Network Working Group	S. Poretsky
Internet-Draft	Allot Communications
Intended status: Informational	B. Imhoff
Expires: August 20, 2011	Juniper Networks
	K. Michielsen
	Cisco Systems
	February 16, 2011

Benchmarking Methodology for Link-State IGP Data Plane Route
Convergence
draft-ietf-bmwg-igp-dataplane-conv-meth-23

Abstract

This document describes the methodology for benchmarking Link-State
Interior Gateway Protocol (IGP) Route Convergence. The methodology is
to be used for benchmarking IGP convergence time through externally
observable (black box) data plane measurements. The methodology can be
applied to any link-state IGP, such as IS-IS and OSPF.

Status of this Memo

Copyright Notice

Table of Contents

[1.](#) Introduction

[1.1.](#) Motivation

Convergence time is a critical performance parameter. Service Providers use IGP convergence time as a key metric of router design and architecture. Fast network convergence can be optimally achieved through deployment of fast converging routers. Customers of Service Providers use packet loss due to Interior Gateway Protocol (IGP) convergence as a key metric of their network service quality. IGP route convergence is a Direct Measure of Quality (DMOQ) when benchmarking the data plane. The fundamental basis by which network users and operators benchmark convergence is packet loss and other packet impairments, which are externally observable events having direct impact on their application performance. For this reason it is important to develop a standard methodology for benchmarking link-state IGP convergence time through externally observable (black-box) data plane measurements. All factors contributing to convergence time are accounted for by measuring on the data plane.

Factors for IGP Route Convergence Time

There are four major categories of factors contributing to the measured IGP convergence time. As discussed in [Vi02], [Ka02], [Fi02], [Al00], [Al02], and [Fr05], these categories are Event Detection, Shortest Path First (SPF) Processing, Link State Advertisement (LSA) / Link State Packet (LSP) Advertisement, and Forwarding Information Base (FIB) Update. These have numerous components that influence the convergence time, including but not limited to the list below:

    *Event Detection

       -Physical Layer failure/recovery indication time

       -Layer 2 failure/recovery indication time

       -IGP Hello Dead Interval

    *SPF Processing

       -SPF Delay Time

       -SPF Hold time

       -SPF Execution time

    *LSA/LSP Advertisement

       -LSA/LSP Generation time

       -LSA/LSP Flood Packet Pacing

       -LSA/LSP Retransmission Packet Pacing

    *FIB Update

       -Tree Build time

       -Hardware Update time

    *Increased Forwarding Delay due to Queueing

The contribution of each of these factors listed above will vary with each router vendors' architecture and IGP implementation. Routers may have a centralized forwarding architecture, in which one forwarding table is calculated and referenced for all arriving packets, or a distributed forwarding architecture, in which the central forwarding table is calculated and distributed to the interfaces for local look-up as packets arrive. The distributed forwarding tables are typically maintained in hardware.

The variation in router architecture and implementation necessitates
the design of a convergence test that considers all of these components
contributing to convergence time and is independent of the Device Under
Test (DUT) architecture and implementation. The benefit of designing a
test for these considerations is that it enables black-box testing in
which knowledge of the routers' internal implementation is not
required. It is then possible to make valid use of the convergence
benchmarking metrics when comparing routers from different vendors.
Convergence performance is tightly linked to the number of tasks a
router has to deal with. As the most impacting tasks are mainly related
to the control plane and the data plane, the more the DUT is stressed
as in a live production environment, the closer performance measurement
results match the ones that would be observed in a live production
environment.

1.3. Use of Data Plane for IGP Route Convergence Benchmarking

Customers of Service Providers use packet loss and other packet
impairments as metrics to calculate convergence time. Packet loss and
other packet impairments are externally observable events having direct
impact on customers' application performance. For this reason it is
important to develop a standard router benchmarking methodology that is
a Direct Measure of Quality (DMOQ) for measuring IGP convergence. An
additional benefit of using packet loss for calculation of IGP Route
Convergence time is that it enables black-box tests to be designed.
Data traffic can be offered to the Device Under Test (DUT), an emulated
network event can be forced to occur, and packet loss and other
impaired packets can be externally measured to calculate the
convergence time. Knowledge of the DUT architecture and IGP
implementation is not required. There is no need to rely on the DUT to
produce the test results. There is no need to build intrusive test
harnesses for the DUT. All factors contributing to convergence time are
accounted for by measuring on the dataplane.
Other work of the Benchmarking Methodology Working Group (BMWG) focuses
on characterizing single router control plane convergence. See [Ma05],
[Ma05t], and [Ma05c].

1.4. Applicability and Scope

The methodology described in this document can be applied to IPv4 and
IPv6 traffic and link-state IGPs such as IS-IS [Ca90][Ho08], OSPF
[Mo98][Co08], and others. IGP adjacencies established over any kind of
tunnel (such as Traffic Engineering tunnels) are outside the scope of
this document. Convergence time benchmarking in topologies with non
point-to-point IGP adjacencies will be covered in a later document.
Convergence from Bidirectional Forwarding Detection (BFD) is outside
the scope of this document. Non-Stop Forwarding (NSF), Non-Stop Routing
(NSR), Graceful Restart (GR), or any other High Availability mechanism
are outside the scope of this document. Fast reroute mechanisms such as

IP Fast-Reroute [Sh10i] or MPLS Fast-Reroute [Pa05] are outside the scope of this document.

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. This document uses much of the terminology defined in [Po11t]. For any conflicting content, this document supersedes [Po11t]. This document uses existing terminology defined in other documents issued by the Benchmarking Methodology Working Group (BMWG). Examples include, but are not limited to:

Throughput	[Ref.[Br91], section 3.17]
Device Under Test (DUT)	[Ref.[Ma98], section 3.1.1]
System Under Test (SUT)	[Ref.[Ma98], section 3.1.2]
Out-of-Order Packet	[Ref.[Po06], section 3.3.4]
Duplicate Packet	[Ref.[Po06], section 3.3.5]
Stream	[Ref.[Po06], section 3.3.2]
Loss Period	[Ref.[Ko02], section 4]
Forwarding Delay	[Ref.[Po06], section 3.2.4]
IP Packet Delay Variation (IPDV)	[Ref.[De02], section 1.2]

3. Test Topologies

3.1. Test topology for local changes

Figure shows the test topology to measure IGP convergence time due to local Convergence Events such as Local Interface failure and recovery (Section 8.1.1), layer 2 session failure and recovery (Section 8.2.1), and IGP adjacency failure and recovery (Section 8.2.2). This topology is also used to measure IGP convergence time due to route withdrawal and readvertisement (Section 8.2.3), and route cost change (Section 8.3.2) Convergence Events. IGP adjacencies MUST be established between Tester and DUT: one on the Ingress Interface, one on the Preferred Egress Interface, and one on the Next-Best Egress Interface. For this purpose the Tester emulates three routers (RTa, RTb, and RTc), each establishing one adjacency with the DUT.

```
                    -------
                    |     | Preferred         .......
                    |     |-----------------. RTb .
    .......    Ingress |     | Egress Interface .......
    . RTa .------------| DUT |
    .......  Interface |     | Next-Best         .......
                    |     |-----------------. RTc .
                    |     | Egress Interface .......
                    -------
```

Figure shows the test topology to measure IGP convergence time due to
local Convergence Events with a non-Equal Cost Multipath (ECMP)
Preferred Egress Interface and Equal Cost Multipath (ECMP) Next-Best
Egress Interfaces (Section 8.1.1). In this topology, the DUT is
configured with each Next-Best Egress interface as a member of a single
ECMP set. The Preferred Egress Interface is not a member of an ECMP
set. The Tester emulates N+2 neighbor routers (N>0): one router for the
Ingress Interface (RTa), one router for the Preferred Egress Interface
(RTb), and N routers for the members of the ECMP set (RTc1...RTcN). IGP
adjacencies MUST be established between Tester and DUT: one on the
Ingress Interface, one on the Preferred Egress Interface, and one on
each member of the ECMP set. When the test specifies to observe the
Next-Best Egress Interface statistics, the combined statistics for all
ECMP members should be observed.

```
                    -------
                    |     | Preferred         .......
                    |     |-----------------. RTb .
                    |     | Egress Interface .......
                    |     |
                    |     | ECMP Set          ........
    .......    Ingress |     |-----------------. RTc1 .
    . RTa .------------| DUT | Interface 1        .......
    .......  Interface |     |        .
                    |     |        .
                    |     |        .
                    |     | ECMP Set          ........
                    |     |-----------------. RTcN .
                    |     | Interface N       ........
                    -------
```

3.2. Test topology for remote changes

Figure shows the test topology to measure IGP convergence time due to
Remote Interface failure and recovery (Section 8.1.2). In this topology
the two routers DUT1 and DUT2 are considered System Under Test (SUT)
and SHOULD be identically configured devices of the same model. IGP
adjacencies MUST be established between Tester and SUT, one on the
Ingress Interface, one on the Preferred Egress Interface, and one on

the Next-Best Egress Interface. For this purpose the Tester emulates
three routers (RTa, RTb, and RTc). In this topology there is a
possibility of a packet forwarding loop that may occur transiently
between DUT1 and DUT2 during convergence (micro-loop, see [Sh10]).

```
                         --------
                         |      |  -------- Preferred       .......
                         |      |--| DUT2 |-----------------. RTb .
.......     Ingress |      |  -------- Egress Interface .......
. RTa .------------| DUT1 |
.......   Interface |      | Next-Best                   .......
                         |      |---------------------------. RTc .
                         |      | Egress Interface          .......
                         --------
```

Figure shows the test topology to measure IGP convergence time due to
remote Convergence Events with a non-ECMP Preferred Egress Interface
and ECMP Next-Best Egress Interfaces (Section 8.1.2). In this topology
the two routers DUT1 and DUT2 are considered System Under Test (SUT)
and MUST be identically configured devices of the same model. Router
DUT1 is configured with the Next-Best Egress Interface an ECMP set of
interfaces. The Preferred Egress Interface of DUT1 is not a member of
an ECMP set. The Tester emulates N+2 neighbor routers (N>0), one for
the Ingress Interface (RTa), one for DUT2 (RTb) and one for each member
of the ECMP set (RTc1...RTcN). IGP adjacencies MUST be established
between Tester and SUT, one on each interface of SUT. For this purpose
each of the N+2 routers emulated by the Tester establishes one
adjacency with the SUT. In this topology there is a possibility of a
packet forwarding loop that may occur transiently between DUT1 and DUT2
during convergence (micro-loop, see [Sh10]). When the test specifies to
observe the Next-Best Egress Interface statistics, the combined
statistics for all members of the ECMP set should be observed.

```
                         --------
                         |      |  -------- Preferred       .......
                         |      |--| DUT2 |-----------------. RTb .
                         |      |  -------- Egress Interface .......
                         |      |
                         |      | ECMP Set                  ........
.......     Ingress |      |---------------------------. RTc1 .
. RTa .------------| DUT1 | Interface 1                ........
.......   Interface |      |      .
                         |      |        .
                         |      |         .
                         |      | ECMP Set                  ........
                         |      |---------------------------. RTcN .
                         |      | Interface N                ........
                         --------
```

[3.3.](#) Test topology for local ECMP changes

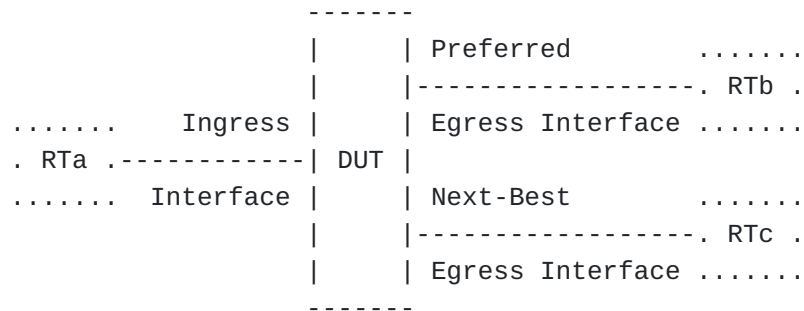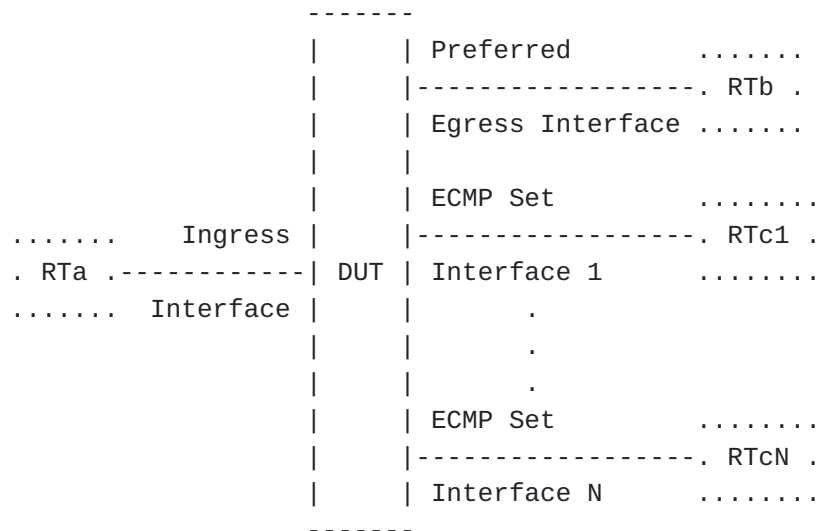Figure shows the test topology to measure IGP convergence time due to
local Convergence Events of a member of an Equal Cost Multipath (ECMP)
set ([Section 8.1.3](#)). In this topology, the DUT is configured with each
egress interface as a member of a single ECMP set and the Tester
emulates N+1 next-hop routers, one for the Ingress Interface (RTa) and
one for each member of the ECMP set (RTb1...RTbN). IGP adjacencies MUST
be established between Tester and DUT, one on the Ingress Interface and
one on each member of the ECMP set. For this purpose each of the N+1
routers emulated by the Tester establishes one adjacency with the DUT.
When the test specifies to observe the Next-Best Egress Interface
statistics, the combined statistics for all ECMP members except the one
affected by the Convergence Event, should be observed.

```
                          -------
                         |       | ECMP Set    ........
                         |       |-------------. RTb1 .
                         |       | Interface 1 ........
.......     Ingress      |       |       .
. RTa .------------------| DUT   |       .
.......     Interface    |       |       .
                         |       | ECMP Set    ........
                         |       |-------------. RTbN .
                         |       | Interface N ........
                          -------
```

[3.4.](#) Test topology for remote ECMP changes

Figure shows the test topology to measure IGP convergence time due to
remote Convergence Events of a member of an Equal Cost Multipath (ECMP)
set ([Section 8.1.4](#)). In this topology the two routers DUT1 and DUT2 are
considered System Under Test (SUT) and MUST be identically configured
devices of the same model. Router DUT1 is configured with each egress
interface as a member of a single ECMP set and the Tester emulates N+1
neighbor routers (N>0), one for the Ingress Interface (RTa) and one for
each member of the ECMP set (RTb1...RTbN). IGP adjacencies MUST be
established between Tester and SUT, one on each interface of SUT. For
this purpose each of the N+1 routers emulated by the Tester establishes
one adjacency with the SUT (N-1 emulated routers are adjacent to DUT1
egress interfaces, one emulated router is adjacent to DUT1 Ingress
Interface, and one emulated router is adjacent to DUT2). In this
topology there is a possibility of a packet forwarding loop that may
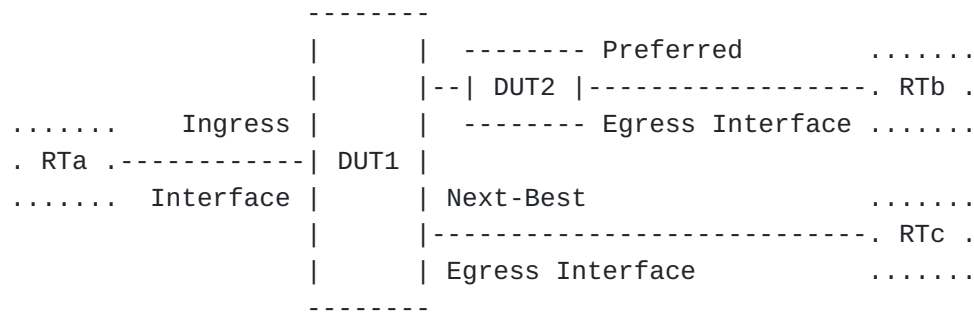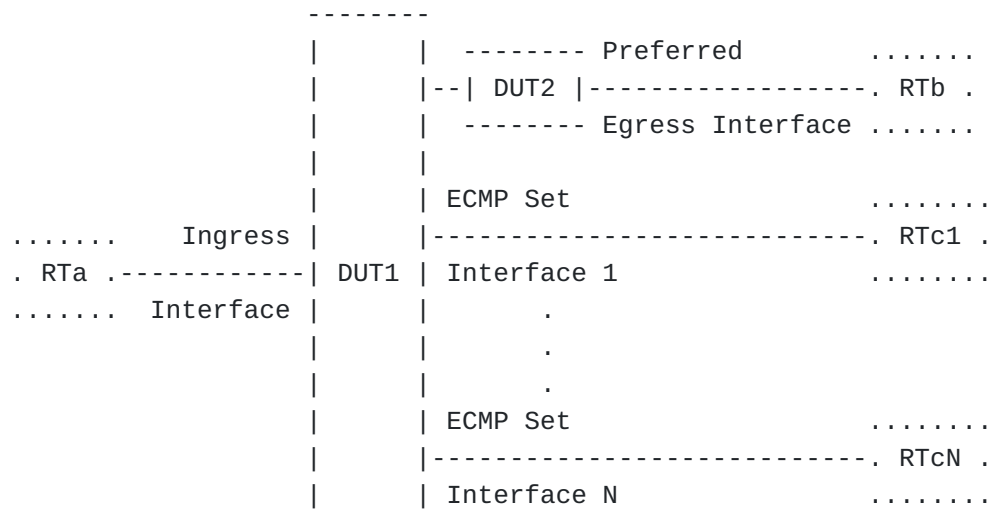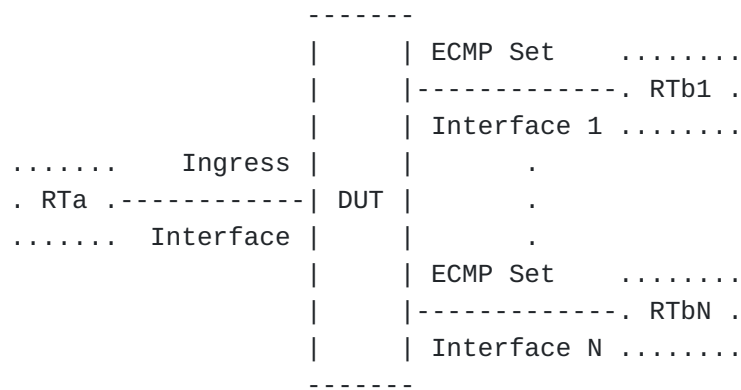occur transiently between DUT1 and DUT2 during convergence (micro-loop,
see [[Sh10]](#)). When the test specifies to observe the Next-Best Egress
Interface statistics, the combined statistics for all ECMP members
except the one affected by the Convergence Event, should be observed.

```
                     --------
                     |      | ECMP Set    --------   ........
                     |      |-------------| DUT2 |---. RTb1 .
                     |      | Interface 1 --------   ........
                     |      |
                     |      | ECMP Set              .......
.......   Ingress    |      |----------------------. RTb2 .
. RTa .--------------| DUT1 | Interface 2           ........
.......  Interface   |      |        .
                     |      |        .
                     |      |        .
                     |      | ECMP Set              ........
                     |      |----------------------. RTbN .
                     |      | Interface N           ........
                     --------
```
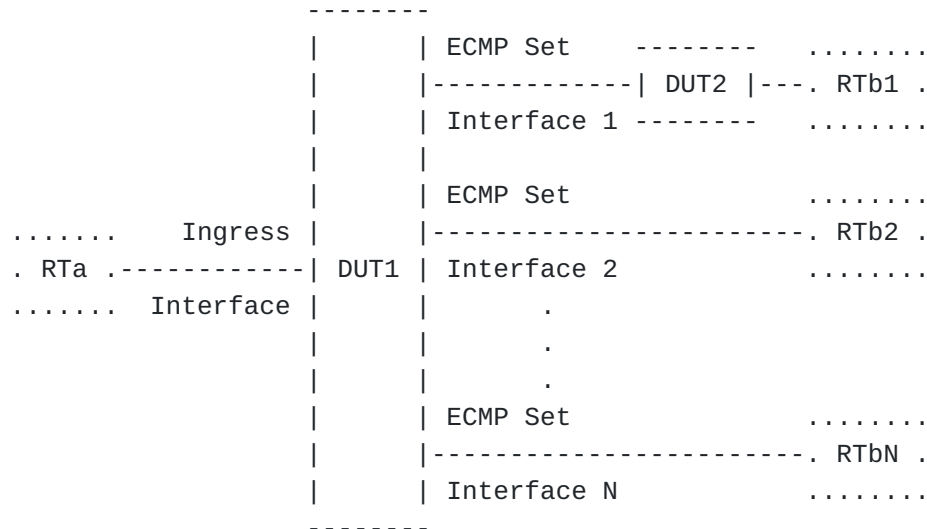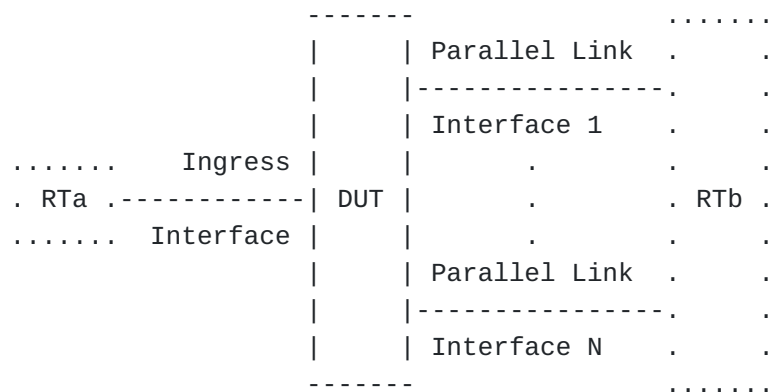
Test topology for Parallel Link changes

Figure shows the test topology to measure IGP convergence time due to
local Convergence Events with members of a Parallel Link (Section
8.1.5). In this topology, the DUT is configured with each egress
interface as a member of a Parallel Link and the Tester emulates two
neighbor routers, one for the Ingress Interface (RTa) and one for the
Parallel Link members (RTb). IGP adjacencies MUST be established on the
Ingress Interface and on all N members of the Parallel Link between
Tester and DUT (N>0). For this purpose the routers emulated by the
Tester establishes N+1 adjacencies with the DUT. When the test
specifies to observe the Next-Best Egress Interface statistics, the
combined statistics for all Parallel Link members except the one
affected by the Convergence Event, should be observed.

```
                     -------            .......
                     |     | Parallel Link  .     .
                     |     |---------------.     .
                     |     | Interface 1   .     .
.......   Ingress    |     |        .      .     .
. RTa .--------------| DUT |        .      . RTb .
.......  Interface   |     |        .      .     .
                     |     | Parallel Link  .     .
                     |     |---------------.     .
                     |     | Interface N   .     .
                     -------            .......
```

Convergence Time and Loss of Connectivity Period

Two concepts will be highlighted in this section: convergence time and
loss of connectivity period.
The Route Convergence [Po11t] time indicates the period in time between
the Convergence Event Instant [Po11t] and the instant in time the DUT

is ready to forward traffic for a specific route on its Next-Best
Egress Interface and maintains this state for the duration of the
Sustained Convergence Validation Time [Po11t]. To measure Route
Convergence time, the Convergence Event Instant and the traffic
received from the Next-Best Egress Interface need to be observed.
The Route Loss of Connectivity Period [Po11t] indicates the time during
which traffic to a specific route is lost following a Convergence Event
until Full Convergence [Po11t] completes. This Route Loss of
Connectivity Period can consist of one or more Loss Periods [Ko02]. For
the testcases described in this document it is expected to have a
single Loss Period. To measure Route Loss of Connectivity Period, the
traffic received from the Preferred Egress Interface and the traffic
received from the Next-Best Egress Interface need to be observed.
The Route Loss of Connectivity Period is most important since that has
a direct impact on the network user's application performance.
In general the Route Convergence time is larger than or equal to the
Route Loss of Connectivity Period. Depending on which Convergence Event
occurs and how this Convergence Event is applied, traffic for a route
may still be forwarded over the Preferred Egress Interface after the
Convergence Event Instant, before converging to the Next-Best Egress
Interface. In that case the Route Loss of Connectivity Period is
shorter than the Route Convergence time.
At least one condition needs to be fulfilled for Route Convergence time
to be equal to Route Loss of Connectivity Period. The condition is that
the Convergence Event causes an instantaneous traffic loss for the
measured route. A fiber cut on the Preferred Egress Interface is an
example of such a Convergence Event.
A second condition applies to Route Convergence time measurements based
on Connectivity Packet Loss [Po11t]. This second condition is that
there is only a single Loss Period during Route Convergence. For the
testcases described in this document this is expected to be the case.

4.1. Convergence Events without instant traffic loss

To measure convergence time benchmarks for Convergence Events caused by
a Tester, such as an IGP cost change, the Tester MAY start to discard
all traffic received from the Preferred Egress Interface at the
Convergence Event Instant, or MAY separately observe packets received
from the Preferred Egress Interface prior to the Convergence Event
Instant. This way these Convergence Events can be treated the same as
Convergence Events that cause instantaneous traffic loss.
To measure convergence time benchmarks without instantaneous traffic
loss (either real or induced by the Tester) at the Convergence Event
Instant, such as a reversion of a link failure Convergence Event, the
Tester SHALL only observe packet statistics on the Next-Best Egress
Interface. If using the Rate-Derived method to benchmark convergence
times for such Convergence Events, the Tester MUST collect a timestamp
at the Convergence Event Instant. If using a loss-derived method to
benchmark convergence times for such Convergence Events, the Tester

MUST measure the period in time between the Start Traffic Instant and
the Convergence Event Instant. To measure this period in time the
Tester can collect timestamps at the Start Traffic Instant and the
Convergence Event Instant.

The Convergence Event Instant together with the receive rate
observations on the Next-Best Egress Interface allow to derive the
convergence time benchmarks using the Rate-Derived Method [Po11t].

By observing packets on the Next-Best Egress Interface only, the
observed Impaired Packet count is the number of Impaired Packets
between Traffic Start Instant and Convergence Recovery Instant. To
measure convergence times using a loss-derived method, the Impaired
Packet count between the Convergence Event Instant and the Convergence
Recovery Instant is needed. The time between Traffic Start Instant and
Convergence Event Instant must be accounted for. An example may clarify
this.

Figure illustrates a Convergence Event without instantaneous traffic
loss for all routes. The top graph shows the Forwarding Rate over all
routes, the bottom graph shows the Forwarding Rate for a single route
Rta. Some time after the Convergence Event Instant, Forwarding Rate
observed on the Preferred Egress Interface starts to decrease. In the
example, route Rta is the first route to experience packet loss at time
Ta. Some time later, the Forwarding Rate observed on the Next-Best
Egress Interface starts to increase. In the example, route Rta is the
first route to complete convergence at time Ta'.
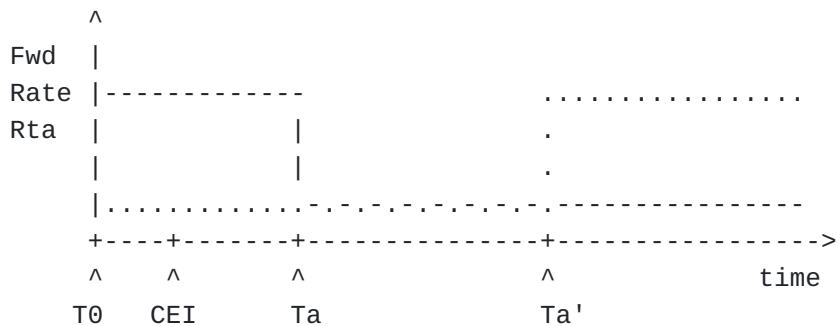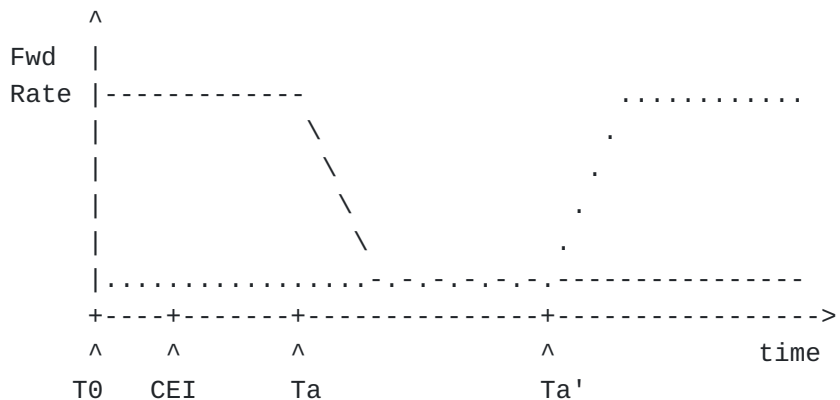
```
      ^
Fwd   |
Rate  |-------------                   ...........
      |             \                      .
      |              \                       .
      |               \                       .
      |                \                        .
      |..................................-.-.-.-.-.-.---------------
      +----+-------+--------------+----------------->
       ^    ^       ^              ^            time
      T0   CEI     Ta            Ta'


      ^
Fwd   |
Rate  |-------------              ................
Rta   |            |                  .
      |            |                  .
      |...........................-.-.-.-...-...-.---------------
      +----+-------+--------------+----------------->
       ^    ^       ^              ^            time
      T0   CEI     Ta            Ta'


      Preferred Egress Interface: ---
      Next-Best Egress Interface: ...
```

With T0 the Start Traffic Instant; CEI the Convergence Event Instant;
Ta the time instant packet loss for route Rta starts; Ta' the time
instant packet impairment for route Rta ends.
If only packets received on the Next-Best Egress Interface are
observed, the duration of the loss period for route Rta can be
calculated from the received packets as in Equation 1. Since the
Convergence Event Instant is the start time for convergence time
measurement, the period in time between T0 and CEI needs to be
subtracted from the calculated result to become the convergence time,
as in Equation 2.

```
Next-Best Egress Interface loss period
    = (packets transmitted
        - packets received from Next-Best Egress Interface) / tx rate
    = Ta' - T0
```

Equation 1

```
convergence time
    = Next-Best Egress Interface loss period - (CEI - T0)
    = Ta' - CEI
```

Equation 2

4.2. Loss of Connectivity (LoC)

Route Loss of Connectivity Period SHOULD be measured using the Route-
Specific Loss-Derived Method. Since the start instant and end instant
of the Route Loss of Connectivity Period can be different for each
route, these can not be accurately derived by only observing global
statistics over all routes. An example may clarify this.
Following a Convergence Event, route Rta is the first route for which
packet impairment starts, the Route Loss of Connectivity Period for
route Rta starts at time Ta. Route Rtb is the last route for which
packet impairment starts, the Route Loss of Connectivity Period for
route Rtb starts at time Tb with Tb>Ta.

```
      ^
Fwd   |
Rate  |--------                      -----------
      |        \                     /
      |         \                   /
      |          \                 /
      |           \               /
      |            ---------------
      +--------------------------------------->
              ^   ^              ^    ^     time
            Ta    Tb           Ta'   Tb'
                               Tb''  Ta''
```
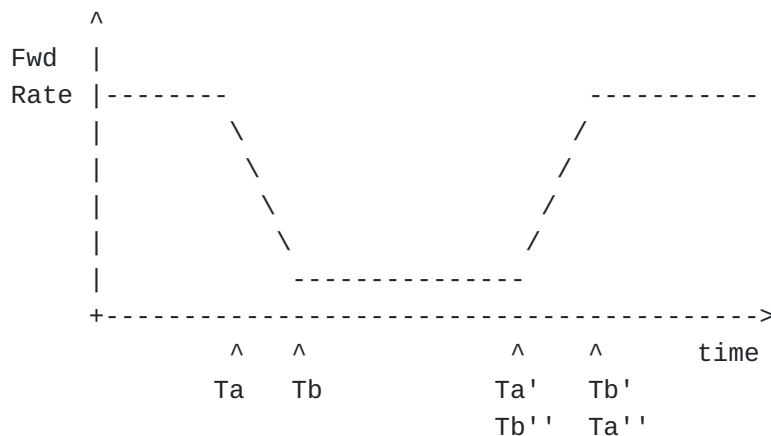
If the DUT implementation were such that route Rta would be the first
route for which traffic loss ends at time Ta' (with Ta'>Tb) and route
Rtb would be the last route for which traffic loss ends at time Tb'
(with Tb'>Ta'). By only observing global traffic statistics over all
routes, the minimum Route Loss of Connectivity Period would be measured
as Ta'-Ta. The maximum calculated Route Loss of Connectivity Period
would be Tb'-Ta. The real minimum and maximum Route Loss of
Connectivity Periods are Ta'-Ta and Tb'-Tb. Illustrating this with the
numbers Ta=0, Tb=1, Ta'=3, and Tb'=5, would give a Loss of Connectivity
Period between 3 and 5 derived from the global traffic statistics,
versus the real Loss of Connectivity Period between 3 and 4.
If the DUT implementation were such that route Rtb would be the first
for which packet loss ends at time Tb'' and route Rta would be the last
for which packet impairment ends at time Ta'', then the minimum and
maximum Route Loss of Connectivity Periods derived by observing only
global traffic statistics would be Tb''-Ta, and Ta''-Ta. The real
minimum and maximum Route Loss of Connectivity Periods are Tb''-Tb and
Ta''-Ta. Illustrating this with the numbers Ta=0, Tb=1, Ta''=5, Tb''=3,
would give a Loss of Connectivity Period between 3 and 5 derived from
the global traffic statistics, versus the real Loss of Connectivity
Period between 2 and 5.
The two implementation variations in the above example would result in
the same derived minimum and maximum Route Loss of Connectivity Periods
when only observing the global packet statistics, while the real Route
Loss of Connectivity Periods are different.

5. Test Considerations

5.1. IGP Selection

The test cases described in Section 8 can be used for link-state IGPs,
such as IS-IS or OSPF. The IGP convergence time test methodology is
identical.

5.2. Routing Protocol Configuration

The obtained results for IGP convergence time may vary if other routing
protocols are enabled and routes learned via those protocols are
installed. IGP convergence times SHOULD be benchmarked without routes
installed from other protocols. Any enabled IGP routing protocol
extension (such as extensions for Traffic Engineering) and any enabled
IGP routing protocol security mechanism must be reported with the
results.

5.3. IGP Topology

The Tester emulates a single IGP topology. The DUT establishes IGP
adjacencies with one or more of the emulated routers in this single IGP
topology emulated by the Tester. See test topology details in Section

[3](). The emulated topology SHOULD only be advertised on the DUT egress
interfaces.
The number of IGP routes and number of nodes in the topology, and the
type of topology will impact the measured IGP convergence time. To
obtain results similar to those that would be observed in an
operational network, it is RECOMMENDED that the number of installed
routes and nodes closely approximate that of the network (e.g.
thousands of routes with tens or hundreds of nodes).
The number of areas (for OSPF) and levels (for IS-IS) can impact the
benchmark results.

5.4. Timers

There are timers that may impact the measured IGP convergence times.
The benchmark metrics MAY be measured at any fixed values for these
timers. To obtain results similar to those that would be observed in an
operational network, it is RECOMMENDED to configure the timers with the
values as configured in the operational network.
Examples of timers that may impact measured IGP convergence time
include, but are not limited to:

    *Interface failure indication

    *IGP hello timer

    *IGP dead-interval or hold-timer

    *Link State Advertisement (LSA) or Link State Packet (LSP)
     generation delay

    *LSA or LSP flood packet pacing

    *route calculation delay

5.5. Interface Types

All test cases in this methodology document can be executed with any
interface type. The type of media may dictate which test cases may be
executed. Each interface type has a unique mechanism for detecting link
failures and the speed at which that mechanism operates will influence
the measurement results. All interfaces MUST be the same media and
Throughput [Br91][Br99] for each test case. All interfaces SHOULD be
configured as point-to-point.

5.6. Offered Load

The Throughput of the device, as defined in [Br91] and benchmarked in
[Br99] at a fixed packet size, needs to be determined over the
preferred path and over the next-best path. The Offered Load SHOULD be
the minimum of the measured Throughput of the device over the primary

path and over the backup path. The packet size is selectable and MUST
be recorded. Packet size is measured in bytes and includes the IP
header and payload.
The destination addresses for the Offered Load MUST be distributed such
that all routes or a statistically representative subset of all routes
are matched and each of these routes is offered an equal share of the
Offered Load. It is RECOMMENDED to send traffic matching all routes,
but a statistically representative subset of all routes can be used if
required.
Splitting traffic flows across multiple paths (as with ECMP or Parallel
Link sets) is in general done by hashing on various fields on the IP or
contained headers. The hashing is typically based on the IP source and
destination addresses, the protocol ID, and higher-layer flow-dependent
fields such as TCP/UDP ports. In practice, within a network core, the
hashing is based mainly or exclusively on the IP source and destination
addresses. Knowledge of the hashing algorithm used by the DUT is not
always possible beforehand, and would violate the black-box spirit of
this document. Therefor it is RECOMMENDED to use a randomly distributed
range of source and destination IP addresses, protocol IDs, and higher-
layer flow-dependent fields for the packets of the Offered Load (see
also [Ne07]). The content of the Offered Load MUST remain the same
during the test. It is RECOMMENDED to repeat a test multiple times with
different random ranges of the header fields such that convergence time
benchmarks are measured for different distributions of traffic over the
available paths.
In the Remote Interface failure testcases using topologies , ,
and there is a possibility of a packet forwarding loop that may occur
transiently between DUT1 and DUT2 during convergence (micro-loop, see
[Sh10]). The Time To Live (TTL) or Hop Limit value of the packets sent
by the Tester may influence the benchmark measurements since it
determines which device in the topology may send an ICMP Time Exceeded
Message for looped packets.
The duration of the Offered Load MUST be greater than the convergence
time plus the Sustained Convergence Validation Time.
Offered load should send a packet to each destination before sending
another packet to the same destination. It is RECOMMENDED that the
packets be transmitted in a round-robin fashion with a uniform
interpacket delay.

5.7. Measurement Accuracy

Since Impaired Packet count is observed to measure the Route
Convergence Time, the time between two successive packets offered to
each individual route is the highest possible accuracy of any Impaired
Packet based measurement. The higher the traffic rate offered to each
route the higher the possible measurement accuracy.
Also see Section 6 for method-specific measurement accuracy.

5.8. Measurement Statistics

The benchmark measurements may vary for each trial, due to the statistical nature of timer expirations, cpu scheduling, etc. Evaluation of the test data must be done with an understanding of generally accepted testing practices regarding repeatability, variance and statistical significance of a small number of trials.

5.9. Tester Capabilities

It is RECOMMENDED that the Tester used to execute each test case have the following capabilities:

1. Ability to establish IGP adjacencies and advertise a single IGP topology to one or more peers.

2. Ability to measure Forwarding Delay, Duplicate Packets and Out-of-Order Packets.

3. An internal time clock to control timestamping, time measurements, and time calculations.

4. Ability to distinguish traffic load received on the Preferred and Next-Best Interfaces [Po11t].

5. Ability to disable or tune specific Layer-2 and Layer-3 protocol functions on any interface(s).

The Tester MAY be capable to make non-data plane convergence observations and use those observations for measurements. The Tester MAY be capable to send and receive multiple traffic Streams [Po06]. Also see Section 6 for method-specific capabilities.

6. Selection of Convergence Time Benchmark Metrics and Methods

Different convergence time benchmark methods MAY be used to measure convergence time benchmark metrics. The Tester capabilities are important criteria to select a specific convergence time benchmark method. The criteria to select a specific benchmark method include, but are not limited to:

Tester capabilities:	Sampling Interval, number of Stream statistics to collect
Measurement accuracy:	Sampling Interval, Offered Load, number of routes
Test specification:	number of routes
DUT capabilities:	Throughput, IP Packet Delay Variation

6.1. Loss-Derived Method

6.1.1. Tester capabilities

To enable collecting statistics of Out-of-Order Packets per flow (See
[Th00], Section 3) the Offered Load SHOULD consist of multiple Streams
[Po06] and each Stream SHOULD consist of a single flow . If sending
multiple Streams, the measured traffic statistics for all Streams MUST
be added together.
In order to verify Full Convergence completion and the Sustained
Convergence Validation Time, the Tester MUST measure Forwarding Rate
each Packet Sampling Interval.
The total number of Impaired Packets between the start of the traffic
and the end of the Sustained Convergence Validation Time is used to
calculate the Loss-Derived Convergence Time.

6.1.2. Benchmark Metrics

The Loss-Derived Method can be used to measure the Loss-Derived
Convergence Time, which is the average convergence time over all
routes, and to measure the Loss-Derived Loss of Connectivity Period,
which is the average Route Loss of Connectivity Period over all routes.

6.1.3. Measurement Accuracy

The actual value falls within the accuracy interval [-(number of
destinations/Offered Load), +(number of destinations/Offered Load)]
around the value as measured using the Loss-Derived Method.

6.2. Rate-Derived Method

6.2.1. Tester Capabilities

To enable collecting statistics of Out-of-Order Packets per flow (See
[Th00], Section 3) the Offered Load SHOULD consist of multiple Streams
[Po06] and each Stream SHOULD consist of a single flow . If sending
multiple Streams, the measured traffic statistics for all Streams MUST
be added together.
The Tester measures Forwarding Rate each Sampling Interval. The Packet
Sampling Interval influences the observation of the different
convergence time instants. If the Packet Sampling Interval is large
compared to the time between the convergence time instants, then the
different time instants may not be easily identifiable from the
Forwarding Rate observation. The presence of IP Packet Delay Variation
(IPDV) [De02] may cause fluctuations of the Forwarding Rate observation
and can prevent correct observation of the different convergence time
instants.
The Packet Sampling Interval MUST be larger than or equal to the time
between two consecutive packets to the same destination. For maximum
accuracy the value for the Packet Sampling Interval SHOULD be as small

as possible, but the presence of IPDV may enforce using a larger Packet
Sampling Interval. The Packet Sampling Interval MUST be reported.
IPDV causes fluctuations in the number of received packets during each
Packet Sampling Interval. To account for the presence of IPDV in
determining if a convergence instant has been reached, Forwarding Delay
SHOULD be observed during each Packet Sampling Interval. The minimum
and maximum number of packets expected in a Packet Sampling Interval in
presence of IPDV can be calculated with Equation 3.

```
number of packets expected in a Packet Sampling Interval
   in presence of IP Packet Delay Variation
     = expected number of packets without IP Packet Delay Variation
       +/-( (maxDelay - minDelay) * Offered Load)
with minDelay and maxDelay the minimum resp. maximum Forwarding Delay
   of packets received during the Packet Sampling Interval
```

Equation 3
To determine if a convergence instant has been reached the number of
packets received in a Packet Sampling Interval is compared with the
range of expected number of packets calculated in Equation 3.

6.2.2. Benchmark Metrics

The Rate-Derived Method SHOULD be used to measure First Route
Convergence Time and Full Convergence Time. It SHOULD NOT be used to
measure Loss of Connectivity Period (see Section 4).

6.2.3. Measurement Accuracy

The measurement accuracy interval of the Rate-Derived Method depends on
the metric being measured or calculated and the characteristics of the
related transition. IP Packet Delay Variation (IPDV) [De02] adds
uncertainty to the amount of packets received in a Packet Sampling
Interval and this uncertainty adds to the measurement error. The effect
of IPDV is not accounted for in the calculation of the accuracy
intervals below. IPDV is of importance for the convergence instants
were a variation in Forwarding Rate needs to be observed (Convergence
Recovery Instant and for topologies with ECMP also Convergence Event
Instant and First Route Convergence Instant).
If the Convergence Event Instant is observed on the dataplane using the
Rate Derived Method, it needs to be instantaneous for all routes (see
Section 4.1). The actual value of the Convergence Event Instant falls
within the accuracy interval [–(Packet Sampling Interval + 1/Offered
Load), +0] around the value as measured using the Rate-Derived Method.
If the Convergence Recovery Transition is non-instantaneous for all
routes then the actual value of the First Route Convergence Instant
falls within the accuracy interval [–(Packet Sampling Interval + time
between two consecutive packets to the same destination), +0] around
the value as measured using the Rate-Derived Method, and the actual

value of the Convergence Recovery Instant falls within the accuracy interval [–(2 * Packet Sampling Interval), –(Packet Sampling Interval - time between two consecutive packets to the same destination)] around the value as measured using the Rate-Derived Method.

The term "time between two consecutive packets to the same destination" is added in the above accuracy intervals since packets are sent in a particular order to all destinations in a stream and when part of the routes experience packet loss, it is unknown where in the transmit cycle packets to these routes are sent. This uncertainty adds to the error.

The accuracy intervals of the derived metrics First Route Convergence Time and Rate-Derived Convergence Time are calculated from the above convergence instants accuracy intervals. The actual value of First Route Convergence Time falls within the accuracy interval [–(Packet Sampling Interval + time between two consecutive packets to the same destination), +(Packet Sampling Interval + 1/Offered Load)] around the calculated value. The actual value of Rate-Derived Convergence Time falls within the accuracy interval [–(2 * Packet Sampling Interval), + (time between two consecutive packets to the same destination + 1/ Offered Load)] around the calculated value.

6.3. Route-Specific Loss-Derived Method

6.3.1. Tester Capabilities

The Offered Load consists of multiple Streams. The Tester MUST measure Impaired Packet count for each Stream separately.

In order to verify Full Convergence completion and the Sustained Convergence Validation Time, the Tester MUST measure Forwarding Rate each Packet Sampling Interval. This measurement at each Packet Sampling Interval MAY be per Stream.

Only the total number of Impaired Packets measured per Stream at the end of the Sustained Convergence Validation Time is used to calculate the benchmark metrics with this method.

6.3.2. Benchmark Metrics

The Route-Specific Loss-Derived Method SHOULD be used to measure Route-Specific Convergence Times. It is the RECOMMENDED method to measure Route Loss of Connectivity Period.

Under the conditions explained in Section 4, First Route Convergence Time and Full Convergence Time as benchmarked using Rate-Derived Method, may be equal to the minimum resp. maximum of the Route-Specific Convergence Times.

6.3.3. Measurement Accuracy

The actual value falls within the accuracy interval [-(number of destinations/Offered Load), +(number of destinations/Offered Load)]

around the value as measured using the Route-Specific Loss-Derived
Method.

7. Reporting Format

For each test case, it is RECOMMENDED that the reporting tables below
be completed and all time values SHOULD be reported with a sufficiently
high resolution.

Parameter	Units
Test Case	test case number
Test Topology	Test Topology Figure number
IGP	(IS-IS, OSPF, other)
Interface Type	(GigE, POS, ATM, other)
Packet Size offered to DUT	bytes
Offered Load	packets per second
IGP Routes advertised to DUT	number of IGP routes
Nodes in emulated network	number of nodes
Number of Parallel or ECMP links	number of links
Number of Routes measured	number of routes
Packet Sampling Interval on Tester	seconds
Forwarding Delay Threshold	seconds
Timer Values configured on DUT:	
Interface failure indication delay	seconds
IGP Hello Timer	seconds
IGP Dead-Interval or hold-time	seconds
LSA/LSP Generation Delay	seconds
LSA/LSP Flood Packet Pacing	seconds
LSA/LSP Retransmission Packet Pacing	seconds
route calculation Delay	seconds

Test Details:

   *Describe the IGP extensions and IGP security mechanisms that are
    configured on the DUT.

   *Describe how the various fields on the IP and contained headers
    for the packets for the Offered Load are generated (Section 5.6).

   *If the Offered Load matches a subset of routes, describe how this
    subset is selected.

*Describe how the Convergence Event is applied; does it cause
  instantaneous traffic loss or not?

The table below should be completed for the initial Convergence Event
and the reversion Convergence Event.

Parameter	Units
Convergence Event	(initial or reversion)
Traffic Forwarding Metrics:	
Total number of packets offered to DUT	number of Packets
Total number of packets forwarded by DUT	number of Packets
Connectivity Packet Loss	number of Packets
Convergence Packet Loss	number of Packets
Out-of-Order Packets	number of Packets
Duplicate Packets	number of Packets
excessive Forwarding Delay Packets	number of Packets
Convergence Benchmarks:	
Rate-Derived Method:	
First Route Convergence Time	seconds
Full Convergence Time	seconds
Loss-Derived Method:	
Loss-Derived Convergence Time	seconds
Route-Specific Loss-Derived Method:	
Route-Specific Convergence Time[n]	array of seconds
Minimum Route-Specific Convergence Time	seconds
Maximum Route-Specific Convergence Time	seconds
Median Route-Specific Convergence Time	seconds
Average Route-Specific Convergence Time	seconds
Loss of Connectivity Benchmarks:	
Loss-Derived Method:	
Loss-Derived Loss of Connectivity Period	seconds
Route-Specific Loss-Derived Method:	
Route Loss of Connectivity Period[n]	array of seconds
Minimum Route Loss of Connectivity Period	seconds
Maximum Route Loss of Connectivity Period	seconds
Median Route Loss of Connectivity Period	seconds

Parameter	Units
Average Route Loss of Connectivity Period	seconds

8. Test Cases

It is RECOMMENDED that all applicable test cases be performed for best characterization of the DUT. The test cases follow a generic procedure tailored to the specific DUT configuration and Convergence Event [Po11t]. This generic procedure is as follows:

1. Establish DUT and Tester configurations and advertise an IGP topology from Tester to DUT.

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is routed correctly. Verify if traffic is forwarded without Impaired Packets [Po06].

4. Introduce Convergence Event [Po11t].

5. Measure First Route Convergence Time [Po11t].

6. Measure Full Convergence Time [Po11t].

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period [Po11t]. At the same time measure number of Impaired Packets [Po11t].

9. Wait sufficient time for queues to drain. The duration of this time period MUST be larger than or equal to the Forwarding Delay Threshold.

10. Restart Offered Load.

11. Reverse Convergence Event.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets [Po11t].

[8.1.](#) **Interface Failure and Recovery**

[8.1.1.](#) **[Convergence Due to Local Interface Failure and Recovery](#)**

Objective
To obtain the IGP convergence measurements for Local Interface failure and recovery events. The Next-Best Egress Interface can be a single interface (Figure) or an ECMP set (Figure). The test with ECMP topology (Figure) is OPTIONAL.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology shown in Figure or .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is forwarded over Preferred Egress Interface.

4. Remove link on the Preferred Egress Interface of the DUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times and Loss-Derived Convergence Time. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore link on the Preferred Egress Interface of the DUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period.At the same time measure number of Impaired Packets.

Objective
To obtain the IGP convergence measurements for Remote Interface failure and recovery events. The Next-Best Egress Interface can be a single interface (Figure) or an ECMP set (Figure). The test with ECMP topology (Figure) is OPTIONAL.
Procedure

1. Advertise an IGP topology from Tester to SUT using the topology shown in Figure or .

2. Send Offered Load from Tester to SUT on ingress interface.

3. Verify traffic is forwarded over Preferred Egress Interface.

4. Remove link on the interface of the Tester connected to the Preferred Egress Interface of the SUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times and Loss-Derived Convergence Time. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore link on the interface of the Tester connected to the Preferred Egress Interface of the SUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Discussion

In this test case there is a possibility of a packet forwarding loop that may occur transiently between DUT1 and DUT2 during convergence (micro-loop, see [Sh10]), which may increase the measured convergence times and loss of connectivity periods.

8.1.3. Convergence Due to ECMP Member Local Interface Failure and Recovery

Objective
To obtain the IGP convergence measurements for Local Interface link failure and recovery events of an ECMP Member.
Procedure

1. Advertise an IGP topology from Tester to DUT using the test setup shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is forwarded over the ECMP member interface of the DUT that will be failed in the next step.

4. Remove link on one of the ECMP member interfaces of the DUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times and Loss-Derived Convergence Time. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore link on the ECMP member interface of the DUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Objective
To obtain the IGP convergence measurements for Remote Interface link failure and recovery events for an ECMP Member.
Procedure

1. Advertise an IGP topology from Tester to DUT using the test setup shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is forwarded over the ECMP member interface of the DUT that will be failed in the next step.

4. Remove link on the interface of the Tester to R2. This is the Convergence Event Trigger.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times and Loss-Derived Convergence Time. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore link on the interface of the Tester to R2.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Discussion
In this test case there is a possibility of a packet forwarding loop that may occur temporarily between DUT1 and DUT2 during convergence

(micro-loop, see [Sh10]), which may increase the measured convergence times and loss of connectivity periods.

8.1.5. Convergence Due to Parallel Link Interface Failure and Recovery

Objective
To obtain the IGP convergence measurements for local link failure and recovery events for a member of a parallel link. The links can be used for data load balancing
Procedure

1. Advertise an IGP topology from Tester to DUT using the test setup shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is forwarded over the parallel link member that will be failed in the next step.

4. Remove link on one of the parallel link member interfaces of the DUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times and Loss-Derived Convergence Time. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore link on the Parallel Link member interface of the DUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Other Failures and Recoveries

Convergence Due to Layer 2 Session Loss and Recovery

Objective
To obtain the IGP convergence measurements for a local layer 2 loss and recovery.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is routed over Preferred Egress Interface.

4. Remove Layer 2 session from Preferred Egress Interface of the DUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore Layer 2 session on Preferred Egress Interface of the DUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Discussion

When removing the layer 2 session, the physical layer must stay up. Configure IGP timers such that the IGP adjacency does not time out before layer 2 failure is detected.

To measure convergence time, traffic SHOULD start dropping on the Preferred Egress Interface on the instant the layer 2 session is removed. Alternatively the Tester SHOULD record the time the instant layer 2 session is removed and traffic loss SHOULD only be measured on the Next-Best Egress Interface. For loss-derived benchmarks the time of the Start Traffic Instant SHOULD be recorded as well. See Section 4.1.

8.2.2. Convergence Due to Loss and Recovery of IGP Adjacency

Objective
To obtain the IGP convergence measurements for loss and recovery of an IGP Adjacency. The IGP adjacency is removed on the Tester by disabling processing of IGP routing protocol packets on the Tester.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is routed over Preferred Egress Interface.

4. Remove IGP adjacency from the Preferred Egress Interface while the layer 2 session MUST be maintained. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Restore IGP session on Preferred Egress Interface of the DUT.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived
    Convergence Time, Route Loss of Connectivity Periods, and Loss-
    Derived Loss of Connectivity Period. At the same time measure
    number of Impaired Packets.

Discussion
Configure layer 2 such that layer 2 does not time out before IGP
adjacency failure is detected.
To measure convergence time, traffic SHOULD start dropping on the
Preferred Egress Interface on the instant the IGP adjacency is removed.
Alternatively the Tester SHOULD record the time the instant the IGP
adjacency is removed and traffic loss SHOULD only be measured on the
Next-Best Egress Interface. For loss-derived benchmarks the time of the
Start Traffic Instant SHOULD be recorded as well. See Section 4.1.

8.2.3. Convergence Due to Route Withdrawal and Re-advertisement

Objective
To obtain the IGP convergence measurements for route withdrawal and re-
advertisement.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology
   shown in Figure . The routes that will be withdrawn MUST be a
   set of leaf routes advertised by at least two nodes in the
   emulated topology. The topology SHOULD be such that before the
   withdrawal the DUT prefers the leaf routes advertised by a node
   "nodeA" via the Preferred Egress Interface, and after the
   withdrawal the DUT prefers the leaf routes advertised by a node
   "nodeB" via the Next-Best Egress Interface.

2. Send Offered Load from Tester to DUT on Ingress Interface.

3. Verify traffic is routed over Preferred Egress Interface.

4. The Tester withdraws the set of IGP leaf routes from nodeA.
   This is the Convergence Event. The withdrawal update message
   SHOULD be a single unfragmented packet. If the routes cannot be
   withdrawn by a single packet, the messages SHOULD be sent using
   the same pacing characteristics as the DUT. The Tester MAY
   record the time it sends the withdrawal message(s).

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived
    Convergence Time, Route Loss of Connectivity Periods, and Loss-
    Derived Loss of Connectivity Period. At the same time measure
    number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Re-advertise the set of withdrawn IGP leaf routes from nodeA
    emulated by the Tester. The update message SHOULD be a single
    unfragmented packet. If the routes cannot be advertised by a
    single packet, the messages SHOULD be sent using the same
    pacing characteristics as the DUT. The Tester MAY record the
    time it sends the update message(s).

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived
    Convergence Time, Route Loss of Connectivity Periods, and Loss-
    Derived Loss of Connectivity Period. At the same time measure
    number of Impaired Packets.

Discussion
To measure convergence time, traffic SHOULD start dropping on the
Preferred Egress Interface on the instant the routes are withdrawn by
the Tester. Alternatively the Tester SHOULD record the time the instant
the routes are withdrawn and traffic loss SHOULD only be measured on
the Next-Best Egress Interface. For loss-derived benchmarks the time of
the Start Traffic Instant SHOULD be recorded as well. See Section 4.1.

8.3. Administrative changes

8.3.1. Convergence Due to Local Interface Adminstrative Changes

Objective
To obtain the IGP convergence measurements for administratively
disabling and enabling a Local Interface.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology
    shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is routed over Preferred Egress Interface.

4. Administratively disable the Preferred Egress Interface of the DUT. This is the Convergence Event.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. Administratively enable the Preferred Egress Interface of the DUT..

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

8.3.2. Convergence Due to Cost Change

Objective
To obtain the IGP convergence measurements for route cost change.
Procedure

1. Advertise an IGP topology from Tester to DUT using the topology shown in Figure .

2. Send Offered Load from Tester to DUT on ingress interface.

3. Verify traffic is routed over Preferred Egress Interface.

4. The Tester, emulating the neighbor node, increases the cost for all IGP routes at Preferred Egress Interface of the DUT so that the Next-Best Egress Interface becomes preferred path. The update message advertising the higher cost MUST be a single unfragmented packet. This is the Convergence Event. The Tester

MAY record the time it sends the update message advertising the higher cost on the Preferred Egress Interface.

5. Measure First Route Convergence Time.

6. Measure Full Convergence Time.

7. Stop Offered Load.

8. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

9. Wait sufficient time for queues to drain.

10. Restart Offered Load.

11. The Tester, emulating the neighbor node, decreases the cost for all IGP routes at Preferred Egress Interface of the DUT so that the Preferred Egress Interface becomes preferred path. The update message advertising the lower cost MUST be a single unfragmented packet.

12. Measure First Route Convergence Time.

13. Measure Full Convergence Time.

14. Stop Offered Load.

15. Measure Route-Specific Convergence Times, Loss-Derived Convergence Time, Route Loss of Connectivity Periods, and Loss-Derived Loss of Connectivity Period. At the same time measure number of Impaired Packets.

Discussion
To measure convergence time, traffic SHOULD start dropping on the Preferred Egress Interface on the instant the cost is changed by the Tester. Alternatively the Tester SHOULD record the time the instant the cost is changed and traffic loss SHOULD only be measured on the Next-Best Egress Interface. For loss-derived benchmarks the time of the Start Traffic Instant SHOULD be recorded as well. See Section 4.1.

9. Security Considerations

Benchmarking activities as described in this memo are limited to technology characterization using controlled stimuli in a laboratory environment, with dedicated address space and the constraints specified in the sections above.

The benchmarking network topology will be an independent test setup and
MUST NOT be connected to devices that may forward the test traffic into
a production network, or misroute traffic to the test management
network.
Further, benchmarking is performed on a "black-box" basis, relying
solely on measurements observable external to the DUT/SUT.
Special capabilities SHOULD NOT exist in the DUT/SUT specifically for
benchmarking purposes. Any implications for network security arising
from the DUT/SUT SHOULD be identical in the lab and in production
networks.

10. IANA Considerations

This document requires no IANA considerations.

11. Acknowledgements

Thanks to Sue Hares, Al Morton, Kevin Dubray, Ron Bonica, David Ward,
Peter De Vriendt, Anuj Dewagan, Julien Meuric, Adrian Farrel, Stewart
Bryant, and the Benchmarking Methodology Working Group for their
contributions to this work.

12. References

12.1. Normative References

[Br91]	Bradner, S., "Benchmarking terminology for network interconnection devices", RFC 1242, July 1991.
[Br97]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
[Br99]	Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.
[Ca90]	Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
[Ma98]	Mandeville, R., "Benchmarking Terminology for LAN Switching Devices", RFC 2285, February 1998.
[Mo98]	Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
[Po06]	Poretsky, S., Perser, J., Erramilli, S. and S. Khurana, "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", RFC 4689, October 2006.
[Co08]	Coltun, R., Ferguson, D., Moy, J. and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
[Ho08]	Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.
[Ko02]	Koodli, R. and R. Ravikanth, "One-way Loss Pattern Sample Metrics", RFC 3357, August 2002.
[Po11t]	Poretsky, S, Imhoff, B and K Michielsen, "Terminology for Benchmarking Link-State IGP Data Plane Route

	Convergence", Internet-Draft draft-ietf-bmwg-igp-dataplane-conv-term-23, January 2011.
[De02]	Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, November 2002.
[Sh10]	Shand, M. and S. Bryant, "A Framework for Loop-Free Convergence", RFC 5715, January 2010.
[Ne07]	Newman, D. and T. Player, "Hash and Stuffing: Overlooked Factors in Network Device Benchmarking", RFC 4814, March 2007.
[Ma05]	Manral, V., White, R. and A. Shaikh, "Benchmarking Basic OSPF Single Router Control Plane Convergence", RFC 4061, April 2005.
[Ma05t]	Manral, V., White, R. and A. Shaikh, "OSPF Benchmarking Terminology and Concepts", RFC 4062, April 2005.
[Ma05c]	Manral, V., White, R. and A. Shaikh, "Considerations When Using Basic OSPF Convergence Benchmarks", RFC 4063, April 2005.
[Sh10i]	Shand, M. and S. Bryant, "IP Fast Reroute Framework", RFC 5714, January 2010.
[Pa05]	Pan, P., Swallow, G. and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.
[Th00]	Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", RFC 2991, November 2000.

12.2. Informative References

[Al00]	Alaettinoglu, C., Jacobson, V. and H. Yu, "Towards Millisecond IGP Convergence", NANOG 20, October 2000.
[Al02]	Alaettinoglu, C. and S. Casner, "ISIS Routing on the Qwest Backbone: a Recipe for Subsecond ISIS Convergence", NANOG 24, February 2002.
[Fi02]	Filsfils, C., "Tutorial: Deploying Tight-SLA Services on an Internet Backbone: ISIS Fast Convergence and Differentiated Services Design", NANOG 25, June 2002.
[Ka02]	Katz, D., "Why are we scared of SPF? IGP Scaling and Stability", NANOG 25, June 2002.
[Vi02]	Villamizar, C., "Convergence and Restoration Techniques for ISP Interior Routing", NANOG 25, June 2002.
[Fr05]	Francois, P., Filsfils, C., Evans, J. and O. Bonaventure, "Achieving SubSecond IGP Convergence in Large IP Networks", ACM SIGCOMM Computer Communication Review v.35 n.3, July 2005.

Authors' Addresses

Scott Poretsky Poretsky Allot Communications 67 South Bedford Street, Suite 400 Burlington, MA 01803 USA Phone: + 1 508 309 2179 EMail: sporetsky@allot.com

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

Kris Michielsen Michielsen Cisco Systems 6A De Kleetlaan Diegem, BRABANT 1831 Belgium EMail: kmichiel@cisco.com