

Network Working Group
Internet Draft

Vishwas Manral
Netplane Systems
Russ White
Cisco Systems
Aman Shaikh

Expiration Date: June 2003

University of California

File Name: [draft-ietf-bmwg-ospfconv-intraarea-03.txt](#)

January 2003

Benchmarking Methodology for Basic OSPF Convergence
draft-ietf-bmwg-ospfconv-intraarea-03.txt

1. 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. Internet Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "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>.

2. Abstract

This draft establishes standards for measuring OSPF convergence performance on a single router (SR-Convergence [4]). Its initial emphasis is on the control plane of single OSPF routers. We do not address forwarding plane performance.

3. Conventions used in this document

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 [1].

4. Motivation

There is a growing interest in routing protocol SR-Convergence testing, with many people looking at various tests to determine how long it takes for a network to converge after various conditions occur. The major problem with this sort of testing is that the framework of the tests has a major impact on the results; for instance, determining when a network is converged, what parts of the router's operation are considered within the testing, and other such things will have a major impact on what apparent performance routing protocols provide.

This document attempts to provide a framework within which Open Shortest Path First (OSPF) [2] performance testing can be placed, and provide some tests with which some aspects of OSPF performance can be measured. The motivation of the draft is to provide a set of tests that can provide the user comparable data from various vendors with which to evaluate the OSPF protocol performance on the devices.

5. Overview & Scope

While this document describes a specific set of tests aimed at characterizing the SR-Convergence performance of OSPF processes in routers or other boxes that incorporate OSPF functionality, a key objective is to propose methodology that will standardize the conducting and reporting of convergence-related measurements.

Things which are outside the scope of this document include:

- o The interactions of SR-Convergence and forwarding; testing is restricted to events occurring within the control plane. Forwarding performance is the primary focus in [3] and it is expected to be dealt with in work that ensues from [6].
- o Inter area routing, AS-external routing, and simultaneous traffic on the control and data paths within the DUT.

Other drafts in the future may cover some of the items noted as not covered in the scope of this draft. For a discussion of the

terminology used in this draft (in relation to the tests themselves), refer to [4]. For a discussion of the applicability of this draft, refer to [8].

6. Test Conditions

In all tests, the following test conditions will be assumed:

- o The link speed should be high enough so that does not become a bottleneck. Link speeds of 10MBps or higher are recommended. The link speed between routers should be specified in the test report.
- o For all point-to-point links, it is assumed that a link failure results in an immediate notification to the operating system, and thus to the OSPF process; this is explained thoroughly in [5].
- o No data traffic will be running between the routers during these tests.
- o Optional capabilities which can reduce performance, such as authentication, should be noted in the test results if they are enabled.
- o Optional changes in the default timer values, such as the SPF, hello, router dead, and other intervals, should be noted in the test results.
- o All places where injecting a set of LSAs is referenced, the set can include varying numbers of LSAs of varying types representing a varying number of reachable destinations. See [4] for further information about issues with LSA sets and network topologies.

Tests should be run more than once, since a single test run cannot be relied on to produce statistically sound results. The number of test runs and any variations between the tests should be recorded in the test results (see [4] for more information on what items should be recorded in the test results).

7. Reference Topologies

Several reference topologies will be used throughout the tests described in the remainder of this document. Rather than repeating these topologies, we've gathered them all in one section.

o Reference Topology 1

```

              (
DUT----Generator----( emulated topology )
              (

```

A simple back-to-back configuration. It's assumed that the link between the generator and the DUT is a point-to-point link, while the connections within the generator represent some emulated topology.

o Reference Topology 2

```

              (
Collector-----DUT-----Generator--( emulated topology )
      \                /              (
      \-----/

```

All routers are connected through point-to-point links. The cost of all links is assumed to be the same unless otherwise noted.

o Reference Topology 3

```

DUT      R1      R2
|         |         |
+-+-----+-----+-----+-----+

```

Any number of routers could be included on the common broadcast network.

o Reference Topology 4

```

  /--(link 1)-----\
DUT              Generator--( emulated topology )
  \--(link 2)-----/

```

In all cases the tests and topologies are designed to allow performance measurements to be taken all on a single device, whether the DUT or some other device in the network. This eliminates the need for synchronized clocks within the test networks.

8. Basic Process Performance Tests

These tests will measure aspects of the OSPF implementation as a process on the device under test, including:

- o Time required to process an LSA
- o Flooding time
- o Shortest Path First computation

8.1. Time required to process an LSA

- o Using reference topology 1, begin with all links up and a full adjacency established between the DUT and the generator.
- o Send an LSA that is already there in the DUT (a duplicate LSA), note the time difference between when the LSA is sent to when the ack is received. This measures the time to propagate the LSA and the ack, as well as processing time of the duplicate LSA. This is dupLSAprocTime.
- o Send a new LSA from the generator to the DUT, followed immediately by a duplicate LSA (LSA that already resides in the database of DUT, but not the same as the one just sent).
- o The DUT will acknowledge this second LSA immediately; note the time of this acknowledgement. This is newLSAprocTime.

The amount of time required for an OSPF implementation to process the new LSA can be computed by subtracting dupLSAprocTime from newLSAprocTime.

Note: The duplicate LSA cannot be the same as the one just sent because of the MinLSInterval restriction.[\[RFC2328\]](#) This test is taken from [\[7\]](#).

8.2. Flooding Time

- o Using reference topology 2, enable OSPF on all links and allow the devices to build full adjacencies. Configure the collector so it will block all flooding towards the DUT, although it continues receiving advertisements from the

DUT.

- o Inject a new set of LSAs from the generator towards the collector and the DUT.
- o On the collector, note the time the flooding is complete across the link to the generator. Also note the time the flooding is complete across the link from the DUT.

Two measurements can be taken from this test:

- o The time between the last LSA is received on the collector from the generator and the time the last LSA is received on the collector from the DUT.
- o The time between the last LSA is received on the collector from the generator and the time the first LSA is received on the collector from the DUT.

Depending on the number of LSAs flooded, the sizes of the LSAs, and the rate of flooding, these numbers could vary by some amount. The settings and variances of these numbers should be reported with the test results.

This time is important in link state protocols, since the loop free nature of the network is reliant on the speed at which revised topology information is flooded.

8.3. Shortest Path First Computation Time (Internal Test)

- o Using reference topology 1, establish a full adjacency between the generator and the DUT.
- o Inject a set of LSAs from the generator towards the DUT. Allow the DUT to stabilize and install all best paths in the routing table, etc.
- o Change the link cost between the DUT and the generator (or the link between the generator and the emulated network it is advertising), such that a full SPF is required to run, although only one piece of information is changed.
- o Measure the amount of time required for the DUT to compute new shortest path tree as a result of the topology changes

injected by the generator. These measurements should be taken using available show and debug information on the DUT.

The test results should note what available commands were used to determine the length of the shortest path first run, and should note any performance implications of monitoring the SPF calculations using information available from the DUT itself.

8.4. Shortest Path First Computation Time (External Test)

- o Use reference topology 1, beginning with the DUT and the generator fully adjacent.
- o The default SPF timer on the DUT should be set to 0, so that any new LSA that arrives, immediately results in the SPF calculation [7].
- o The generator should inject a set of LSAs towards the DUT; the DUT should be allowed to converge and install all best paths in the local routing table, etc..
- o Send an LSA that is already there in the DUT (a duplicate LSA), note the time difference between when the LSA is sent to when the ack is received. This measures the time to propagate the LSA and the ack, as well as processing time of the duplicate LSA. This is dupLSAprocTime.
- o Change the link cost between the generator and the emulated network it is advertising.
- o Immediately inject another LSA which is a duplicate of some other LSA the generator has previously injected (preferably a stub network someplace within the emulated network).
- o Measure the time between transmitting the second (duplicate) LSA and the acknowledgement for that LSA; this is the totalSPFtime. The total time required to run SPF can be computed by subtracting dupLSAprocTime from totalSPFtime.

The accuracy of this test is crucially dependant on the amount of time between the transmission of the first and second LSAs. If there is too much time between them, the test

is meaningless because the SPF run will complete before the second (duplicate) LSA is received. If there is too little time between the LSAs being generated, then they will both be handled before the SPF run is scheduled and started, and thus the measurement would only be for the handling of the duplicate LSA.

This test is also specified in [7].

Note: This test may not be accurate on systems which implement OSPF as a multithreaded process, where the flooding takes place in a separate process (or on a different processor) than shortest path first computations.

9. Basic Intra-Area OSPF tests

These tests measure the performance of an OSPF implementation for basic intra-area tasks, including:

- o Forming Adjacencies on Point-to-Point Link Initialization
- o Forming Adjacencies on Point-to-Point Links
- o Link Up with Information Already in the Database
- o Initial SR-Convergence Time on a DR electing (broadcast) network
- o Point-to-point Link Down with Layer 2 Detection
- o Broadcast Media Link Down with Layer 3 Detection
- o Designated Router Election Time on A Broadcast Network

9.1. Forming Adjacencies on Point-to-Point Link Initialization

- o Use reference topology 1, beginning with the link between the generator and DUT disabled on the DUT. OSPF should be configured and operating on both devices.
- o Inject a set of LSAs from the generator towards the DUT.
- o Bring the link up at the DUT, noting the time that the link carrier is established on the generator.

- o Note the time the acknowledgement for the last LSA transmitted from the DUT is received on the generator.

The time between the carrier establishment and the acknowledgement for the last LSA transmitted by the generator should be taken as the total amount of time required for the OSPF process on the DUT to react to a link up event with the set of LSAs injected, including the time required for the operating system to notify the OSPF process about the link up, etc.. The acknowledgement for the last LSA transmitted is used instead of the last acknowledgement received in order to prevent timing skews due to retransmitted acknowledgements or LSAs.

9.2. Forming Adjacencies on Point-to-Point Links

- o Using reference topology 1, configure the DUT and the generator so traffic can be passed along the link between them.
- o Configure the generator so OSPF is running on the point-to-point link towards the DUT, and inject a set of LSAs.
- o Configure the DUT so OSPF is initialized, but not running on the point-to-point link between the DUT and the generator.
- o Enable OSPF on the interface between the DUT and the generator on the DUT.
- o Note the time of the first hello received from the DUT on the generator.
- o Note the time of the acknowledgement from the DUT for the last LSA transmitted on the generator.

The time between the first hello received and the acknowledgement for the last LSA transmitted by the generator should be taken as the total amount of time required for the OSPF process on the DUT to build a FULL neighbor adjacency with the set of LSAs injected. The acknowledgement for the last LSA transmitted is used instead of the last acknowledgement received in order to prevent timing skews due to retransmitted acknowledgements or LSAs.

9.3. Link Up with Information Already in the Database

- o Using reference topology 2, configure all three devices to run OSPF.
- o Configure the DUT so the link between the DUT and the generator is disabled .
- o Inject a set of LSAs into the network from the generator; the DUT should receive these LSAs through normal flooding from the collector.
- o Enable the link between the DUT and the generator.
- o Note the time of the first hello received from the DUT on the generator.
- o Note the time of the acknowledgement from the DUT for the last LSA transmitted on the generator.

The time between the hello received from the DUT by the generator and the acknowledgement for the last LSA transmitted by the generator should be taken as the total amount of time required for the OSPF process on the DUT to build a FULL neighbor adjacency with the set of LSAs injected. In this test, the DUT is already aware of the entire network topology, so the time required should only include the processing of each LSA from the generator and transmitting an acknowledgement. The acknowledgement for the last LSA transmitted is used instead of the last acknowledgement received in order to prevent timing skews due to retransmitted acknowledgements or LSAs.

9.4. Initial SR-Convergence Time on a DR electing (broadcast) network

- o Using reference topology 3, begin with the DUT connected to the network with OSPF enabled. OSPF should be enabled on R1, but the broadcast link should be disabled.
- o Enable the broadcast link between R1 and the DUT. Note the time of the first hello received by R1.
- o Note the time the first network LSA is flooded by the DUT at R1.

- o The differential between the first hello and the first network LSA is the time required by the DUT to converge on this new topology.

This test assumes that the DUT will be the designated router on the broadcast link. A similar test could be designed to test the SR-Convergence time when the DUT is not the designated router as well.

This test may be performed with varying numbers of devices attached to the broadcast network, and varying sets of LSAs being advertised to the DUT from the routers attached to the broadcast network. Variations in the LSA sets and other factors should be noted in the test results.

9.5. Point-to-point Link Down with Layer 2 Detection

- o Using reference topology 4, begin with OSPF in the full state between the generator and the DUT. Both links should be point-to-point links with the ability to notify the operating system immediately upon link failure.
- o Disable link 1; this should be done in such a way that the keepalive timers at the data link layer will have no impact on the DUT recognizing the link failure (the operating system in the DUT should recognize this link failure immediately). Disconnecting the cable on the generator end would be one possibility, or shutting the link down.
- o Note the time of the link failure on the generator.
- o At the generator, note the time of the receipt of the new router LSA from the DUT notifying the generator of the link 2 failure.

The difference in the time between the initial link failure and the receipt of the LSA on the generator across link 2 should be taken as the time required for an OSPF implementation to recognize and process a link failure.

9.6. Broadcast Media Link Down with Layer 3 Detection

- o Using reference topology 4, begin with OSPF in the full state between the generator and the DUT.
- o Disable OSPF processing on link 1 from the generator. This should be done in such a way so it does not affect link status; the DUT must note the failure of the adjacency through the dead interval.
- o At the generator, note the time of the receipt of the new router LSA from the DUT notifying the generator of the link 2 failure.

The difference in the time between the initial link failure and the receipt of the LSA on the generator across link 2 should be taken as the time required for an OSPF implementation to recognize and process an adjacency failure.

9.7. Designated Router Election Time on A Broadcast Network

- o Using reference topology 3, configure R1 to be the designated router on the link, and the DUT to be the backup designated router.
- o Enable OSPF on the common broadcast link on all the routers in the test bed.
- o Disable the broadcast link on R1.
- o Note the time of the last hello received from R1 on R2.
- o Note the time of the first network LSA generated by the DUT as received on R2.

The time between the last hello received on R2 and the first network LSA generated by the DUT should be taken as the amount of time required for the DUT to complete a designated router election computation. Note this test includes the dead interval timer at the DUT, so this time can be factored out, or the hello and dead intervals reduced to make these timers impact the overall test times less. All changed timers, the number of routers connected to the link, and other variable factors should be noted in the test results.

10. Acknowledgements

Thanks to Howard Berkowitz, (hcb@clark.net), for his encouragement and support. Thanks also to Gurpreet Singh (Gurpreet.Singh@SpirentCom.COM) for his help as well.

11. References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [RFC2119](#), March 1997.
- [2] Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.
- [3] Bradner, S., McQuaid, J., "Benchmarking Methodology for Network Interconnect Devices", [RFC2544](#), March 1999.
- [4] Manral, V., "OSPF Convergence Testing Terminology and Concepts", [draft-ietf-bmwg-ospfconv-term](#), January 2003
- [5] Alaettinoglu C., et al., "Towards Milli-Second IGP Convergence" [draft-alaettinoglu-isis-convergence](#)
- [6] Trotter, G., "Terminology for Forwarding Information Base (FIB) based Router Performance", [RFC3222](#), October 2001.
- [7] Shaikh, Aman, Greenberg, Albert, "Experience in Black-Box OSPF measurement"
- [8] Manral, V., "Benchmarking Applicability for Basic OSPF Convergence", [draft-ietf-bmwg-ospfconv-applicability](#)

12. Authors' Addresses

Vishwas Manral
Netplane Systems
189 Prashasan Nagar
Road number 72
Jubilee Hills
Hyderabad, India

vmanral@netplane.com

Russ White
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709

riw@cisco.com

Aman Shaikh
University of California
School of Engineering
1156 High Street
Santa Cruz, CA 95064

aman@soe.ucsc.edu

