Redundancy with OSPF Sham-Link

Mohsin Khan

Faculty of Telecommunication Engineering and Environment, Birmingham City University England, UK

mohsin.bcu@gmail.com

Abstract– **In modern world, network technology has witnessed a great progress. The enterprise networks are now being connected to public internet to benefit from it and to increase their business productivity. Internet Service Providers provide high bandwidth connection to enterprise to connect all of its sites, however, a disconnection for any reason with the service provider can be disastrous for enterprise. Clever business enterprises always make sure that they have a backup connection to keep all of their sites connected in case of any loss of connection to the main service provider. In this research work we address the issues related to redundancy between the business sites running OSPF routing protocols as their internal routing protocol and BGP on the MPLS based service provider network. For quick implementation and better understanding of readers, the proposed network design is implemented with Dynagen Simulator. The research work aims to provide some helpful guidelines to students and network engineers on redundancy in scalable enterprise networks.**

Index Terms– **OSPF Sham-Link, BGP, MPLS and Frame Relay**

I. INTRODUCTION

THE purpose of this little research work is to identify, investigate and to get around with the redundancy issue in the network running Open Shortest Path First (OSPF). OSPF [1] is an open standard protocol which is defined in RFC 2328. It can support different vendors. OSPF is a link state, interior gateway routing protocol which uses bandwidth as cost; it selects the shortest path based on the bandwidth of the link. When a change occurs in network, OSPF sends LSA (link state advertisement) to all routers in that specific area which describes the change. LSA is OSPF packet which contains information about the change. OSPF propagates LSA to all neighbor routers. The neighbor routers then update their link state database (LSDB). OSPF [2] then runs Dijkstra algorithm on LSDB of each router to calculate the best (shortest) paths.

These best paths are selected in the routing table. OSPF is an event-driven [7] routing protocol which means it sends updates only when a change occurs to the topology of the network. Moreover, it does not send the entire routing table when there occurs a change like a link or nod failure in the network but it sends information only about the change.

OSPF is very scalable because it provides hierarchal structure [3]. OSPF splits the large network (autonomous system) into smaller and manageable segments called Areas [4]. Each router within an area have identical topology database. These areas can exchange routing information with each other. This is called inter-area routing. Each area enjoys a level of autonomy. For example, if a link fails on an area in an autonomous system, the router on the other area in the same autonomous system will not be troubled to recalculate the SPF. Area 0 [5] is backbone area; all other areas in the autonomous system must have to be connected to it either physically or virtually.

Hierarchal structure of OSPF gives following benefits:

➢ Minimum entries in routing table
➢ Reduces frequency of SPF calculations
➢ Bounds the impact of any topological change to the area only
➢ Reduces link state updates (LSU) overhead
➢ Very fast convergence
➢ OSPF uses less bandwidth for routing updates

OSPF uses following Link State Advertisement (LSA) [6] types to exchange routing information:

Type 1 or Router LSA: these LSAs are propagated by the routers with the scope of within area.

Type 2 or Network LSA: these LSAs are generated by DR router and flooded with in area.

Type 3 or Summary LSA: ABR generates summary LSAs. Summary LSAs contain information about other areas. It contains summery of the OSPF domain.

Type 4 or ASBR Summary LSA: these LSAs are also generated by ABRs but they contain information about ASBRs in other areas.

Type 5 or External LSA: these LSAs are generated by ASBRs, and they contain network information of the outside OSPF domain.

Knowing all the aforementioned benefits of OSPF, most of the enterprise networks use OSPF as their interior gateway routing protocol. But OSPF shows some signs of trouble when the business enterprise runs a backup connection
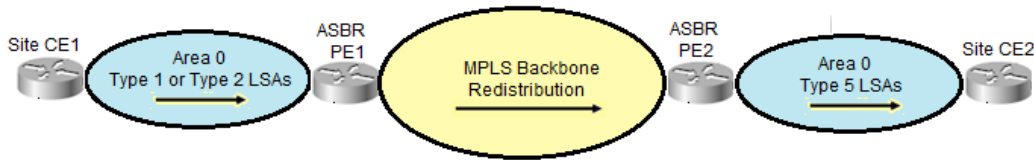
Fig. 2: Reception of Type 5 LSAs due to redistribution of OSPF routes into BGP and back to OSPF

between its sites besides the main connection through service provider network.

II. REDUNDANCY ISSUE WITH OSPF

To explain the link redundancy issue in network running OSPF as routing protocol, we design a service provider network shown in Fig. 1.

There are two Customer Edge sites i.e. CE1 and CE2 which are connected to each other through an MPLS service provider network at Port Edges i.e., PE1 and PE2. Table 1 show the addressing scheme used in the network model.

MPLS Backbone is connecting both sites of the customer i.e., CE1 and CE2 with each other through a high speed OC48 connection which gives a bandwidth of 2.5 Gbps. Since we are using OSPF on the customer sites and BGP on MPLS backbone, the routes of OSPF (Type 1 or Type 2 LSAs) are redistributed into BGP at the ingress PE router and then again they are redistributed into OSPF (Type 5 LSAs) at the egress PE router. Since MPLS backbone is placed between the two customer sites, both the PE routers serves as ASBR routers in terms of OSPF because the Type 1 or Type 2 LSA routes learnt by PE1 about CE1 are advertised as Type 5 routes by PE2 to CE2 and similarly the Type 1 or Type 2 LSA routes learnt by PE2 about CE2 are redistributed into BGP and then PE1 redistributes them into OSPF and advertises them as Type 5 LSAs to CE1. This is explained in Fig. 2.

The network model with customer sites connected to each other through service provider OC48 connection works fine, but there occurs a problem when we need redundancy and we create a direct backup connection between CE1 and CE2 as shown in Fig. 3.

The backup connection is a Frame Relay T1 connection which is desirably needed in case of service provider failure. FR backup connection is configured as OSPF area 0. Now there are two connections between the sites of Customer B. First they are connected through MPLS backbone network, which the faster OC48 link and supposed to be the primary link and secondly they are connected through a backdoor direct link, which is the slower T1 link and it is supposed to be the secondary link. The Customer sites (CE1 and CE2) learn each other's routes as intra-area routes (O) through backdoor Frame Relay connection because they are in the same area (area 0) whereas on service provider MPLS backbone network, they learn routes as external (O E2) routers. Since OSPF according to shortest path first algorithm (SPF) calculation, always prefers intra-area routes over
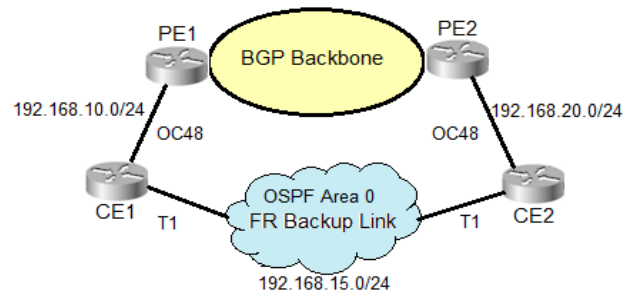

Fig. 1: Topology of the service provider network

Fig. 3: Network topology with backdoor Frame Relay Link

inter-area (IA O) and external (O E2) routers, Network traffic between CE1 and CE2 will Flow through Frame Relay backup link which is much slower than link going through MPLS backbone. OSPF provides no scalable solution to avoid this by changing administrative cost of the links [9]. The only solution to get around this issue is to configure OSPF Sham-Link.

III. OSPF SHAM-LINK

In an OSPF AS, every area has to connect to area 0 either through a direct or through a virtual link; OSPF Super-backbone feature overcomes this requirement. In OSPF Super-Backbone [8], service provider MP-BGP becomes the Super-Backbone and all areas has to connect to it instead of area 0. Doing so PE routers become Area Border Routers (ABR) what were ASBR previously when area 0 was backbone. Even now for CE1 and CE2, preferable link is backdoor Frame Relay because routes learnt through service provider are Inter-Area (IA O) and where as those learnt through FR are intra- area routes (O).

Since MPLS backbone link is faster than the backup link, it is desired that OSPF prefer MPLS backbone over backup link.

Sham Link is configured on PE routers to make OSPF prefer the MPLS VPN routes over intra-area backup route. Sham link simply just makes an intra-area link between PE routers [10]. To make MPLS VPN preferred we can configure metrics for sham-link and we also raise the cost of OSPF for backup routes on CE routers on right interfaces to make MPLS VPN a preferred path.

A sham-link makes a relation between two VRFs [10] configured on the PE routers. Each of these VRFs is associated with a loopback address which is called end-point address, these end-point addresses are then advertised by BGP.

A) Configuration of OSPF Sham-Link

To make OSPF prefer MPLS backbone path over FR backdoor path, OSPF sham-Link must have to be configured between PE1 and PE2. Sham-Link actually creates an intra-area link between PE routers. Fig. 4 shows the configuration entered on PE1 and PE2 to established OSPF Sham-Link between them. Following commands creates Sham-Link between PE1 and PE2.

We associate loopback addresses to the VRFs of Customers sites on PE router, these loopback addresses are called Sham-Link end point addresses.

These end point addresses are advertises by PE routers to each other through BGP so that they may ensure that the end point addresses are reachable [11].

The actual Sham-Link is configured under router OSPF mode. Area 0 is the area ID on both the PE routers and a lower OSPF cost (cost 2) is configured so that OSPF may prefer MPLS backbone link over backdoor link. We also configured high OSPF costs (cost 100) on interfaces on CE sites connecting backdoor link.

B) Verification

We use show commands to verify successful configuration and desired results of OSPF Sham-Link. Figure 5 shows all routes learnt by CE1 and CE2 routers. It is clear that despite the fact that connection through MPLS backbone is faster, with no OSPF Sham-Link configured on PE routers of the service provider, the communication between CE1 and CE2 takes place through backup FR link because OSPF prefers intra-area routers over external routes. This is shown in the Fig. 5.

The traffic between CE1 and CE2 is forwarded through external routes (O E2) via 192.168.15.0/24 subnet, which is backup FR connection.

To fix this problem we have configured sham link between PE1 and PE2. After configuring sham link, MPLS backbone network is preferred over backdoor FR link. It is shown in the Fig. 6.

Now since, OSPF inserts an intra-area link between PE1 and PE2 routers. CE1 and CE2 advertise their routes as intra-area routes to each other. Site CE1 now forwards its traffic to site CE2 via 192.168.10.2, hence through MPLS network and customer site CE2 forwards its traffic via 192.168.20.2, hence MPLS network.

Fig. 7 shows the sham link between PE1 and PE2. For PE1 the sham link end point address is 10.20.20.20 and for PE2 the end point address is 10.10.10.10.

Fig. 8 shows successful connectivity between CE1 and CE2 through MPLS service provider network.

IV. CONCLUSION

In this research work firstly, we presented an overview of OSPF and redundancy issue caused by OSPF. Secondly, we engineered a network model to explain problem rose by redundant link in OSPF network and then fixed it by configuring OSPF Sham-Link on the required PE routers of the service provider.

REFERENCES

[1]   Cisco, 2005. Introduction. http://www.cisco.com/en/US/tech/tk365/technologies_white_paper09186a0080094e9e.shtml [Accessed: 20 Jan 2013]

[2]   Cisco, 2005. Shortest Path First Algorithm. [Online] http://www.cisco.com/en/US/tech/tk365/technologies_white_paper09186a0080094e9e.shtml#t5 [Accessed: 22 Jan 2013]

[3]   Charles M. Kozierok, 2005. OSPF Hierarchical Topology [Online] http://www.tcpipguide.com/free/t_OSPFHierarchicalTopologyAreasandRouterRoles-3.htm [Accessed: 23 Jan 2013]

[4]   Stretch, 2008. OSPF Area Types. [Online] http://packetlife.net/blog/2008/jun/24/ospf-area-types/ [Accessed : 25 Jan 2013]

[5]   Microsoft Technet, 2013. OSPF Area Types [Online] http://technet.microsoft.com/en-us/library/cc957882.aspx [Accessed : 24 Jan 2013]

[6]   CCIETalk, 2008. OSPF LSA Types [Online] http://www.ccietalk.com/2008/07/13/ospf-lsa-types [Accessed On 27 Jan 2013]

[7]   Allan Reid, Jim Lorenz, Sheryl, 2008. Routing Using OSPF Protocol [Online]

[8]   James Reagon, 2002. CCIP MPLS Study Guide, pp 303. SYBEX Inc., Marina Village Parkway, Alameda, CA, USA.

[9]   Bryan Bartik, 2010. Introduction to OSPF-Sham Link [online] http://blog.ipexpert.com/2010/01/20/introduction-to-ospf-sham-link/ [Accessed: 28 Jan 2013]

[10]   Amit . Bhagat, 2011. OSPF Sham-Link Support for MPLS VPN. [Online] https://sites.google.com/site/amitsciscozone/home/important-tips/mpls-wiki/ospf-sham-link-support-for-mpls-vpn [Accessed on 30 Jan 2013]

[11]   Cisco, 2012. Sham-Link Configuration Example [Online] http://www.cisco.com/en/US/docs/ios/12_2t/12_2t8/feature/guide/ospfshmk.html#wp1024995 [Accessed: 30 Jan 2013]

Table 1: Addressing scheme used at PE and CE routers

IP Addressing in Implemented Network Model			
PE1	Serial 1/0 10.0.1.1/24	PE2	Serial 1/0 10.0.10.1/24
	Serial 1/1 192.168.10.2/24		Serial 1/1 192.168.20.2/24
	Loopback 1 10.10.10.10/32		Loopback 1 10.20.20.20/32
CE1	Serial 1/0 192.168.10.1/24	CE2	Serial 1/0 192.168.20.1/24
	Serial 1/1 192.168.15.1/24		Serial 1/1 192.168.15.2/24
	Loopback 0 192.168.150.1/24		Loopback 0 192.168.160.1/24
	Loopback 1 192.168.151.1/24		Loopback 1 192.168.161.1/24

PE1	PE2
! interface Loopback1 ip vrf forwarding cus-1-B ip address 10.10.10.10 255.255.255.255 ! router bgp 50 ! address-family ipv4 vrf cus-1-B network 10.10.10.10 mask 255.255.255.255 exit-address-family ! router ospf 10 vrf cus-1-B area 0 sham-link 10.10.10.10 10.20.20.20 cost 2 !	! interface Loopback1 ip vrf forwarding cus-2-B ip address 10.20.20.20 255.255.255.255 ! router bgp 50 ! address-family ipv4 vrf cus-2-B network 10.20.20.20 mask 255.255.255.255 exit-address-family ! router ospf 10 vrf cus-2-B area 0 sham-link 10.20.20.20 10.10.10.10 cost 2 !

Fig. 4: Configuration entered for Sham-Link on PE1 and PE2 routers

```
CE1#show ip route
Gateway of last resort is not set

C     192.168.151.0/24 is directly connected, Loopback1
C     192.168.15.0/24 is directly connected, Serial1/1
C     192.168.150.0/24 is directly connected, Loopback0
C     192.168.10.0/24 is directly connected, Serial1/0
      192.168.160.0/32 is subnetted, 1 subnets
O        192.168.160.1 [110/201] via 192.168.15.2, 00:00:36, Serial1/1
      192.168.161.0/32 is subnetted, 1 subnets
O        192.168.161.1 [110/201] via 192.168.15.2, 00:00:36, Serial1/1
O     192.168.20.0/24 [110/264] via 192.168.15.2, 00:00:42, Serial1/1
      10.0.0.0/32 is subnetted, 2 subnets
O E2     10.10.10.10 [110/1] via 192.168.10.2, 00:00:42, Serial1/0
O E2     10.20.20.20 [110/1] via 192.168.15.2, 00:00:42, Serial1/1


CE2#show ip route
Gateway of last resort is not set

      192.168.151.0/32 is subnetted, 1 subnets
O        192.168.151.1 [110/201] via 192.168.15.1, 00:00:54, Serial1/1
C     192.168.15.0/24 is directly connected, Serial1/1
      192.168.150.0/32 is subnetted, 1 subnets
O        192.168.150.1 [110/201] via 192.168.15.1, 00:00:54, Serial1/1
O     192.168.10.0/24 [110/264] via 192.168.15.1, 00:00:54, Serial1/1
C     192.168.160.0/24 is directly connected, Loopback0
C     192.168.161.0/24 is directly connected, Loopback1
C     192.168.20.0/24 is directly connected, Serial1/0
      10.0.0.0/32 is subnetted, 2 subnets
O E2     10.10.10.10 [110/1] via 192.168.15.1, 00:01:00, Serial1/1
O E2     10.20.20.20 [110/1] via 192.168.20.2, 00:01:00, Serial1/0
```

Fig. 5: Traffic between CE1 and CE2 is forwarded through undesired FR backup link

```
CE1#show ip route

Gateway of last resort is not set

C    192.168.151.0/24 is directly connected, Loopback1
C    192.168.15.0/24 is directly connected, Serial1/1
C    192.168.150.0/24 is directly connected, Loopback0
C    192.168.10.0/24 is directly connected, Serial1/0
     192.168.160.0/32 is subnetted, 1 subnets
O       192.168.160.1 [110/131] via 192.168.10.2, 00:00:07, Serial1/0
     192.168.161.0/32 is subnetted, 1 subnets
O       192.168.161.1 [110/131] via 192.168.10.2, 00:00:07, Serial1/0
O    192.168.20.0 [110/130] via 192.168.10.2, 00:00:07, Serial1/0
     10.0.0.0/32 is subnetted, 2 subnets
O E2    10.10.10.10 [110/1] via 192.168.10.2, 00:00:08, Serial1/0
O E2    10.20.20.20 [110/1] via 192.168.10.2, 00:00:08, Serial1/0

CE2#show ip route
Gateway of last resort is not set

     192.168.151.0/32 is subnetted, 1 subnets
O       192.168.151.1 [110/131] via 192.168.20.2, 00:00:35, Serial1/0
C    192.168.15.0/24 is directly connected, Serial1/1
     192.168.150.0/32 is subnetted, 1 subnets
O       192.168.150.1 [110/131] via 192.168.20.2, 00:00:35, Serial1/0
O    192.168.10.0/24 [110/130] via 192.168.20.2, 00:00:35, Serial1/0
C    192.168.160.0/24 is directly connected, Loopback0
C    192.168.161.0/24 is directly connected, Loopback1
C    192.168.20.0/24 is directly connected, Serial1/0
     10.0.0.0/32 is subnetted, 2 subnets
O E2    10.10.10.10 [110/1] via 192.168.20.2, 00:00:36, Serial1/0
O E2    10.20.20.20 [110/1] via 192.168.20.2, 00:00:36, Serial1/0
```

Fig. 6: Traffic between CE1 and CE2 is forwarded through Service Provider after configuration of OSPF Sham-Link between PE1 and PE2

```
PE1#show ip ospf sham-links
Sham Link OSPF_SL0 to address 10.20.20.20 is up
Area 0 source address 10.10.10.10
  Run as demand circuit
  DoNotAge LSA allowed. Cost of using 2 State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40,
    Hello due in 00:00:00
    Adjacency State FULL (Hello suppressed)
    Index 2/2, retransmission queue length 0, number of retransmission 0
    First 0x0(0)/0x0(0) Next 0x0(0)/0x0(0)
    Last retransmission scan length is 0, maximum is 0
    Last retransmission scan time is 0 msec, maximum is 0 msec
PE2#
..............................................................

PE2#sho ip ospf sham-links
Sham Link OSPF_SL0 to address 10.10.10.10 is up
Area 0 source address 10.20.20.20
  Run as demand circuit
  DoNotAge LSA allowed. Cost of using 2 State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40,
    Hello due in 00:00:00
    Adjacency State FULL (Hello suppressed)
    Index 2/2, retransmission queue length 0, number of retransmission 0
    First 0x0(0)/0x0(0) Next 0x0(0)/0x0(0)
    Last retransmission scan length is 0, maximum is 0
    Last retransmission scan time is 0 msec, maximum is 0 msec
PE2#
```

Fig. 7: OSPF Sham-Link between PE1 and PE2

```
CE1#ping 192.168.160.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.160.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 92/140/212 ms
....................................................................
CE2#ping 192.168.150.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.150.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 4/35/60 ms
```

Fig. 8: Successful connectivity between CE1 and CE2