

Expiration Date: September 2003

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

OSPF Benchmarking Terminology and Concepts
draft-ietf-bmwg-ospfconv-term-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 explains the terminology and concepts used in [[BENCHMARK](#)] and future OSPF benchmarking drafts, within the context of those drafts. While some of these terms may be defined elsewhere, and we will refer the reader to those definitions in some cases, we also include discussions concerning these terms as they relate specifically to the tasks involved in benchmarking the OSPF protocol.

3. Motivation

This draft is a companion to [[BENCHMARK](#)], which describes basic Open Shortest Path First [OSPF] testing methods. This draft explains terminology and concepts used in OSPF Testing Framework Drafts, such as [[BENCHMARK](#)].

4. Definitions

o Internal Measurements

- Definition

Also known as White Box Measurements; internal measurements are measurements taken on the Device Under Test (DUT) itself.

- Discussion

These measurements rely on output and event recording, along with the clocking and timestamping available on the DUT itself. Taking measurements on the DUT may impact the actual outcome of the test, since it can increase processor loading, memory utilization, and timing factors. Some devices may not have the required output readily available for taking internal measurements, as well.

Note: Internal measurements can be influenced by the vendor's implementation of the various timers and processing models. Whenever possible, internal measurements should be compared to external measurements to verify and validate them.

o External Measurements

- Definition

Also known as Black Box Measurements; external measurements infer the performance of the DUT through observation of its communications with other devices.

- Discussion

One example of an external measurement is when a downstream device receives complete routing information from the DUT, it can be inferred that the DUT has transmitted all the routing information available. External measurements suffer in that they include not just the protocol action times, but also propagation delays, queuing delays, and other such factors.

For the purposes of this paper, external techniques are more readily applicable.

- o Multi-device Measurements

- Definition

Multi-device measurements require the measurement of events occurring on multiple devices within the testbed.

- Discussion

For instance, the timestamp on a device generating an event could be used as the marker for the beginning of a test, while the timestamp on the DUT or some other device might be used to determine when the DUT has finished processing the event.

These sorts of measurements are the most problematic, and are to be avoided where possible, since the timestamps of the devices in the test bed must be synchronized within milliseconds for the test results to be meaningful. Given the state of network time protocol implementation, expecting the timestamps on several devices to be within milliseconds of each other is highly optimistic.

- o Point-to-Point links

- Definition

See [OSPF], [Section 1.2](#).

- Discussion

A point-to-point link can take lesser time to converge than a broadcast link of the same speed because it does not have the overhead of DR election. Point-to-point links can be either numbered or unnumbered. However in the context of [[BENCHMARK](#)] and [OSPF], the two can be regarded the same.

- o Broadcast Link

- Definition

See [OSPF], [Section 1.2](#).

- Discussion

The adjacency formation time on a broadcast link can be more than that on a point-to-point link of the same speed, because DR election has to take place. All routers on a broadcast network form adjacency with the DR and BDR.

Async flooding also takes place thru the DR. In context of convergence, it may take more time for an LSU to be flooded from one DR-other router to another DR-other router, because the LSA has to be first processed at the DR.

- o Shortest Path First Time

- Definition

The time taken by a router to complete the SPF process, as described in [OSPF].

- Discussion

This does not include the time taken by the router to give routes to the forwarding engine.

- o Measurement Units

The SPF time is generally measured in milliseconds.

- o Hello Interval

- Definition

- See [OSPF], [Section 7.1](#).

- Discussion

- The hello interval should be the same for all routers on a network.

- Decreasing the hello interval can allow the router dead interval (below) to be reduced, thus reducing convergence times in those situations where the router dead interval timing out causes an OSPF process to notice an adjacency failure. Further discussion on small hello intervals is given in [[CONGESTION](#)] and [[MARKING](#)].

- o Router Dead interval

- Definition

- See [OSPF], [Section 7.1](#).

- Discussion

- This is advertised in the router's Hello Packets in the RouterDeadInterval field. The router dead interval should be some multiple of the HelloInterval (say 4 times the hello interval), and must be the same for all routers attached to a common network.

- o Incremental SPF

- Definition

- The ability to recalculate a small portion of the SPF tree, rather than the entire SPF tree, on receiving

notification of a change in the network topology.

- Discussion

At worst, incremental SPF should perform no worse than a full SPF. In better situations, an incremental SPF run will rebuild the SPF tree in much shorter time than a full SPF run.

5. Concepts

5.1. The Meaning of Control Plane Convergence

A network is termed to be converged when all of the devices within the network have a loop free path to each possible destination. Since we are not testing network convergence, but performance for a particular device within a network, however, this definition needs to be narrowed somewhat to fit within a single device view.

In this case, convergence will mean the point in time when the DUT has performed all actions needed to react to the change in topology represented by the test condition; for instance, an OSPF device must flood any new information it has received, rebuild its shortest path first (SPF) tree, and install any new paths or destinations in the local routing information base (RIB, or routing table).

Note that the word convergence has two distinct meanings; the process of a group of individuals meeting the same place, and the process of a single individual meeting in the same place as an existing group. This work focuses on the second meaning of the word, so we consider the time required for a single device to adapt to a network change to be SR-Convergence, or Single Router Convergence.

This concept does not include the time required for the control plane of the device to transfer the information required to forward packets to the data plane, nor the amount of time between the data plane receiving that information and being able to actually forward traffic.

5.2. Measuring Convergence

Obviously, there are several elements to convergence, even under the definition given above for a single device, including (but not limited to):

- o The time it takes for the DUT to pass the information about a network event on to its neighbors.
- o The time it takes for the DUT to process information about a network event and calculate a new Shortest Path Tree (SPT).
- o The time it takes for the DUT to make changes in its local rib reflecting the new shortest path tree.

5.3. Types of Network Events

A network event is an event which causes a change in the network topology.

- o Link or Neighbor Device Up

The time needed for an OSPF implementation to recognize a new link coming up on the device, build any necessarily adjacencies, synchronize its database, and perform all other needed actions to converge.

- o Initialization

The time needed for an OSPF implementation to be initialized, recognize any links across which OSPF must run, build any needed adjacencies, synchronize its database, and perform other actions needed to converge.

- o Adjacency Down

The time needed for an OSPF implementation to recognize a link down/adjacency loss based on hello timers alone, propagate any information as necessary to its remaining adjacencies, and perform other actions needed to converge.

- o Link Down

The time needed for an OSPF implementation to recognize a link down based on layer 2 provided information, propagate any information as needed to its remaining adjacencies, and perform other actions needed to converge.

6. Acknowledgements

The authors would like to thank Howard Berkowitz (hcb@clark.net), Kevin Dubray, (kdubray@juniper.net), and Randy Bush (randy@psg.com) for their discussion, ideas, and support.

7. Normative References

[BENCHMARK]

Manral, V., "Benchmarking Methodology for Basic OSPF Convergence", [draft-bmwg-ospfconv-intraarea-04](#), March 2003

[OSPF]Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.

8. Informative References

[CONGESTION]

Ash, J., "Proposed Mechanisms for Congestion Control/Failure Recovery in OSPF & ISIS Networks", October, 2001

[MARKING]

Choudhury, G., et al, "Explicit Marking and Prioritized Treatment of Specific IGP Packets for Faster IGP Convergence and Improved Network Scalability and Stability", [draft-ietf-ospf-scalability](#), April 2002

9. Authors' Addresses

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

`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`

