DC-HA

Overview:

In any SDWAN solutions, availability of DC is always more critical. As an interim solution, we support only ACTIVE-PASSIVE mode in DC. In this design, there will be only two DC's per tenant in the HA cluster and there will be a Gateway node through which both the DC's reach their LAN network.

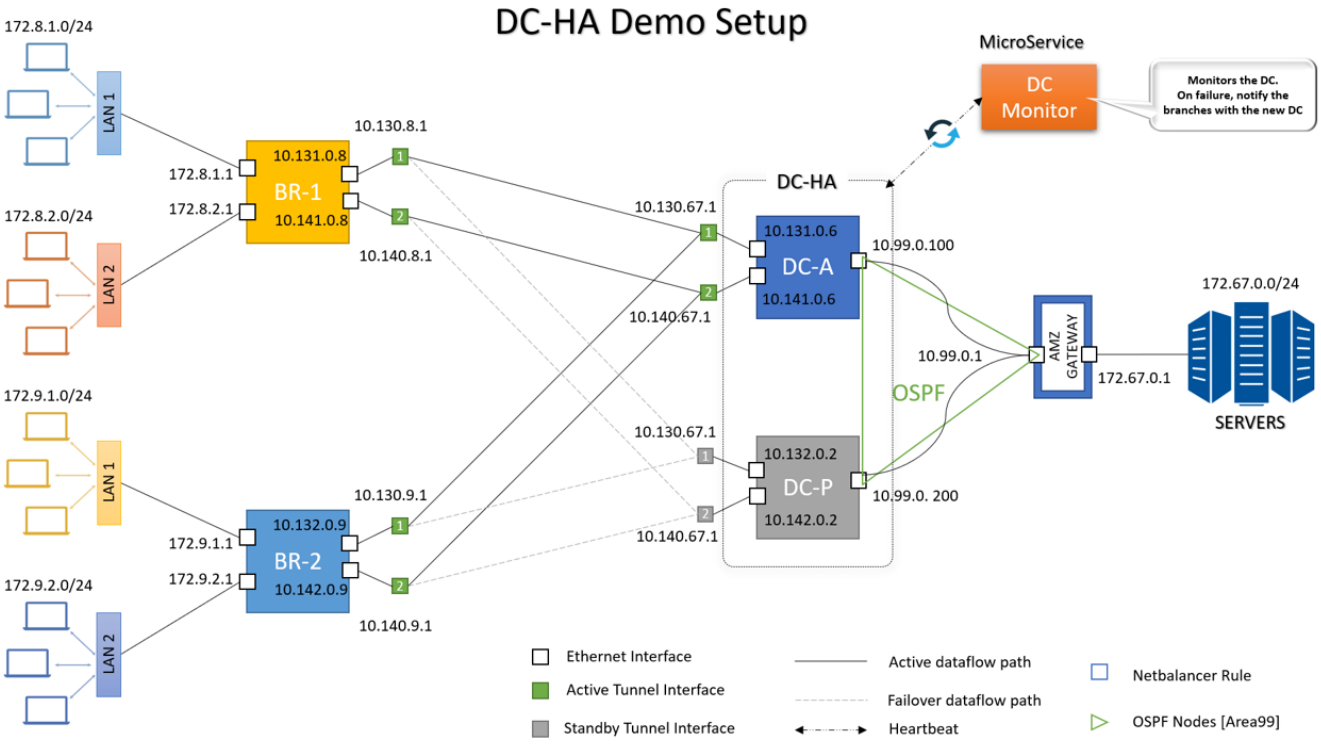
All the DC's will have public IP [WAN] as an endpoint through which all the branches will establish their IPSEC tunnels. And the tunnels are usually one to many from DC. So, if the ACTIVE DC fails that failure needs to be informed to all of their branches and reestablish the connection from the PASSIVE DC.

As of today, topology has the all the details about both the branches and their DCs w.r.t tunnels.

So, we have created a new micro-service named ***dc-monitor***, that checks the livelyness of ACTIVE DC and collects the topology for every few [*say 20*] seconds. If the ***dc-monitor*** detects any failure, then it checks for any passive DC is available to takeover. If that is available, then the ***dc-monitor*** informs all the branches to connect to the PASSIVE DC and change its state to new ACTIVE.

During these switchover, the LAN behind the DC's should reach the branches through new ACTIVE DC. Someone needs to inject the reverse route here, we are achieving this by having our vCPE as Gateway node that connects both the ACTIVE and PASSIVE DCs and all the DC-LAN traffic is reached through that Gateway. Also OSPF is configured between DC-ACTIVE, DC-PASSIVE and Gateway node, so that the reverse routed is injected seamlessly without any isuses even during DC failover and failback.

Please see the overview of our setup as shown in the *figure* below:



Setup Configuration:

Make sure that your setup is configured as shown in the *figure* above. It has 5 main nodes:

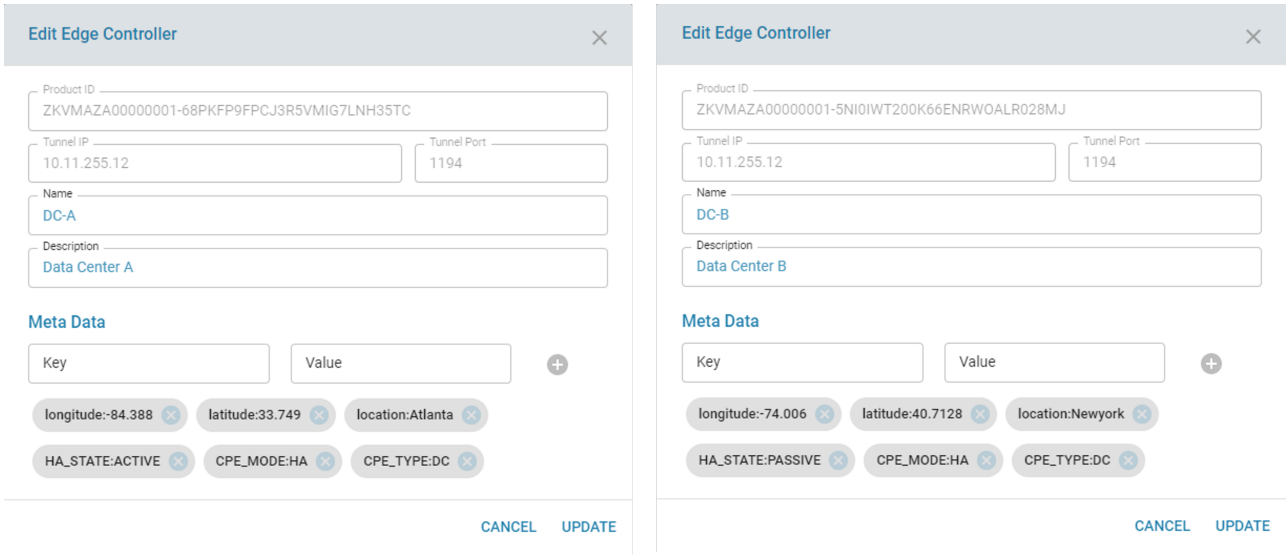
- DC-A (HA_DC1 runs in ACTIVE mode)
- DC-P (HA_DC2 runs in Passive mode)
- AMZ-Gateway (AMZ_LAN_GW node behind DC nodes)
- BR-1 (Branch 1)
- BR-2 (Branch 2)

1. Onboard AMZ_LAN_GW, HA_DC1, HA_DC2, BR1 & BR2 CPE's from secure shell. After onboarded, the Alpsee UI should list like shown below:

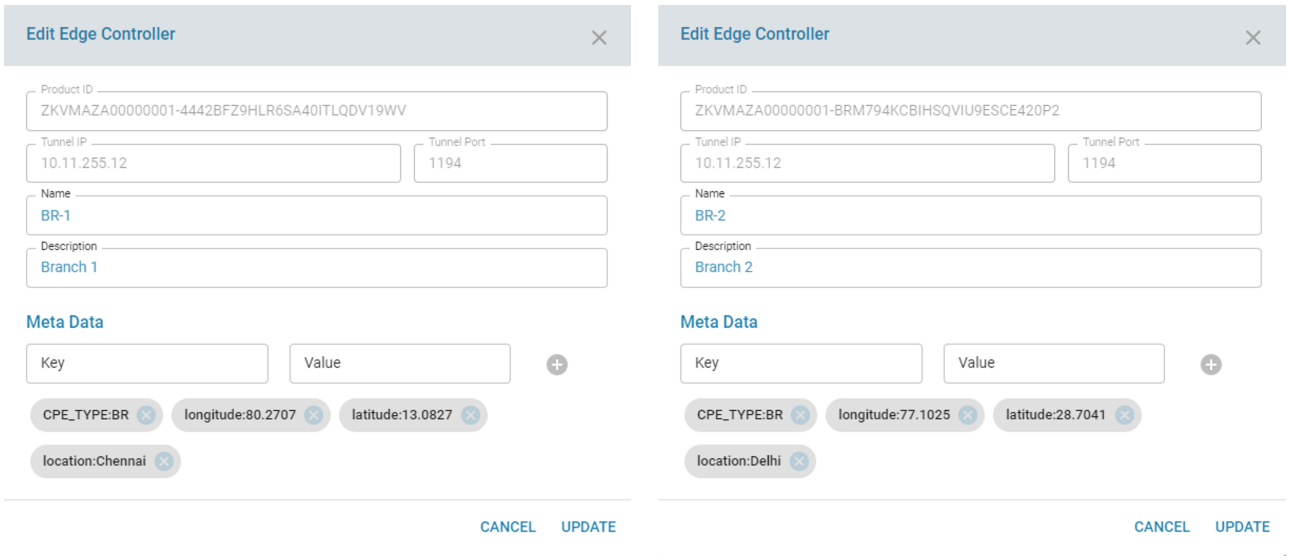
   NOTE: *Make sure that **dc_monitor** micro-service is running in the onboarding provider, without that HA will not function.*

   - While onboarding, add the **metadata** to differentiate DC and BR CPE's.

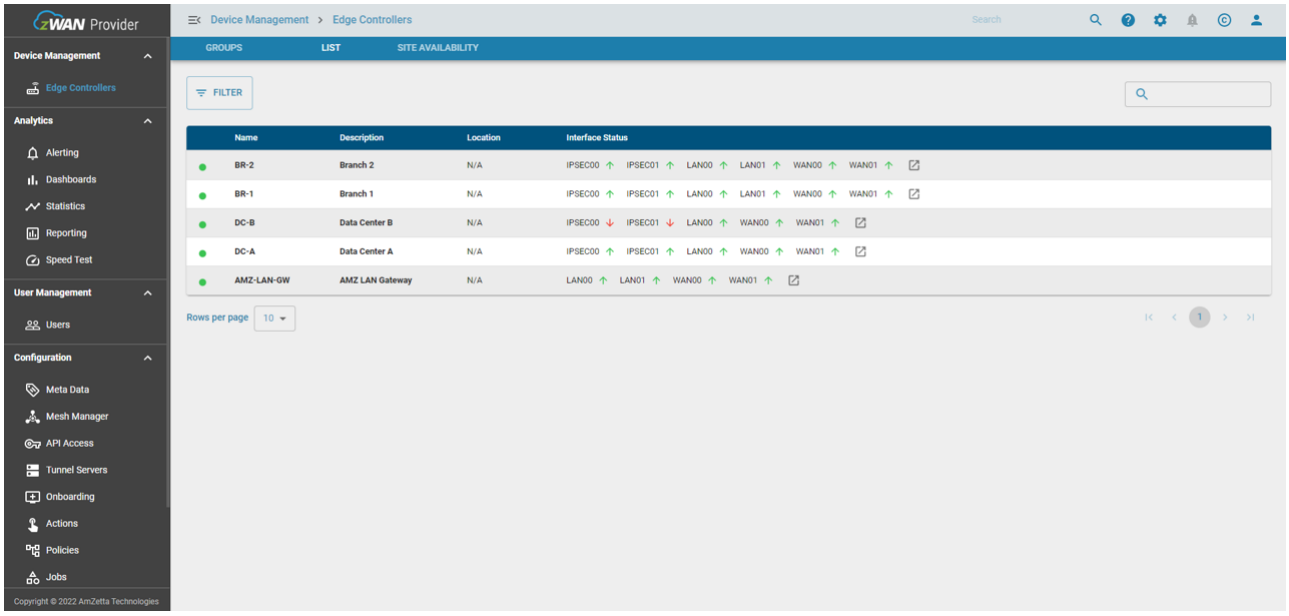
```
For HA_DC1 --> CPE_TYPE = DC, CPE_MODE = HA, HA_STATE = ACTIVE
For HA_DC2 --> CPE_TYPE = DC, CPE_MODE = HA, HA_STATE = PASSIVE
```




```
For both BR --> CPE_TYPE = BR
```



- After onboarding everything, **list** all the edge controllers

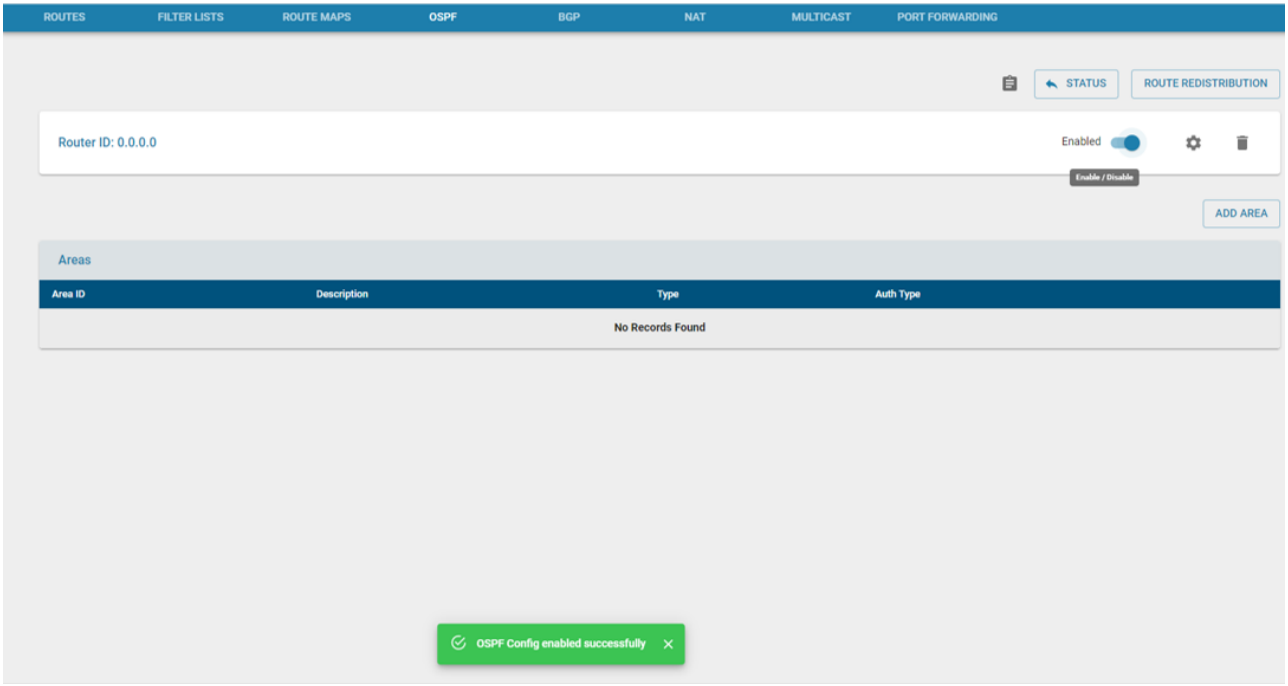


2. Now configure the OSPF in AMZ_LAN_GW, HA_DC1 & HA_DC2 CPE's [*for reverse route injection*]

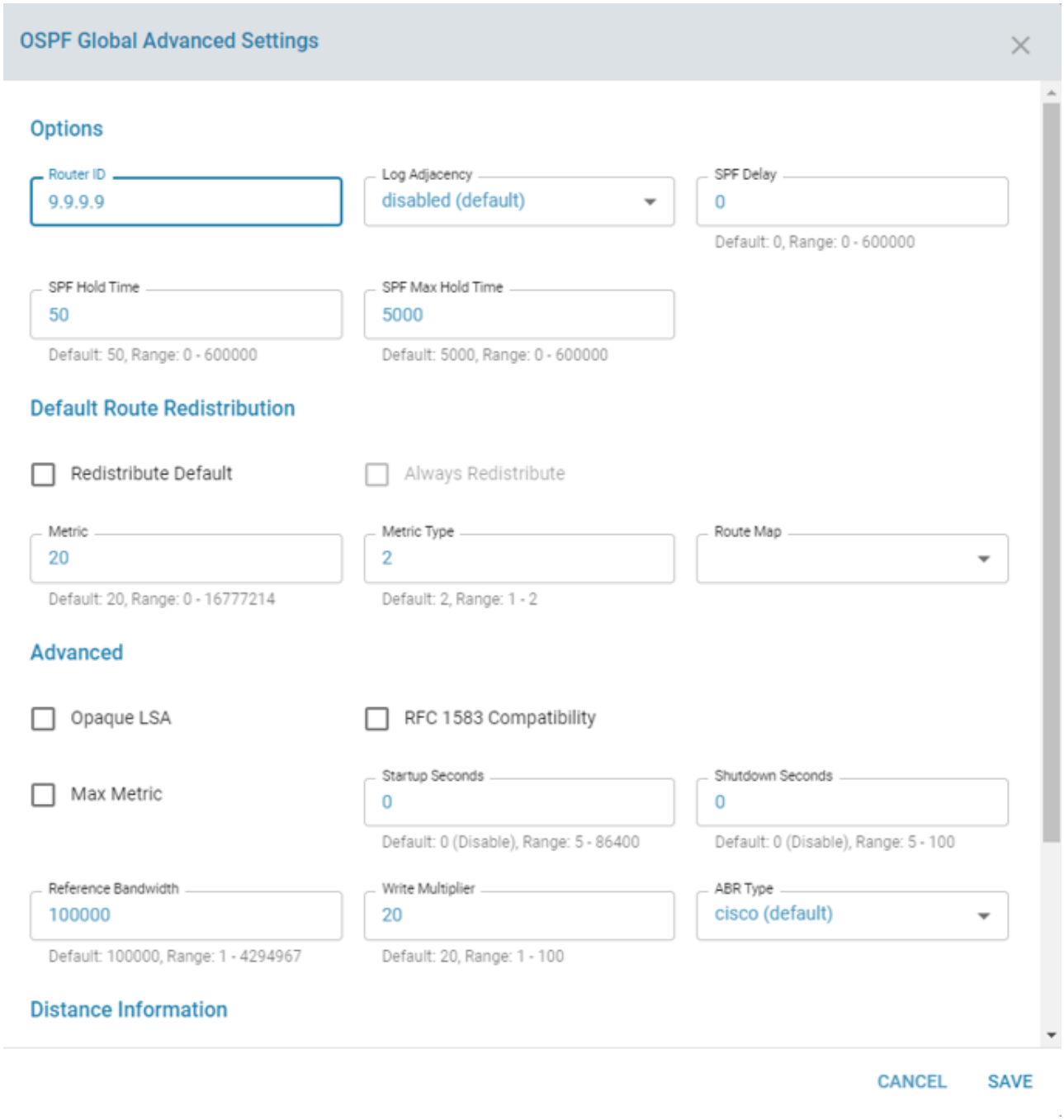
OSPF Configuration

```
Please Configure OSPF on AMZ_GATEWAY with the below mentioned steps:
```

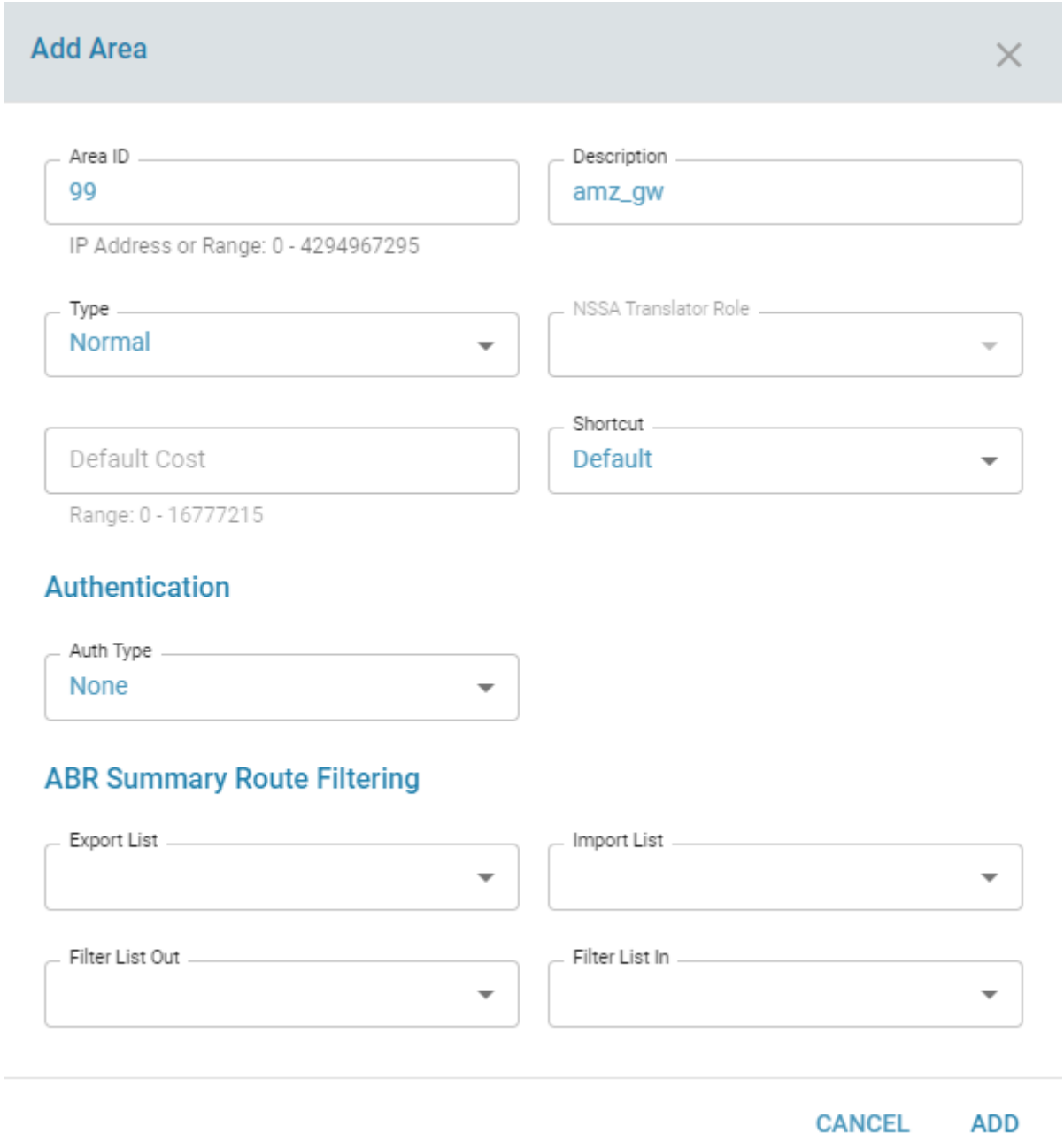
1. Enable OSPF



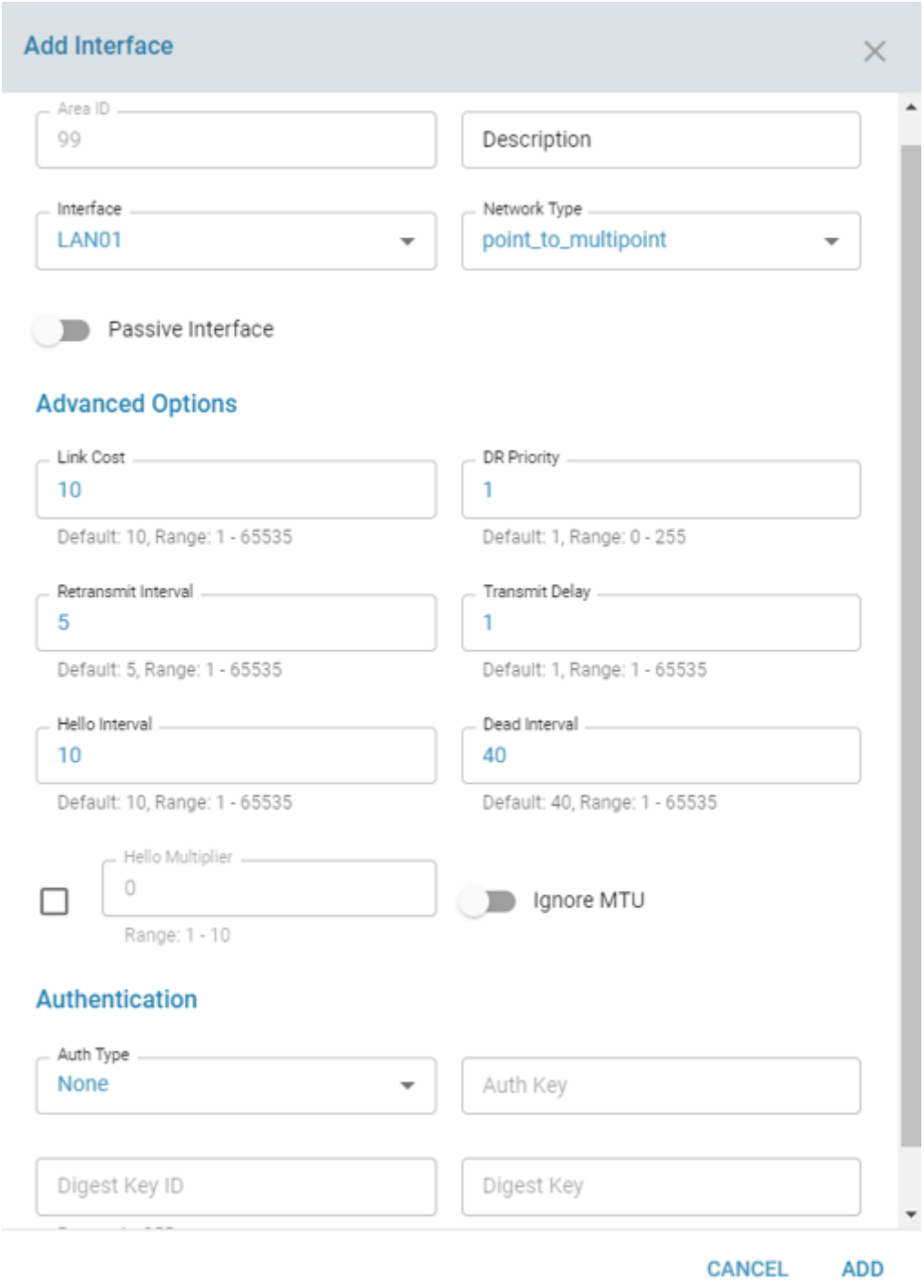
2. Add the RouterID



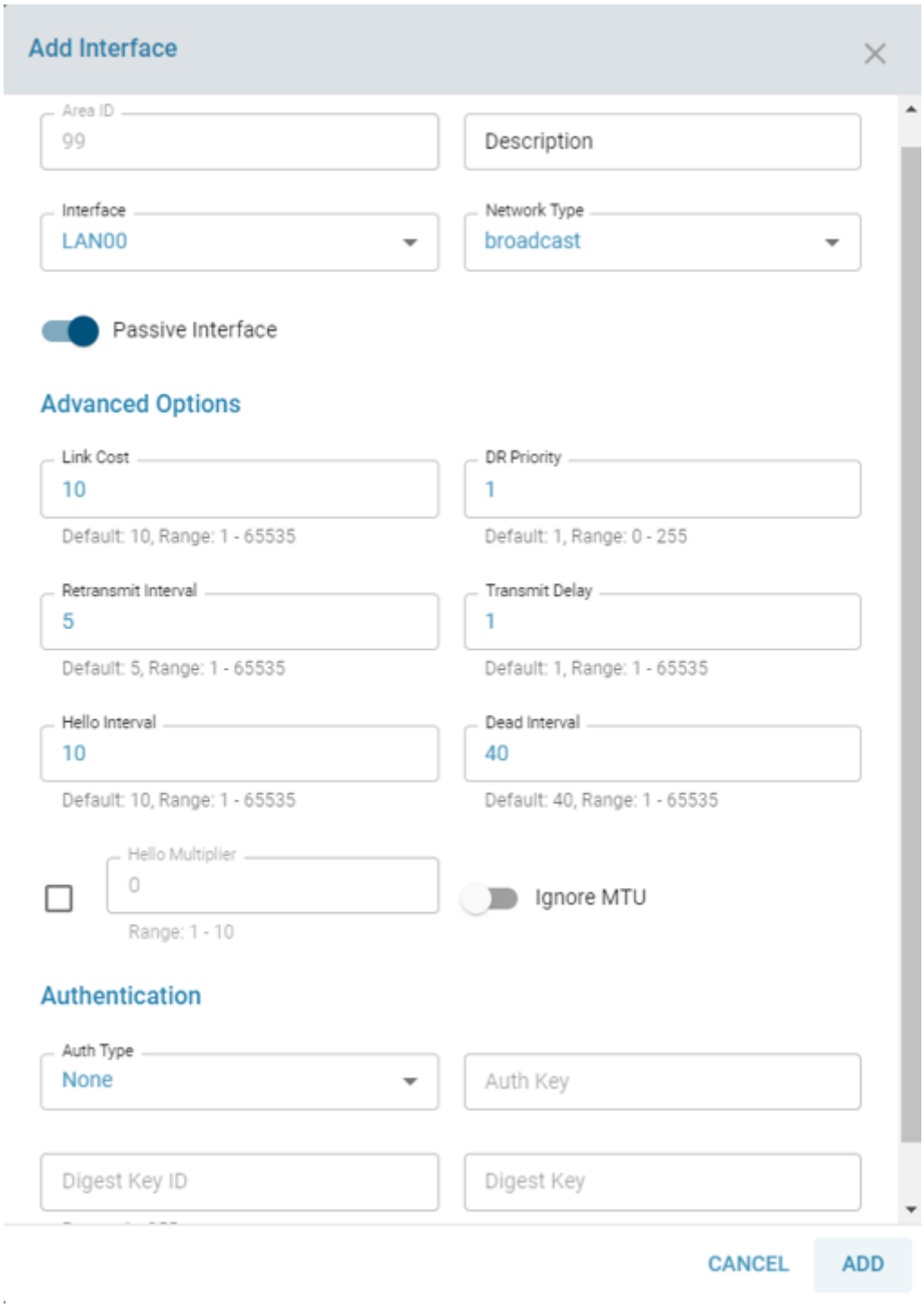
3. Add new area



4. Add active interface through which the routes are learned and select the network type as ***point_to_multicast***

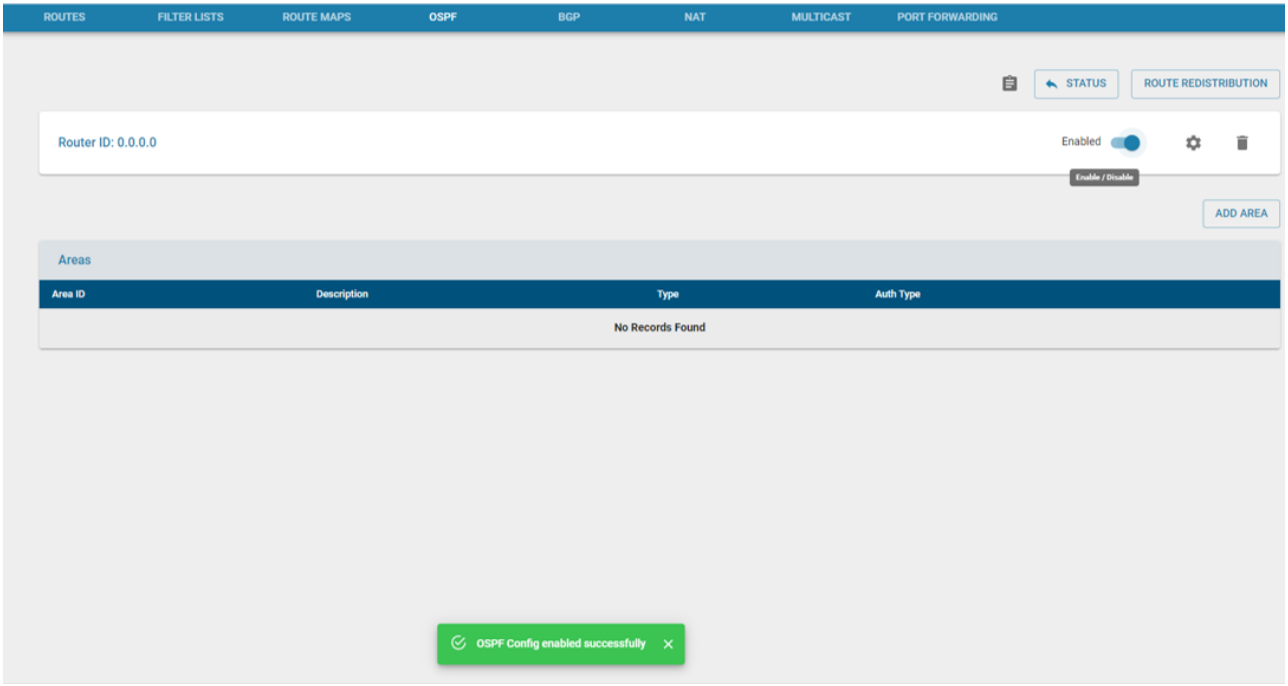


5. Add passive interface through which the DC LAN servers are reachable and select the network type as ***broadcast***

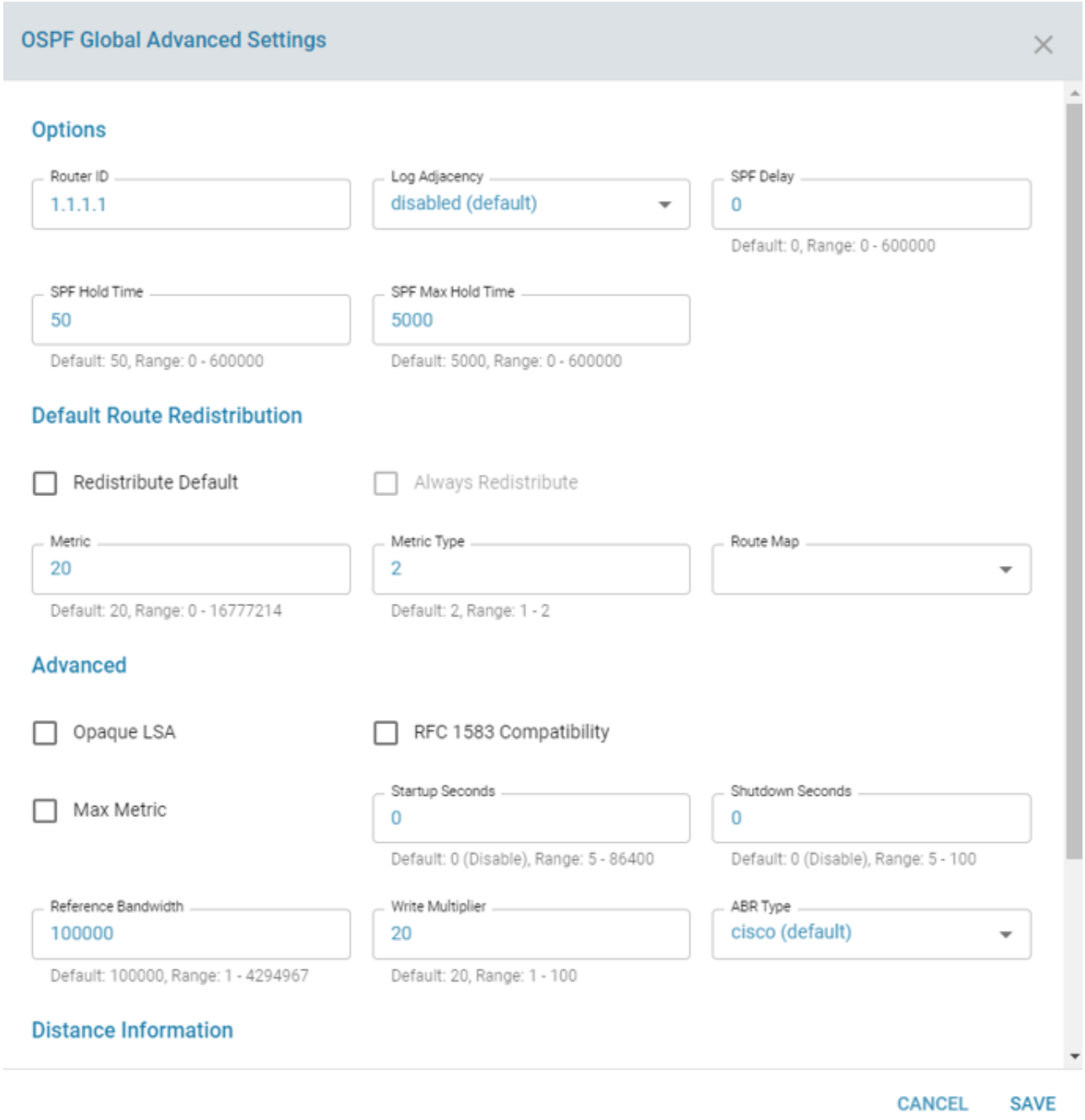



```
Please Configure OSPF on DC-A with the below mentioned steps:
```

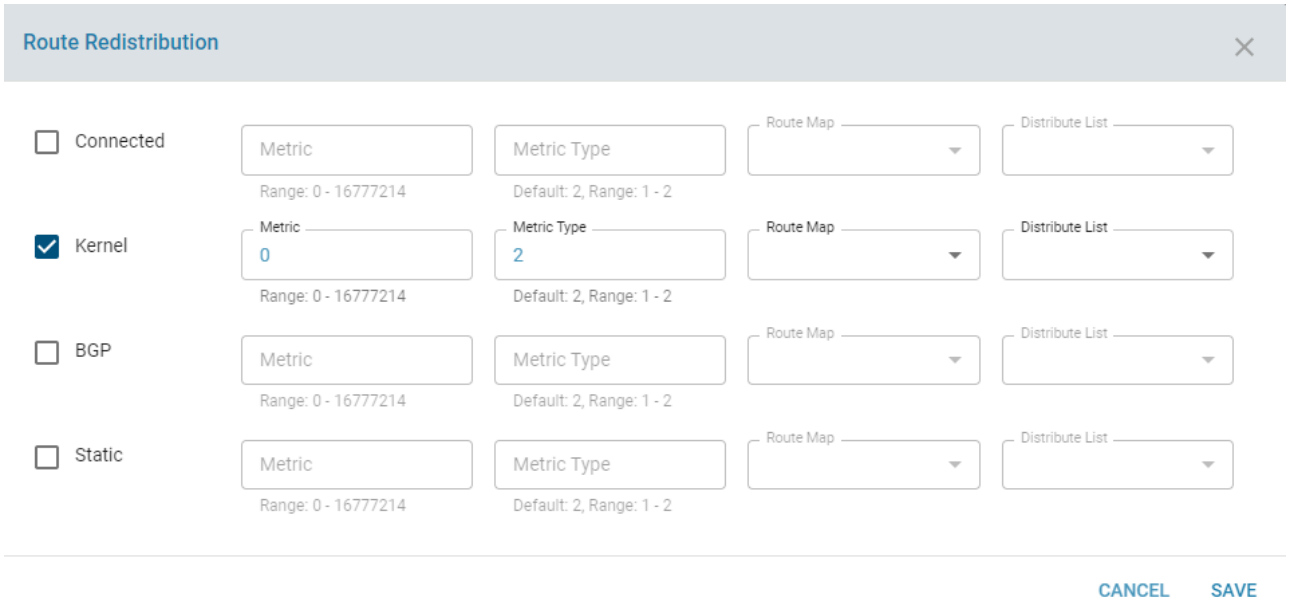
1. Enable OSPF



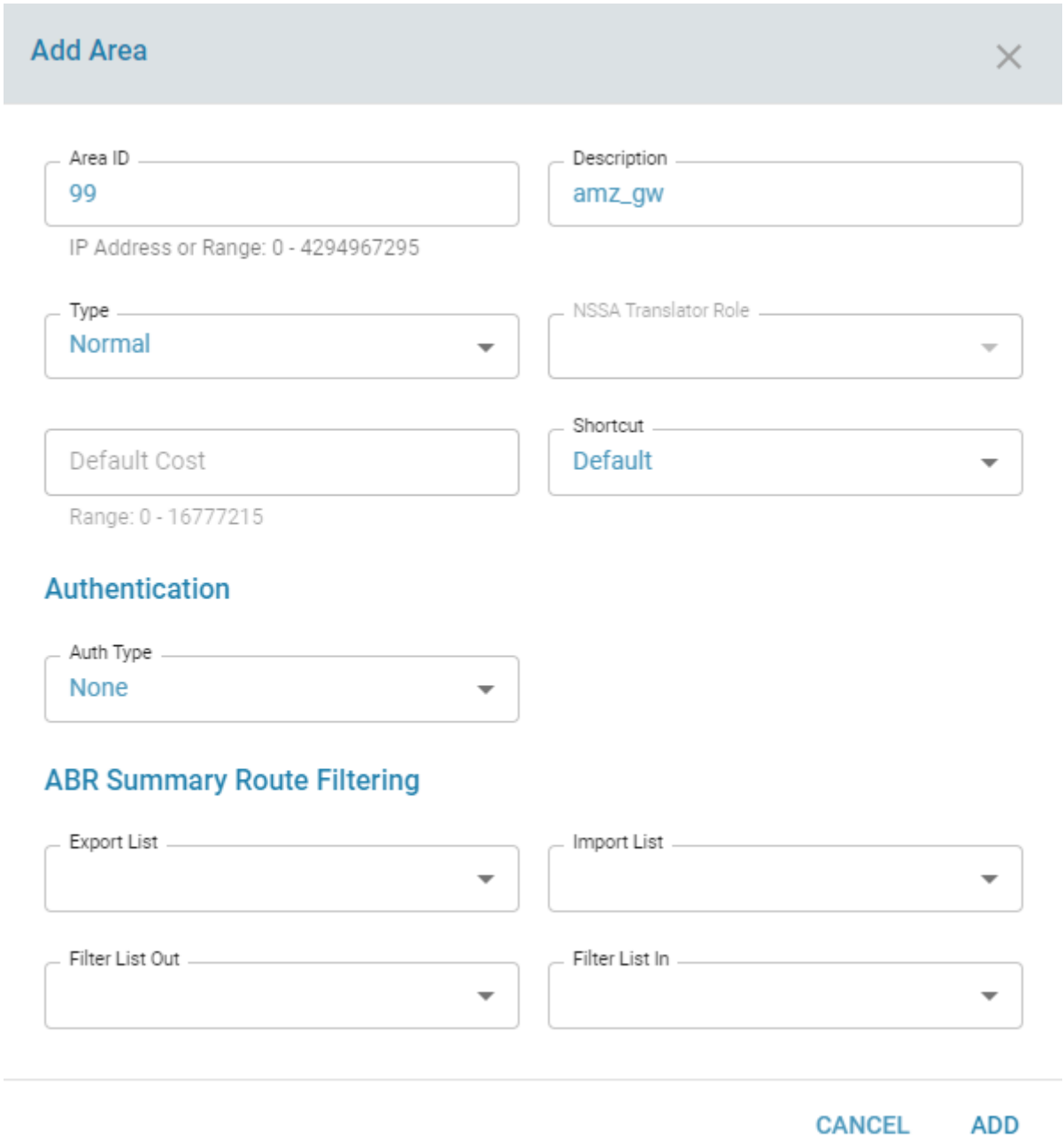
2. Add the RouterID



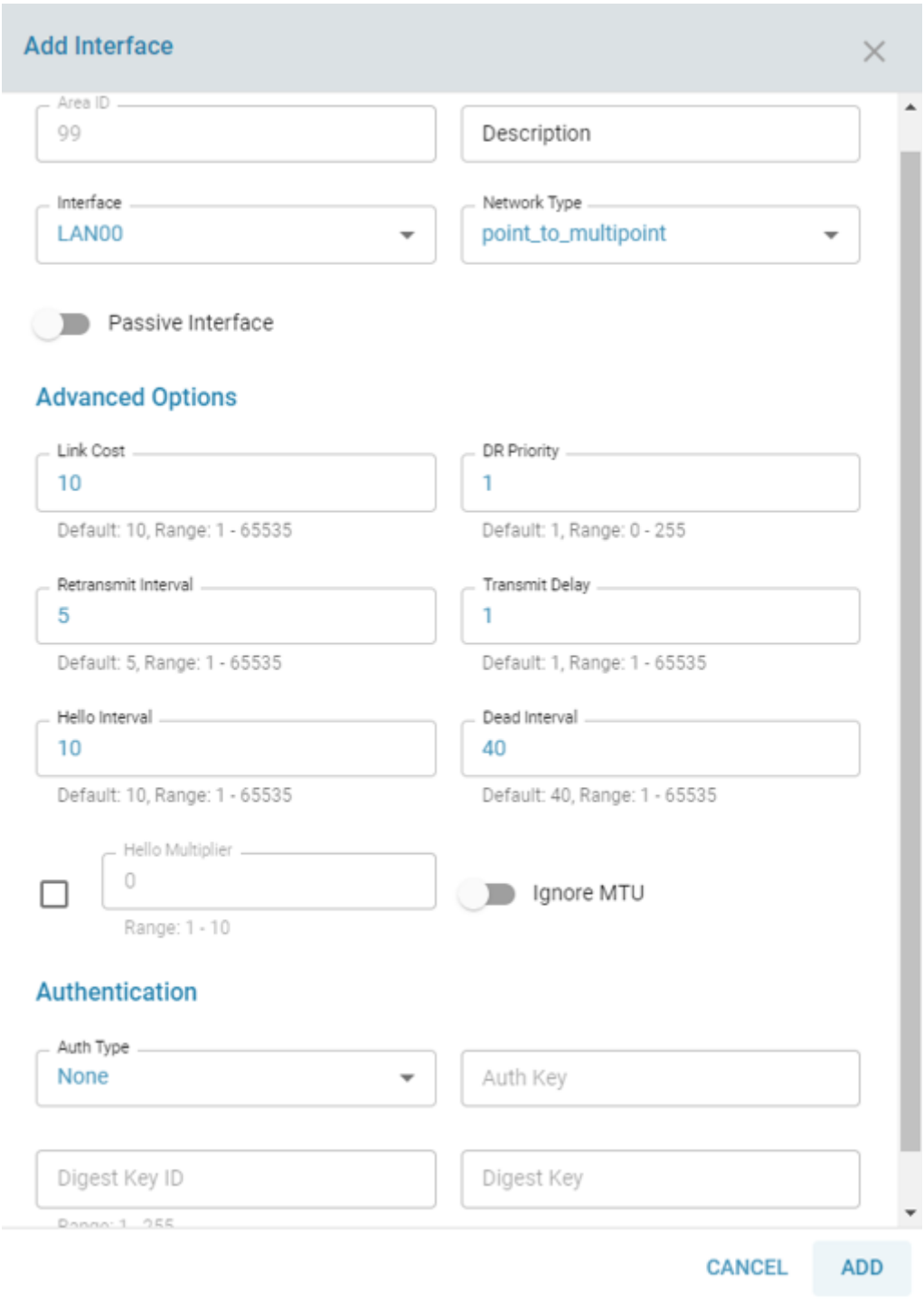
3. Add the redistribution for all kernel routes



4. Add new area

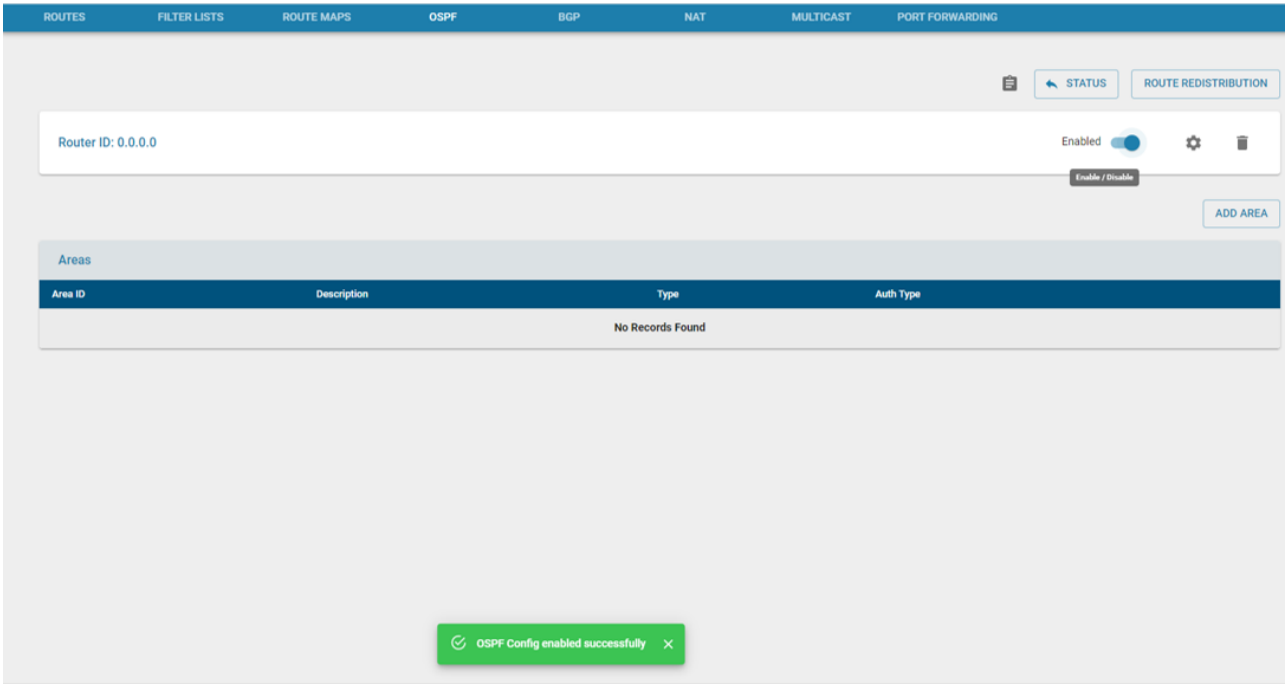


5. Add active interface through which the routes are learned and select the network type as ***point_to_multicast***

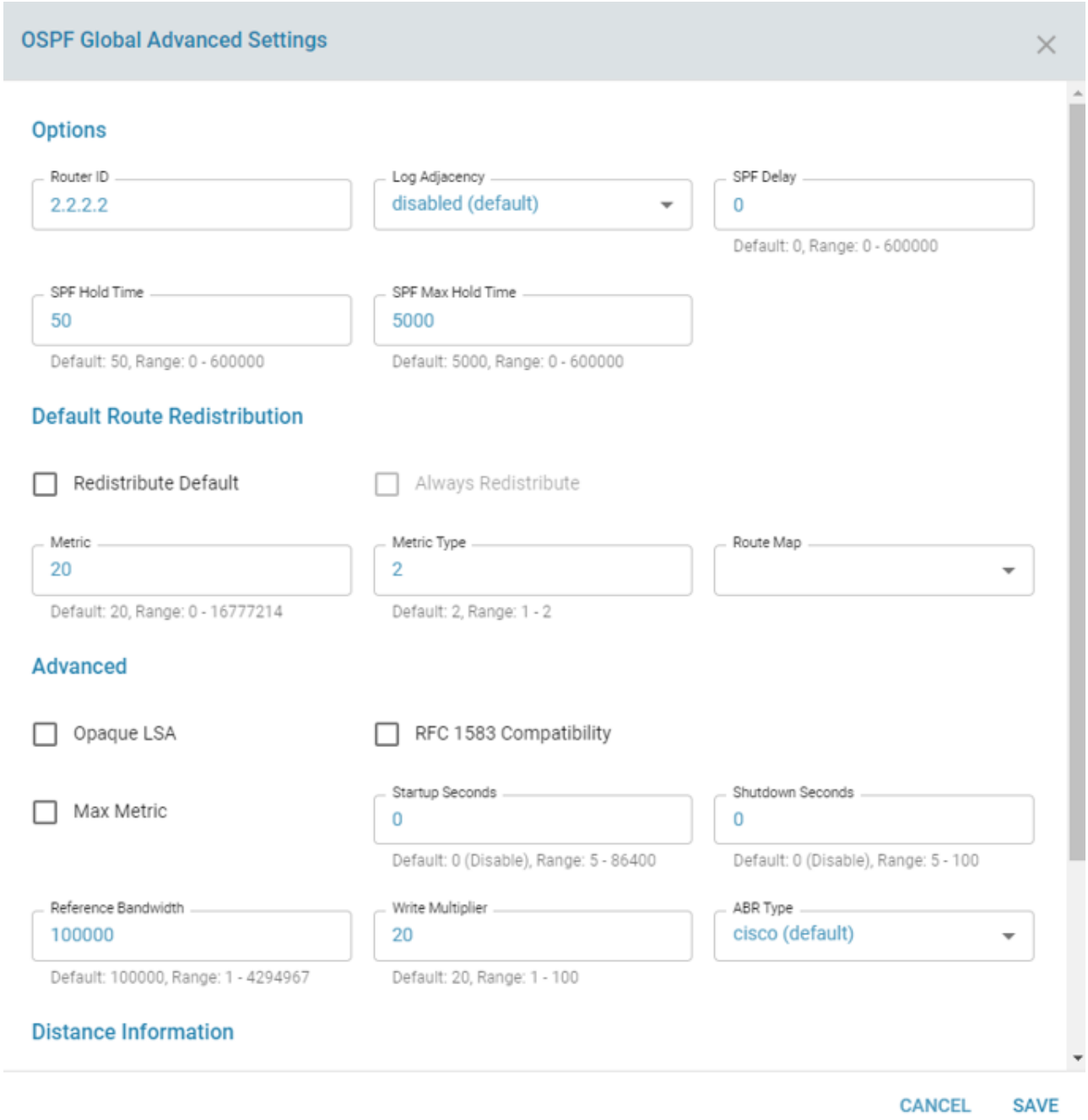



```
Please Configure OSPF on DC-P with the below mentioned steps:
```

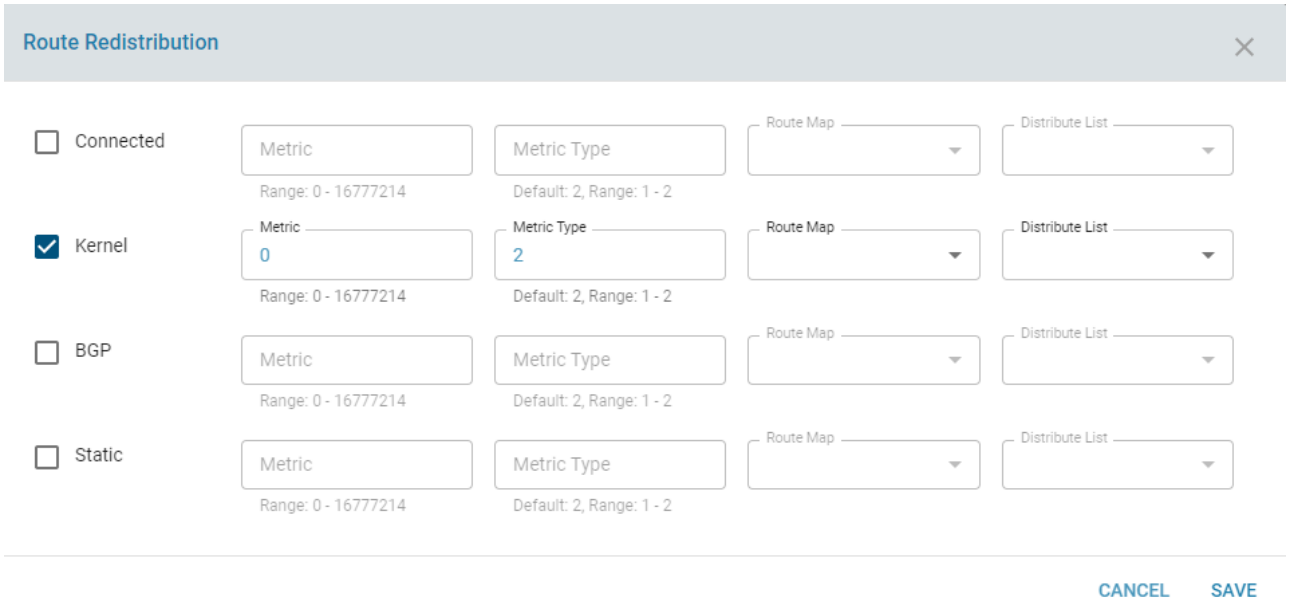
1. Enable OSPF



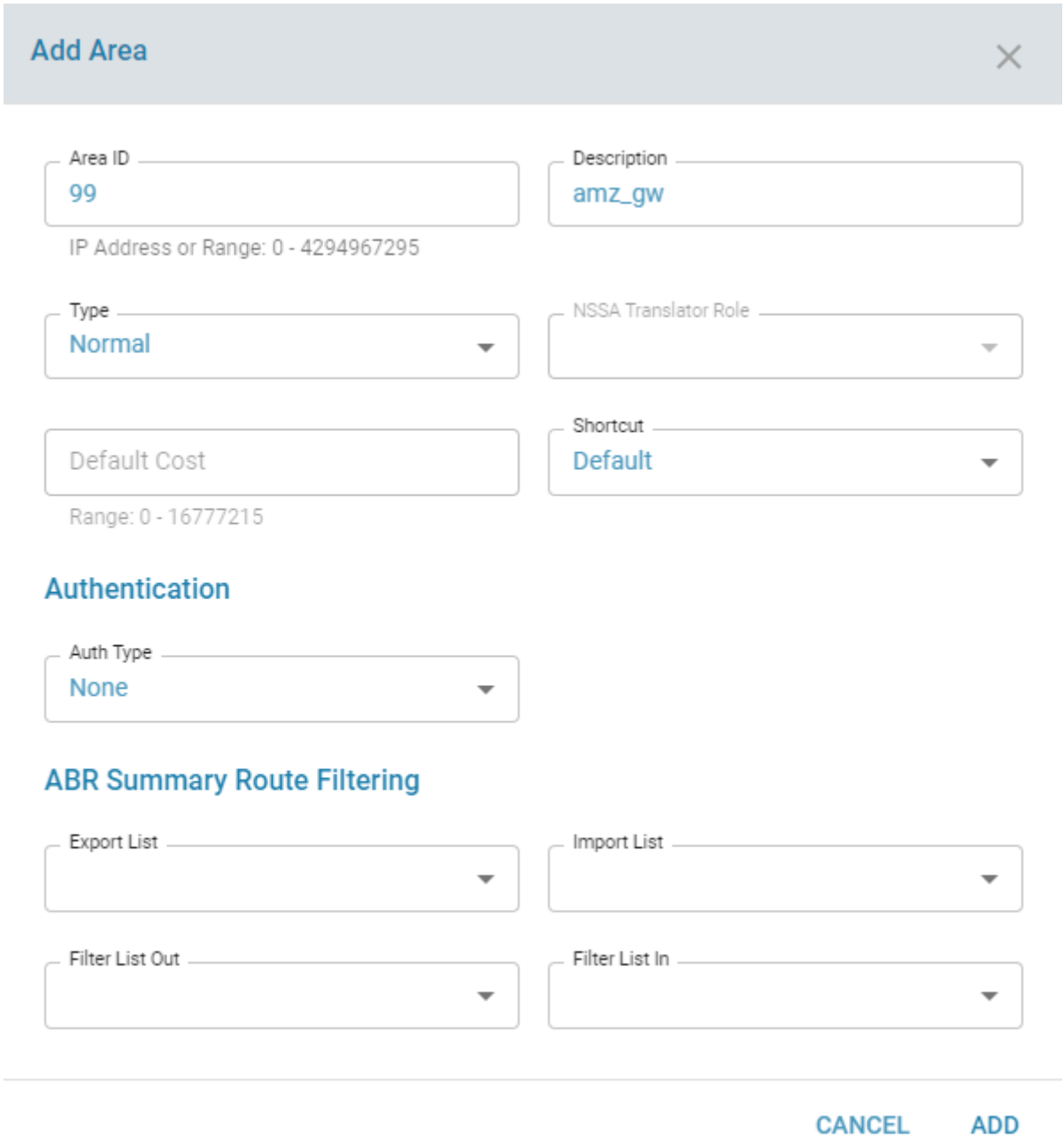
2. Add the RouterID



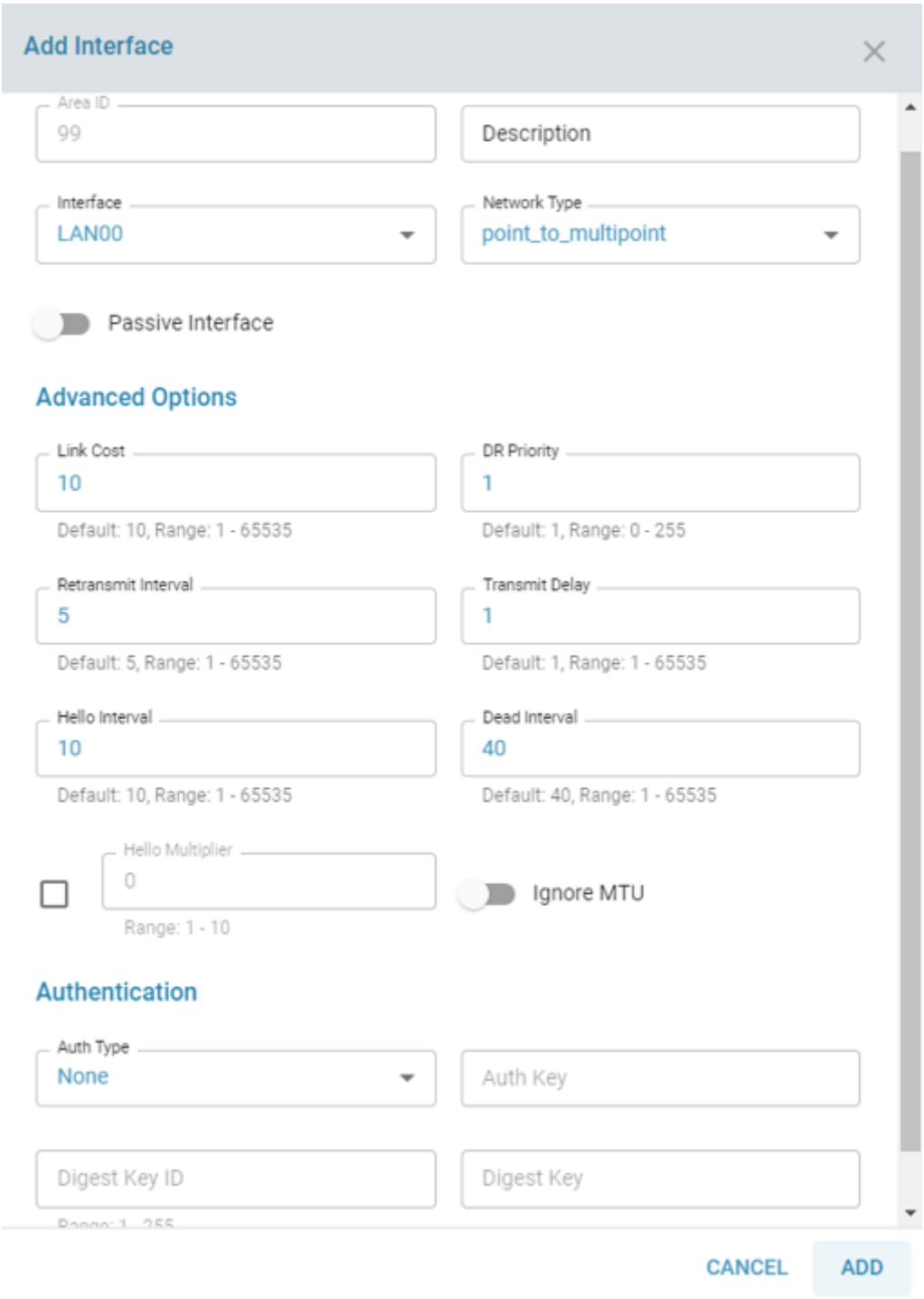
3. Add the redistribution for all kernel routes



4. Add new area



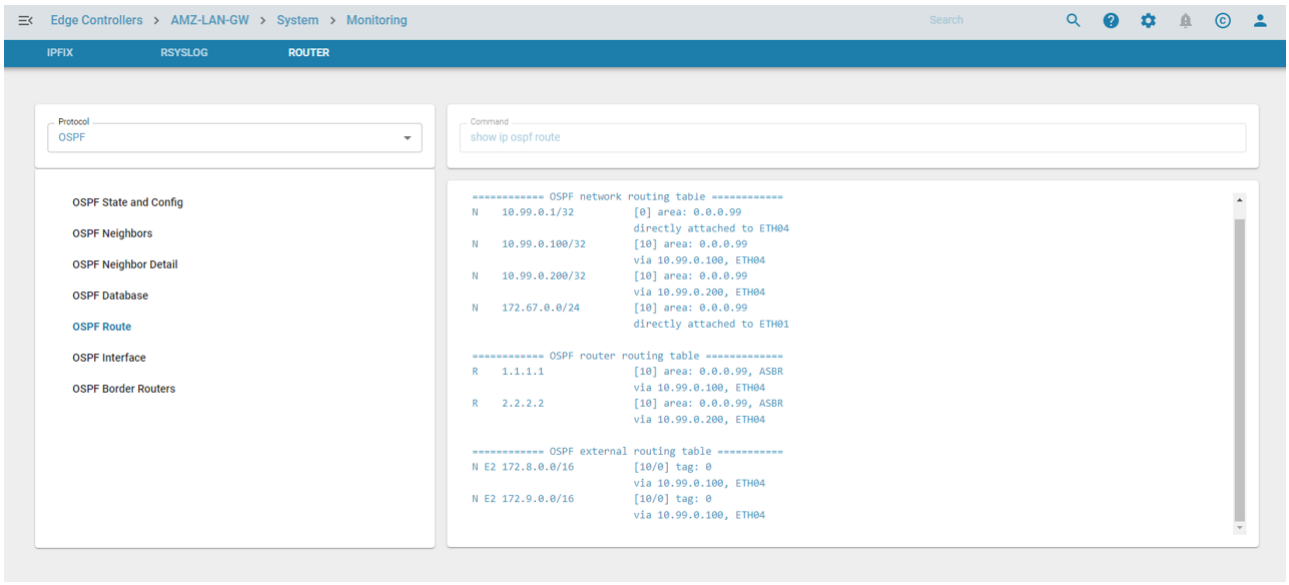
5. Add active interface through which the routes are learned and select the network type as ***point_to_multicast***



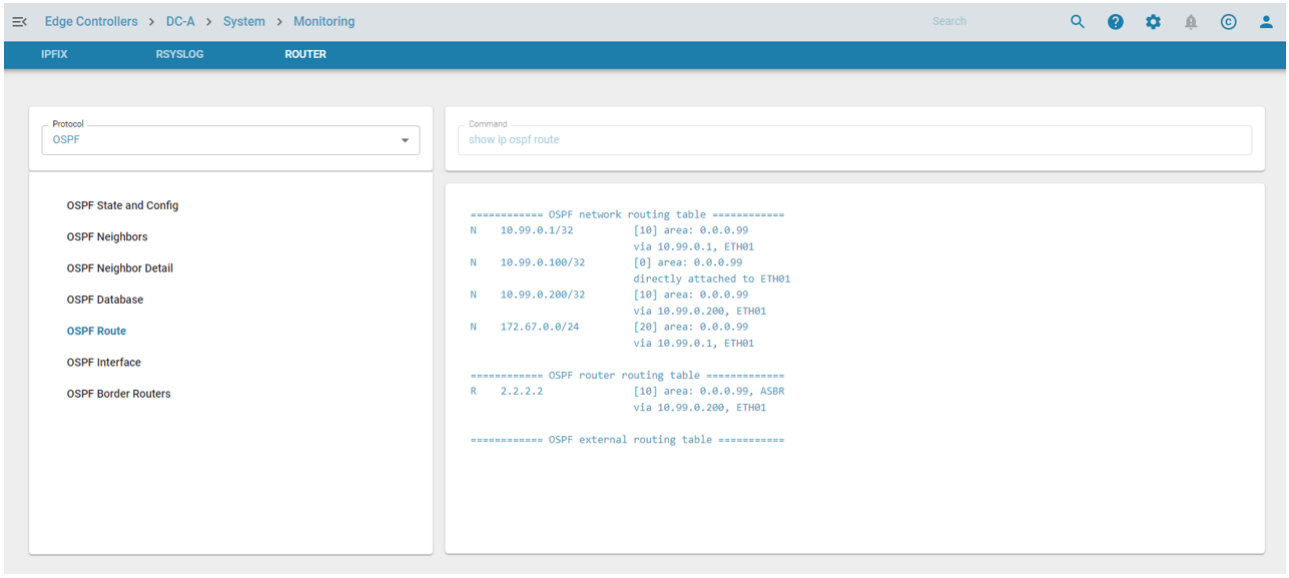
Verify the OSPF config

As shown in the overview, both the branches connects to DC-A, so all the DC Servers should go through DC-A to reach any branch or branch lans.

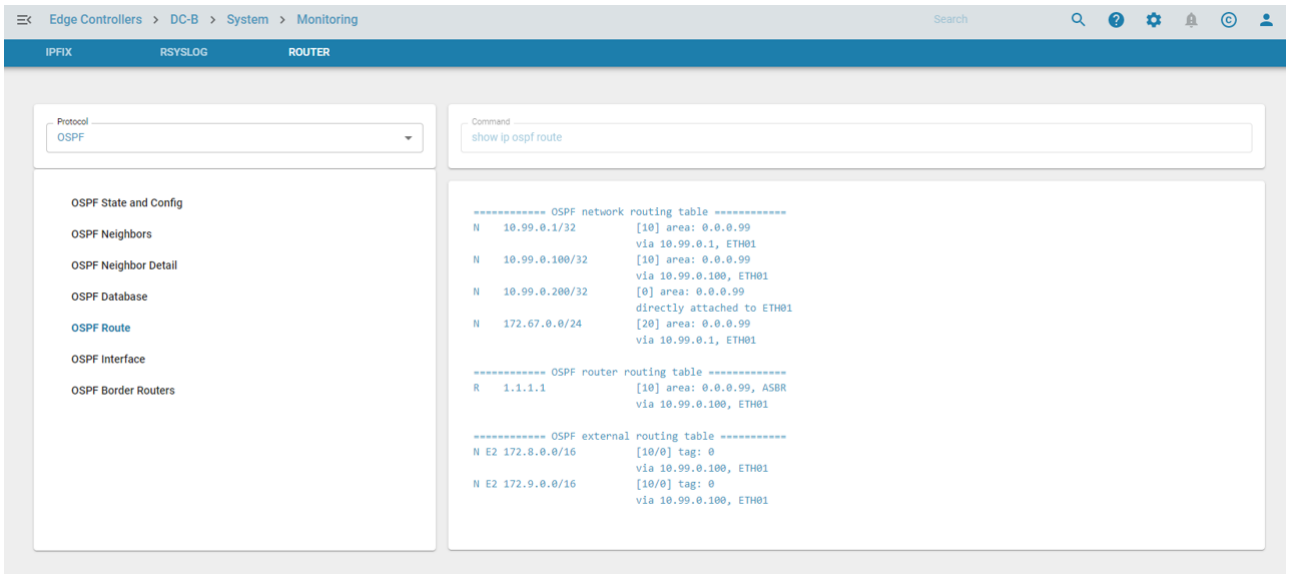
- OSPF route in AMZ_GATEWAY [External routes are learned from DC-A]



- OSPF route in DC-A [External routes will be empty, as DC-A is active right now]

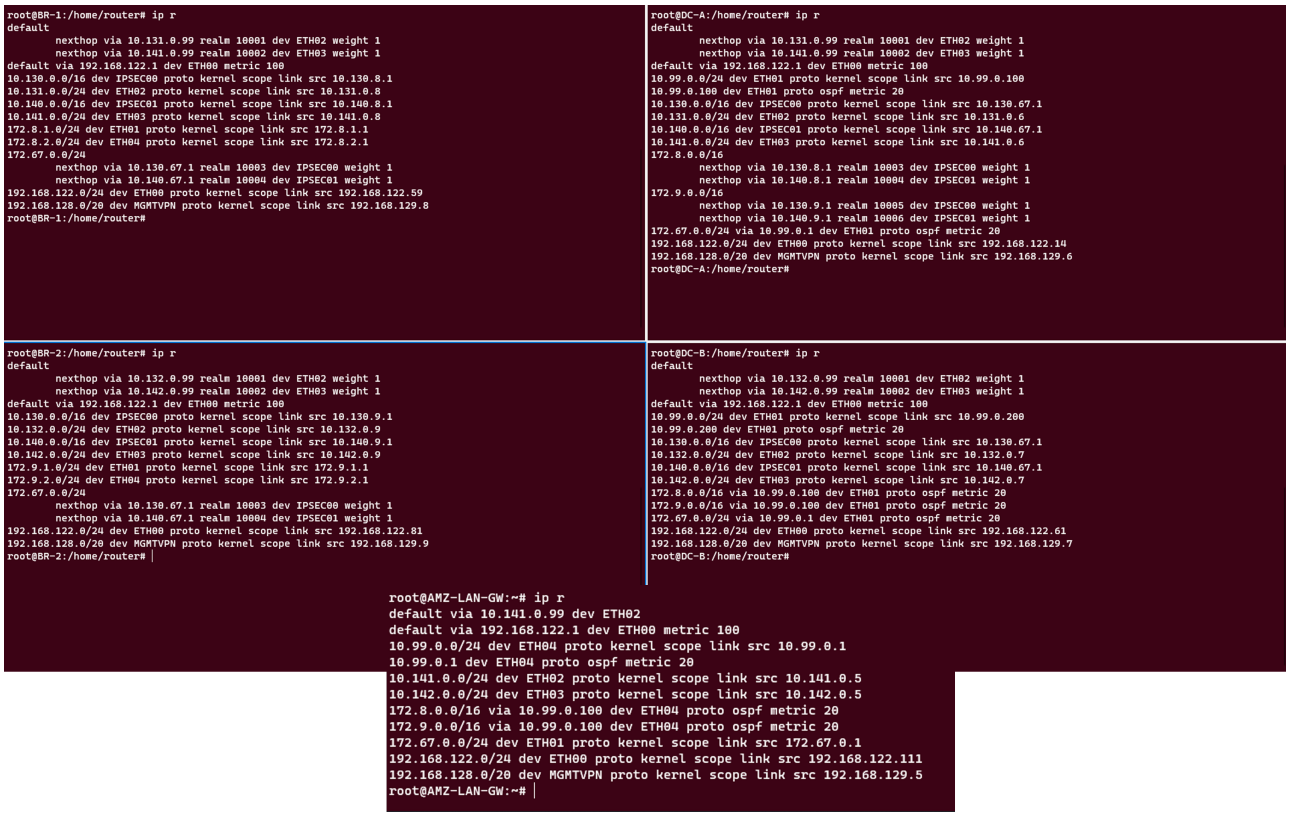


- OSPF route in DC-B [External routes are learned from DC-A]



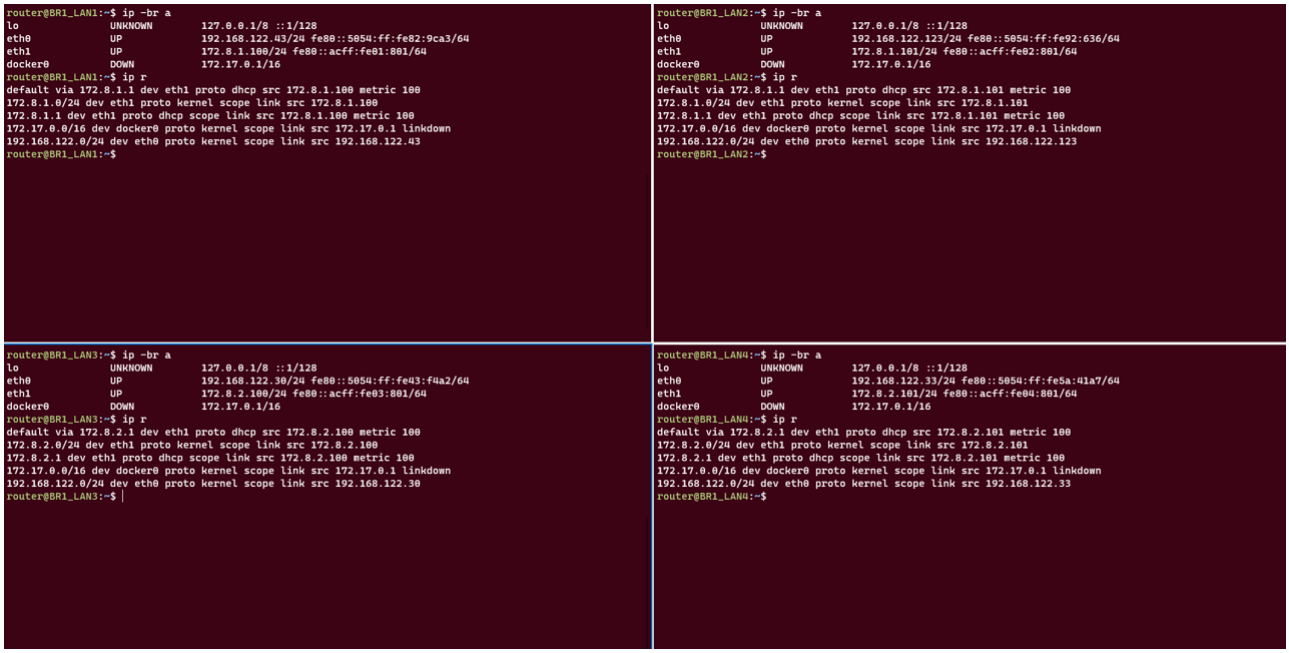
Make sure the routes are fine by running IOs

- Check the routes in all the CPEs [AMZ_GATEWAY, DC-A, DC-B, BR-1 & BR-2]

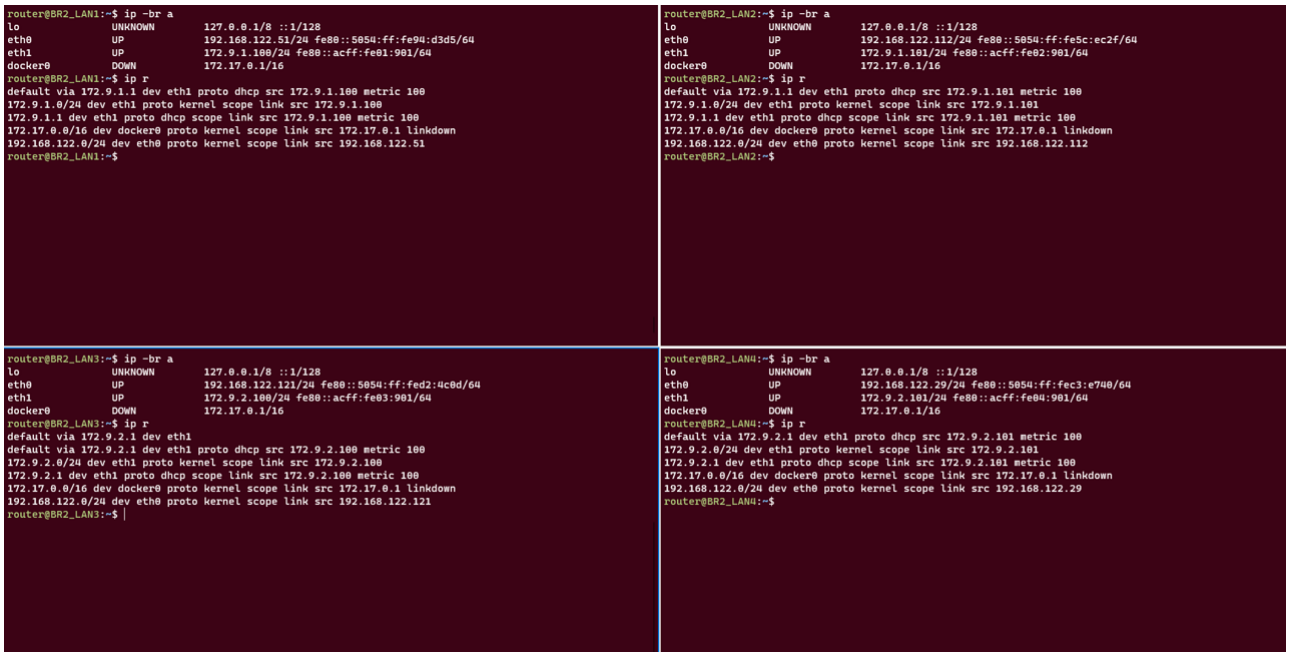


- Login to each LAN-PCs and make sure that all got the IPS leased from CPEs

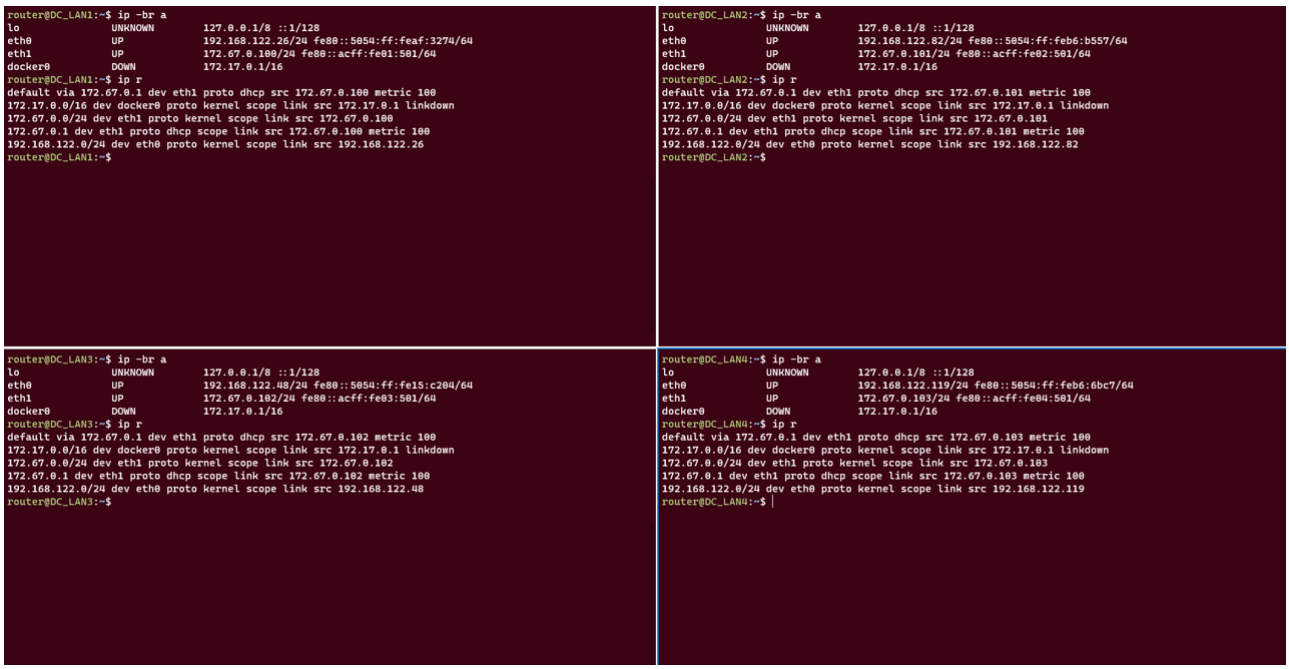
BR-1 LAN PCs:



BR-2 LAN PCs:

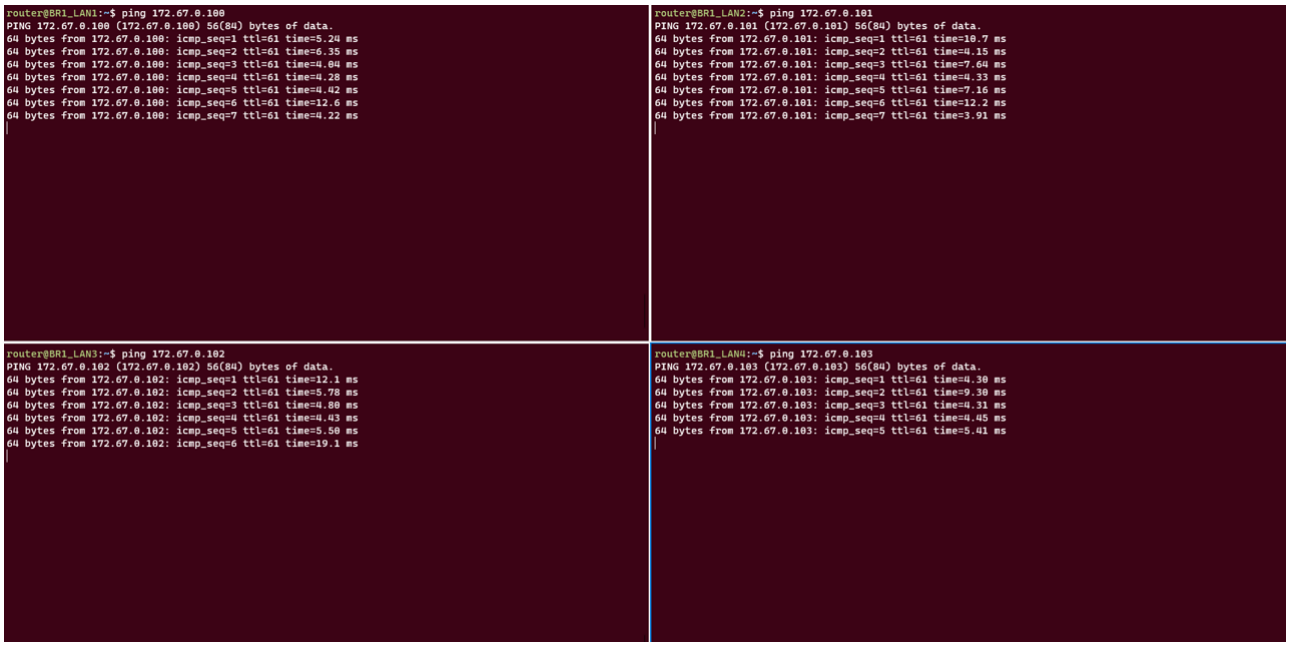


DC LAN Servers:

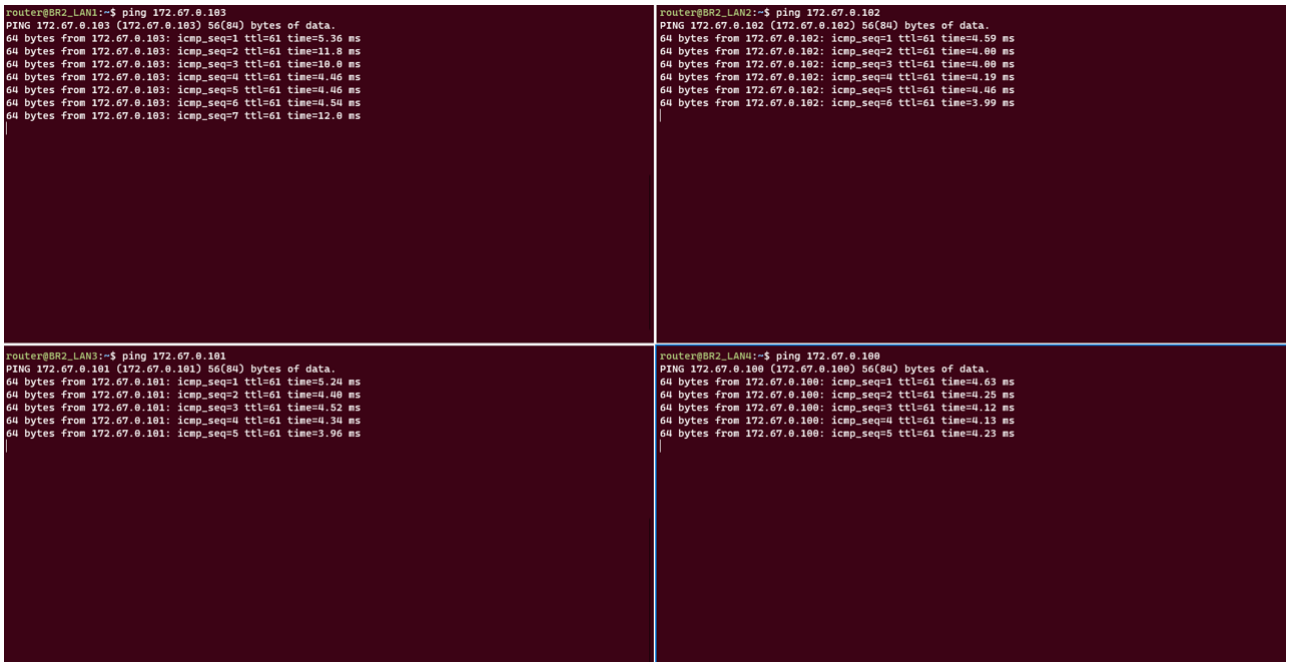


- Start the IOs from all the branch lan pcs and monitor the same using tcpdump in dc lan servers

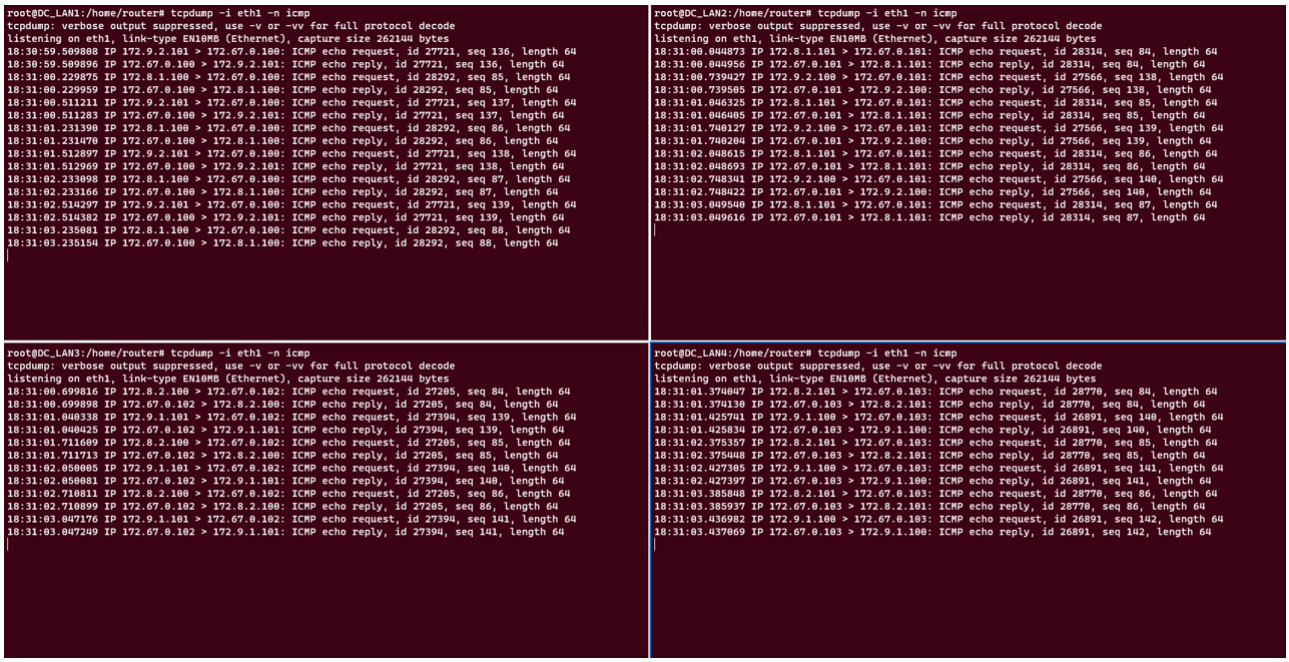
BR-1 LAN running IOs:



BR-2 LAN running IOs:



DC LAN running tcpdumps:


root@DC_LAN1:/home/router# tcpdump -i eth1 -n icmp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1, link-type EN10MB (Ethernet), capture size 262144 bytes
18:30:59.509808 IP 172.9.2.101 > 172.67.0.100: ICMP echo request, id 27721, seq 136, length 64
18:30:59.509896 IP 172.67.0.100 > 172.9.2.101: ICMP echo reply, id 27721, seq 136, length 64
18:31:00.229875 IP 172.8.1.100 > 172.67.0.100: ICMP echo request, id 28292, seq 85, length 64
18:31:00.229959 IP 172.67.0.100 > 172.8.1.100: ICMP echo reply, id 28292, seq 85, length 64
18:31:00.511211 IP 172.9.2.101 > 172.67.0.100: ICMP echo request, id 27721, seq 137, length 64
18:31:00.511283 IP 172.67.0.100 > 172.9.2.101: ICMP echo reply, id 27721, seq 137, length 64
18:31:01.231390 IP 172.8.1.100 > 172.67.0.100: ICMP echo request, id 28292, seq 86, length 64
18:31:01.231470 IP 172.67.0.100 > 172.8.1.100: ICMP echo reply, id 28292, seq 86, length 64
18:31:01.512897 IP 172.9.2.101 > 172.67.0.100: ICMP echo request, id 27721, seq 138, length 64
18:31:01.512969 IP 172.67.0.100 > 172.9.2.101: ICMP echo reply, id 27721, seq 138, length 64
18:31:02.233098 IP 172.8.1.100 > 172.67.0.100: ICMP echo request, id 28292, seq 87, length 64
18:31:02.233166 IP 172.67.0.100 > 172.8.1.100: ICMP echo reply, id 28292, seq 87, length 64
18:31:02.514297 IP 172.9.2.101 > 172.67.0.100: ICMP echo request, id 27721, seq 139, length 64
18:31:02.514382 IP 172.67.0.100 > 172.9.2.101: ICMP echo reply, id 27721, seq 139, length 64
18:31:03.235081 IP 172.8.1.100 > 172.67.0.100: ICMP echo request, id 28292, seq 88, length 64
18:31:03.235154 IP 172.67.0.100 > 172.8.1.100: ICMP echo reply, id 28292, seq 88, length 64
root@DC_LAN2:/home/router# tcpdump -i eth1 -n icmp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1, link-type EN10MB (Ethernet), capture size 262144 bytes
18:31:00.044873 IP 172.8.1.101 > 172.67.0.101: ICMP echo request, id 28314, seq 84, length 64
18:31:00.044956 IP 172.67.0.101 > 172.8.1.101: ICMP echo reply, id 28314, seq 84, length 64
18:31:00.739427 IP 172.9.2.100 > 172.67.0.101: ICMP echo request, id 27566, seq 138, length 64
18:31:00.739505 IP 172.67.0.101 > 172.9.2.100: ICMP echo reply, id 27566, seq 138, length 64
18:31:01.046325 IP 172.8.1.101 > 172.67.0.101: ICMP echo request, id 28314, seq 85, length 64
18:31:01.046405 IP 172.67.0.101 > 172.8.1.101: ICMP echo reply, id 28314, seq 85, length 64
18:31:01.740127 IP 172.9.2.100 > 172.67.0.101: ICMP echo request, id 27566, seq 139, length 64
18:31:01.740204 IP 172.67.0.101 > 172.9.2.100: ICMP echo reply, id 27566, seq 139, length 64
18:31:02.048615 IP 172.8.1.101 > 172.67.0.101: ICMP echo request, id 28314, seq 86, length 64
18:31:02.048693 IP 172.67.0.101 > 172.8.1.101: ICMP echo reply, id 28314, seq 86, length 64
18:31:02.748341 IP 172.9.2.100 > 172.67.0.101: ICMP echo request, id 27566, seq 140, length 64
18:31:02.748422 IP 172.67.0.101 > 172.9.2.100: ICMP echo reply, id 27566, seq 140, length 64
18:31:03.049540 IP 172.8.1.101 > 172.67.0.101: ICMP echo request, id 28314, seq 87, length 64
18:31:03.049616 IP 172.67.0.101 > 172.8.1.101: ICMP echo reply, id 28314, seq 87, length 64
root@DC_LAN3:/home/router# tcpdump -i eth1 -n icmp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1, link-type EN10MB (Ethernet), capture size 262144 bytes
18:31:00.699816 IP 172.8.2.100 > 172.67.0.102: ICMP echo request, id 27205, seq 84, length 64
18:31:00.699898 IP 172.67.0.102 > 172.8.2.100: ICMP echo reply, id 27205, seq 84, length 64
18:31:01.040338 IP 172.9.1.101 > 172.67.0.102: ICMP echo request, id 27394, seq 139, length 64
18:31:01.040425 IP 172.67.0.102 > 172.9.1.101: ICMP echo reply, id 27394, seq 139, length 64
18:31:01.711689 IP 172.8.2.100 > 172.67.0.102: ICMP echo request, id 27205, seq 85, length 64
18:31:01.711713 IP 172.67.0.102 > 172.8.2.100: ICMP echo reply, id 27205, seq 85, length 64
18:31:02.050005 IP 172.9.1.101 > 172.67.0.102: ICMP echo request, id 27394, seq 140, length 64
18:31:02.050081 IP 172.67.0.102 > 172.9.1.101: ICMP echo reply, id 27394, seq 140, length 64
18:31:02.710811 IP 172.8.2.100 > 172.67.0.102: ICMP echo request, id 27205, seq 86, length 64
18:31:02.710899 IP 172.67.0.102 > 172.8.2.100: ICMP echo reply, id 27205, seq 86, length 64
18:31:03.047176 IP 172.9.1.101 > 172.67.0.102: ICMP echo request, id 27394, seq 141, length 64
18:31:03.047249 IP 172.67.0.102 > 172.9.1.101: ICMP echo reply, id 27394, seq 141, length 64
root@DC_LAN4:/home/router# tcpdump -i eth1 -n icmp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1, link-type EN10MB (Ethernet), capture size 262144 bytes
18:31:01.374047 IP 172.8.2.101 > 172.67.0.103: ICMP echo request, id 28770, seq 84, length 64
18:31:01.374130 IP 172.67.0.103 > 172.8.2.101: ICMP echo reply, id 28770, seq 84, length 64
18:31:01.425741 IP 172.9.1.100 > 172.67.0.103: ICMP echo request, id 26891, seq 140, length 64
18:31:01.425834 IP 172.67.0.103 > 172.9.1.100: ICMP echo reply, id 26891, seq 140, length 64
18:31:02.375357 IP 172.8.2.101 > 172.67.0.103: ICMP echo request, id 28770, seq 85, length 64
18:31:02.375448 IP 172.67.0.103 > 172.8.2.101: ICMP echo reply, id 28770, seq 85, length 64
18:31:02.427305 IP 172.9.1.100 > 172.67.0.103: ICMP echo request, id 26891, seq 141, length 64
18:31:02.427397 IP 172.67.0.103 > 172.9.1.100: ICMP echo reply, id 26891, seq 141, length 64
18:31:03.385848 IP 172.8.2.101 > 172.67.0.103: ICMP echo request, id 28770, seq 86, length 64
18:31:03.385937 IP 172.67.0.103 > 172.8.2.101: ICMP echo reply, id 28770, seq 86, length 64
18:31:03.436982 IP 172.9.1.100 > 172.67.0.103: ICMP echo request, id 26891, seq 142, length 64
18:31:03.437069 IP 172.67.0.103 > 172.9.1.100: ICMP echo reply, id 26891, seq 142, length 64