Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 15, 2012

R. Papneja
Isocore
S. Vapiwala
J. Karthik
Cisco Systems
S. Poretsky
Allot Communications
S. Rao
Qwest Communications
J. Roux
France Telecom

September 12, 2011

Methodology for benchmarking MPLS protection mechanisms

draft-ietf-bmwg-protection-meth-08.txt

Abstract

This draft describes the methodology for benchmarking MPLS Protection mechanisms for link and node protection as defined in [MPLS-FRR-EXT]. This document provides test methodologies and testbed setup for measuring failover times while considering all dependencies that might impact faster recovery of real-time applications bound to MPLS based traffic engineered tunnels. The benchmarking terms used in this document are defined in [TERM-ID].

Table of Contents

1. Introduction

This draft describes the methodology for benchmarking MPLS based protection mechanisms. The new terminology that this document introduces is defined in [TERM-ID].

MPLS based protection mechanisms provide fast recovery of real-time services from a planned or an unplanned link or node failures. MPLS protection mechanisms are generally deployed in a network infrastructure where MPLS is used for provisioning of point-to- point traffic engineered tunnels (tunnel). MPLS based protection mechanisms promise to improve service disruption period by minimizing recovery time from most common failures.

Network elements from different manufacturers behave differently to network failures, which impacts the network's ability and performance for failure recovery. It therefore becomes imperative for service providers to have a common benchmark to understand the performance behaviors of network elements.

There are two factors impacting service availability: frequency of failures and duration for which the failures persist. Failures can be classified further into two types: correlated and uncorrelated. Correlated and uncorrelated failures may be planned or unplanned. Planned failures are predictable. Network implementations should be able to handle both planned and unplanned failures and recover gracefully within a time frame to maintain service assurance. Hence, failover recovery time is one of the most important benchmark that a service provider considers in choosing the building blocks for their network infrastructure.

A correlated failure is the simultaneous occurrence of two or more failures. A typical example is failure of a logical resource (e.g. layer-2 links) due to a dependency on a common physical resource (e.g. common conduit) that fails. Within the context of MPLS protection mechanisms, failures that arise due to Shared Risk Link Groups (SRLG) [MPLS-FRR-EXT] can be considered as correlated failures. Not all correlated failures are predictable in advance, for example, those caused by natural disasters.

MPLS Fast Re-Route (MPLS-FRR) allows for the possibility that the Label Switched Paths can be re-optimized in the minutes following Failover. IP Traffic would be re-routed according to the preferred path for the post-failure topology. Thus, MPLS-FRR includes an additional step to the General model:

(1) Failover Event - Primary Path (Working Path) fails
(2) Failure Detection- Failover Event is detected
(3)
    a. Failover - Working Path switched to Backup path
    b. Re-Optimization of Working Path (possible change from Backup Path)
(4) Restoration - Primary Path recovers from a Failover Event
(5) Reversion (optional) - Working Path returns to Primary Path

2. Document Scope

This document provides detailed test cases along with different topologies and scenarios that should be considered to effectively benchmark MPLS protection mechanisms and failover times on the Data Plane. Different Failover Events and scaling considerations are also provided in this document.

All benchmarking testcases defined in this document apply to both facility backup and local protection enabled in detour mode. The test cases cover all possible failure scenarios and the associated procedures benchmark the performance of the Device Under Test (DUT) to recover from failures. Data plane traffic is used to benchmark failover times.

Benchmarking of correlated failures is out of scope of this document. Protection from Bi-directional Forwarding Detection (BFD) is outside the scope of this document.

As described above, MPLS-FRR may include a Re-optimization of the Working Path, with possible packet transfer impairments. Characterization of Re-optimization is beyond the scope of this memo.

3. Existing Definitions and 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 [Br97]. 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.

The reader is assumed to be familiar with the commonly used MPLS terminology, some of which is defined in [MPLS-FRR-EXT].

This document uses much of the terminology defined in [TERM-ID].

This document also uses existing terminology defined in other BMWG work. Examples include, but are not limited to:

Throughput	[Ref.[Br91], section 3.17]
Device Under Test (DUT)	[Ref.[Ma98], section 3.1.1]
System Under Test (SUT)	[Ref.[Ma98], section 3.1.2]
Out-of-order Packet	[Ref.[Po06], section 3.3.2]
Duplicate Packet	[Ref.[Po06], section 3.3.3]

4. General Reference Topology

Figure 1 illustrates the basic reference testbed and is applicable to all the test cases defined in this document. The Tester is comprised of a Traffic Generator (TG) & Test Analyzer (TA). A Tester is directly connected to the DUT. The Tester sends and receives IP traffic to the tunnel ingress and performs signaling protocol emulation to simulate real network scenarios in a lab environment. The Tester may also support MPLS-TE signaling to act as the ingress node to the MPLS tunnel.

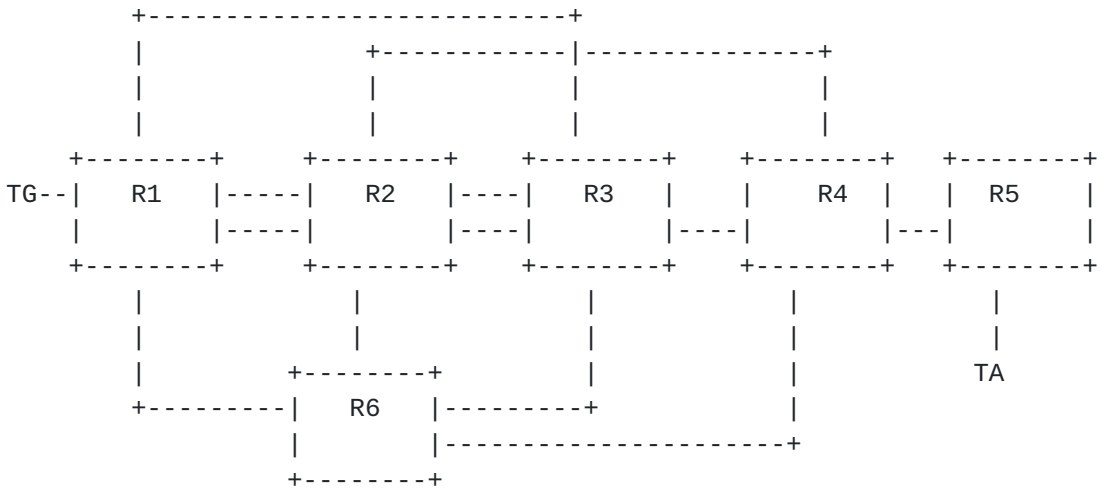

Fig. 1 Fast Reroute Topology

The tester MUST record the number of lost, duplicate, and reordered packets. It should further record arrival and departure times so that Failover Time, Additive Latency, and Reversion Time can be measured. The tester may be a single device or a test system emulating all the different roles along a primary or backup path. The label stack is dependent of the following 3 entities:

(1) Type of protection (Link Vs Node)
(2) # of remaining hops of the primary tunnel from the PLR
(3) # of remaining hops of the backup tunnel from the PLR

Due to this dependency, it is RECOMMENDED that the benchmarking of failover times be performed on all the topologies provided in section 6.

5. Test Considerations

This section discusses the fundamentals of MPLS Protection testing:

(1) The types of network events that causes failover
(2) Indications for failover
(3) the use of data traffic
(4) Traffic generation
(5) LSP Scaling
(6) Reversion of LSP
(7) IGP Selection

5.1. Failover Events [TERM-ID]

The failover to the backup tunnel is primarily triggered by either link or node failures observed downstream of the Point of Local repair (PLR). Some of these failure events are listed below.

Link Failure Events

- Interface Shutdown on PLR side with POS Alarm
- Interface Shutdown on remote side with POS Alarm
- Interface Shutdown on PLR side with RSVP hello enabled
- Interface Shutdown on remote side with RSVP hello enabled
- Interface Shutdown on PLR side with BFD
- Interface Shutdown on remote side with BFD
- Fiber Pull on the PLR side (Both TX & RX or just the TX)
- Fiber Pull on the remote side (Both TX & RX or just the RX)
- Online insertion and removal (OIR) on PLR side
- OIR on remote side
- Sub-interface failure (e.g. shutting down of a VLAN)
- Parent interface shutdown (an interface bearing multiple sub-interfaces

Node Failure Events

- A System reload initiated either by a graceful shutdown or by a power failure.
- A system crash due to a software failure or an assert.

5.2. Failure Detection [TERM-ID]

Link failure detection time depends on the link type and failure detection protocols running. For SONET/SDH, the alarm type (such as LOS, AIS, or RDI) can be used. Other link types have layer-two alarms, but they may not provide a short enough failure detection time. Ethernet based links do not have layer 2 failure indicators, and therefore relies on layer 3 signaling for failure detection. However for directly connected devices, remote fault indication in the ethernet auto-negotiation scheme could be considered as a type of layer 2 link failure indicator.

MPLS has different failure detection techniques such as BFD, or use of RSVP hellos. These methods can be used for the layer 3 failure indicators required by Ethernet based links, or for some other non-Ethernet based links to help improve failure detection time.

The test procedures in this document can be used for a local failure or remote failure scenarios for comprehensive benchmarking and to evaluate failover performance independent of the failure detection techniques.

5.3. Use of Data Traffic for MPLS Protection benchmarking

Currently end customers use packet loss as a key metric for Failover Time [TERM-ID]. Failover Packet Loss [TERM-ID] is an externally observable event and has direct impact on application performance. MPLS protection is expected to minimize the packet loss in the event

of a failure. For this reason it is important to develop a standard router benchmarking methodology for measuring MPLS protection that uses packet loss as a metric. At a known rate of forwarding, packet loss can be measured and the failover time can be determined. Measurement of control plane signaling to establish backup paths is not enough to verify failover. Failover is best determined when packets are actually traversing the backup path.

An additional benefit of using packet loss for calculation of failover time is that it allows use of a black-box test environment. Data traffic is offered at line-rate to the device under test (DUT) an emulated network failure event is forced to occur, and packet loss is externally measured to calculate the convergence time. This setup is independent of the DUT architecture.

In addition, this methodology considers the packets in error and duplicate packets that could have been generated during the failover process. The methodologies consider lost, out-of-order, and duplicate packets to be impaired packets that contribute to the Failover Time.

5.4. LSP and Route Scaling

Failover time performance may vary with the number of established primary and backup tunnel label switched paths (LSP) and installed routes. However the procedure outlined here should be used for any number of LSPs (L) and number of routes protected by PLR(R). The amount of L and R must be recorded.

5.5. Selection of IGP

The underlying IGP could be ISIS-TE or OSPF-TE for the methodology proposed here. See [IGP-METH] for IGP options to consider and report.

5.6. Restoration and Reversion [TERM-ID]

Fast Reroute provides a method to return or restore an original primary LSP upon recovery from the failure (Restoration) and to switch traffic from the Backup Path to the restored Primary Path (Reversion). In MPLS-FRR, Reversion can be implemented as Global Reversion or Local Reversion. It is important to include Restoration and Reversion as a step in each test case to measure the amount of packet loss, out of order packets, or duplicate packets that is produced.

Note: In addition to restoration and reversion, re-optimization can take place while the failure is still not recovered but it depends on

the user configuration, and re-otimization timers.

5.7. Offered Load

It is suggested that there be one or more traffic streams as long as there is a steady and constant rate of flow for all the streams. In order to monitor the DUT performance for recovery times, a set of route prefixes should be advertised before traffic is sent. The traffic should be configured towards these routes.

At least 16 flows should be used, and more if possible. Prefix-dependency behaviors are key in IP and tests with route-specific flows spread across the routing table will reveal this dependency. Generating traffic to all of the prefixes reachable by the protected tunnel (probably in a Round-Robin fashion, where the traffic is destined to all the prefixes but one prefix at a time in a cyclic manner) is not recommended. The reason why traffic generation is not recommended in a Round-Robin fashion to all the prefixes, one at a time is that if there are many prefixes reachable through the LSP the time interval between 2 packets destined to one prefix may be significantly high and may be comparable with the failover time being measured which does not aid in getting an accurate failover measurement.

5.8. Tester Capabilities

It is RECOMMENDED that the Tester used to execute each test case have the following capabilities:

1. Ability to establish MPLS-TE tunnels and push/pop labels.
2. Ability to produce Failover Event [TERM-ID].
3. Ability to insert a timestamp in each data packet's IP payload.
4. An internal time clock to control timestamping, time measurements, and time calculations.
5. Ability to disable or tune specific Layer-2 and Layer-3 protocol functions on any interface(s).
6. Ability to react upon the receipt of path error from the PLR

The Tester MAY be capable to make non-data plane convergence observations and use those observations for measurements.

6. Reference Test Setup

In addition to the general reference topology shown in figure 1, this section provides detailed insight into various proposed test setups that should be considered for comprehensively benchmarking the failover time in different roles along the primary tunnel
This section proposes a set of topologies that covers all the scenarios for local protection. All of these topologies can be mapped to the reference topology shown in Figure 1. Topologies provided in this section refer to the testbed required to benchmark failover time when the DUT is configured as a PLR in either Headend or midpoint role. Provided with each topology below is the label stack at the PLR. Penultimate Hop Popping (PHP) MAY be used and must be reported when used.
Figures 2 thru 9 use the following convention:

a) HE is Headend
b) TE is Tail-End
c) MID is Mid point
d) MP is Merge Point
e) PLR is Point of Local Repair
f) PRI is Primary Path
g) BKP denotes Backup Path and Nodes

6.1. Link Protection

6.1.1. Link Protection - 1 hop primary (from PLR) and 1 hop backup TE tunnels

Figure 2.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	0	0
Layer3 VPN (PE-PE)	1	1
Layer3 VPN (PE-P)	2	2
Layer2 VC (PE-PE)	1	1
Layer2 VC (PE-P)	2	2

Mid-point LSPs	0	0

6.1.2. Link Protection - 1 hop primary (from PLR) and 2 hop backup TE tunnels

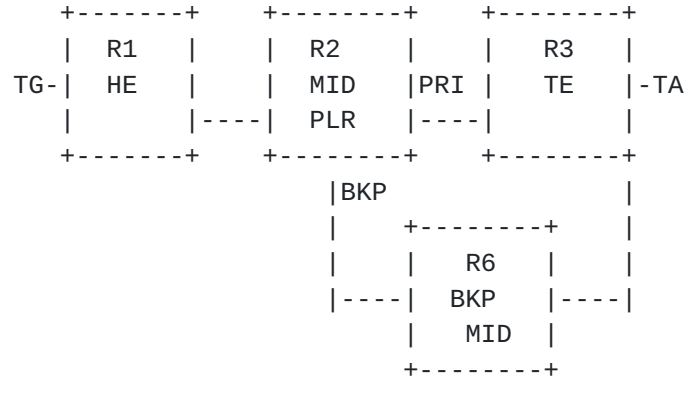

Figure 3.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	0	1
Layer3 VPN (PE-PE)	1	2
Layer3 VPN (PE-P)	2	3
Layer2 VC (PE-PE)	1	2
Layer2 VC (PE-P)	2	3
Mid-point LSPs	0	1

6.1.3. Link Protection - 2+ hop (from PLR) primary and 1 hop backup TE tunnels

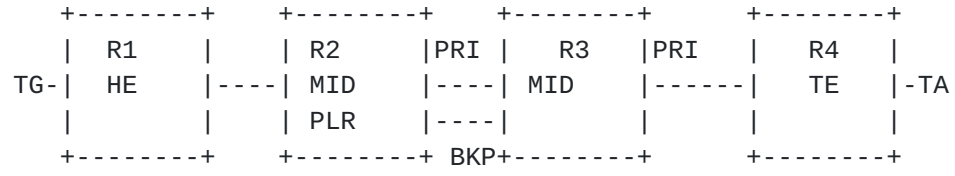

Figure 4.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	1
Layer3 VPN (PE-PE)	2	2
Layer3 VPN (PE-P)	3	3
Layer2 VC (PE-PE)	2	2
Layer2 VC (PE-P)	3	3
Mid-point LSPs	1	1

6.1.4. Link Protection - 2+ hop (from PLR) primary and 2 hop backup TE tunnels

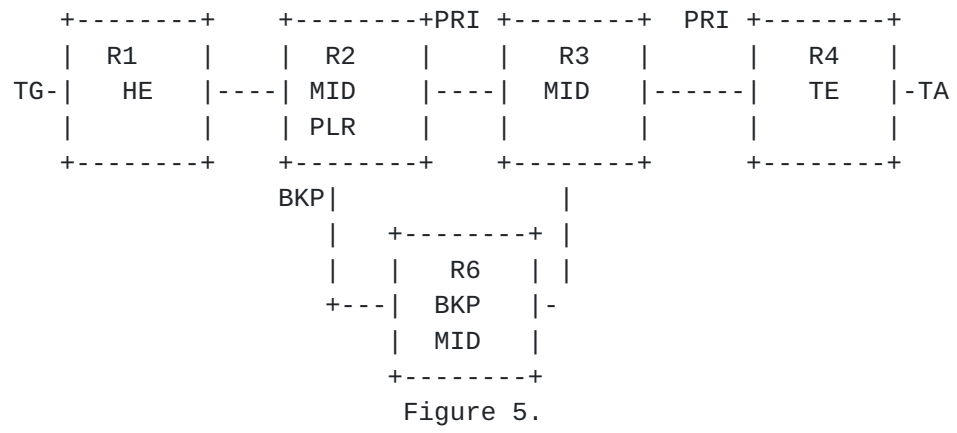

Figure 5.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	2
Layer3 VPN (PE-PE)	2	3
Layer3 VPN (PE-P)	3	4
Layer2 VC (PE-PE)	2	3
Layer2 VC (PE-P)	3	4
Mid-point LSPs	1	2

6.2. Node Protection

6.2.1. Node Protection - 2 hop primary (from PLR) and 1 hop backup TE tunnels

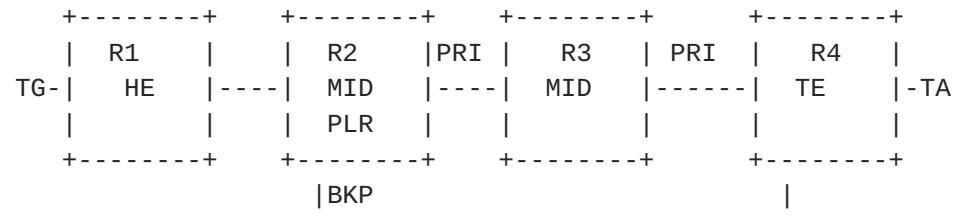

Figure 6.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	0
Layer3 VPN (PE-PE)	2	1
Layer3 VPN (PE-P)	3	2
Layer2 VC (PE-PE)	2	1
Layer2 VC (PE-P)	3	2
Mid-point LSPs	1	0

6.2.2. Node Protection - 2 hop primary (from PLR) and 2 hop backup TE tunnels

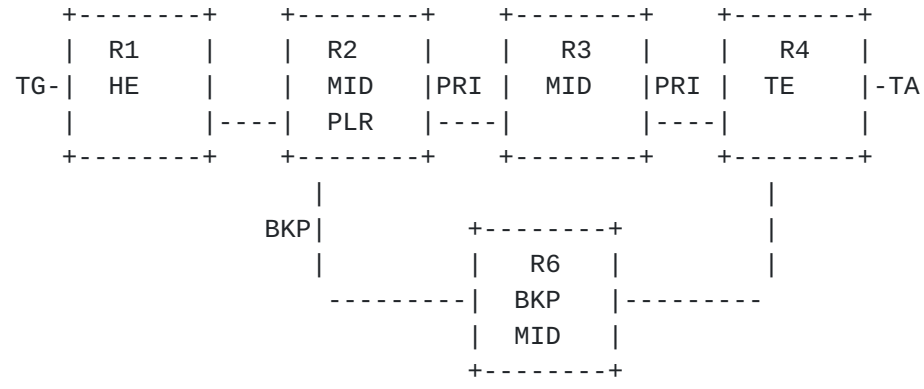

Figure 7.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	1
Layer3 VPN (PE-PE)	2	2
Layer3 VPN (PE-P)	3	3
Layer2 VC (PE-PE)	2	2
Layer2 VC (PE-P)	3	3
Mid-point LSPs	1	1

6.2.3. Node Protection - 3+ hop primary (from PLR) and 1 hop backup TE tunnels

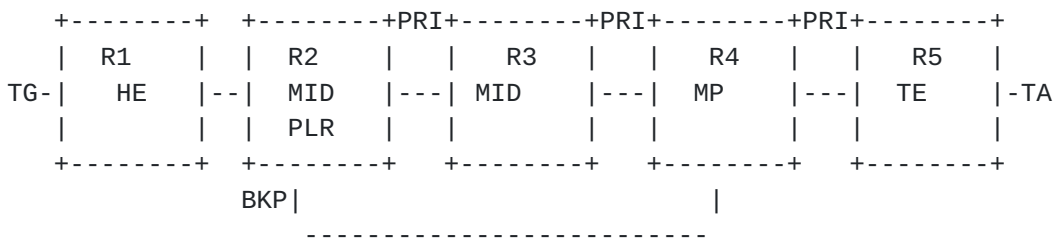

Figure 8.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	1
Layer3 VPN (PE-PE)	2	2
Layer3 VPN (PE-P)	3	3
Layer2 VC (PE-PE)	2	2
Layer2 VC (PE-P)	3	3
Mid-point LSPs	1	1

6.2.4. Node Protection - 3+ hop primary (from PLR) and 2 hop backup TE tunnels

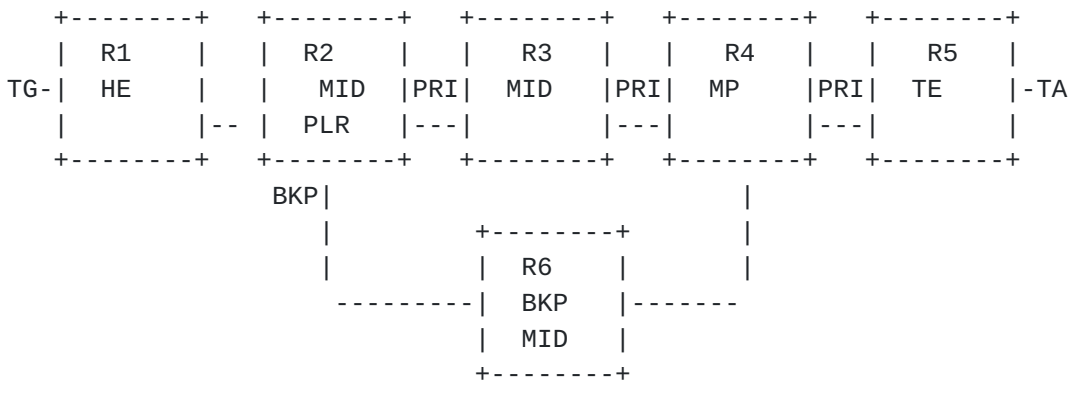

Figure 9.

Traffic	Num of Labels before failure	Num of labels after failure
IP TRAFFIC (P-P)	1	2
Layer3 VPN (PE-PE)	2	3
Layer3 VPN (PE-P)	3	4
Layer2 VC (PE-PE)	2	3
Layer2 VC (PE-P)	3	4
Mid-point LSPs	1	2

7. Test Methodology

The procedure described in this section can be applied to all the 8 base test cases and the associated topologies. The backup as well as the primary tunnels are configured to be alike in terms of bandwidth usage. In order to benchmark failover with all possible label stack depth applicable as seen with current deployments, it is RECOMMENDED to perform all of the test cases provided in this section. The forwarding performance test cases in section 7.1 MUST be performed prior to performing the failover test cases.

The considerations of Section 4 of [RFC2544] are applicable when evaluating the results obtained using these methodologies as well.

7.1. MPLS FRR Forwarding Performance

Benchmarking Failover Time [TERM-ID] for MPLS protection first requires baseline measurement of the forwarding performance of the test topology including the DUT. Forwarding performance is benchmarked by the Throughput as defined in [MPLS-FWD] and measured in units pps. This section provides two test cases to benchmark forwarding performance. These are with the DUT configured as a Headend PLR, Mid-Point PLR, and Egress PLR.

7.1.1. Headend PLR Forwarding Performance

Objective:

To benchmark the maximum rate (pps) on the PLR (as headend) over primary LSP and backup LSP.

Test Setup:

A. Select any one topology out of the 8 from section 6.
B. Select overlay technologies (e.g. IGP, VPN, or VC) with DUT as Headend PLR.
C. The DUT will also have 2 interfaces connected to the traffic Generator/analyzer. (If the node downstream of the PLR is not a simulated node, then the Ingress of the tunnel should have one link connected to the traffic generator and the node downstream to the PLR or the egress of the tunnel should have a link connected to the traffic analyzer).

Procedure:

1. Establish the primary LSP on R2 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection is enabled and ready.

5. Setup traffic streams as described in section 5.7.
6. Send MPLS traffic over the primary LSP at the Throughput supported by the DUT.
7. Record the Throughput over the primary LSP.
8. Trigger a link failure as described in section 5.1.
9. Verify that the offered load gets mapped to the backup tunnel and measure the Additive Backup Delay.
10. 30 seconds after Failover, stop the offered load and measure the Throughput, Packet Loss, Out-of-Order Packets, and Duplicate Packets over the Backup LSP.
11. Adjust the offered load and repeat steps 6 through 10 until the Throughput values for the primary and backup LSPs are equal.
12. Record the Throughput. This is the offered load that will be used for the Headend PLR failover test cases.

7.1.2. Mid-Point PLR Forwarding Performance

Objective:

To benchmark the maximum rate (pps) on the PLR (as mid-point) over primary LSP and backup LSP.

Test Setup:

A. Select any one topology out of the 8 from section 6.
B. Select overlay technologies (e.g. IGP, VPN, or VC) with DUT as Mid-Point PLR.
C. The DUT will also have 2 interfaces connected to the traffic generator.

Procedure:

1. Establish the primary LSP on R1 required by the topology selected.

2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection is enabled and ready.
5. Setup traffic streams as described in section 5.7.
6. Send MPLS traffic over the primary LSP at the Throughput supported by the DUT.
7. Record the Throughput over the primary LSP.
8. Trigger a link failure as described in section 5.1.
9. Verify that the offered load gets mapped to the backup tunnel and measure the Additive Backup Delay.
10. 30 seconds after Failover, stop the offered load and measure the Throughput, Packet Loss, Out-of-Order Packets, and Duplicate Packets over the Backup LSP.
11. Adjust the offered load and repeat steps 6 through 10 until the Throughput values for the primary and backup LSPs are equal.
12. Record the Throughput. This is the offered load that will be used for the Mid-Point PLR failover test cases.

7.1.3. Egress PLR Forwarding Performance

Objective:

To benchmark the maximum rate (pps) on the PLR (as egress) over primary LSP and backup LSP.

Test Setup:

A. Select any one topology out of the 8 from section 6.
B. Select overlay technologies (e.g. IGP, VPN, or VC) with DUT as Egress PLR.

C. The DUT will also have 2 interfaces connected to the traffic generator.

Procedure:

1. Establish the primary LSP on R1 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection is enabled and ready.
5. Setup traffic streams as described in section 5.7.
6. Send MPLS traffic over the primary LSP at the Throughput supported by the DUT.
7. Record the Throughput over the primary LSP.
8. Trigger a link failure as described in section 5.1.
9. Verify that the offered load gets mapped to the backup tunnel and measure the Additive Backup Delay.
10. 30 seconds after Failover, stop the offered load and measure the Throughput, Packet Loss, Out-of-Order Packets, and Duplicate Packets over the Backup LSP.
11. Adjust the offered load and repeat steps 6 through 10 until the Throughput values for the primary and backup LSPs are equal.
12. Record the Throughput. This is the offered load that will be used for the Egress PLR failover test cases.

7.2. Headend PLR with Link Failure

Objective:

To benchmark the MPLS failover time due to link failure events described in section 5.1 experienced by the DUT which is the Headend PLR.

Test Setup:

A. Select any one topology out of the 8 from section 6.
B. Select overlay technology for FRR test (e.g. IGP, VPN, or VC).
C. The (Headend PLR) DUT will also have 2 interfaces connected to Generator/analyzer. (If the node downstream of the PLR is not the traffic a simulated node, then the Ingress of the tunnel should have one link connected to the traffic generator and the node downstream to the PLR or the egress of the tunnel should have a link connected to the traffic analyzer).

Test Configuration:

1. Configure the number of primaries on R2 and the backups on R2 as required by the topology selected.
2. Configure the test setup to support Reversion.
3. Advertise prefixes (as per FRR Scalability Table described in Appendix A) by the tail end.

Procedure:

Test Case "7.1.1. Headend PLR Forwarding Performance" MUST be completed first to obtain the Throughput to use as the offered load.

1. Establish the primary LSP on R2 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection is enabled and ready.

5. Setup traffic streams for the offered load as described in section 5.7.
6. Provide the offered load from the tester at the Throughput [Br91] level obtained from test case 7.1.1.
7. Verify traffic is switched over Primary LSP without packet loss.
8. Trigger a link failure as described in section 5.1.
9. Verify that the offered load gets mapped to the backup tunnel and measure the Additive Backup Delay.
10. 30 seconds after Failover [TERM-ID], stop the offered load and measure the total Failover Packet Loss [TERM-ID].
11. Calculate the Failover Time [TERM-ID] benchmark using the selected Failover Time Calculation Method (TBLM, PLBM, or TBM) [TERM-ID].
12. Restart the offered load and restore the primary LSP to verify Reversion [TERM-ID] occurs and measure the Reversion Packet Loss [TERM-ID].
13. Calculate the Reversion Time [TERM-ID] benchmark using the selected Failover Time Calculation Method (TBLM, PLBM, or TBM) [TERM-ID].
14. Verify Headend signals new LSP and protection should be in place again.

IT is RECOMMENDED that this procedure be repeated for each of the link failure triggers defined in section 5.1.

7.3. Mid-Point PLR with Link Failure

Objective:

To benchmark the MPLS failover time due to link failure events described in section 5.1 experienced by the DUT which is the Mid-Point PLR.

Test Setup:

A. Select any one topology out of the 8 from section 6.
B. Select overlay technology for FRR test as Mid-Point LSPs.
C. The DUT will also have 2 interfaces connected to the traffic generator.

Test Configuration:

1. Configure the number of primaries on R1 and the backups on R2 as required by the topology selected.
2. Configure the test setup to support Reversion.
3. Advertise prefixes (as per FRR Scalability Table described in Appendix A) by the tail end.

Procedure:

Test Case "7.1.2. Mid-Point PLR Forwarding Performance" MUST be completed first to obtain the Throughput to use as the offered load.

1. Establish the primary LSP on R1 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Perform steps 3 through 14 from section 7.2 Headend PLR with Link Failure.

IT is RECOMMENDED that this procedure be repeated for each of the link failure triggers defined in section 5.1.

7.4. Headend PLR with Node Failure

Objective:

To benchmark the MPLS failover time due to Node failure events described in section 5.1 experienced by the DUT which is the Headend PLR.

Test Setup:

A. Select any one topology from section 6.
B. Select overlay technology for FRR test (e.g. IGP, VPN, or VC).
C. The DUT will also have 2 interfaces connected to the traffic generator/analyzer.

Test Configuration:

1. Configure the number of primaries on R2 and the backups on R2 as required by the topology selected.
2. Configure the test setup to support Reversion.
3. Advertise prefixes (as per FRR Scalability Table described in Appendix A) by the tail end.

Procedure:

Test Case "7.1.1. Headend PLR Forwarding Performance" MUST be completed first to obtain the Throughput to use as the offered load.

1. Establish the primary LSP on R2 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection.
5. Setup traffic streams for the offered load as described in section 5.7.
6. Provide the offered load from the tester at the Throughput [Br91] level obtained from test case 7.1.1.

7. Verify traffic is switched over Primary LSP without packet loss.
8. Trigger a node failure as described in section 5.1.
9. Perform steps 9 through 14 in 7.2 Headend PLR with Link Failure.

IT is RECOMMENDED that this procedure be repeated for each of the node failure triggers defined in section 5.1.

7.5. Mid-Point PLR with Node Failure

Objective:

To benchmark the MPLS failover time due to Node failure events described in section 5.1 experienced by the DUT which is the Mid-Point PLR.

Test Setup:

A. Select any one topology from section 6.1 to 6.2.
B. Select overlay technology for FRR test as Mid-Point LSPs.
C. The DUT will also have 2 interfaces connected to the traffic generator.

Test Configuration:

1. Configure the number of primaries on R1 and the backups on R2 as required by the topology selected.
2. Configure the test setup to support Reversion.
3. Advertise prefixes (as per FRR Scalability Table described in Appendix A) by the tail end.

Procedure:

Test Case "7.1.1. Mid-Point PLR Forwarding Performance" MUST be completed first to obtain the Throughput to use as the offered load.

1. Establish the primary LSP on R1 required by the topology selected.
2. Establish the backup LSP on R2 required by the selected topology.
3. Verify primary and backup LSPs are up and that primary is protected.
4. Verify Fast Reroute protection.
5. Setup traffic streams for the offered load as described in section 5.7.
6. Provide the offered load from the tester at the Throughput [Br91] level obtained from test case 7.1.1.
7. Verify traffic is switched over Primary LSP without packet loss.
8. Trigger a node failure as described in section 5.1.
9. Perform steps 9 through 14 in 7.2 Headend PLR with Link Failure.

IT is RECOMMENDED that this procedure be repeated for each of the node failure triggers defined in section 5.1.

8. Reporting Format

For each test, it is recommended that the results be reported in the following format.

Parameter	Units
IGP used for the test	ISIS-TE/ OSPF-TE
Interface types	Gige,POS,ATM,VLAN etc.
Packet Sizes offered to the DUT	Bytes (at layer 3)
Offered Load	packets per second

```
            IGP routes advertised              Number of IGP routes
            Penultimate Hop Popping            Used/Not Used
            RSVP hello timers                  Milliseconds
            Number of Protected tunnels        Number of tunnels
            Number of VPN routes installed     Number of VPN routes
            on the Headend
            Number of VC tunnels               Number of VC tunnels
            Number of mid-point tunnels        Number of tunnels
            Number of Prefixes protected by    Number of LSPs
            Primary
            Topology being used                Section number, and
                                               figure reference
            Failover Event                     Event type
              Re-optimization                        Yes/No
        Benchmarks (to be recorded for each test case):
        Failover-
            Failover Time                       seconds
            Failover Packet Loss                packets
            Additive Backup Delay               seconds
            Out-of-Order Packets                packets
            Duplicate Packets                   packets
            Failover Time Calculation Method    Method Used
        Reversion-
            Reversion Time                      seconds
            Reversion Packet Loss               packets
            Additive Backup Delay               seconds
            Out-of-Order Packets                packets
            Duplicate Packets                   packets
            Failover Time Calculation Method    Method Used
   Failover Time suggested above is calculated using one of the
   following three methods
```

1. Packet-Loss Based method (PLBM): (Number of packets dropped/ packets per second * 1000) milliseconds. This method could also be referred as Loss-Derived method.
2. Time-Based Loss Method (TBLM): This method relies on the ability of the Traffic generators to provide statistics which reveal the duration of failure in milliseconds based on when the packet loss occurred (interval between non-zero packet loss and zero loss).
3. Timestamp Based Method (TBM): This method of failover calculation is based on the timestamp that gets transmitted as payload in the packets originated by the generator. The Traffic Analyzer records the timestamp of the last packet received before the failover event and the first packet after the failover and derives the time based on the difference between these 2 timestamps. Note: The payload could also contain sequence numbers for out-of-order packet calculation and duplicate packets.

The timestamp based method method would be able to detect Reversion impairments beyond loss, thus it is RECOMMENDED method as a Failover Time method.

9. 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.

Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the DUT/SUT.

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.

10. IANA Considerations

This draft does not require any new allocations by IANA.

11. References

11.1. Informative References

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

11.2. Normative References

[I-D.ietf-bmwg-igp-dataplane-conv-term] Poretsky, S., Imhoff, B., and K. Michielsen, "Terminology for Benchmarking Link-State IGP Data Plane Route Convergence", draft-ietf-bmwg-igp-dataplane-conv-term-23 (work in progress), February 2011.

[I-D.ietf-bmwg-protection-term] Papneja, R., Poretsky, S., Vapiwala, S., and J. Karthik, "Benchmarking Terminology for Protection Performance", draft-ietf-bmwg-protection-term-08 (work in progress), December 2009.

[RFC1242] Bradner, S., "Benchmarking terminology for network interconnection devices", RFC 1242, July 1991.

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

[RFC2285] Mandeville, R., "Benchmarking Terminology for LAN Switching Devices", RFC 2285, February 1998.

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

[RFC4689] Poretsky, S., Perser, J., Erramilli, S., and S. Khurana, "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", RFC 4689, October 2006.

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

Appendix A. Acknowledgements

We would like to thank Jean Philip Vasseur for his invaluable input to the document and Curtis Villamizar his contribution in suggesting text on definition and need for benchmarking Correlated failures.

Additionally we would like to thank Al Morton, Arun Gandhi, Amrit Hanspal, Karu Ratnam, Raveesh Janardan, Andrey Kiselev, and Mohan Nanduri for their formal reviews of this document.

Appendix B. Fast Reroute Scalability Table

This section provides the recommended numbers for evaluating the scalability of fast reroute implementations. It also recommends the typical numbers for IGP/VPNv4 Prefixes, LSP Tunnels and VC entries. Based on the features supported by the device under test (DUT), appropriate scaling limits can be used for the test bed.

A1. FRR IGP Table

No. of Headend TE Tunnels	IGP Prefixes
1	100
1	500
1	1000
1	2000
1	5000
2 (Load Balance)	100
2 (Load Balance)	500
2 (Load Balance)	1000
2 (Load Balance)	2000
2 (Load Balance)	5000
100	100
500	500
1000	1000
2000	2000

A2. FRR VPN Table

No. of Headend TE Tunnels	VPNv4 Prefixes
1	100
1	500
1	1000
1	2000
1	5000
1	10000
1	20000
1	Max
2 (Load Balance)	100
2 (Load Balance)	500
2 (Load Balance)	1000
2 (Load Balance)	2000
2 (Load Balance)	5000
2 (Load Balance)	10000
2 (Load Balance)	20000
2 (Load Balance)	Max

A3. FRR Mid-Point LSP Table

No of Mid-point TE LSPs could be configured at recommended levels - 100, 500, 1000, 2000, or max supported number.

A2. FRR VC Table

No. of Headend TE Tunnels	VC entries
1	100
1	500
1	1000
1	2000
1	Max
100	100
500	500
1000	1000
2000	2000

Appendix C. Abbreviations

BFD - Bidirectional Fault Detection
BGP - Border Gateway protocol
CE - Customer Edge
DUT - Device Under Test
FRR - Fast Reroute
IGP - Interior Gateway Protocol
IP - Internet Protocol
LSP - Label Switched Path
MP - Merge Point
MPLS - Multi Protocol Label Switching
N-Nhop - Next - Next Hop
Nhop - Next Hop
OIR - Online Insertion and Removal
P - Provider
PE - Provider Edge
PHP - Penultimate Hop Popping
PLR - Point of Local Repair
RSVP - Resource reSerVation Protocol
SRLG - Shared Risk Link Group
TA - Traffic Analyzer
TE - Traffic Engineering
TG - Traffic Generator
VC - Virtual Circuit
VPN - Virtual Private Network

Authors' Addresses

Rajiv Papneja
Isocore
12359 Sunrise Valley Dr. STE100
Reston, VA 20191
USA
Email: rpapneja@isocore.com
Samir Vapiwala
Cisco Systems
300 Beaver Brook Road
Boxborough, MA 01719
USA
Email: svapiwal@cisco.com
Jay Karthik
Cisco Systems
300 Beaver Brook Road
Boxborough, MA 01719
USA
Email: jkarthik@cisco.com
Scott Poretsky
Allot Communications
USA
Email: sporetsky@allot.com
Shankar Rao
Qwest Communications
950 17th Street
Suite 1900
Denver, CO 80210
USA
Email: shankar.rao@qwest.com

Jean-Louis Le Roux
France Telecom
2 av Pierre Marzin
22300 Lannion
France
Email: jeanlouis.leroux@orange-ft.com