

Network Working Group
Request for Comments: 4062
Category: Informational

V. Manral
SiNett Corp.
R. White
Cisco Systems
A. Shaikh
AT&T Labs (Research)
April 2005

OSPF Benchmarking Terminology and Concepts

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

This document explains the terminology and concepts used in OSPF benchmarking. Although some of these terms may be defined elsewhere (and we will refer the reader to those definitions in some cases) we include discussions concerning these terms, as they relate specifically to the tasks involved in benchmarking the OSPF protocol.

1. Introduction

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

2. 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] key words in this document are used to ensure methodological control, which is very important in the specification of benchmarks. This document does not specify a network-related protocol.

3. Common Definitions

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

o White Box (Internal) Measurements

- Definition

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

- Discussion

These measurements 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.

Note: White box measurements can be influenced by the vendor's implementation of 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 for comparison in benchmarks. This has been a guiding principle of the 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 are included only as they apply to [BENCHMARK].

- o Point-to-Point Links

- Definition

See [OSPF], Section 1.2.

- Discussion

A point-to-point link can take less 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 as the same.

- o Broadcast Link

- Definition

See [OSPF], Section 1.2.

- Discussion

The adjacency formation time on a broadcast link can be greater 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 through the DR. In the context of convergence, it may take more time for an LSA to be flooded from one DR-other router to another because the LSA first has to be 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 install routes in 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 must 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's timing out causes an OSPF process to notice an adjacency failure. Further discussion of 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 (perhaps 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 as converged when all the devices within the network have a loop-free path to each possible destination. However, because we are not testing network convergence but testing performance for a particular device within a network, this definition needs to be streamlined 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 in order to react to the change in the 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 at the same place, and the process of an individual coming to 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. It also does not include the amount of time between when the data plane receives that information and when it is able to 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) the following:

- 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 to 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 that 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, to build any necessary adjacencies, to synchronize its database, and to perform all other actions necessary to converge.

- o Initialization

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

- o Adjacency Down

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

- o Link Down

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

6. Security Considerations

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

7. 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.

8. Normative References

- [BENCHMARK] Manral, V., White, R., and A. Shaikh, "Benchmarking Basic OSPF Single Router Control Plane Convergence", RFC 4061, April 2005.
- [OSPF] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9. Informative References

- [OSPF-SCALING] Choudhury, Gagan L., Editor, "Prioritized Treatment of Specific OSPF Packets and Congestion Avoidance", Work in Progress, August 2003.

Authors' Addresses

Vishwas Manral,
SiNett Corp,
Ground Floor,
Embassy Icon Annexe,
2/1, Infantry Road,
Bangalore, India

EMail: vishwas@sinett.com

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

EMail: riw@cisco.com

Aman Shaikh
AT&T Labs (Research)
180 Park Av, PO Box 971
Florham Park, NJ 07932

EMail: ashaikh@research.att.com

Full Copyright Statement

Copyright (C) The Internet Society (2005).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any 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.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.