Network Working Group   S. Poretsky
Internet Draft   NextPoint Networks
Expires: August 2008
Intended Status: Informational   R. Papneja
                                                     Isocore

                                                     J. Karthik
                                                     Cisco Systems

                                                     S. Vapiwala
                                                     Cisco Systems

February 25, 2008

Benchmarking Terminology for Protection Performance

<draft-ietf-bmwg-protection-term-04.txt >

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 protection 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), Stateful High Availability (HA), 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 Sub-IP Protection 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-FRR) [8].

Benchmarking terminology have been defined for IP-layer route convergence [7]. 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. The sequence of events is as follows:

1. Failover Event - Primary Path fails
2. Failure Detection- Failover Event is detected
3. Failover - Backup Path becomes the Working Path due to Failover Event
4. Restoration - Primary Path recovers from a Failover Event
5. Reversion (optional) - Primary Path becomes the Working Path

These terms are further defined in this document. Figures 1 through 5 show fundamental models that MAY be used in benchmarking Sub-IP Protection mechanisms. Sub-IP Protection mechanisms MUST use a Protection Switching System that consists of a minimum of two Protection-Switching Nodes, an Ingress Node known as the Headend Node and an Egress Node known as the Merge Node. The protection MAY be provided with either a Primary Path and Backup Path, as shown in Figures 1 through 4, or a Primary Node and Standby Node, as shown in Figure 5.

A Protection Switching System may provide link protection, node protection, path protection, local link protection, and high availability, as shown in Figures 1 through 5 respectively. A Failover Event occurs along the Primary Path or at the Primary

Node. The Working Path is the Primary Path prior to the Failover Event and the Backup Path after the Failover Event. A Tester is set outside the two paths or nodes as it sends and receives IP traffic along the Working Path. The tester MUST record the IP packet sequence numbers, departure time, and arrival time so that the metrics of Failover Time, Additive Latency, Packet Reordering, Duplicate Packets, and Reversion Time can be measured. The Tester may be a single device or a test system. If Reversion is supported then the Working Path is the Primary Path after Restoration (Failure Recovery) of the Primary Path.

Link Protection, as shown in Figure 1, provides protection when a Failover Event occurs on the link between two nodes along the Primary Path. Node Protection, as shown in Figure 2, provides protection when a Failover Event occurs at a Node along the Primary Path. Path Protection, as shown in Figure 3, provides protection for link or node failures for multiple hops along the Primary Path. Local Link Protection, as shown in Figure 4, provides Sub-IP Protection of a link between two nodes, without a Backup Node. An example of such a Sub-IP Protection mechanism is SONET APS. High Availability Protection, as shown in Figure 5, provides protection of a Primary Node with a redundant Standby Node. State Control is provided between the Primary and Standby Nodes. Failure of the Primary Node is detected at the Sub-IP layer to force traffic to switch to the Standby Node, which has state maintained for zero or minimal packet loss.

```
               +-----------+
+--------------|  Tester   |<----------------------+
|              +-----------+                       |
| IP Traffic        | Failover           IP Traffic |
|                   |  Event                       |
|                   |                              |
|     ------------  |                ----------    |
+--->|  Ingress/  | V               | Egress/  |---+
     |Headend Node|------------------|Merge Node|  Primary
      ------------                   ----------    Path
         |                              ^
         |         ---------            |  Backup
         +--------| Backup  |-----------+   Path
                  |  Node   |
                   ---------
```

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

```
                            +-----------+
       +--------------------|  Tester   |<-----------------+
       |                    +-----------+                  |
       | IP Traffic               | Failover    IP Traffic |
       |                          | Event                  |
       |                          V                        |
       |     ------------      --------      ----------    |
       +--->|  Ingress/  |    |MidPoint|    | Egress/  |---+
            |Headend Node|----|  Node  |----|Merge Node|  Primary
             ------------      --------      ----------    Path
                |                                ^
                |         ---------              |  Backup
                +--------| Backup  |-------------+   Path
                         |  Node   |
                          ---------
```

Figure 2. System Under Test (SUT) for Sub-IP Node Protection

```
                               +-----------+
    +---------------------------|  Tester   |<----------------------+
    |                           +-----------+                       |
    | IP Traffic                      | Failover         IP Traffic |
    |                                 | Event                       |
    |                Primary Path     |                             |
    |     ------------      --------  |  --------     ----------    |
    +--->|  Ingress/  |    |MidPoint| V |Midpoint|   | Egress/  |---+
         |Headend Node|----|  Node  |---|  Node  |---|Merge Node|
          ------------      --------     --------     ----------
                |                                         ^
                |         ---------      --------         | Backup
                +--------| Backup  |----| Backup |--------+  Path
                         |  Node   |    |  Node  |
                          ---------      --------
```

Figure 3. System Under Test (SUT) for Sub-IP Path Protection

```
                                  +-----------+
             +--------------------|  Tester   |<-------------------+
             |                    +-----------+                    |
             | IP Traffic               | Failover      IP Traffic |
             |                          | Event                    |
             |              Primary     |                          |
             |    +--------+  Path      v            +--------+    |
             |    |        |------------------------>|        |    |
             +--->| Ingress|                         | Egress |----+
                  |  Node  |- - - - - - - - - - - - >|  Node  |
                  +--------+      Backup Path        +--------+
                  ^                                           ^
```

```
                  |            IP-Layer Forwarding            |
                  +-------------------------------------------+
```

Figure 4. System Under Test (SUT) for Sub-IP Local Link Protection

```
                         +-----------+
       +-----------------|  Tester   |<--------------------+
       |                 +-----------+                     |
       | IP Traffic            | Failover       IP Traffic |
       |                       | Event                     |
       |                       V                           |
       |     ---------      --------      ----------       |
       +--->| Ingress |    |Primary |    | Egress/  |------+
            |   Node  |----|  Node  |----|Merge Node|  Primary
             ---------      --------      ----------    Path
               |        State |Control       ^
               |    Interface |(Optional)    |
               |            ---------        |
               +---------| Standby |---------+
                         |  Node   |
                           ---------
```

Figure 5. System Under Test (SUT) for Sub-IP Redundant Node Protection

2. Existing definitions

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

Latency	[Ref.[2], section 3.8]
Frame Loss Rate	[Ref.[2], section 3.6]
Throughput	[Ref.[2], section 3.17]
Device Under Test (DUT)	[Ref.[3], section 3.1.1]
System Under Test (SUT)	[Ref.[3], section 3.1.2]
Out-of-order Packet	[Ref.[4], section 3.3.2]
Duplicate Packet	[Ref.[4], section 3.3.3]
Forwarding Delay	[Ref.[4], section 3.2.4]
Jitter	[Ref.[4], section 3.2.5]
Packet Loss	[Ref.[7], Section 3.5]
Packet Reordering	[Ref.[10], section 3.3]

This document has the following frequently used acronyms:

DUT Device Under Test
SUT System Under Test

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

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 [5]. 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. Paths

3.1.1 Path

Definition:
A unidirectional sequence of nodes, <R1, ..., Rn>, and links <L12,... L(n-1)n> with the following properties:

a. R1 is the ingress node and forwards IP packets, which input into DUT/SUT, to R2 as sub-IP frames over link L12.

b. Ri is a node which forwards data frames to R[i+1] over Link Li[i+1] for all i, 1<i<n, based on information in the sub-IP layer.

c. 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 [1]. 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 [9] will traverse one or more paths. The path is unidirectional. Example paths are the SONET/SDH path and the label switched path for MPLS.

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:
Working Path
Primary Path
Backup Path

3.1.2. Working Path

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

Discussion:
A Primary Path is the 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

3.1.3. Primary Path

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

Discussion:
The Primary Path is the Path that traffic traverses prior to a Failover Event.

Measurement units:
n/a

Issues:
None

See Also:
Path
Failover Event

3.1.4. Protected Primary Path

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

Discussion:
A Protected Primary Path MUST include at least one Protection Switching Node.

Measurement units:
n/a

Issues: None

See Also:
Path
Primary Path

3.1.5. Backup Path

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

Discussion:
The Backup Path SHALL be the Working Path upon a Failover Event. A Path MAY have one or more Backup Paths. A Backup Path MAY protect one or more Primary Paths. There are various types of Backup Paths:

a. dedicated recovery Backup Path (1+1), which has 100% redundancy for a specific ordinary path,

b. shared Backup Path (1:N), which is dedicated to the protection for more than one specific Primary Path

c. associated shared Backup Path (M:N) for which a specific set of Backup Paths protects a specific set of more than one Primary Path.

A Backup Path may be signaled or unsignaled. The Backup Path MUST be created prior to the Failover Event. A new Path computed after the Failover Event is simply Convergence [7] to a new Primary Path.

Measurement units:
n/a

Issues:

See Also:
Path
Working Path
Primary Path

3.1.6. Standby Backup Path

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

Discussion:
The Standby Backup Path and Dynamic Backup Path provide protection, but are established at different times.

Measurement units: n/a

Issues: None

See Also:
Backup Path
Primary Path
Failover Event

3.1.7. Dynamic Backup Path

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

Discussion:
The Standby Backup Path and Dynamic Backup Path provide protection, but are established at different times.

Measurement units: n/a

Issues: None

See Also:
Backup Path
Standby Backup Path
Failover Event

3.1.8. Disjoint Paths

Definition:
A pair of paths is 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: None

See Also:
Path
Primary Path
SRLG

3.1.9. Point of Local Repair (PLR)

Definition:
A node along the Primary Path that uses a Backup Path to protect another node or link.

Discussion:
Based on the functionality of the PLR, its role is defined based on the type of method used. If the one-to-one backup method is used, the PLR is responsible for computing a separate Backup Path for each Primary Path. In the case the facility backup method is used, the PLR creates a single Backup Path that can be used to protect multiple Primary Paths. Any node from the ingress node to the penultimate egress node MAY be a PLR. If the PLR is at the ingress, the Backup Path is a Disjoint Path from the ingress to egress.

Measurement units: n/a

Issues: None

See Also:
Primary Path
Backup Path
Failover

3.1.10. Shared Risk Link Group (SRLG)

Definition:
SRLG is a set of links which share a physical resource.

Discussion:
SRLG is considered the set of links to be avoided when the primary and secondary paths are considered disjoint. The SRLG will fail as a group if the shared resource fails.

Measurement units: n/a

Issues: None

See Also:

Path
Primary Path

3.2. Protection

3.2.1. Link Protection

Definition:
A Backup Path that is signaled to at least one Backup Node to protect for failure of interfaces and links along a Primary Path.

Discussion:
Link Protection may or may not protect the entire Primary Path. Link protection is shown in Figure 1.

Measurement units: n/a

Issues: None

See Also:
Primary Path
Backup Path

3.2.2. Node Protection

Definition:
A Backup Path that is signaled to at least one Backup Node to protect for failure of interfaces, links, and nodes along a Primary Path.

Discussion:
Node Protection may or may not protect the entire Primary Path. Node Protection also provides Link Protection. Node Protection is shown in Figure 2.

Measurement units: n/a

Issues: None

See Also:
Link Protection

3.2.3. Path Protection

Definition:
A Backup Path that is signaled to at least one Backup Node to provide protection along 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. Path Protection is shown in Figure 3.

Measurement units: n/a

Issues: None

See Also:
Primary Path
Backup Path
Node Protection
Link protection

3.2.4. Backup Span

Definition:
The number of hops used by a Backup Path.

Discussion:
The Backup Span is an integer obtained by counting the number of nodes along the Backup path

Measurement units:
number of nodes

Issues:
None

See Also:
Primary Path
Backup Path

3.2.5. Local Link Protection

Definition:
A Backup Path that is a redundant path between two nodes which does not use a Backup Node.

Discussion:
Local Link Protection MUST be provided as a Backup Path between two nodes along the Primary Path without the use of a Backup Node. Local Link Protection is provided by Protection Switching Systems such as SONET APS. Local Link Protection is shown in Figure 4.

Measurement units: None

Issues: None

See Also:
Backup path
Headend

3.2.6. Redundant Node Protection

Definition:
A Protection Switching System with a Primary Node protected by a Standby Node along the Primary Path.

Discussion:
Redundant Node Protection is provided by Protection Switching Systems such as VRRP and HA. The protection mechanisms occur at Sub-IP layers to switch traffic from a Primary Node to Backup Node upon a Failover Event at the Primary Node. Traffic continues to traverse the Primary Path through the Standby Node. The failover MAY be stateful, in which the state information MAY be exchanged in-band or over an out-of-band state control interface. The Standby Node MAY be active or passive. Redundant Node Protection is shown in Figure 5.

Measurement units: None

Issues: None

See Also:
Primary Path
Primary Node
Backup Node

3.2.7. State Control Interface

Definition:
An out-of-band control interface used to exchange state information between the Primary Node and Standby Node.

Discussion:
The State Control Interface MAY be used for Redundant Node Protection. The State Control Interface MUST be out-of-band. It is possible to have Redundant Node Protection in which there is no state control or state control is provided in-band. The State Control Interface between the Primary and Standby Node MAY be one or more hops.

Measurement units: None

Issues: None

See Also:
Primary Node
Standby Node

3.2.8. Protected Interface

Definition:
An interface along the Primary Path that is protected by a Backup Path.

Discussion:
A Protected Interface is an interface protected by a Protection Switching Systems that provides Link Protection, Node Protection, Path Protection, Local Link Protection, and Redundant Node Protection.

Measurement units: None

Issues: None

See Also:
Primary Path
Backup Path

3.3. Protection Switching

3.3.1. Protection Switching System

Definition:
A DUT/SUT that is capable of Failure Detection and Failover from a Primary Path to a Backup Path or Standby Node when a Failover Event occurs.

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: n/a

Issues: None

See Also:
Primary Path
Backup Path
Failover

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.

Failover Events may results from failures such as link failure or router failure. The change in path after Failover MAY have a Backup Span of one or more nodes. Failover Events are distinguished from routing changes and Convergence Events [7] by the detection of the failure and subsequent protection switching at a sub-IP layer. Failover occurs at a Point of Local Repair (PLR) or Primary Node.

Measurement units:
n/a

Issues:

See Also:
Path
Failure Detection
Disjoint Path

3.3.3. Failure Detection

Definition:
The process to identify at a sub-IP layer a Failover Event at a Primary Node or along the Primary Path.

Discussion:
Failure Detection occurs at the Primary Node or 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

3.3.4. Failover

Definition:
The process to switch data traffic from the Protected Primary Path to the Backup Path upon Failure Detection of a Failover Event.

Discussion:
Failover to a Backup Path provides Link Protection, Node Protection, or Path Protection. Failover is complete when Packet Loss [7], Out-of-order Packets [4], and Duplicate Packets [4] are no longer observed. Forwarding Delay [4] may continue to be observed.

Measurement units:
n/a

Issues:

See Also:
Primary Path
Backup Path
Failover Event

3.3.5. Restoration

Definition:
The state of failover recovery in which the Primary Path has recovered from a Failover Event, but is not yet forwarding packets because the Backup Path remains the Working Path.

Discussion:
Restoration MUST occur while the Backup Path is the Working Path. The Backup Path is maintained as the Working Path during Restoration. Restoration produces a Primary Path that is recovered from failure, but is not yet forwarding traffic. Traffic is still being forwarded by the Backup Path functioning as the Working Path.

Measurement units:
n/a

Issues:

See Also:
Primary Path
Failover Event

Failure Recovery
Working Path
Backup Path

3.3.6. Reversion

Definition:
The state of failover recovery in which the Primary Path has become the Working Path so that it is forwarding packets.

Discussion:
Protection Switching Systems may or may not support Reversion. Reversion, if supported, MUST occur after Restoration. Packet forwarding on the Primary Path resulting from Reversion may occur either fully or partially over the Primary Path. A potential problem with Reversion is the discontinuity in end to end delay when the Forwarding Delays [4] along the Primary Path and Backup Path are different, possibly causing Out of Order Packets [4], Duplicate Packets [4], and increased Jitter [4].

Measurement units: n/a

Issues: None

See Also:
Protection Switching System
Working Path
Primary Path

3.4. Nodes

3.4.1. Protection-Switching Node

Definition:
A node that is capable of participating in a Protection Switching System.

Discussion:
The Protection Switching Node MAY be an ingress or egress for a Primary Path or Backup Path, such as used for MPLS Fast Reroute configurations. The Protection Switching Node MAY provide Redundant Node Protection as a Primary Node in a Redundant chassis configuration with a Standby Node, such as used for VRRP and HA configurations.

Measurement units:
n/a

Issues:

See Also:
Protection Switching System

3.4.2. Non-Protection Switching Node

Definition:
A node that is 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. Headend Node

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

Discussion:
The Headend 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. The Headend Failover Node is always a PLR.

Measurement units: n/a

Issues:

See Also:
Primary Path
Point of Local Repair
Failover

3.4.4. Backup Node

Definition:
A node along the Backup Path.

Discussion:
The Backup Node can be any node along the Backup Path. There MAY be one or more Backup Nodes along the Backup Path. A Backup Node MAY be the ingress, mid-point, or egress of the Backup Path. If the Backup Path has only one Backup

Node, then that Backup Node is the ingress and egress of the Backup Path.

Measurement units: n/a

Issues:

See Also:
Backup Path

3.4.5. Merge Node

Definition:
A Node along the primary path where backup path terminates.

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

3.4.6. Primary Node

Definition:
A node along the Primary Path that is capable of Failover to a redundant Standby Node.

Discussion:
The Primary Node MAY be used for Protection Switching Systems that provide Redundant Node Protection, such as VRRP and HA

Measurement units: n/a

Issues:

See Also:
Protection Switching System
Redundant Node Protection
Standby Node

3.4.7. Standby Node

Definition:
A redundant node to a Primary Node that forwards traffic along the Primary Path upon Failure Detection of the Primary Node.

Discussion:
The Standby Node MUST be used for Protection Switching Systems that provide Redundant Node Protection, such as VRRP and HA. The Standby Node MUST provide protection along the same Primary Path. If the failover is to a Disjoint Path then it is a Backup Node. The Standby Node MAY be configured for 1:1 or N:1 protection.

The communication between the Primary Node and Standby Node MAY be in-band or across an out-of-band State Control interface. The Standby Node MAY be geographically dispersed from the Primary Node. When geographically dispersed, the number of hops of separation may increase failover time.

The Standby Node MAY be passive or active. The Passive Standby Node is not offered traffic and does not forward traffic until Failure Detection of the Primary Node. Upon Failure Detection of the Primary Node, traffic offered to the Primary Node is instead offered to the Passive Standby Node. The Active Standby Node is offered traffic and forwards traffic along the Primary Path while the Primary Node is also active. Upon Failure Detection of the Primary Node, traffic offered to the Primary Node is switched to the Active Standby Node.

Measurement units: n/a

Issues:

See Also:
Primary Node
State Control Interface

3.5. Benchmarks

3.5.1. Failover Packet Loss

Definition:
The amount of packet loss produced by a Failover Event until Failover completes, where the measurement begins when the last unimpaired packet is received by the Tester on the Protected Primary Path and ends when the first unimpaired packet is received by the Tester on the Backup Path.

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, reordered, or delayed. Failover Packet Loss MAY reach 100% of the offered load.

Measurement units:
Number of Packets

Issues: None

See Also:
Failover Event
Failover

3.5.2. Reversion Packet Loss

Definition:
The amount of packet loss produced by Reversion, where the measurement begins when the last unimpaired packet is received by the Tester on the Backup Path and ends when the first unimpaired packet is received by the Tester on the Protected Primary Path .

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, reordered, or delayed. Reversion Packet Loss MAY reach 100% of the offered load.

Measurement units: Number of Packets

Issues: None

See Also:
Reversion

3.5.3. Failover Time

Definition:
The amount of time it takes for Failover to successfully complete.

Discussion:
Failover Time can be calculated using the Time-Based Loss Method (TBLM), Packet-Loss Based Method (PLBM), or Timestamp-Based Method (TBM). It is RECOMMENDED that the TBM is used.

Measurement units:
milliseconds

Issues: None

See Also:
Failover
Failover Time
Time-Based Loss Method (TBLM)
Packet-Loss Based Method (PLBM)
Timestamp-Based Method (TBM)

3.5.4. 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 using the Time-Based Loss Method (TBLM), Packet-Loss Based Method (PLBM), or Timestamp-Based Method (TBM). It is RECOMMENDED that the TBM is used.

Measurement units:
milliseconds

Issues: None

See Also:
Reversion
Primary Path
Working Path
Reversion Packet Loss
Time-Based Loss Method (TBLM)
Packet-Loss Based Method (PLBM)
Timestamp-Based Method (TBM)

3.5.5. Additive Backup Delay

Definition:
The amount of increased Forwarding Delay [4] resulting from data traffic traversing the Backup Path instead of the Primary Path.

Discussion:
Additive Backup Delay is calculated using Equation 1 as shown below:

```
(Equation 1)
Additive Backup Delay =
          Forwarding Delay(Backup Path) -
          Forwarding Delay(Primary Path).
```

Measurement units:
milliseconds

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 Failover Time Calculation Methods

3.6.1 Time-Based Loss Method (TBLM)

Definition:
The method to calculate Failover Time (or Reversion Time) using a time scale on the Tester to measure the interval of Failover Packet Loss.

Discussion:
The Tester MUST provide statistics which show the duration of failure on a time scale to granularity of milliseconds based on occurrence of packet loss on a time scale. This is indicated by the duration of non-zero packet loss. The TBLM includes failure detection time and time for data traffic to begin traversing the Backup Path. Failover Time and Reversion Time are calculated using the TBLM as shown in Equation 2:

(Equation 2)
(Equation 2a)
TBLM Failover Time = Time(Failover) - Time(Failover Event)

(Equation 2b)
TBLM Reversion Time = Time(Reversion) - Time(Restoration)

Measurement units:
milliseconds

Issues:
None

See Also:
Failover
Packet-Loss Based Method

3.6.2 Packet-Loss Based Method (PLBM)

Definition:
The method used to calculate Failover Time (or Reversion Time) from the amount of Failover Packet Loss.

Discussion:
PLBM includes failure detection time and time for data traffic to begin traversing the Backup Path. Failover Time can be calculated using PLBM from the amount Failover Packet Loss as shown below in Equation 3:

```
(Equation 3)
     (Equation 3a)
     PLBM Failover Time =
        Number of packets lost /
                 (Offered Load rate * 1000)

     (Equation 3b)
     PLBM Restoration Time =
        Number of packets lost /
                 (Offered Load rate * 1000)

     Units are packets/(packets/second) = seconds
```

Measurement units:
milliseconds

Issues:
None

See Also:
Failover
Time-Based Loss Method

3.6.3 Timestamp-Based Method (TBM)

Definition:
The method to calculate Failover Time (or Reversion Time) using a time scale to quantify the interval between unimpaired packets arriving in the test stream.

Discussion:
The purpose of this method is to quantify the duration of failure or reversion on a time scale with granularity of milliseconds based on the observation of unimpaired packets, using Equation 2 with the difference being that the time values are obtained from the timestamp in the packet payload rather than from the Tester.

Unimpaired packets are normal packets that are not lost, reordered, or duplicated. A reordered packet is defined in

[10, section 3.3]. A duplicate packet is defined in [4, section 3.3.3]. A lost packet is defined in [7, Section 3.5]. Unimpaired packets may be detected by checking a sequence number in the payload, where the sequence number equals the next expected number for an unimpaired packet. A sequence gap or sequence reversal indicates impaired packets.

For calculating Failover Time, the TBM includes failure detection time and time for data traffic to begin traversing the Backup Path. For calculating Reversion Time, the TBM includes Reversion Time and time for data traffic to begin traversing the Primary Path.

Measurement units:
milliseconds

Issues: None

See Also:
Failover
Failover Time
Reversion
Reversion Time

4. Acknowledgements

We would like thank the BMWG and particularly Al Morton and Curtis Villamizar for their reviews, comments, and contributions to this work.

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] Poretsky, S., et al., "Terminology for Benchmarking

Network-layer Traffic Control Mechanisms", RFC 4689,
November 2006.

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

[6] Paxson, V., et al., "Framework for IP Performance Metrics", RFC 2330, May 1998.

[7] Poretsky, S., Imhoff, B., "Benchmarking Terminology for IGP Convergence", draft-ietf-bmwg-igp-dataplane-conv-term-15, work in progress, February 2008.

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

[9] Nichols, K., et al, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[10] Morton, A., et al, "Packet Reordering Metrics", RFC 4737, November 2006.

7.2. Informative References

None

8. Author's Address

Scott Poretsky
NextPoint Networks
3 Federal Street
Billerica, MA 01821
USA
Phone: + 1 508 439 9008
EMail: sporetsky@nextpointnetworks.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

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

Acknowledgement

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