

Network Working Group  
Internet Draft

Vishwas Manral  
Netplane Networks  
Russ White  
Cisco Systems  
December 2001

Expiration Date: June 2002  
File Name: [draft-manral-ospfconv-term-00.txt](#)

OSPF Benchmarking Terminology and Concepts  
draft-manral-ospfconv-term-00.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 [\[2\]](#) and future OSPF benchmarking drafts.

### [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

This draft is a companion to [\[2\]](#), which describes basic Open Shortest Path First (OSPF [\[3\]](#)) testing methods. This draft explains terminology and concepts used in OSPF Testing Framework Drafts, such as [\[2\]](#).

### [5.](#) Definitions

#### o Internal Measurements

##### - Definition

Internal measurements are measurements taken on the Device Under Test (DUT) itself.

##### - Discussion

These measurement rely on output and event recording, along with the clocking and timestamping available on the DUT itself. Internal measurements are preferred for all tests that can be completely contained on the DUT (which is very rare).

#### o External Measurements

##### - Definition

External measurements infer the performance of the DUT through observation of its communications with other dev-

ices.

- 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.

For the purposes of this paper, external technique are more readily applicable. However, external measurements have their own problems because they include the time to advertise the new route downstream and transmission times for the advertisement within the device under test.

- 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

A network that joins a single pair of routers is called a point-to-point link. For OSPF [\[3\]](#), point-to-point links are those on which a designated router are not elected.

- Discussion

A point-to-point link will 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 [\[2\]](#), the two can be regarded the same.

- o Broadcast Link

- Definition

Networks supporting many (more than two) attached routers, together with the capability to address a single physical message to all of the attached routers (broadcast). In the context of [\[2\]](#) and [\[3\]](#), broadcast links are taken as those on which a designated router is elected.

- 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.

- Discussion

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

- o Measurement Units

- The LSA time is generally measured in milliseconds.

- o Hello Interval

- Definition

- The length of time, in seconds, between the Hello Packets that the router sends on the interface.

- Discussion

- The hello interval should be the same for all routers on the 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. Very small router dead intervals accompanied by very small hello intervals can produce more problems than they resolve, as described in [4] & [5].

- o Router Dead interval

- Definition

- After ceasing to hear a router's Hello Packets, the number of seconds before its neighbors declare the router down.

- 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.

## [6.](#) Concepts

[6.1.](#) 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).

## [6.2.](#) Measuring Convergence

Obviously, there are several elements to convergence, even under the definition given above for a single device. We will try to provide tests to measure each of these:

- 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.

## [6.3.](#) Types of Network Events

- 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 information as needed to its remaining adjacencies, and perform other actions needed to converge.

#### [6.4.](#) LSA and Destination mix

In many OSPF benchmark tests, a generator injecting a number of LSAs is called for. There are several areas in which injected LSAs can be varied in testing:

- o The number of destinations represented by the injected LSAs

Each destination represents a single reachable IP network; these will be leaf nodes on the shortest path tree. The primary impact to performance should be the time required to insert destinations in the local routing table and handling the memory required to store the data.

- o The types of LSAs injected

There are several types of LSAs which would be acceptable under different situations; within an area, for instance, type 1, 2, 3, 4, and 5 are likely to be received by a router. Within a not-so-stubby area, however, type 7 LSAs would replace the type 5 LSAs received. These sorts of characterizations are important to note in any test results.

- o The Number of LSAs injected



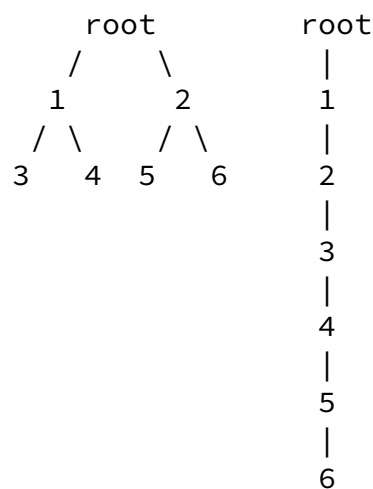
Within any injected set of information, the number of each type of LSA injected is also important. This will impact the shortest path algorithms ability to handle large numbers of nodes, large shortest path first trees, etc.

- o The Order of LSA Injection

The order in which LSAs are injected should not favor any given data structure used for storing the LSA database on the device under test. The ordering can be changed in various tests to provide insight on the efficiency of storage within the DUT. Any such changes in ordering should be noted in test results.

#### [6.5.](#) Tree Shape and the SPF Algorithm

The shortest path first algorithm is a simple algorithm which handles complexity by breaking the problem of finding the shortest paths through a network into smaller parts and recursing (calling itself) to compute the best path within each smaller part. Because of this, moving along a single level of the tree, along the tree's width, is fundamentally different than moving along the depth of the tree.



For instance, the shortest path first algorithm would go through two recursions when finding the shortest paths on the left topology, with an average of two nodes processed per level. The topology on the right would produce five recursions, with one node processed per recursion. While this may not produce dramatically different test results, there may be some apparent difference between the two.

In general, those benchmarking link state protocols which use the shortest path first algorithm to compute the best paths through the network need to be aware that the construction of the tree may impact the performance of the algorithm. Best practice would be to try and make any emulated network look as much like a real network as possible, especially in the area of the tree depth, the meshiness of the network, the number of stub links verses transit links, and the number of connections and nodes to process at each recursion level.

## 7. Route Generation

As the size of networks grows, it becomes more and more difficult to actually create a large scale network on which to test the properties of routing protocols and their implementations. In general, network emulators are used to provide emulated topologies which can be advertised to a device with varying conditions. Route generators either tend to be a specialized device, a piece of software which runs on a router, or a process that runs on another operating system, such as Linux or another variant of Unix.

Some of the characteristics of this device should be:

- o The ability to connect to the several devices using both point-to-point and broadcast high speed media. Point-to-point links can be emulated with high speed Ethernet as long as there is no hub or other device in between the DUT and the route generator, and the link is configured as a point-to-point link within OSPF.
- o The ability to create a set of LSAs which appear to be a logical, realistic topology. For instance, the generator should be able to mix the number of point-to-point and broadcast links within the emulated topology, and should be able to inject varying numbers of externally reachable destinations.
- o The ability to withdraw and add routing information into and from the emulated topology to emulate links flapping.
- o The ability to randomly order the LSAs representing the emulated topology as they are advertised.

- o The ability to log or otherwise measure the time between packets transmitted and received.

- o The ability to change the rate at which OSPF LSAs are transmitted.

## [8. Acknowledgements](#)

The authors would like to thank Aman Shaikh (ashaikh@research.att.com) for his comments and help on this draft.

## [9. References](#)

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [RFC2119](#), March 1997.
- [2] Manral, V., "Benchmarking Methodology for Basic OSPF Convergence", [draft-manral-ospconv-intraarea-00](#), November 2001
- [3] Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.
- [4] [draft-ash-ospf-isis-congestion-control-01.txt](#)
- [5] [draft-ietf-ospf-scalability-00.txt](#)

## [10. Authors' Addresses](#)

Vishwas Manral,  
Netplane Networks,  
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](mailto:riw@cisco.com)

Manral, et. all

[Page 10]