University of New Hampshire
InterOperability Laboratory

Routing Consortium

Open Shortest Path First
Guide

Technical Document

Revision 1.0

University of New Hampshire
InterOperability Laboratory
Routing Consortium
http://www.iol.unh.edu

121 Technology Drive, Suite 2
Durham, NH 03824-3525
Phone: +1-603-862-3941
Fax: +1-603-862-4181

MODIFICATION RECORD

Version 1.0 July 21, 2008

- Initial Release taken from OSPF Testers Guide

ACKNOWLEDGMENTS

The University of New Hampshire would like to acknowledge the efforts of the following individuals in the development of this test suite.

David Bond University of New Hampshire

INTRODUCTION

Overview

The University of New Hampshire's InterOperability Laboratory (IOL) is an institution designed to improve the interoperability of standards based products by providing an environment where a product can be tested against other implementations of a standard. This Guide describes how the Open Shortest Path First Protocol. This is part of a larger internal document describing how OSPF is tested at the IOL.

Acronyms

Acronyms used in this Test Suite:

TR: **T**esting **R**outer
TN: **T**esting **N**ode

When several entities of the same type are present in a test configuration, a number is appended to the acronym to yield a label for each entity. For example, if there were three testing routers in the test configuration, they would be labeled TR1, TR2 and TR3.

REFERENCES

The following documents are referenced in this text:

- Request for Comments 2328 – OSPF, Version 2
- Request for Comments 1583 – OSPF, Version 2
- Request for Comments 3101 – The OSPF Not-So-Stubby Area (NSSA) Option
- Request for Comments 2740 – OSPF for IPv6
- UNH-IOL Open Shortest Path First (OSPF) Operations Test Suite
- UNH-IOL Open Shortest Path First (OSPF) Interoperability Test Suite
- UNH-IOL Open Shortest Path First (OSPF) NSSA Conformance Test Suite
- UNH-IOL Open Shortest Path First Version 3 (OSPFv3) Interoperability Test Suite

TABLE OF CONTENTS

OSPF Protocol Overview

Routing Protocols Introduction

So, you want to test the **Open Shortest Path First Protocol**, or as it is abbreviated, **OSPF,** this is a short guide on the basics of OSPF and how to test it here at the Interoperability Lab. To begin let's look at how OSPF operates. OSPF is what is called an **Interior Gateway Protocol** (**IGP**): a protocol that routes data within a single **autonomous system** (**AS**). This is in contrast to protocols such as **BGP** that route data between autonomous systems.

There are two major types of IGPs, distance vector and link state routing protocols. **Distance vector routing protocols** are best described by thinking of routers as a bunch of cities, with the roads between those cities being the connections. Someone traveling between those cities could represent a packet. At each city there is a cross roads with many signs giving the distances and direction to the other cities. Distance vector routing protocols include **RIP** and IGRP.

OSPF on the other hand is what is called a **Link state routing protocol**. Link state routing protocols differ from distance vector in that each node in a Link state routing protocol knows 'everything' about the network topology while in distance vector protocols the nodes simply know how far it is to the other nodes in the network. Back to the city example when the traveler arrives at a city in a link state routing protocol he goes to a big map of the land and by looking at the map, the decision on which way to travel is then calculated. Just like how the distance vector routing protocol keeps all the road signs up to date, the link state routing protocol keeps this map up to date. To do this they use what's called **Dijkstra's algorithm**. The details of this algorithm will be left for your own curiosity. OSPF and **ISIS** are both examples of link state routing protocols.

OSPF Areas

Moving onto the technical side, each OSPF AS is divided into areas. For simplicity each area is given an identifier, which looks like an IPv4 address such as 2.2.2.2 or 34.46.124.4. Previously it was said that in a link state routing protocol every router knows everything about the entire AS. OSPF cheats a little here, and for good reason. If every router had to know everything about a huge network, there would be an enormous amount of data to be maintained.

This is where areas come into play. They divide the AS into logical segments where all routers within that area know the topology of that area while the topology is hidden from router in other areas. The easiest way to think of areas is that they tell other areas a subset of the information maintained within the area itself.

One area, **0.0.0.0**, is reserved for what is called the **backbone** of OSPF. Just as it sounds, the backbone is the central connection between all of the various areas in a given topology. Every area must have a connection, whether physical or virtual, to the backbone. Don't worry about the virtual connection just yet. The backbone is responsible for distributing the routing information between non-backbone areas. The backbone also must be contiguous whether or not this is maintained physically or by virtual connections.

Neighbor Initialization

Each system in an OSPF topology maintains what is called a **Link State Database** (**LSDB**). The process of forming this database begins as OSPF is enabled on the systems throughout a network. When a system has OSPF enabled on a network it starts to send a **Hello Packet** every **Hello Interval**. This allows routers to see one another and begin the process. Each network has what is called a **Designated Router** (**DR**) and **Backup Designated Router** (**BDR**). The DR and BDR serve to reduce the amount of traffic sent on a single network by having them act as a single point of synchronization between routers. Receiving hello packets from another router causes the sending router to become neighbors with the receiving. The new neighbor is added to a list of neighbors within the receiving router's packets.

DR Election Process

The routers now go through what is called the **DR Election Process**. This process is where they choose the DR and BDR If a router is enabled on a network and if the router does not see any neighbors for more then **Router Dead Interval** it claims itself to be DR for that network by setting the appropriate field it its hello messages. A value of 0.0.0.0 in either the DR or BDR fields indicates this router has not elected one or the other yet. If two routers are both enabled at the same time the router with the higher **Router ID** becomes the DR. Election of the BDR follows the same line of reasoning: if no other router is enabled on the network save for the DR the new router becomes BDR, if there is another router that is enabled but that is not DR then

once again the router with the higher Router ID becomes the BDR. Another field present in OSPF hello packets is called **priority**. This was left out above but as one might assume the router with the highest priority becomes DR or BDR first.

One might ask what happens if a third router comes onto the scene. Well, if the DR and/or BDR are already chosen on any given network, this third router will not override the others even if it has a higher priority or router ID. Any additional routers on a network will become what are called a **DR Other**. It is also important to note that if a router has a priority of zero it will never become DR or BDR.

Database Description Process

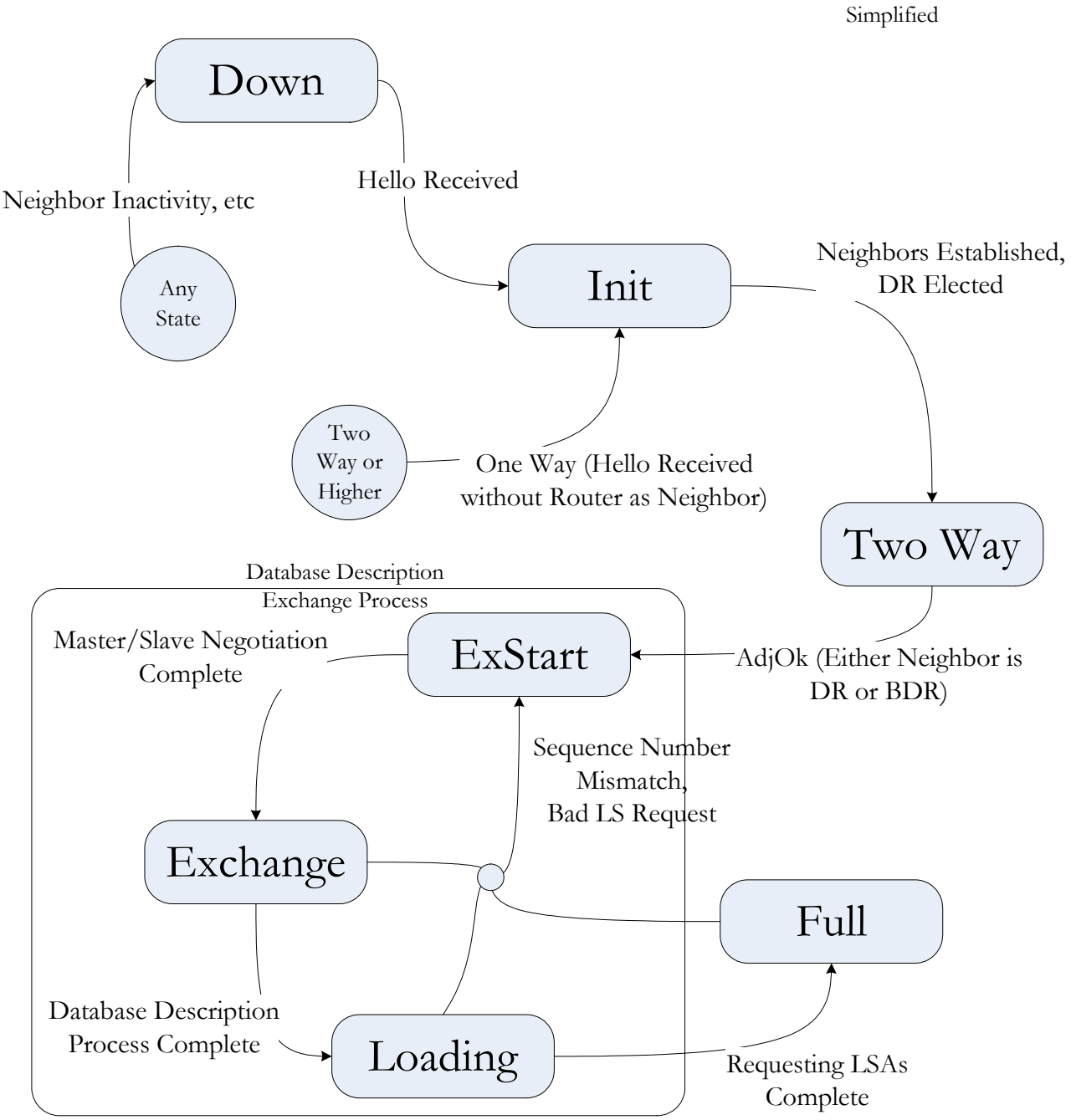
Once a router has gone through the DR Election Process, it must now go through the **Database Description Process (DD Process)** with the DR and BDR. This is where the router tells the other routers what it knows about the network topology. Notice how each router only goes through this process with the DR and BDR. This reduces the amount of traffic sent on the network since otherwise the router would have to synchronize its database with every other router on the network. The DD Process begins between two routers when they each send out a **Database Description Initialization Packet (DD Init)**. This packet tells the neighbor that the router is ready to describe its database. The router with the higher router ID then claims itself to be **Master** while the device with the lower router ID is the **Slave**. The Master will then start describing its database by sending a non-initial **Database Description Packet** (**DD**) and incrementing a field called the **sequence number**, a field that keeps the packets in the correct order. This DD packet will have a set of brief entries with **headers** advertising portions of the topology this router knows about. For instance, a router might have a route to 2.2.2.0/24. This route most likely contains additional information such as metric and tags, but in this header, the router just say that it has the route. The slave will then also describe what it knows about the topology in the same manner inside a DD packet with the same sequence number of the Master. This process of the Master sending a DD packet and the Slave replying with a DD process of its own may continue until the both routers have described their entire LSDBs. If one of the routers has finished describing its LSDB while the other router has not, the finished router will continue to send empty DD packets with the **more bit** clear, indicating it is done describing what it has.

The router now knows what the other router has. It looks through the headers advertised and checks what, within those headers, it does not know about. For these it sends out **request packets**, which simply ask the advertising router to tell the device more about these headers. The advertising router then sees these requests and sends an **update packet**. An update packet tells the other router all the details about this header. Once this process of requests and updates is done, the DD process is finished.

State Diagram

One key skill for testing OSPF is to understand the state diagram. As you look at the diagram below, note the following states: **ExStart**, **Exchange**, **Loading**, **Full,** and **Two Way**. The ExStart state is when the device has sent its DD Init packet and is waiting for the other devices DD Init. The Exchange state is when the devices are describing their databases to one another. Loading is when the devices are sending exchanges and requests to one another. Full is the final state when the devices are gone through the DD process successfully. The Two Way state is the state Two DR Others will be in with one another. Remember a device only goes through the DD Process with the DR and BDR. This means once two DR Others become neighbors they only go to state Two Way instead of state Ex-Start.

OSPF Neighbor State Diagram



Link State Advertisements

Up until now, we have simply said we have these headers and some additional information about the headers. These headers and the additional information are called **Link State Advertisements** (**LSA**). Testers will be working with several types of LSAs. First there is a **Type 1 LSA** or as it is called, a **Router LSA**. A Router LSA lists the links a device has to other networks or routers along with the metrics to those links. As explained earlier, areas represent a division of the information sharing amongst the entire OSPF domain. Router LSAs help to reduce the amount of information each router must store. Router LSAs are not sent, or **flooded**, outside of the area in which the device is present.

The next type of LSA is a **Type 2** or **Network LSA**. A Network LSA describes a broadcast segment. In other words as it sounds these describe a network and the devices connected to them. This includes information such as which router is the DR. These are also only described within the area of origination. A **Type 3** or **Summary LSAs** hearken back to a distance vector protocol. Routers that lay between areas, aka **Area Border Routers** (**ABR**), generate these to provide routing information in a limited scale to other areas. This removes all the detailed topology information that is flooded throughout an area while preserving the essential routes. For instance, a router might generate one of these for 10.10.10.0/24, which says it has a route to that network. As one might guess these are flooded throughout the entire OSPF domain. A router must know how to get to the ABR that is advertising this route, in other words the router must have received a Router LSA for the ABR.

Type 4 or **ASBR-Summary-LSAs** are the next form of LSAs. These provide information about an ASBR Router and are flooded throughout the OSPF domain except for stub areas. **Type 5** or **AS-External-LSAs** are similar to Type 3s. These routes are coming from outside of the given OSPF AS. The meaning of 'outside' is very broad. It could be routes coming from static configuration, non-OSPF interfaces on the device, BGP, RIP, other routing protocols, or even other OSPF instances on the device itself. These typically come from the redistribute command that will be talked about later on. A Type 5 LSA has a field called the **e-bit**. When this is set high the LSA is considered a **Type 2** AS-External LSA, while when it is low, the AS-External LSA is a **Type 1**. Just as a router needs the Router LSA for an ABR advertising a Type 3 route a router needs a ASBR-Summary-LSA for an ASBR advertising a Type 5 LSA. The final

LSA which we will cover is a **Type 7** LSA. This is very similar to a Type 5 LSA with the major difference being it is used within a 'Not-So-Stubby-Area'.

One important note about LSAs and the LSDB is that when a neighbor goes down its LSAs are not removed from the LSDB. They are simply made inactive due to the neighbor adjacency no longer being active. If the neighbor adjacency comes back up these LSAs are active again, until they reach **max age**. From either having a neighbor manually max age them or by having their age increment to that age.

Area Types

With that being said its time to cover the various area 'types.' So far, the only special type of area discussed has been the backbone area. There are also three other types of areas: a **transit area**, a **stub area** and a **Not-So-Stubby-Area** (**NSSA area**). A transit area is simply an area that can carry backbone traffic. Any area that has a virtual link in is a transit area. A stub area on the other hand is an area that does not import AS-External LSAs. This reduces the routes the devices within the stub area must maintain. To ensure the stub area has a route to the AS-External LSAs devices often import a default route into stub areas with a **default stub cost** as the default route's metric.

An NSSA area is the final type. As the name humorously implies this type of area is a cross between a stub and transit area. Just as stub areas do not import Type 5 LSAs, a NSSA area also does not. Instead, in a NSSA area these are translated into Type 7 LSAs. The basic reason for having this additional area type is that it allows some topologies that were previously not configurable. Read the first portion of RFC 3101 if you would like to see a full explanation of this.

ABRs that border a NSSA area have a **NSSA Translator Role**. When this is set to **Always** the router will always **translate** Type 5 LSAs into the network as Type 7 LSAs. When this is set to **Candidate**, the device maintains a list of ABRs on the network and if it has the highest router id amongst these routers, it will translate the LSA. When set to **Disabled** the router does not translate at all.

Within a Router LSA, all the connected networks to a router are described. These networks can be one of four types: two of which are **stub** and **transit networks**. It is important

to note these are different then stub and transit areas. A stub network is an OSPF network the device is connected to, on which there are no neighbors. A transit network is an OSPF network the device is connected to, on which there are one or more neighbors. The other two types of networks are **point-to-point** and virtual. A point-to-point network is simply a special case where the device is connected to another through a dedicated line. Point-to-point links are not tested here at the lab.

Virtual Links

A virtual connection, or more appropriately a **virtual link** was mentioned earlier. All areas must maintain a connection to the backbone; however, at times this is not possible. To account for this a virtual link maybe configured between two devices that have a connection to a segment of the backbone. In other words virtual links serve to make a non-contiguous backbone topology appear as if it were contiguous since virtual links provides a connection through a transit area. The main idea of virtual links is to allow a network engineer to alter the topology to better route traffic. One example might be if a company used OSPF to route between its offices. At one of the larger offices, they might have two areas in use with one of the areas not having backbone connectivity. If this area is given a virtual link to the backbone, the area will be able to route inter-area traffic since it would then have a connection to the backbone.

Virtual links are configured on a per non-backbone area basis. When a router receives a router LSA in an area, which the virtual link was configured for the router will begin sending **unicast hellos** to this neighbor. These hellos travel over the underlying OSPF network, which may include going over many hops, to the neighbor. The path these unicast hello packets must take is the intra-area path to the neighbor. The neighbors then proceed through the database description process. The DR and BDR election is skipped for the virtual link synchronization. The routers will then update their transit area LSAs with a high **v-bit** to indicate the existence of the route. They will also update their backbone Router LSAs to include the virtual connection as a link.

Area Ranges

Most of the major concepts in OSPF have been covered now, however there are a few more to go before we are ready to begin testing. **Area ranges** are a simple way to reduce the size

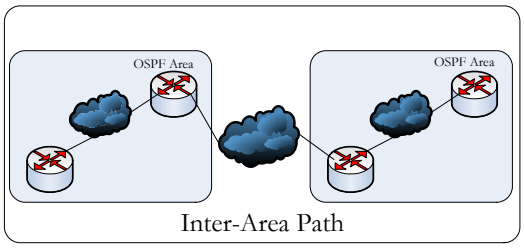
of an OSPF routing database. If a device receives two routes, 10.10.10.0/24 and 10.10.11.0/24, it is desirable to subsume these routes within 10.10.10.0/23. As one can see this means, half the memory will be used. OSPF allows ranges to be configured on an area basis. They are used whenever a route is present that is contained within the range. The OSPF specification also defines a flag, advertise or do not advertise, which simply turns a range on or off. Address ranges are configured on an area whereas any routes flooded from this area are summaries out of, not into. The metric for the range becomes the highest of its component routes.

Authentication

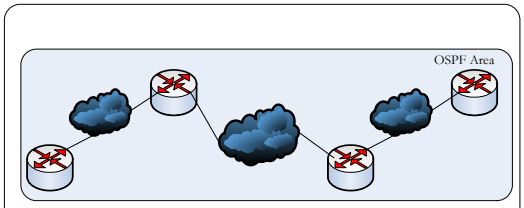
Often it desirable to prevent devices from becoming neighbors with unauthorized devices. Perhaps an ISP is running OSPF and they do not want their customer devices that are running OSPF connecting to their networks and potentially destroying connectivity for other users. To prevent this OSPF defines three **authentication** modes: **none**, **Type 1** (**simple**), and **Type 2** (**MD5**). Authentication type none is just as it sounds, the devices do not authenticate. Type 1 adds a **simple plain text password** to the packets. The packets must match for the devices to become neighbors, and while this does not prevent malicious attacks, it can prevent many problems. Type 2 or MD5 authentication adds another layer of protection. Users can still read any data from the OSPF packets, but unless they know the **md5 password** and **md5 key,** they will not be able to become neighbors since the password is encrypted.

Route Calculation

There is one key set of concepts left to cover in OSPF, that being the route calculation. The details of Dijkstra's algorithm are not important, but the general idea is a device routes a packet based on the lowest cost route to its destination address. OSPF tries to keep packets within one area, and so **Intra-area** paths, paths that only cross through one area, are always preferred over **Inter-area** paths, paths that cross through multiple areas. As previously mentioned AS-External-LSAs are either Type 1 or 2. Type 2 LSAs are considered infinitely more expensive then any route within the OSPF domain and so non-type



Inter-Area Path



Intra-Area Path

2 routes will be preferred over the Type-2 AS-External-LSA route. Type 1 AS-External-LSAs are considered comparable to OSPF domain routes and the standard route comparison rules are used. On Inter-area paths, a transit area route can only be used if a higher cost backbone route is available. This is why all areas must maintain a connection to the backbone. These are the general idea behind route calculations in OSPF. Of course, this is very simplified but it gives a general overview. Devices may also have a **RFC-1583 compatibility flag** that can be configured. This changes the route calculation method to match an older RFC.

Conclusion

In conclusion, OSPF is a link state routing protocol that populates the routing table of a router so that packets can be routed around the Internet.

OSPF Packet Reference

Here is a general OSPF Packet Reference with some useful packet details. As a note there are three e-bits' in OSPF. The e-bit in a type 5 LSA indicates internal versus external metrics. The e-bit in a router LSA signifies an ASBR/Non-ASBR Router. The e-bit in any OSPF packet headers indicate whether the area the packet is being generated within is a stub. Also worth a note, the b-bit in a router LSA indicates an ABR while the links within a network LSA consist of the RID's of neighboring routers.

Link State Acknowledgement Types

- Type 1, Router LSA
  - Originated by: Each Router
  - Flooding Scope: Throughout a single area
  - LSID: Originator's RID
  - AdvRtr: Originator's RID
- Type 2, Network LSA
  - Originated by: DR of the network
  - Flooding Scope: Throughout a single area
  - LSID: Interface IP address of DR
  - AdvRtr: RID of DR
- Type 3, Summary Network(IP) LSA
  - Originated by: ABR
  - Flooding Scope: Throughout the OSPF AS
  - LSID: Destination Network IP Address
  - AdvRtr: ABR RID
- Type 4, Summary ABSR LSA
  - Originated by: ABR
  - Flooding Scope: Throughout the OSPF AS save for stub/NSSA areas
  - LSID: RID of ASBR
  - AdvRtr: ABR RID
- Type 5, AS External LSA
  - Originated by: ASBR

- Flooding Scope: Throughout the OSPF AS save for stub/NSSA areas
- LSID: Destination Network IP Address
- AdvRtr: Originating ASBR RID

- Type 7
  - Originated by: ASBR
  - Flooding Scope: Throughout the OSPF AS save for stub areas
  - LSID: Destination Network IP Address
  - AdvRtr: Originating ASBR RID

Router LSA Link Types

- Type 2, Transit
  - Link ID: IP of DR
  - Link Data: Router's Interface IP Address
- Type 3, Stub
  - Link ID: IP of network
  - Link Data: Netmask
- Type 4, Virtual Link
  - Link ID: Neighbor RID
  - Link Data: Router's Interface IP Address