Network Working Group
INTERNET-DRAFT
Expires in: April 2004

Scott Poretsky
Quarry Technologies

Brent Imhoff
Wiltel Communications

October 2003

Benchmarking Methodology for IGP Data Plane Route Convergence

<draft-ietf-bmwg-igp-dataplane-conv-meth-01.txt>

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

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.

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 accounting for all of the factors contributing to route 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].

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.

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, PPP 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).

```
                  -----                         -----------
                  |   | 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 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.   For 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.

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
SONET 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
Figure 5 shows a graph model of Convergence Time as measured from the data plane. Refer to [2] for definitions of the terms used. Rate-Derived Convergence Time and Loss-Derived Convergence Time are the two metrics for convergence time. An offered Load of maximum forwarding rate at a fixed packet size is recommended for accurate measurement. The test duration must be greater than the convergence time.

Ideally, Convergence Event Transition and Convergence Recovery Transition are instantaneous so that the
Rate-Derived Convergence Time = Loss-Derived Convergence Time.

When the Convergence Event Transition and Convergence Recovery Transition are not instantaneous so that there is a slope, as shown in Figure 5, the accuracy of the Rate-Derived Convergence Time and Loss-Derived Convergence Time are dependent upon the Packet Sampling Interval.

Under this condition and the Packet Sampling Interval <= 100 millisecond, the Rate-Derived Convergence Time > Loss-Derived Convergence Time and Rate-Derived Convergence Time is the preferred metric. Under this condition and the Packet Sampling Interval > 100 millisecond the Rate-Derived Convergence Time < Loss-Derived Convergence Time and Loss-Derived Convergence Time is the better metric. For all test cases, the Rate-Derived Convergence Time and Loss-Derived Convergence Time must be recorded.

```
                            Recovery  Convergence Event   Time = 0sec
        Maximum              ^           ^                  ^
        Forwarding Rate--> ----\    Packet   /---------------
                                    \    Loss   /<----Convergence
              Convergence------->\           /           Event
Transition
        Recovery Transition       \         /
                                           \_____/<------100% Packet
Loss

        X-axis = Time
        Y-axis = Forwarding Rate
```

Figure 5. Convergence Graph

3.2.6 Packet Sampling Interval

Selection of the Packet Sampling Interval on the Test Equipment impacts the measured Rate-Derived Convergence Time. Packet Sampling Interval time is that is too large exaggerates the slope of the Convergence Event Transition and Convergence Recovery Transition producing a larger than the actual Rate-Derived Convergence Time. This impact is greater as routers achieve millisecond convergence times. The recommended value for the Packet Sampling Interval is 100 millisecond. It is possible to have commercially available test equipment with a minimum configurable Packet Sampling Interval of 1 second.

3.2.7 Interface Types

All test cases in this methodology document may be executed with any interface type. SONET is recommended and specifically mentioned in the procedures because it can be configured to have no or negligible affect on the measured convergence time. Ethernet (10Mb, 100Mb, 1Gb, and 10Gb) is not preferred since broadcast media are unable to detect loss of host and rely upon IGP Hellos to detect session loss.

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                                             bytes
        IGP Routes                                              number
of IGP routes
        Packet Sampling Interval                        seconds or
milliseconds
        IGP Timer Values
                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
      Results
                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 traffic at maximum forwarding rate to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.

4. Remove SONET on DUT's Local Interface [2] by performing an administrative shutdown of the interface.
5. Measure Rate-Derived Convergence Time [2] and Loss-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. Restore SONET on DUT's Local Interface by administratively enabling the interface.
7. 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 SONET 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 traffic at maximum forwarding rate to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Remove SONET on Tester's Neighbor Interface [2] connected to DUT' s Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] and Loss-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. Restore SONET on Tester's Neighbor Interface connected to DUT's Preferred Egress Interface.
7. 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 SONET 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. NOTE: All routers in the SUT must be the same model and identically configured.
2. Send traffic at maximum forwarding rate 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 SONET on Tester's Neighbor Interface [2] connected to

SUT' s Preferred Egress Interface.

5. Measure Rate-Derived Convergence Time [2] and Loss-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. Restore SONET on Tester's Neighbor Interface connected to SUT's Preferred Egress Interface.
7. Measure Restoration Convergence Time [2] as SUT detects the link up event and converges all IGP routes and traffic over the Preferred Egress Interface.

Results
The measured IGP Convergence time is influenced by the SONET 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 PPP Session Failure

Objective
To obtain the IGP Route Convergence due to a Local PPP 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 traffic at maximum forwarding rate to destinations matching all IGP routes from Tester to DUT on Ingress Interface [2].
3. Verify traffic routed over Preferred Egress Interface.
4. Remove PPP session from Tester's Neighbor Interface [2] connected to Preferred Egress Interface.
5. Measure Rate-Derived Convergence Time [2] and Loss-Derived Convergence Time [2] as DUT detects the PPP session down event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Restore PPP 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 PPP

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 traffic at maximum forwarding rate 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] and Loss-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. Restore IGP 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 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 traffic at maximum forwarding rate 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] and Loss-Derived Convergence Time [2] as DUT processes the route withdrawal event and converges all IGP routes and traffic over the Next-Best Egress Interface.
6. Re-advertise IGP routes to DUT's Preferred Egress Interface.
7. 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 traffic at maximum forwarding rate 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 Inerface has lower cost and becomes preferred path.
5. Measure Rate-Derived Convergence Time [2] and Loss-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. Re-advertise IGP routes to DUT's Preferred Egress Interface with original lower cost metric.
7. 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 traffic at maximum forwarding rate 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 SONET on Tester's Neighbor Interface [2] connected to one of the DUT's ECMP member interfaces.
6. Measure Rate-Derived Convergence Time [2] and Loss-Derived Convergence Time [2] as DUT detects the link down event and converges all IGP routes and traffic over the other ECMP members.
7. Restore SONET on Tester's Neighbor Interface connected to DUT's ECMP member interface.
8. 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 the Local SONET 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.

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 traffic at maximum forwarding rate 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 SONET on Tester's Neighbor Interface [2] connected to one of the DUT's Parallel Link member interfaces.
6. Measure Rate-Derived Convergence Time [2] and Loss-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. Restore SONET on Tester's Neighbor Interface connected to DUT's Parallel Link member interface.
8. 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 SONET indication, Tree Build Time, and Hardware Update Time.

5. 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.

6. References

[1] Poretsky, S., "Benchmarking Applicability for IGP Convergence", draft-ietf-bmwg-igp-dataplane-conv-app-01, work in progress, October 2003.

[2] Poretsky, S., Imhoff, B., "Benchmarking Terminology for IGP Convergence", draft-ietf-bmwg-igp-dataplane-conv-term-01, work in progress, October 2003.

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

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

7. Author's Address

Scott Poretsky
Quarry Technologies
8 New England Executive Park
Burlington, MA 01803
USA

Phone: + 1 781 395 5090
EMail: sporetsky@quarrytech.com

Brent Imhoff
WilTel Communications
3180 Rider Trail South
Bridgeton, MO 63045
USA

Phone: +1 314 595 6853
EMail: brent.imhoff@wcg.com

8. Full Copyright Statement