Network Working Group
Internet Draft
Expires: April 2007

S. Poretsky
Reef Point Systems

R. Papneja
Isocore

J. Karthik
Cisco Systems

October 2006

Benchmarking Terminology for Protection Performance
<draft-ietf-bmwg-protection-term-00.txt >

Intellectual Property Rights (IPR) statement:
By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Status of this Memo

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Copyright Notice
Copyright (C) The Internet Society (2006).

Abstract

This document provides common terminology and metrics for benchmarking the performance of sub-IP layer protection mechanisms. The performance benchmarks are measured at the IP-Layer, so avoid dependence on specific sub-IP protections mechanisms. The benchmarks and terminology can be applied in methodology documents for different sub-IP layer protection mechanisms such as Automatic Protection Switching (APS), Virtual Router Redundancy Protocol (VRRP), and Multi-Protocol Label Switching Fast Reroute (MPLS-FRR).

Table of Contents

1. Introduction

The IP network layer provides route convergence to protect data traffic against planned and unplanned failures in the internet. Fast convergence times are critical to maintain reliable network connectivity and performance. Technologies that function at sub-IP layers can be enabled to provide further protection of IP traffic by providing the failure recovery at the sub-IP layers so that the outage is not observed at the IP-layer. Such technologies include High Availability (HA) stateful failover. Virtual Router Redundancy Protocol (VRRP), Automatic Link Protection (APS) for SONET/SDH, Resilient Packet Ring (RPR) for Ethernet, and Fast Reroute for Multi-Protocol Label Switching (MPLS).

Benchmarking terminology and methodology have been defined for IP-layer route convergence [7,8,9]. New terminology and methodologies specific to benchmarking sub-IP layer protection mechanisms are required. This will enable different implementations of the same protection mechanisms to be benchmarked and evaluated. In addition, different protection mechanisms can be benchmarked and evaluated. The metrics for benchmarking the performance of sub-IP protection mechanisms are measured at the IP layer, so that the results are always measured in reference to IP and independent of

the specific protection mechanism being used. The purpose of this document is to provide a single terminology for benchmarking sub-IP protection mechanisms. It is intended that there can exist unique methodology documents for each sub-IP protection mechanism.

Figure 1 shows the fundamental model that is to be used in benchmarking sub-IP protection mechanisms. Protection Switching consists of a minimum of two Protection-Switching Nodes with a Primary Path and a Backup Path. A Failover Event occurs along the Primary Path. A tester is set outside the two nodes as it sends and receives IP traffic along the Working Path. The Working Path is the Primary Path prior to the Failover Event and the Backup Path following the Failover Event. If Reversion is supported then the

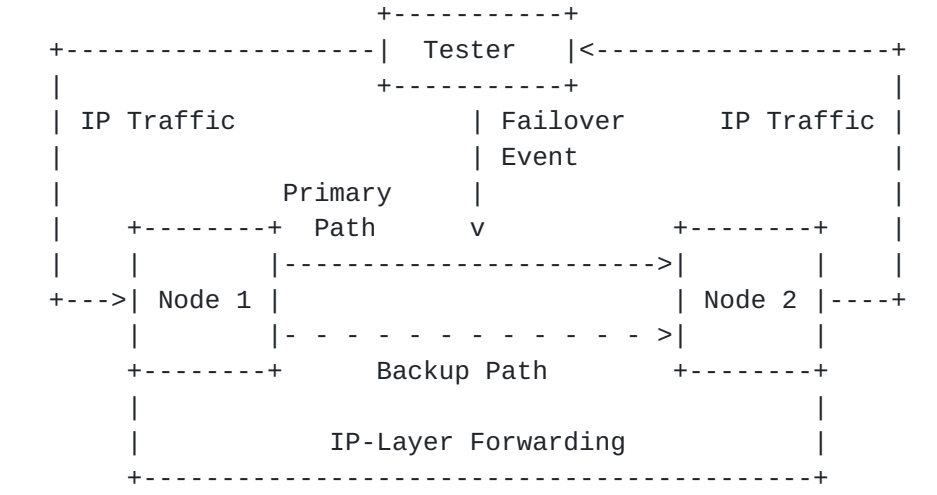

Figure 1. System Under Test (SUT) for Sub-IP Protection Mechanisms

Working Path is the Primary Path after Failure Recovery. The tester MUST record the IP packet sequence numbers, departure time, and arrival time so that the metrics of Failover Time, Additive Latency, and Reversion Time can be measured. The Tester may be a single device or a test system.

2. Existing definitions

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

Latency	[RFC 1242, section 3.8]
Frame Loss Rate	[RFC 1242, section 3.6]
Throughput	[RFC 1242, section 3.17]
Device Under Test (DUT)	[RFC 2285, section 3.1.1]
System Under Test (SUT)	[RFC 2285, section 3.1.2]
Out-of-order Packet	[Ref.[4], section 3.3.2]
Duplicate Packet	[Ref.[4], section 3.3.3]

This document adopts the definition format in Section 2 of RFC 1242.

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. 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.

3. Test Considerations

3.1. Path

3.1.1 Path

Definition:
A sequence of nodes, <R1, ..., Rn>, with the following properties:
- R1 is the ingress node and forwards IP packets, which input into DUT/SUT, to R2 as sub-IP frames.
- Ri is a node which forwards data frames to R[i+1] for all i, 1<i<n, based on information in the sub-IP layer.
- Rn is the egress node and it outputs sub-IP frames from DUT/SUT as IP packets.

Discussion:
The path is defined in the sub-IP layer in this document, unlike an IP path in RFC 2026. For example, the SONET/SDH path, the label switched path for MPLS, and optical path. One path may be regarded as being equivalent to one IP link between two IP nodes, i.e., R1 and Rn. The two IP nodes may have multiple paths for protection. A packet will travel on only one path between the nodes. Packets belonging to a micro flow (RFC 2474) will transverse one or more paths. The path is unidirectional.

Measurement units:
n/a

Issues:
"A bidirectional path", which transmits traffic in both directions along the same nodes, consists of two unidirectional paths. Therefore, the two unidirectional paths belonging to "one bidirectional path" will be treated independently when benchmarking for "a bidirectional path".

See Also:

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. Path

3.1.1. Path

Definition:
A sequence of nodes, <R1, ..., Rn>, with the following properties:
- R1 is the ingress node and forwards IP packets, which input into DUT/SUT, to R2 as sub-IP frames.
- Ri is a node which forwards data frames to R[i+1] for all i, 1<i<n, based on information in the sub-IP layer.
- Rn is the egress node and it outputs sub-IP frames from DUT/SUT as IP packets.

Discussion:
The path is defined in the sub-IP layer in this document, unlike an IP path in RFC 2026. For example, the SONET/SDH path, the label switched path for MPLS, and optical path. One path may be regarded as being equivalent to one IP link between two IP nodes, i.e., R1 and Rn. The two IP nodes may have multiple paths for protection. A packet will travel on only one path between the nodes. Packets belonging to a microflow (RFC 2474) will transverse one or more paths. The path is unidirectional.

Measurement units:
n/a

3.1.2. Tunnel

Definition:
Tunnel is a collection of related Paths.

Discussion:
A tunnel is used to carry a specific flow of traffic which is generally large aggregation of microflows, but may be any flow defined by a classifier at the ingress. A Tunnel may include two primary paths during the MPLS make-before-break reroute.

Measurement units:
n/a

Issues:

See Also:
Path
Primary Path
Backup Path

3.1.3. Working Path

Definition:
The path that the DUT/SUT is currently using to forward packets.

Discussion:
A Primary Path is a Working Path before occurrence of a Failover Event. A Backup Path becomes the Working Path after a Failover Event.

Measurement units:
n/a

Issues:

See Also:
Path
Primary Path
Backup Path

3.1.4. Primary Path

Definition:
The preferred path for forwarding traffic between two or more nodes.

Discussion:

Measurement units:
n/a

Issues:

See Also:
Path

3.1.5. Protected Primary Path

Definition:
The Primary Path that is protected with a Backup Path.

Discussion:

Measurement units:
n/a

Issues:

See Also:
Path
Primary Path

3.1.6. Backup Path

Definition:
A path that exists to carry data traffic only if a Failover Event occurs.

Discussion:
The Backup Path is the Working Path upon a Failover Event. There are various types of Backup Paths: a dedicated recovery path (1+1), which has 100% redundancy for a specific ordinary path, a shared Backup Path (1:N), which is dedicated to the protection for more than one specific Primary Path, and an associated shared Backup Path (M:N) for which a specific set of Backup Paths protects a specific set of more than one Primary Path. Backup path is always computed before the failover event. A new path computed after the failover event is simply reroute of the primary path. A backup may be signaled or unsignalled.

Measurement units:
n/a

Issues:

See Also:
Path
Working Path
Primary Path

3.1.7. Standby Backup Path

Definition:
A Backup Path that is established prior to a Failover Event to protect a Primary Path.

Discussion:

Measurement units:
n/a

Issues:

See Also:
Path
Working Path
Primary Path
Failover Event

3.1.8. Dynamic Backup Path

Definition:
A Backup Path that is established upon occurrence of a Failover Event.

Discussion:

Measurement units:
n/a

Issues:

See Also:
Path
Working Path
Primary Path
Failover Event

3.1.9. Disjoint Paths

Definition:
A pair of paths are considered disjoint if they do not share a common link.

Discussions:

Paths that protect a segment of a path may merge beyond the segment being protected and are considered disjoint if they do not use a link from the set of links in the protected segment. A path is node disjoint if it does not share a common node other than the ingress and egress.

Measurement units:
n/a

Issues:

See Also:
Path
Primary Path
SRLG

3.1.10. Shared Risk Link Group (SRLG)

Definition:
SRLG is a set of links which are likely to fail concurrently due to sharing a physical resource.

Discussion:
SRLG are considered the set of links to be avoided when the primary and secondary paths are considered disjoint.

Measurement units:
n/a

Issues:

See Also:
Path
Primary Path
Disjoint Path

3.2. Protection

3.2.1. Protection Switching System

Definition:
A SUT that is capable of Failover from a Primary to a Backup Path.

Discussion:
The Protection Switching System MUST have a Primary Path and a Backup Path. The Backup Path MAY be a Standby Backup Path or a dynamic Backup Path. The Protection Switching System includes the mechanisms for both Failure Detection and Failover.

Measurement units:

Issues:

See Also:
Primary Path
Backup Path
Failure Detection
Failover

3.2.2. Link Protection

Definition:
A Backup Path that provides protection for link failure.

Discussion:
Link Protection may or may not protect the entire Primary Path.

Measurement units: n/a

Issues:

See Also:
Primary Path
Backup Path

3.2.3. Node Protection

Definition:
A Backup Path that provides protection for failure of a single node and its directly connected links.

Discussion:
Node Protection may or may not protect the entire Primary Path. Node Protection also provides Link Protection.

Measurement units: n/a

Issues:

See Also:
Primary Path
Backup Path

3.2.4. Path Protection

Definition:
A Backup Path that provides protection for the entire Primary Path.

Discussion:
Path Protection provides Node Protection and Link Protection for every node and link along the Primary Path. A Backup Path providing Path Protection MUST have the same ingress node as the Primary Path.

Measurement units:
n/a

Issues:

See Also:
Primary Path
Backup Path
Node Protection
Link protection

3.2.5. Backup Span

Definition:
The number of nodes in the Primary Path that are protected by a Backup Path.

Discussion:

Measurement units:
number of nodes

Issues:

See Also:
Primary Path
Backup Path

3.3. Failure

3.3.1. Failure Detection

Definition:
To identify a Primary Path failure at a sub-IP layer.

Discussion:
Failure Detection occurs at the ingress node of the Primary Path. Failure Detection occurs via a sub-IP mechanism such as detection of a link down event or timeout for receipt of a control packet. A failure may be completely isolated. A failure may affect a set of links which share a single SRLG (e.g. port with many sub-interfaces). A failure may affect multiple links that are not part of SRLG.

Measurement units:
n/a

Issues:

See Also:
Primary Path

3.3.2. Failover Event

Definition:
The occurrence of a planned or unplanned action in the network that results in a change in the Path that data traffic traverses.

Discussion:
Failover Events include, but are not limited to, link failure and router failure. Routing changes are considered Convergence Events [7] and are not Failover Events. This restricts Failover Events to sub-IP layers. Failover may be at the PLR or at the ingress. If the failover is at the ingress it is generally on a disjoint path from the ingress to egress.

Measurement units:
n/a

Issues:

See Also:
Path
Failure Detection
Disjoint Path

3.3.3. Failover

Definition:
To switch data traffic from the Primary Path to the Backup Path upon a Failover Event.

Discussion:
Failover to a Backup Path provides Link Protection, Node Protection, or Path Protection. Failover is complete when Lost Packets, Out-of-Order Packets, and Duplicate Packets are no longer observed.

Measurement units:
n/a

Issues:

See Also:
Primary Path
Backup Path
Failover Event

3.3.4. Restoration

Definition:
The act of Failover Recovery in which the Primary Path is restored following a Failover Event.

Discussion:
Failure Recovery MUST occur when the Backup Path is the Working Path. The Backup Path is maintained as the Working Path during Failure Recovery. This implies that the service is either restored fully or partially. Usually, FRR restoration can cause congestion, but primary paths rerouting avoid restoration. An unavoidable problem in any restoration is the discontinuity in end to end delay when the primary and backup path delays differ significantly. If the backup path has a shorter delay out of order delivery may occur if restoration is fast. If the backup path is longer then a sudden increase in delay will occur which can affect real time applications which use playback buffers to remove limited jitter.

Measurement units:

Issues:

See Also:
Primary Path
Failover Event
Failure Recovery

Working Path
Backup Path

3.3.5. Reversion

Definition:
The act of restoring the Primary Path as the Working Path.

Discussion:
Protection Switching Systems may or may not support Reversion. Reversion, if supported, MUST occur after Failure Recovery.

Measurement units:
n/a

Issues:

See Also:
Protection Switching System
Working Path
Primary Path

3.4. Nodes

3.4.1. Protection-Switching Node

Definition:
A node that is capable to participate in a Protection Switching System.

Discussion:
The Protection Switching Node MAY be an ingress or egress for a Primary Path or Backup Path.

Measurement units:
n/a

Issues:

See Also:
Protection Switching System
Primary Path
Backup Path

3.4.2. Non-Protection Switching Node

Definition:

A node that not capable of participating in a Protection Switching System, however it MAY exist along the Primary Path or Backup Path.

Discussion:

Measurement units:
n/a

Issues:

See Also:
Protection Switching System
Primary Path
Backup Path

3.4.3. Failover Node

Definition:
A node along the Primary Path that is capable of Failover.

Discussion:
The Failover Node can be any node along the Primary Path except the egress node of the Primary Path. There can be multiple Failover Nodes along a Primary Path. The Failover Node MUST be the ingress to the Backup Path. The Failover Node MAY also be the ingress of the Primary Path.

Measurement units:
n/a

Issues:

See Also:
Primary Path
Backup Path
Failover

3.4.4. Merge Node

Definition:
A node along the Primary Path that is also the egress node of the Backup Path.

Discussion:
The Merge Node can be any node along the Primary Path except the ingress node of the Primary Path. There can be multiple Merge Nodes along a Primary Path. A Merge Node can be the egress node for a single or multiple Backup Paths. The Merge Node MUST be the egress to the Backup Path. The Merge Node MAY also be the egress of the Primary Path or point of local repair (PLR).

Measurement units:
n/a

Issues:

See Also:
Primary Path
Backup Path
PLR
Failover

3.4.5. Point of Local repair (PLR)

Definition:
The head-end LSR of a backup tunnel or a detour LSP.

Discussion:
Based on the functionality of the PLR, its role is defined based on the type of method used. If it is one-to-one backup method, the PLR is responsible for computing a separate backup LSP, called a detour LSP for each LSP that PLR is protecting. And in case if facility backup method is used, the PLR creates a single bypass tunnel that can be used to protect multiple LSPs.

Measurement units: n/a

Issues:

See Also:
Primary Path
Backup Path
Failover

3.4.6. Head-end Failover Node

Definition:
A node that is ingress to the Primary Path that is capable of Failover.

Discussion:
Based on the functionality of the Head-end, its role is defined to be as the ingress of the signaled LSP. It could also occur, that this node happens to be a PLR. In this scenario the term head-end failover node is defined.

Measurement units: n/a

Issues:

See Also:
Primary Path
Backup Path
Failover

3.5. Metrics

3.5.1. Failover Packet Loss

Definition:
The amount of packet loss produced by a Failover Event until Failover completes.

Discussion:
Packet loss can be observed as a reduction of forwarded traffic from the maximum forwarding rate. Failover Packet Loss includes packets that were lost and packets that were delayed due to buffering. Failover Packet Loss MAY reach 100% of the offered load.

Measurement units: Number of Packets

Issues:

See Also:
Failover Event
Failover

3.5.2. Reversion Packet Loss

Definition:
The amount of packet loss produced by Reversion.

Discussion:
Packet loss can be observed as a reduction of forwarded traffic from the maximum forwarding rate. Reversion Packet Loss includes packets that were lost and packets that were delayed due to buffering. Reversion Packet Loss MAY reach 100% of the offered load.

Measurement units: Number of Packets

Issues:

See Also:
Reversion

3.5.3. Primary Path Latency

Definition:
Latency [2] measured along the Primary Path.

Discussion:

Measurement units:
seconds

Issues:

See Also:
Primary Path

3.5.4. Backup Path Latency

Definition:
Latency [2] measured along the Backup Path.

Discussion:

Measurement units:
seconds

Issues:

See Also:
Backup Path

3.6. Benchmarks

3.6.1. Failover Time

Definition:
The amount of time it takes for Failover to complete so that the Backup Path is the Working Path.

Discussion:
Failover Time can be calculated from Failover Packet Loss that occurs due to a Failover Event and Failover as shown below in Equation 1:

(eq 1) Failover Time =
Failover Packets Loss / Offered Load
NOTE: Units for this measurement are
packets / packets/second = seconds

Failover Time includes failure detection time and time for data traffic to begin traversing the Backup Path.

Measurement units:
Seconds

Issues:

See Also:
Failover
Failover Packet loss
Working Path
Backup Path

3.6.2. Additive Backup Latency

Definition:
The amount of increased latency resulting from data traffic traversing the Backup Path instead of the Primary Path.

Discussion:
Additive Backup Latency is calculated using Equation 2 as shown below:

(eq 2) Additive Backup Latency =
Backup Path Latency - Primary Path Latency.

Measurement units:
Seconds

Issues:
Additive Backup Latency MAY be a negative result. This is theoretically possible, but could be indicative of a sub-optimum network configuration .

See Also:
Primary Path
Backup Path
Primary Path Latency
Backup Path Latency

3.6.3. Reversion Time

Definition:
The amount of time it takes for Reversion to complete so that the Primary Path is restored as the Working Path.

Discussion:
Reversion Time can be calculated from Reversion Packet Loss that occurs due to a Failure Recovery as shown below in Equation 3:

(eq 3) Reversion Time =
Reversion Packets Loss / Offered Load
NOTE: Units for this measurement are
packets / packets/second = seconds

Reversion Time starts upon completion of Failure Recovery and includes the time for data traffic to begin traversing the Primary Path.

Measurement units:
Seconds

Issues:

See Also:
Reversion
Primary Path
Working Path
Reversion Packet Loss
Failure Recovery

4. Acknowledgements

We would like thank Curtis Villamizar for providing input to the existing definitions, and proposing text for the new definitions on the BMWG mailing list.

5. IANA Considerations

This document requires no IANA considerations.

6. Security Considerations

This document only addresses terminology for the performance benchmarking of protection systems, and the information contained in this document has no effect on the security of the Internet.

7. References

7.1. Normative References

[1] Bradner, S., "The Internet Standards Process -- Revision 3", RFC 2026, October 1996.

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

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

[4] Perser, J., et al., "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", Internet Draft, Work in Progress, draft-ietf-bmwg-dsmterm-13.txt, July 2006.

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

[6] Paxson, V., et al., "Framework for IP Performance Metrics", RFC 2026, May 1998.

[7] Poretsky, S., Imhoff, B., "Benchmarking Terminology for IGP Convergence", draft-ietf-bmwg-igp-dataplane-conv-term-09, work in progress, January 2006.

[8] P. Pan., et al., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

7.2. Informative References

None.

8. Author's Address

Scott Poretsky
Reef Point Systems
8 New England Executive Park
Burlington, MA 01803
USA
Phone: + 1 508 439 9008
EMail: sporetsky@reefpoint.com

Rajiv Papneja
Isocore
12359 Sunrise Valley Drive
Reston, VA 22102
USA
Phone: 1 703 860 9273
Email: rpapneja@isocore.com

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

Full Copyright Statement

Copyright (C) The Internet Society (2006).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.