Open Shortest Path First

Xiang Mei

March 17, 2022

1 Prologue

This is a personal note for OSPF. OSPF can discover neighbors, perform reliable flooding to transfer the information, compute the shortest path to the arbitrary router, and detect new changes in the system. Also, the networking system on OSPF is dynamic and it works on the IP layer (Protocol Number 89). The OSPF is a complex protocol and I would only introduce some basic and important procedures and concepts of OSPF. Besides, I would like to attach some screenshots of real OSPF packages to help understand.

2 OSPF

In this section, I gonna introduce some simple concepts of OSPF.

In the OSPF subnet, we have lots of routers and they run OSPF to dynamically build the shortest path to arbitrary hosts on the subnet. They collect the link state by sending link state packages to their neighbors so that they can use these link-state packages to build a database and compute some shortest-path trees.

By building these threes, every node knows the shortest way to reach a node. And in OSPF, we can detect the changes in the subnets and rebuild the forwarding table. We also split a big subnet into several small subnets, I'll explore the reasons behind these mechanisms.

For this passage, I can't ensure all my analysis is correct, please feel free to correct me.

3 Area

Splitting a big subnet to different areas is an important concept. In OSPF, we would have a backbone area. This area would abut all other areas and all the ABR(Area Border Router) are at the border of the backbone area, which is used to connect other areas to the backbone area. So why do we need to split a big subnet into several areas?

As we know, every router in the OSPF subnet would have a database to store the link-state information. These records If all the routers on the earth are in one area, every router has to store all the links' state that would take too much memory. And in a big subnet, the link-state would keep changing so that all the routers have to keep computing the new shortest path.

4 Link-State

In OSPF, every node has a full topology and every node could run some algorithm, such as Dijkstra, to compute the shortest path to an arbitrary node on the subnet. If some changes on the subnet are detected, such as down nodes and broken links, the node would inform other nodes so that every node could know the new topology and rebuild the shortest paths.

Metric. How do the nodes evaluate the costs of a specific link? In OSPF, the nodes use bandwidth to computer the costs.

$$Cost = \frac{100Mbps}{LinkSpeed}$$

For example, we have two routers A and B, and the link from A to B is 10Mbps. We can get the cost by above formula $Cost = 100/10 \rightarrow 10$.The faster, the less cost. That makes sense choosing the least costs would achieve the best performance.

5 Multicast

Unicast is the base of point2point communication and using broadcast could send messages to all the hosts in a broadcast domain while multicast could send packages to a group of hosts. Multicast has some advantages, such as it only duplicates the message when in need and provides efficient control of a group of hosts.

Multicast uses Class D IP address space. "Could a switch recognize the IP? and only sends the packages to the hosts in a specific group?" Obviously, the answer is no, the switch works on layer two and it even doesn't know what's the "IP". Actually, similar to the broadcast MAC address, we have special reserved MAC addresses for multicast. And the switch would default forward the multicast packets to all the outports. And IGMP is used to eliminate unecessary flooding.

Things are similar in IP, we also reserve some addresses for multicast. 224.0.0.0 - 239.255.255.255, more specifically 224.0.0.0 - 224.0.0.255 are for routing protocols. And there are some very important addresses for the OSPF.

- 224.0.0.1 -> all the hosts on subnet
- 225.0.0.2 -> all the routers on subnet
- 224.0.0.5 -> all routers on subnet who run OSPF
- 224.0.0.22 -> IGMP Traffic

6 States

For the routers running OSPF, there are 7 states. And the router moves from one to another during their working time.

Down, **Init** and **Two-Way**. In the beginning, each node doesn't know the existence of other nodes, they are in a "down" state. Once they get a hello packet from others, they are in init state. At this time the receiver knows the existence of the sender because they can find it in the OSPF header. And then they would send back a hello packet so that the sender moves to the two-way state from down or init as the sender knows that the receiver's IP and the receiver also know his IP.

And the nodes at the "two-way" state start to select the DR/BDR and this state is called **Exstart** state. After having a DR/BDR, the nodes would move to **Exchange** state. In this state, the nodes would exchange their link-state databases. They will use DBD packets to exchange the information.

After collecting database description packets, the nodes would compare the received database and the local database to confirm the lost items and send LSRequest packets to ask for some records. And other nodes would send LSUpdate packets to help build a full topology. That's the **Loading** state. After this state, the routers have enough information to generate the shortest paths to every node. Then, they only need to periodically send Hello packages to keep **Full** state.

7 Link State Packets

We just talked that the OSPF routers need to exchange information to build a full topology. Link State Packages are used to transfer this information. There are 5 types of packets, including Hello, DBD, LSU, LSR and LSAck. I'll go through them with real examples

And I also captured some real OSPF in Wireshark. You can check theseelements in the real packets.

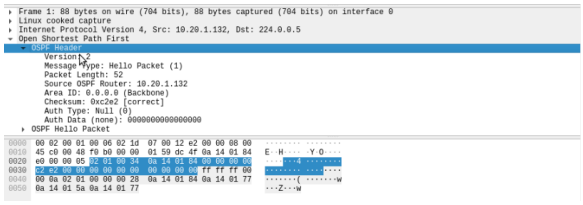

Figure 1: OSPF Packets Header in Wireshark

As we can see in figure 1, the type is LSAck and we can see the source router's ID, area id, and the authentication information in the header.

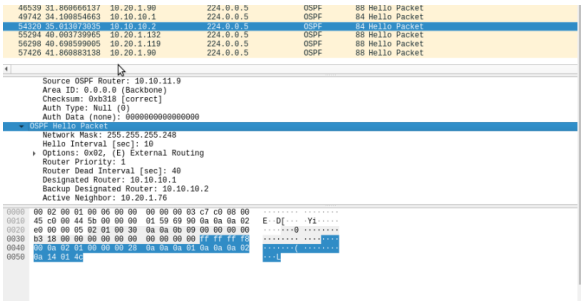

Figure 2: Hello

Hello Hello packets are the most common packet in an OSPF subnet. Nodes use this kind of packet to find new neighbors and keep adjacency. In the packet, we can see the mask, hello (dead) interval, DR, BDR, and active neighbor(s). These kinds of packets are sent periodically and that's the reason why the node could detect the changes in the subnet. If the node can't receive the hello packets from a neighbor more than the dead interval. The node would mark the neighbor as dead and the node would inform other neighbors.

DBD,LSR,LSU, nad LSAck Database Description Packets are used to synchronize the databases in the OSPF network. As we can see in the above figure, this DBD packet has three LSA. 2 type3-LSA and 1 type1-LSA, we would talk about LSA in the next section.

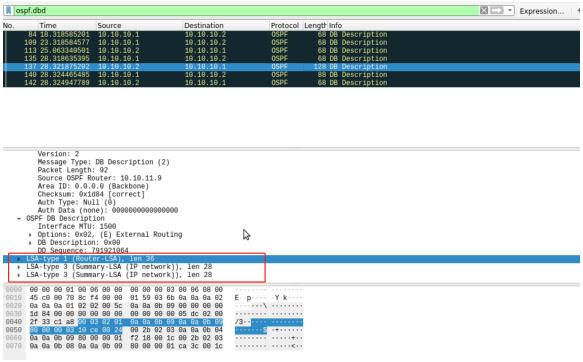

Figure 3: DBD

If the node finds that its local database could get some new LSA from the sender, it would use LSR to request the sender to give the missing LSA. figure 3 is the shown DBD's corresponding LSR. As we can see, there are three LSA requests. And a fun fact is the DBD and LSR packets are unicast rather than multicast. But, for the corresponding LSU packet, it's the multicast. I have also attached that here, you can check the LSU packet in figure 5.

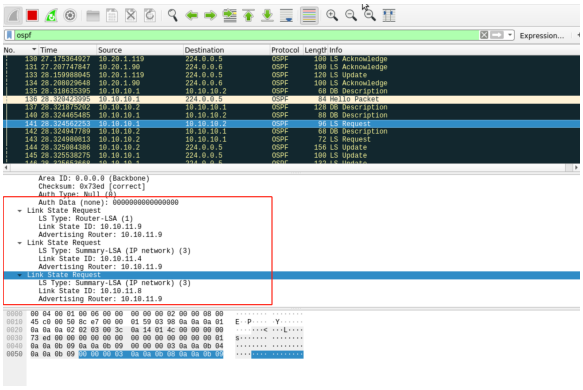

Figure 4: LSR

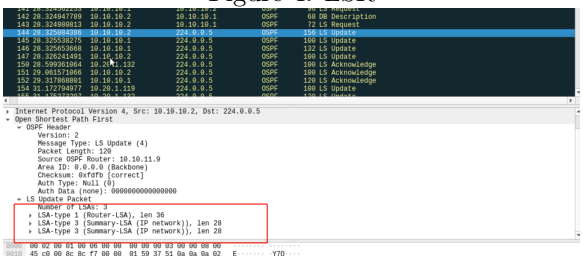

Figure 5: LSU

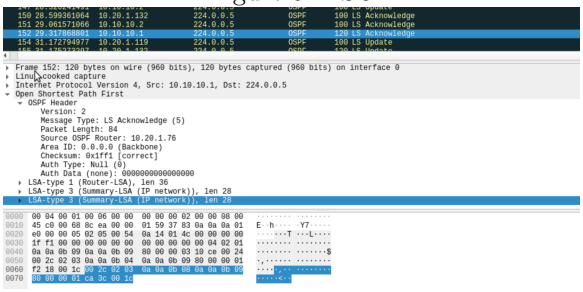

Figure 6: LSAck

The last packet to finish the synchronising procedure is the acknowledge packet. It Acknowledges that the update packets are successfully received.

The LSP exchanges the information of different routers so that the routers could have the same topology. And it also could help to detect the changes in the topology.

8 LSA

There are 8 types of LSA. But I don't want to dive too deep into one topic so I would shortly introduce the most common 3 types(1,2,3).

Type1(Router LSA), this type of LSA is used to inform other routers in the current area, would attach routers connected to the sender and lists the IP with the matric.

Type2(Network LSA), includes several attached routers and the sender's netmask.

Type3(Network Summary LSA), LSA's type 1(Router LSA), and 2(Network LSA) can't be used to transfer information from area to area. We use LSA type3 Network Summary LSA to inform routers out of the current area.