Network Working Group
Internet Draft
Intended status: Informational
Expires: August 2008

R. Papneja
Isocore
S. Vapiwala
J. Karthik
Cisco Systems
S. Poretsky
Reef Point
S. Rao
Qwest Communications
Jean-Louis Le Roux
France Telecom
February 19, 2008

Methodology for benchmarking MPLS Protection mechanisms

<draft-ietf-bmwg-protection-meth-03.txt>

Status of this Memo

Copyright Notice

Poretsky, Rao, Le Roux

Abstract

This draft describes the methodology for benchmarking MPLS Protection mechanisms for link and node protection as defined in [MPLS-FRR-EXT]. The benchmarking and terminology [TERM-ID] are to be used for benchmarking MPLS based protection mechanisms [MPLS-FRR-EXT]. This document provides test methodologies and test-bed setup for measuring failover times while considering all dependencies that might impact faster recovery of real time services riding on MPLS based primary tunnel. The terms used in the procedures included in this document are defined in [TERM-ID].

Table of Contents

1. Introduction..3
2. Existing definitions..6
3. Test Considerations...6
3.1. Failover Events...6
3.2. Failure Detection [TERM-ID]...................................7
3.3. Use of Data Traffic for MPLS Protection Benchmarking......8
3.4. LSP and Route Scaling...8
3.5. Selection of IGP..8
3.6. Reversion [TERM-ID]...9
3.7. Traffic generation..9
3.8. Motivation for topologies.....................................9
4. Test Setup...10
4.1. Link Protection with 1 hop primary (from PLR) and 1 hop backup...11
TE tunnels...11
4.2. Link Protection with 1 hop primary (from PLR) and 2 hop backup TE tunnels..11
4.3. Link Protection with 2+ hop (from PLR) primary and 1 hop backup TE tunnels..12
4.4. Link Protection with 2+ hop (from PLR) primary and 2 hop backup TE tunnels..12

4.5. Node Protection with 2 hop primary (from PLR) and 1 hop
backup TE tunnels...13

4.6. Node Protection with 2 hop primary (from PLR) and 2 hop backup TE tunnels..14
4.7. Node Protection with 3+ hop primary (from PLR) and 1 hop backup TE tunnels..15
4.8. Node Protection with 3+ hop primary (from PLR) and 2 hop backup TE tunnels..16
5. Test Methodology..16
5.1. Headend as PLR with link failure.........................16
5.2. Mid-Point as PLR with link failure.......................18
5.3. Headend as PLR with Node failure.........................19
5.4. Mid-Point as PLR with Node failure.......................20
5.5. MPLS FRR Forwarding Performance Test Cases...............22
5.5.1. PLR as Headend.....................................22
5.5.2. PLR as Mid-point...................................23
6. Reporting Format..24
7. IANA Considerations...25
This document requires no IANA considerations..................25
8. Security Considerations.....................................25
9. Acknowledgements..26
10. References...26
10.1. Normative References.....................................26
10.2. Informative References...................................27
11. Authors' Addresses...27
Intellectual Property Statement................................29
Appendix A: Fast Reroute Scalability Table.....................30

1. Introduction

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

MPLS based protection mechanisms provide faster recovery of real time services in case of an unplanned link or node failure in the network core, where MPLS is used as a signaling protocol to setup point-to-point traffic engineered tunnels. MPLS based protection mechanisms improve service availability by minimizing the duration of the most common failures. There are generally two factors impacting service availability. One is the frequency and the other is the duration of the failure. Unexpected correlated failures are less common. Correlated failures mean co-occurrence of two or more failures simultaneously.

These failures are often observed when two or more logical resources (for e.g. layer-2 links), relying on a common physical resource (for e.g. common transport) fail. Common transport may include TDM and WDM links providing multiplexing at layer-2 and layer-1. Within the context of MPLS protection mechanisms, Shared Risk Link Groups [MPLS-FRR-EXT] encompass correlations failures.

Not all correlated failures can be anticipated in advance of their occurrence. Failures due to natural disasters or planned failures are the most notable causes. Due to the frequent occurrences of such failures, it is necessary that implementations can handle these faults gracefully, and recover the services affected by failures very quickly.

Some routers recover faster as compared to the others, hence benchmarking this type of failures become very useful. Benchmarking of unexpected correlated failures should include measurement of restoration with and without the availability of IP fallback. This document provides detailed test cases focusing on benchmarking MPLS protection mechanisms. Benchmarking of unexpected correlated failures is currently out of scope of this document.

A link or a node failure could occur either at the head-end or at the mid point node of a primary tunnel. The backup tunnel could offer either link or node protection following a failure along the path of the primary tunnel. The time lapsed in transitioning primary tunnel traffic to the backup tunnel is a key measurement that ensures the service level agreements. Failover time depends upon many factors such as the number of prefixes bound to a tunnel, services (such as IGP, BGP, Layer 3/ Layer 2 VPNs) that are bound to the tunnel, number of primary tunnels affected by the failure event, number of primary tunnels protected by backup, the type of failure and the physical media on which the failover occurs. This document describes all different topologies and scenarios that should be considered to effectively benchmark MPLS protection mechanisms and failover times. Different failure scenarios and scaling considerations are also provided in this document. In addition the document provides a reporting format for the observed results.

To benchmark the failover time, data plane traffic is used as defined in [IGP-METH]. Traffic loss is the key component in a black-box type test and is used to measure convergence.

All benchmarking test cases 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 ability of the DUT to perform recovery from failures within target failover time.

Figure 1 represents the basic reference test bed and is applicable to all the test cases defined in this document. TG & TA represents Traffic Generator & Analyzer respectively. A tester is connected to the DUT and it sends and receives IP traffic along with the working Path, run protocol emulations simulating real world peering scenarios.

```
              ---------------------------
            |               ------------|---------------
            |              |            |               |
            |              |            |               |
        --------       --------      --------      --------     --------
    TG-|   R1   |-----|   R2   |----|   R3   |    |    R4  |   |  R5    |-
TA
       |        |-----|        |----|        |----|        |---|        |
        --------       --------      --------      --------     --------
              |            |              |           |
              |            |              |           |
              |          --------         |           |
               ---------|   R6   |--------            |
                        |        |--------------------
                         --------
```

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.

2. Existing definitions

For the sake of clarity and continuity this RFC adopts the template for definitions set out in Section 2 of RFC 1242. Definitions are indexed and grouped together in sections for ease of reference.

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

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

3. Test Considerations

This section discusses the fundamentals of MPLS Protection testing:

-The types of network events that causes failover
-Indications for failover
-the use of data traffic
-Traffic generation
-LSP Scaling
-Reversion of LSP
-IGP Selection

3.1. Failover Events

The failover to the backup tunnel is primarily triggered by either a 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
- Interface Shutdown on remote side with RSVP hello
- 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 is initiated either by a graceful shutdown or by a power failure. A system crash is referred to as a software failure or an assert.

- Reload protected Node, when RSVP Hello is enabled
- Crash Protected Node, when RSVP Hello is enable
- Reload Protected Node, when BFD is enable
- Crash Protected Node, when BFD is enable

3.2. Failure Detection [TERM-ID]

Local failures can be detected via SONET/SDH failure with directly connected LSR. Failure indication may vary with the type of alarm - LOS, AIS, or RDI. Failures on Ethernet links such as Gigabit Ethernet rely upon Layer 3 signaling indication for failure.

Different MPLS protection mechanisms and different implementations use different failure detection techniques such as RSVP hellos, BFD etc. Ethernet technologies such as Gigabit Ethernet rely upon layer 3 failure indication mechanisms since there is no Layer 2 failure indication mechanism. The failure detection time may not always be negligible and it could impact the overall failover 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.

3.3. Use of Data Traffic for MPLS Protection Benchmarking

Currently end customers use packet loss as a key metric for failover time. Packet loss is an externally observable event and has direct impact on customers' applications. MPLS protection mechanism 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 tests environment. Data traffic is offered at line-rate to the device under test (DUT), and 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. In scenarios, where separate measurement of packets in error and duplicate packets is difficult to obtain, these packets should be attributed to lost packets.

3.4. LSP and Route Scaling

Failover time performance may vary with the number of established primary and backup tunnels (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). Number of L and R must be recorded.

3.5. Selection of IGP

The underlying IGP could be ISIS-TE or OSPF-TE for the methodology proposed here.

3.6. Reversion [TERM-ID]

Fast Reroute provides a method to return or restore a backup path to original primary LSP upon recovery from the failure. This is referred to as Reversion, which can be implemented as Global Reversion or Local Reversion. In all test cases listed here Reversion should not produce any packet loss, out of order or duplicate packets. Each of the test cases in this methodology document provides a check to confirm that there is no packet loss.

3.7. Traffic generation

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.

A typical example would be configuring the traffic generator to send the traffic to the first, middle and last of the advertised routes. (First, middle and last could be decided by the numerically smallest, median and the largest respectively of the advertised prefix). 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.

3.8. Motivation for topologies

Given that the label stack is dependent of the following 3 entities it is recommended that the benchmarking of failover time be performed on all the 8 topologies provided in section 4

- Type of protection (Link Vs Node)

- # of remaining hops of the primary tunnel from the PLR

- # of remaining hops of the backup tunnel from the PLR

4. Test Setup

Topologies to be used for benchmarking the failover time:

This section proposes a set of topologies that covers all the scenarios for local protection. All of these 8 topologies shown (figure 2- figure 9) can be mapped to the reference topology shown in figure 1. Topologies provided in sections 4.1 to 4.8 refer to test-bed required to benchmark failover time when DUT is configured as a PLR in either head-end or midpoint role. The labels stack provided with each topology is at the PLR.

The label stacks shown below each figure in section 4.1 to 4.9 considers enabling of Penultimate Hop Popping (PHP).

Figures 2-9 uses the following convention:

a) HE is Head-End

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

g) BKP denotes Backup Node

4.1. Link Protection with 1 hop primary (from PLR) and 1 hop backup TE tunnels

```
         -------    -------- PRI  --------
        |  R1   |  |   R2   |    |   R3   |
     TG-|  HE   |--|  MID   |----|    TE  |-TA
        |       |  |  PLR   |----|        |
         -------    -------- BKP  --------
```

Figure 2: Represents the setup for section 4.1

Traffic	No of Labels before failure	No 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

4.2. Link Protection with 1 hop primary (from PLR) and 2 hop backup TE tunnels

```
         -------      --------      --------
        |  R1   |    |  R2    |    |   R3   |
     TG-|  HE   |    |  MID   |PRI |   TE   |-TA
        |       |----|  PLR   |----|        |
         -------      --------      --------
                        |BKP                |
                        |     --------      |
                        |    |   R6   |     |
                        |----|  BKP   |----|
                             |   MID  |
                              --------
```

Figure 3: Representing setup for section 4.2

Traffic	No of Labels before failure	No 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

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

```
        --------      --------      --------        --------
       |  R1    |    | R2     |PRI |   R3   |PRI    |   R4   |
   TG-|  HE    |----| MID    |----| MID    |------|   TE   |-TA
       |        |    | PLR    |----|        |      |        |
        --------      -------- BKP  --------        --------
```

Figure 4: Representing setup for section 4.3

Traffic	No of Labels before failure	No 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

4.4. Link Protection with 2+ hop (from PLR) primary and 2 hop backup TE tunnels

```
        --------      -------- PRI  --------  PRI   --------
       |  R1    |    |  R2    |    |   R3   |      |   R4   |
```

```
       TG-|   HE   |----| MID    |----|  MID   |------|   TE   |-TA
          |        |    | PLR    |    |        |      |        |
           --------      --------      --------         --------
                        BKP|              |
                           |    --------  |
                           |   |   R6   | |
                            ---|  BKP   |-
                               |  MID   |
                                --------
```

Figure 5: Representing the setup for section 4.4

Traffic	No of Labels before failure	No 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

4.5. Node Protection with 2 hop primary (from PLR) and 1 hop backup TE tunnels

```
           --------      --------      --------         --------
          |  R1    |    |  R2    |PRI |   R3   | PRI  |   R4   |
       TG-|   HE   |----|  MID   |----|  MID   |------|  TE    |-TA
          |        |    |  PLR   |    |        |      |        |
           --------      --------      --------         --------
                          |BKP                          |
                           -----------------------------
```

Figure 6: Representing the setup for section 4.5

Traffic	No of Labels before failure	No 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

4.6. Node Protection with 2 hop primary (from PLR) and 2 hop backup TE tunnels

```
          --------      --------      --------      --------
         |  R1    |    |  R2    |    |   R3   |    |   R4   |
      TG-|  HE    |    |  MID   |PRI |  MID   |PRI |  TE    |-TA
         |        |----|  PLR   |----|        |----|        |
          --------      --------      --------      --------
                          |                            |
                       BKP|          --------          |
                          |         |   R6   |         |
                           ---------|  BKP   |---------
                                    |  MID   |
                                     --------
```

Figure 7: Representing setup for section 4.6

Traffic	No of Labels before failure	No 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

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

```
      --------    -------- PRI -------- PRI -------- PRI --------
     |  R1    |  |  R2    |    |   R3   |    |   R4   |    |   R5   |
  TG-|   HE   |--|  MID   |---| MID    |---|  MP    |---|  TE    |-TA
     |        |  |  PLR   |    |        |    |        |    |        |
      --------    --------     --------     --------     --------
                BKP|                           |
                    --------------------------
```

Figure 8: Representing setup for section 4.7

Traffic	No of Labels before failure	No 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

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

```
       --------     --------     --------     --------     --------
      |  R1     |   |  R2     |   |  R3     |   |  R4     |   |  R5     |
   TG-|  HE     |   |   MID   |PRI|  MID    |PRI|  MP     |PRI|  TE     |-TA
      |         |-- |  PLR    |---|         |---|         |---|         |
       --------     --------     --------     --------     --------
                  BKP|                           |
                     |          --------         |
                     |         |  R6     |       |
                      ---------|  BKP    |-------
                               |  MID    |
                                --------
```

Figure 9: Representing setup for section 4.8

Traffic	No of Labels before failure	No 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

5. 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 tunnel 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 suggested that the methodology includes all the scenarios listed here

5.1. Headend as PLR with link failure

Objective

To benchmark the MPLS failover time due to Link failure events described in section 3.1 experienced by the DUT which is the point of local repair (PLR).

Test Setup

- select any one topology out of 8 from section 4
- select overlay technology for FRR test e.g. IGP,VPN,or VC
- 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).

Test Configuration

1. Configure the number of primaries on R2 and the backups on R2 as required by the topology selected.
2. Advertise prefixes (as per FRR Scalability table describe in Appendix A) by the tail end.

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 3.7
6. Send IP traffic at maximum Forwarding Rate to DUT.
7. Verify traffic switched over Primary LSP.
8. Trigger any choice of Link failure as describe in section 3.1
9. Verify that primary tunnel and prefixes gets mapped to backup tunnels
10. Stop traffic stream and measure the traffic loss.

11. Failover time is calculated as defined in section 6, Reporting format.
12. Start traffic stream again to verify reversion when protected interface comes up. Traffic loss should be 0 due to make before break or reversion.
13. Enable protected interface that was down (Node in the case of NNHOP)
14. Verify head-end signals new LSP and protection should be in place again

5.2. Mid-Point as PLR with link failure

Objective

To benchmark the MPLS failover time due to Link failure events described in section 3.1 experienced by the device under test which is the point of local repair (PLR).

Test Setup

- select any one topology out of 8 from section 4
- select overlay technology for FRR test as Mid-Point lsps
- 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. Advertise prefixes (as per FRR Scalability table describe in Appendix A) by the tail end.

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
5. Setup traffic streams as described in section 3.7
6. Send IP traffic at maximum Forwarding Rate to DUT.
7. Verify traffic switched over Primary LSP.
8. Trigger any choice of Link failure as describe in section 3.1
9. Verify that primary tunnel and prefixes gets mapped to backup tunnels
10. Stop traffic stream and measure the traffic loss.
11. Failover time is calculated as per defined in section 6, Reporting format.
12. Start traffic stream again to verify reversion when protected interface comes up. Traffic loss should be 0 due to make before break or reversion
13. Enable protected interface that was down (Node in the case of NNHOP)
14. Verify head-end signals new LSP and protection should be in place again

5.3. Headend as PLR with Node failure

Objective

To benchmark the MPLS failover time due to Node failure events described in section 3.1 experienced by the device under test which is the point of local repair (PLR).

Test Setup

- select any one topology from section 4.5 to 4.8
- select overlay technology for FRR test e.g. IGP,VPN,or VC
- The DUT will also have 2 interfaces connected to the traffic generator.

Test Configuration

1. Configure the number of primaries on R2 and the backups on R2 as required by the topology selected
2. Advertise prefixes (as per FRR Scalability table describe in Appendix A) by the tail end.

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
5. Setup traffic streams as described in section 3.7
6. Send IP traffic at maximum Forwarding Rate to DUT.
7. Verify traffic switched over Primary LSP.
8. Trigger any choice of Node failure as describe in section 3.1
9. Verify that primary tunnel and prefixes gets mapped to backup tunnels
10. Stop traffic stream and measure the traffic loss.
11. Failover time is calculated as per defined in section 6, Reporting format.
12. Start traffic stream again to verify reversion when protected interface comes up. Traffic loss should be 0 due to make before break or reversion
13. Boot protected Node that was down.
14. Verify head-end signals new LSP and protection should be in place again

5.4. Mid-Point as PLR with Node failure

Objective

To benchmark the MPLS failover time due to Node failure events described in section 3.1 experienced by the device under test which is the point of local repair (PLR).

Test Setup

- select any one topology from section 4.5 to 4.8
- select overlay technology for FRR test as Mid-Point lsps
- 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. Advertise prefixes (as per FRR Scalability table describe in Appendix A) by the tail end.

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
5. Setup traffic streams as described in section 3.7
6. Send IP traffic at maximum Forwarding Rate to DUT.
7. Verify traffic switched over Primary LSP.
8. Trigger any choice of Node failure as describe in section 3.1
9. Verify that primary tunnel and prefixes gets mapped to backup tunnels
10. Stop traffic stream and measure the traffic loss.
11. Failover time is calculated as per defined in section 6, Reporting format.
12. Start traffic stream again to verify reversion when protected interface comes up. Traffic loss should be 0 due to make before break or reversion
13. Boot protected Node that was down

14. Verify head-end signals new LSP and protection should be in place again

5.5. MPLS FRR Forwarding Performance Test Cases

For the following MPLS FRR Forwarding Performance Benchmarking cases, Test the maximum PPS rate allowed by given hardware

5.5.1. PLR as Headend

Objective

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

Test Setup

- select any one topology out of 8 from section 4
- select overlay technology for FRR test e.g. IGP,VPN,or VC
- 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 3.7
6. Send IP traffic at maximum forwarding rate (pps) that the device under test supports over the primary LSP
7. Record maximum PPS rate forwarded over primary LSP

8. Stop traffic stream
9. Trigger any choice of Link failure as describe in section 3.1
10. Verify that primary tunnel and prefixes gets mapped to backup tunnels
11. Send IP traffic at maximum forwarding rate (pps) that the device under test supports over the primary LSP
12. Record maximum PPS rate forwarded over backup LSP

5.5.2. PLR as Mid-point

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

Test Setup

- select any one topology out of 8 from section 4
- select overlay technology for FRR test as Mid-Point lsps
- 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 3.7
6. Send IP traffic at maximum forwarding rate (pps) that the device under test supports over the primary LSP
7. Record maximum PPS rate forwarded over primary LSP
8. Stop traffic stream

9. Trigger any choice of Link failure as describe in section 3.1
10. Verify that primary tunnel and prefixes gets mapped to backup tunnels
11. Send IP traffic at maximum forwarding rate (pps) that the device under test supports over the backup LSP
12. Record maximum PPS rate forwarded over backup LSP

6. 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
Forwarding rate	number of packets
IGP routes advertised	number of IGP routes
RSVP hello timers configured (if any)	milliseconds
Number of FRR tunnels configured	number of tunnels
Number of VPN routes in head-end	number of VPN routes
Number of VC tunnels	number of VC tunnels
Number of BGP routes	number of BGP routes
Number of mid-point tunnels	number of tunnels
Number of Prefixes protected by Primary	number of prefixes
Number of LSPs being protected	number of LSPs
Topology being used	Section number
Failure Event	Event type

Benchmarks	
Minimum failover time	milliseconds
Mean failover time	milliseconds
Maximum failover time	milliseconds
Minimum reversion time	milliseconds

Mean reversion time	milliseconds
Maximum reversion time	milliseconds

Failover time suggested above is calculated using one of the following 3 methods

1. Packet-Based Loss method (PBLM): (Number of packets dropped/packets per second * 1000) milliseconds. This method could also be referred as Rate 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.

Note: If the primary is configured to be dynamic, and if the primary is to reroute, make before break should occur from the backup that is in use to a new alternate primary. If there is any packet loss seen, it should be added to failover time.

7. IANA Considerations

This document requires no IANA considerations.

8. 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.

The isolated nature of the benchmarking environments and the fact that no special features or capabilities, other than those used in operational networks, are enabled on the DUT/SUT requires no security considerations specific to the benchmarking process.

9. 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 Arun Gandhi, Amrit Hanspal, Karu Ratnam and for their input to the document.

10. References

10.1. Normative References

[MPLS-FRR-EXT]   Pan, P., Atlas, A., Swallow, G., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090.

10.2. Informative References

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

[TERM-ID] Poretsky S., Papneja R., Karthik J., Vapiwala S., "Benchmarking Terminology for Protection Performance", draft-ietf-bmwg-protection-term-02.txt, work in progress.

[MPLS-FRR-EXT] Pan P., Swollow G., Atlas A., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090.

[IGP-METH] S. Poretsky, B. Imhoff, "Benchmarking Methodology for IGP Data Plane Route Convergence, "draft-ietf-bmwg-igp-dataplane-conv-meth-12.txt", work in progress.

11. Authors' Addresses

Rajiv Papneja
Isocore
12359 Sunrise Valley Drive, STE 100
Reston, VA 20190
USA
Phone: +1 703 860 9273
Email: rpapneja@isocore.com

Samir Vapiwala
Cisco System
300 Beaver Brook Road
Boxborough, MA 01719
USA
Phone: +1 978 936 1484
Email: svapiwal@cisco.com

Jay Karthik
Cisco Systems,
300 Beaver Brook Road
Boxborough, MA 01719
USA
Phone: + 1 978 936 0533
Email: jkarthik@cisco.com

Scott Poretsky
Reef Point Systems
8 New England Executive Park
Burlington, MA 01803
USA
Phone: + 1 781 395 5090
EMail: sporetsky@reefpoint.com

Shankar Rao
Qwest Communications,
950 17th Street
Suite 1900
Denver, CO 80210
USA
Phone: + 1 303 437 6643
Email: shankar.rao@qwest.com

Jean-Louis Le Roux
France Telecom
2 av Pierre Marzin
22300 Lannion
France
Phone: 00 33 2 96 05 30 20
Email: jeanlouis.leroux@orange-ft.com

Acknowledgment

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

Appendix A: 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, appropriate scaling limits can be used for the test bed.

A 1. FRR IGP Table

No of Headend TE LSPs	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

A 2. FRR VPN Table

No of Headend TE LSPs	VPNv4 Prefixes

Poretsky, Rao, Le Roux

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

A 3. FRR Mid-Point LSP Table

No of Mid-point TE LSPs could be configured at the following recommended levels

100
500
1000
2000
Max supported number

A 4. FRR VC Table

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

100	100
500	500
1000	1000
2000	2000

Appendix B: 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