive ports Bp).

The DUT must advertise the IP prefix(es) along with the MPLS label(s) via a label distribution protocol to the testool on tool transmit ports Ap.

The test tool then must use the learned MPLS label values and learned IP prefix values in MPLS packets transmitted on ports Ap.

Discussion

The DUT's MPLS forwarding table must contain an aggregate outgoing label and IP forwarding table must contain a valid entry for the learned prefix, resulting in MPLS-to-IP forwarding operation (i.e. MPLS header removal followed by IP lookup). The testool must receive IP packets on receive ports Bp (from DUT).

Procedure

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

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

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.

6.2. Latency Measurement

This measures the time taken by the DUT to forward the MPLS packet during various MPLS switching paths such as IP-to-MPLS or MPLS-to-MPLS or MPLS-to-IP involving one or more MPLS headers.

The forwarding delay measurement requires the accurate propagation delay measurement as a prerequisite.

One of the propagation delay measurement mechanisms is to connect test transmit port such as A1 and test receive port such as B1 with the wire length=X (bypass DA1 and DB1) and measure the time (t1) taken by the packet to reach from A1 to B1.

Once the time t1 has been recorded, then the DUT should be inserted such that the test port A1 connects to DA1 and B1 connects to DB1, and the sum of A1-DA1 wire length and B1-DB1 wire length equals X.

The packet should be sent from A1 to B1 such that the packet is received by DA1, which after consulting with its forwarding table, forwards the packet to B1 via DB1. The time (t2) taken by the packet to reach B1 (from A1) is recorded.

The difference of time t2-t1 would provide the ballpark measurement of DUT's forwarding delay.

The measurement for t2 should be performed under each of three forwarding operations (IP-to-MPLS, MPLS-to-MPLS, MPLS-to-IP) and measured accordingly.

Objective

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

Test Setup

Follow the test setup guidelines established for each of three MPLS forwarding operations in section 6.1.1 (for IP-to-MPLS), 6.1.2 (for MPLS-to-MPLS) and 6.1.3 (for MPLS-to-IP) one by one.

Procedure

Please refer to RFC2544. Additionally, follow the associated procedure for each MPLS forwarding operation -

IP-to-MPLS forwarding	(Imposition)	Section 6.1.1
MPLS-to-MPLS forwarding	(Swap)	Section 6.1.2

MPLS-to-IP forwarding (Disposition) Section 6.1.3

MPLS-to-IP forwarding (Aggregate) Section 6.1.4

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

6.3. Frame Loss Rate Measurement (FLR)

This measures the percentage of MPLS frames that were not forwarded during various switching paths such as IP-to-MPLS (imposition) or MPLS-to-IP (swap) or MPLS-IP (disposition) by the DUT under overloaded state.

Please refer to RFC2544 section 26.3 for more details.

Objective

To obtain the frame loss rate, as defined in RFC1242, for each of 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 three MPLS forwarding operations in section 6.1.1 (for IP-to-MPLS), 6.1.2 (for MPLS-to-MPLS) and 6.1.3 (for MPLS-to-IP) and procedure one by one.

Procedure

Please refer to RFC2544.

Additionally, follow the associated procedure (and test Setup) for each MPLS forwarding operation one-by-one -

IP-to-MPLS forwarding	(Imposition)	Section 6.1.1
MPLS-to-MPLS forwarding	(Swap)	Section 6.1.2
MPLS-to-IP forwarding	(Disposition)	Section 6.1.3
MPLS-to-IP forwarding	(Aggregate)	Section 6.1.4

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

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 three MPLS forwarding operations in section 6.1.1 (for IP-to-MPLS), 6.1.2 (for MPLS-to-MPLS) and 6.1.3 (for MPLS-to-IP) and procedure one by one.

Procedure

Please refer to RFC2544 section 26.5.

Additionally, follow the associated procedure (and test Setup) for each MPLS forwarding operation one-by-one -

IP-to-MPLS forwarding (Imposition) Section 6.1.1

MPLS-to-MPLS forwarding (Swap) Section 6.1.2

MPLS-to-IP forwarding (Disposition) Section 6.1.3

MPLS-to-IP forwarding (Aggregate) Section 6.1.4

Reporting Format

Same as RFC2544, in addition to parameters in Section 4.

6.5. Reset

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 three MPLS forwarding operations in section 6.1.1 (for IP-to-MPLS), 6.1.2 (for MPLS-to-MPLS) and 6.1.3 (for MPLS-to-IP) and procedure one by one.

For this test, all graceful-restart features MUST be disabled.

Procedure

Please refer to RFC2544 section 26.5. Examples of hardware and software resets are:

hardware reset - forwarding module resetting (e.g. OIR).

software reset - reset initiated through a CLI (e.g. reload).

Additionally, follow the associated procedure (and test Setup) for each MPLS forwarding operation one-by-one -

IP-to-MPLS forwarding	(Imposition)	Section 6.1.1
MPLS-to-MPLS forwarding	(Swap)	Section 6.1.2
MPLS-to-IP forwarding	(Disposition)	Section 6.1.3
MPLS-to-IP forwarding	(Aggregate)	Section 6.1.4

Reporting Format

Same as RFC2544, in addition to parameters in Section 4, and the specific kind of reset performed.

7. Security Considerations

During the course of test, the test topology must be disconnected from devices that may forward the test traffic into a production environment.

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

8. IANA Considerations

There are no considerations for IANA at this time.

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.

[RFC3031] Rosen et al., ''Multiprotocol Label Switching Architecture'', Rosen et al., RFC 3031, August 1999.

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

[RFC5036] Andersson, L., Doolan, P., Feldman, N., Fredette, A. and B. Thomas, "LDP Specification", RFC 5036, January 2001.

9.2. Informative References

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

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

[RFC5180] Popoviciu, C., et al, "IPv6 Benchmarking Methodology for Network Interconnect Devices", RFC 5180, May 2008.

[RFC5151] Farrel, et al, "Inter-Domain MPLS and GMPLS Traffic Engineering --Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, Feb 2008.

Author's Addresses

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

Phone: 919 392 2564
Email: aakhter@cisco.com

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

Phone: 919 392 8558
Email: rajiva@cisco.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Special thanks to Scott Bradner for his very insightful comments delivered on very short notice.

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).