Network Working Group Vishwas Manral
Internet Draft Netplane Systems
Russ White
Cisco Systems
Aman Shaikh
Expiration Date: December 2004 University of California
File Name: draft-bmwg-ospfconv-term-09.txt June 2004

OSPF Benchmarking Terminology and Concepts
draft-bmwg-ospfconv-term-09.txt

Status of this Memo

Abstract

This draft explains the terminology and concepts used in OSPF benchmarking. 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.

1. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119]. [RFC2119] keywords in this document are used to assure methodological control, which is very important in the specification of benchmarks. This document does not specify a network related protocol.

2. Introduction

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].

3. Common Definitions

Definitions in this section are well known industry and benchmarking terms which may be defined elsewhere.

- o White Box (Internal) Measurements

  - Definition

    White Box measurements are measurements reported and collected on the Device Under Test (DUT) itself.

  - Discussion

    These measurement rely on output and event recording, along with the clocking and time stamping 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: White box 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.

Because of the potential for variations in collection and presentation methods across different DUTs, white box measurements MUST NOT be used as a basis of comparison in benchmarks. This has been a guiding principal of Benchmarking Methodology Working Group.

- o Black Box (External) Measurements

  - Definition

    Black Box measurements infer the performance of the DUT through observation of its communications with other devices.

  - Discussion

    One example of a black box 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 of internal operations may 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 [BENCHMARK], external techniques are more readily applicable.

- o Multi-device Measurements

  - Measurements assessing communications (usually in combination with internal operations) between two or more DUTs.

Multi-device measurements may be internal or external.

4. Terms Defined Elsewhere

Terms in this section are defined elsewhere, and included only to include a discussion of those terms in reference to [BENCHMARK].

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.

  Asynchronous flooding also takes place thru the DR. In context of convergence, it may take more time for an LSA 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 Execution 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.

  Some implementations may force two intervals, the SPF hold time and the SPF delay, between successive SPF calculations. If an SPF hold time exists, it should be subtracted from the total SPF execution time. If an SPF delay exists, it should be noted in the test results.

- 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 [OSPF-SCALING].

o Router Dead interval

- Definition

  See [OSPF], Section 7.1.

- Discussion

  This is advertised in the router's Hello Packets in the Router-DeadInterval 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.

5. Concepts

5.1. The Meaning of Single Router 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 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.

- The time it takes for the DUT to process information about a network event and calculate a new Shortest Path Tree (SPT).

- 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.

- 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.

- 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.

- 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 per-form other actions needed to converge.

- 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. IANA Considerations

This document requires no IANA considerations.

7. Security Considerations

This document does not modify the underlying security considerations in [OSPF].

8. Acknowledgements

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

9. Normative References

[BENCHMARK]
Manral, V., "Benchmarking Basic OSPF Single Router Control Plane Convergence", draft-bmwg-ospfconv-intraarea-09, May 2004.

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

[RFC2119]
Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997

10. Informative References

[OSPF-SCALING]
Choudhury, Gagan L., Editor, "Prioritized Treatment of Specific OSPF Packets and Congestion Avoidance", draft-ietf-ospf-scalability-06.txt, August 2003.

11. 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

12. Full Copyright Statement

13. Intellectual Property

copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.