Network Working Group
INTERNET-DRAFT
Expires in: June 2006

Scott Poretsky
Reef Point Systems

Brent Imhoff
Juniper Networks

January 2006

Benchmarking Methodology for IGP Data Plane Route Convergence

<draft-ietf-bmwg-igp-dataplane-conv-meth-09.txt>

ABSTRACT
This dpcument describes the methodology for benchmarking IGP Route Convergence as described in Applicability document [1] and Terminology document [2]. The methodology and terminology are to be used for benchmarking route convergence and can be applied to any link-state IGP such as ISIS [3] and OSPF [4]. The terms used in the procedures provided within this document are defined in [2].

Table of Contents

1. Introduction

This draft describes the methodology for benchmarking IGP Route Convergence. The applicability of this testing is described in [1] and the new terminology that it introduces is defined in [2]. Service Providers use IGP Convergence time as a key metric of router design and architecture. Customers of Service Providers observe convergence time by packet loss, so IGP Route Convergence is considered a Direct Measure of Quality (DMOQ). The test cases in this document are black-box tests that emulate the network events that cause route convergence, as described in [1]. The black-box test designs benchmark the data plane and account for all of the factors contributing to convergence time, as discussed in [1]. The methodology (and terminology) for benchmarking route convergence can be applied to any link-state IGP such as ISIS [3] and OSPF [4]. These methodologies apply to IPv4 and IPv6 traffic as well as IPv4 and IPv6 IGPs.

2. Existing definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [Br97]. RFC 2119 defines the use of these key words to help make the intent of standards track documents as clear as possible. While this document uses these keywords, this document is not a standards track document. The term Throughput is defined in RFC 2544.

3. Test Setup

3.1 Test Topologies

Figure 1 shows the test topology to measure IGP Route Convergence due to local Convergence Events such as SONET Link Failure, Layer 2 Session Failure, IGP Adjacency Failure, Route Withdrawal, and route cost change. These test cases discussed in section 4 provide route convergence times that account for the Event Detection time, SPF Processing time, and FIB Update time. These times are measured by observing packet loss in the data plane.

```
     ---------       Ingress Interface         ---------
     |       |<--------------------------------|       |
     |       |                                 |       |
     |       |    Preferred Egress Interface   |       |
     |  DUT  |-------------------------------->| Tester|
     |       |                                 |       |
     |       |~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>|       |
     |       |    Next-Best Egress Interface   |       |
     ---------                                 ---------
```

Figure 1. IGP Route Convergence Test Topology for Local Changes

Figure 2 shows the test topology to measure IGP Route Convergence time due to remote changes in the network topology. These times are measured by observing packet loss in the data plane. In this topology the three routers are considered a System Under Test (SUT). NOTE: All routers in the SUT must be the same model and identically configured.

```
             -----                       ---------
             |   | Preferred             |       |
     -----   |R2 |---------------------->|       |
     |   |-->|   | Egress Interface      |       |
     |   |   -----                       |       |
     |R1 |                               |Tester |
     |   |   -----                       |       |
     |   |-->|   |   Next-Best           |       |
     -----   |R3 |~~~~~~~~~~~~~~~~~~~~~~>|       |
       ^     |   |   Egress Interface    |       |
       |     -----                       ---------
       |                                     |
       |--------------------------------------
                   Ingress Interface
```

Figure 2. IGP Route Convergence Test Topology for Remote Changes

Figure 3 shows the test topology to measure IGP Route Convergence time with members of an Equal Cost Multipath (ECMP) Set. These times are measured by observing packet loss in the data plane. In this topology, the DUT is configured with each Egress interface as a member of an ECMP set and the Tester emulates multiple next-hop routers (emulates one router for each member).

```
     ---------       Ingress Interface         ---------
     |       |<--------------------------------|       |
     |       |                                 |       |
     |       |     ECMP Set Interface 1        |       |
     |  DUT  |-------------------------------->| Tester|
     |       |                .                |       |
     |       |                .                |       |
     |       |                .                |       |
     |       |~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>|       |
     |       |     ECMP Set Interface N        |       |
     ---------                                 ---------
```

     Figure 3. IGP Route Convergence Test Topology
                     for ECMP Convergence

Figure 4 shows the test topology to measure IGP Route Convergence time with members of a Parallel Link. These times are measured by observing packet loss in the data plane. In this topology, the DUT is configured with each Egress interface as a member of a Parallel Link and the Tester emulates the single next-hop router.

```
     ---------       Ingress Interface         ---------
     |       |<--------------------------------|       |
     |       |                                 |       |
     |       |     Parallel Link Interface 1   |       |
     |  DUT  |-------------------------------->| Tester|
     |       |                .                |       |
     |       |                .                |       |
     |       |                .                |       |
     |       |~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>|       |
     |       |     Parallel Link Interface N   |       |
     ---------                                 ---------
```

     Figure 4. IGP Route Convergence Test Topology
                 for Parallel Link Convergence

3.2 Test Considerations

3.2.1 IGP Selection

The test cases described in section 4 can be used for ISIS or OSPF. The Route Convergence test methodology for both is identical. The IGP adjacencies are established on the Preferred Egress Interface and Next-Best Egress Interface.

3.2.2 BGP Configuration

The obtained results for IGP Route Convergence may vary if

BGP routes are installed. It is recommended that the IGP Convergence times be benchmarked without BGP routes installed.

3.2.3 IGP Route Scaling

The number of IGP routes will impact the measured IGP Route Convergence because convergence for the entire IGP route table is measured. To obtain results similar to those that would be observed in an operational network, it is recommended that the number of installed routes closely approximate that for routers in the network. The number of areas (for OSPF) and levels (for ISIS) can impact the benchmark results.

3.2.4 Timers

There are some timers that will impact the measured IGP Convergence time. The following timers should be configured to the minimum value prior to beginning execution of the test cases:

Timer	Recommended Value
Link Failure Indication Delay	<10milliseconds
IGP Hello Timer	1 second
IGP Dead-Interval	3 seconds
LSA Generation Delay	0
LSA Flood Packet Pacing	0
LSA Retransmission Packet Pacing	0
SPF Delay	0

3.2.5 Convergence Time Metrics

The recommended value for the Packet Sampling Interval [2] is 100 milliseconds. Rate-Derived Convergence Time [2] is the preferred benchmark for IGP Route Convergence. This benchmark must always be reported when the Packet Sampling Interval [2] <= 100 milliseconds. If the test equipment does not permit the Packet Sampling Interval to be set as low as 100 msec, then both the Rate-Derived Convergence Time and Loss-Derived Convergence Time [2] must be reported. The Packet Sampling Interval value MUST be reported as the smallest measurable convergence time.

3.2.6 Interface Types

All test cases in this methodology document may be executed with any interface type. All interfaces MUST be the same media and Throughput [5,6] for each test case. Media and protocols MUST be configured for minimum failure detection delay to minimize the contribution to the measured Convergence time. For example, configure SONET with minimum carrier-loss-delay or Bi-directional Forwarding Detection (BFD).

3.2.7 Offered Load

The offered Load MUST be the Throughput of the device as defined in [5] and benchmarked in [6] at a fixed packet size. Packet size is measured in bytes and includes the IP header and payload. The packet size is selectable and MUST be recorded. The Throughput MUST be measured at the Preferred Egress Interface and the Next-Best Egress Interface. The duration of offered load MUST be greater than the convergence time. The destination addresses for the offered load MUST be distributed such that all routes are matched. This enables Full Convergence [2] to be observed.

3.3 Reporting Format

For each test case, it is recommended that the following reporting format be completed:

```
Parameter                              Units
---------                              -----
IGP                                    (ISIS or OSPF)
Interface Type                         (GigE, POS, ATM, etc.)
Packet Size offered to DUT             bytes
IGP Routes advertised to DUT           number of IGP routes
Packet Sampling Interval on Tester     seconds or milliseconds
IGP Timer Values configured on DUT
    SONET Failure Indication Delay   seconds or milliseconds
    IGP Hello Timer                  seconds or milliseconds
    IGP Dead-Interval                seconds or milliseconds
    LSA Generation Delay             seconds or milliseconds
    LSA Flood Packet Pacing          seconds or milliseconds
    LSA Retransmission Packet Pacing seconds or milliseconds
    SPF Delay                        seconds or milliseconds
Benchmarks
      Rate-Derived Convergence Time  seconds or milliseconds
      Loss-Derived Convergence Time  seconds or milliseconds
      Restoration Convergence Time   seconds or milliseconds
```

4. Test Cases

4.1 Convergence Due to Link Failure

4.1.1 Convergence Due to Local Interface Failure

Objective

To obtain the IGP Route Convergence due to a local link failure event at the DUT's Local Interface.

Procedure

1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Remove Preferred Egress link on DUT's Local Interface [2] by performing an administrative shutdown of the interface.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the link down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Restore Preferred Egress link on DUT's Local Interface by administratively enabling the interface.
8. Measure Restoration Convergence Time [2] as DUT detects the link up event and converges all IGP routes and traffic back to the Preferred Egress Interface.

Results

The measured IGP Convergence time is influenced by the Local link failure indication, SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time.

4.1.2 Convergence Due to Neighbor Interface Failure

Objective

To obtain the IGP Route Convergence due to a local link failure event at the Tester's Neighbor Interface.

Procedure

1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to

DUT on Ingress Interface [2].

3. Verify traffic routed over Preferred Egress Interface.
4. Remove link on Tester's Neighbor Interface [2] connected to DUT' s Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the link down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Restore link on Tester's Neighbor Interface connected to DUT's Preferred Egress Interface.
8. Measure Restoration Convergence Time [2] as DUT detects the link up event and converges all IGP routes and traffic back to the Preferred Egress Interface.

Results
The measured IGP Convergence time is influenced by the Local link failure indication, SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time.

4.1.3 Convergence Due to Remote Interface Failure

Objective
To obtain the IGP Route Convergence due to a Remote Interface Failure event.

Procedure
1. Advertise matching IGP routes from Tester to SUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 2. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic is routed over Preferred Egress Interface.
4. Remove link on Tester's Neighbor Interface [2] connected to SUT' s Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] as SUT detects the link down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Restore link on Tester's Neighbor Interface connected to DUT's Preferred Egress Interface.
8. Measure Restoration Convergence Time [2] as DUT detects the link up event and converges all IGP routes and traffic back to the Preferred Egress Interface.

Results
The measured IGP Convergence time is influenced by the link failure failure indication, LSA/LSP Flood Packet Pacing, LSA/LSP Retransmission Packet Pacing, LSA/LSP Generation time, SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time. The additional convergence time contributed by LSP Propagation can be obtained by subtracting the Rate-Derived Convergence Time measured in 4.1.2 (Convergence Due to Neighbor Interface Failure) from the Rate-Derived Convergence Time measured in this test case.

4.2 Convergence Due to Layer 2 Session Failure

Objective
To obtain the IGP Route Convergence due to a Local Layer 2 Session failure event.

Procedure
1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the IGP routes along the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Remove Layer 2 session from Tester's Neighbor Interface [2] connected to Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the Layer 2 session down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Restore Layer 2 session on DUT's Preferred Egress Interface.
7. Measure Restoration Convergence Time [2] as DUT detects the session up event and converges all IGP routes and traffic over the Preferred Egress Interface.

Results
The measured IGP Convergence time is influenced by the Layer 2 failure indication, SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time.

4.3 Convergence Due to IGP Adjacency Failure

Objective
To obtain the IGP Route Convergence due to a Local IGP Adjacency failure event.

Procedure
1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Remove IGP adjacency from Tester's Neighbor Interface [2] connected to Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the IGP session failure event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Restore IGP session on DUT's Preferred Egress Interface.
8. Measure Restoration Convergence Time [2] as DUT detects the session up event and converges all IGP routes and traffic over the Preferred Egress Interface.

Results
The measured IGP Convergence time is influenced by the IGP Hello Interval, IGP Dead Interval, SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time.

4.4 Convergence Due to Route Withdrawal

Objective
To obtain the IGP Route Convergence due to Route Withdrawal.

Procedure
1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].

3. Verify traffic routed over Preferred Egress Interface.
4. Tester withdraws all IGP routes from DUT's Local Interface on Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the Layer 2 session down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Re-advertise IGP routes to DUT's Preferred Egress Interface.
8. Measure Restoration Convergence Time [2] as DUT converges all IGP routes and traffic over the Preferred Egress Interface.

Results
The measured IGP Convergence time is the SPF Processing and FIB Update time as influenced by the SPF delay, SPF Holdtime, SPF Execution Time, Tree Build Time, and Hardware Update Time.

4.5 Convergence Due to Cost Change

Objective
To obtain the IGP Route Convergence due to route cost change.

Procedure
1. Advertise matching IGP routes from Tester to DUT on Preferred Egress Interface [2] and Next-Best Egress Interface [2] using the topology shown in Figure 1. Set the cost of the routes so that the Preferred Egress Interface is the preferred next-hop.
2. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Tester increases cost for all IGP routes at DUT's Preferred Egress Interface so that the Next-Best Egress Interface has lower cost and becomes preferred path.
5. Measure Rate-Derived Convergence Time [2] as DUT detects the cost change event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
7. Re-advertise IGP routes to DUT's Preferred Egress Interface with original lower cost metric.
8. Measure Restoration Convergence Time [2] as DUT converges all IGP routes and traffic over the Preferred Egress Interface.

Results
There should be no measured packet loss for this case.

4.6 Convergence Due to ECMP Member Interface Failure

Objective

To obtain the IGP Route Convergence due to a local link failure event of an ECMP Member.

Procedure

1. Configure ECMP Set as shown in Figure 3.
2. Advertise matching IGP routes from Tester to DUT on each ECMP member.
3. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
4. Verify traffic routed over all members of ECMP Set.
5. Remove link on Tester's Neighbor Interface [2] connected to one of the DUT's ECMP member interfaces.
6. Measure Rate-Derived Convergence Time [2] as DUT detects the link down event and converges all IGP routes and traffic over the other ECMP members.
7. Stop offered load. Wait 30 seconds for queues to drain. Restart Offered Load.
8. Restore link on Tester's Neighbor Interface connected to DUT's ECMP member interface.
9. Measure Restoration Convergence Time [2] as DUT detects the link up event and converges IGP routes and some distribution of traffic over the restored ECMP member.

Results

The measured IGP Convergence time is influenced by Local link failure indication, Tree Build Time, and Hardware Update Time.

4.7 Convergence Due to Parallel Link Interface Failure

Objective

To obtain the IGP Route Convergence due to a local link failure event for a Member of a Parallel Link. The links can be used for data Load Balancing

Procedure

1. Configure Parallel Link as shown in Figure 4.
2. Advertise matching IGP routes from Tester to DUT on each Parallel Link member.
3. Send offered load at measured Throughput with fixed packet size to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
4. Verify traffic routed over all members of Parallel Link.
5. Remove link on Tester's Neighbor Interface [2] connected to one of the DUT's Parallel Link member interfaces.
6. Measure Rate-Derived Convergence Time [2] as DUT detects the link down event and converges all IGP routes and traffic over

the other Parallel Link members.

7. Stop offered load. Wait 30 seconds for queues to drain.
   Restart Offered Load.

8. Restore link on Tester's Neighbor Interface connected to DUT's Parallel Link member interface.
9. Measure Restoration Convergence Time [2] as DUT detects the link up event and converges IGP routes and some distribution of traffic over the restored Parallel Link member.

Results
The measured IGP Convergence time is influenced by the Local link failure indication, Tree Build Time, and Hardware Update Time.

5. IANA Considerations

This document requires no IANA considerations.

6. Security Considerations

Documents of this type do not directly affect the security of the Internet or corporate networks as long as benchmarking is not performed on devices or systems connected to operating networks.

7. References

7.1 Normative References

[1] Poretsky, S., "Benchmarking Applicability for IGP Convergence", draft-ietf-bmwg-igp-dataplane-conv-app-09, work in progress, January 2006.

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

[3] Callon, R., "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", RFC 1195, IETF, December 1990.

[4] Moy, J., "OSPF Version 2", RFC 2328, IETF, April 1998.

[5] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, IETF, January 1991.

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

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

Brent Imhoff
Juniper Networks
1194 North Mathilda Ave
Sunnyvale, CA 94089
USA
Phone: + 1 314 378 2571
EMail: bimhoff@planetspork.com

specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

Acknowledgement

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