# OSPF Flooding Reduction in MSDC

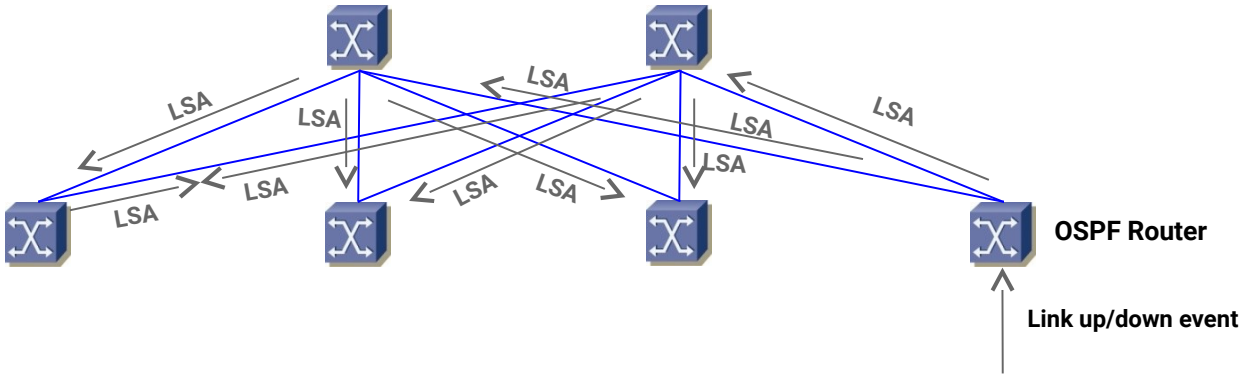
## draft-xu-ospf-flooding-reduction-in-msdc-00

**Xiaohu Xu (Huawei)**

**IETF98, Chicago**

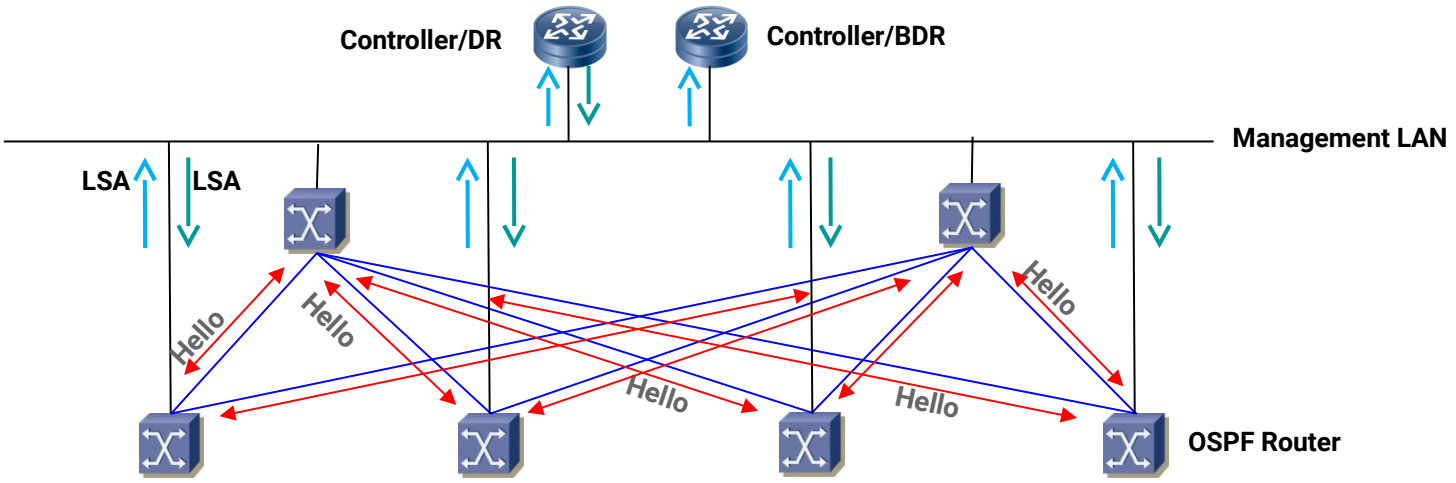
# Problem Statement

- **OSPF is commonly used as a underlay routing protocol for Massively Scalable Data Center (MSDC) networks where CLOS is the most popular topology.**
- **Within the CLOS topology, a given OSPF router would receive multiple copies of exactly the same LSA from multiple OSPF neighbors. The unnecessary link-state information flooding wastes the precious process resource of OSPF routers greatly and therefore OSPF could not scale very well in MSDC networks.**



# Solution Overview

- **Mix of centralized link-state information distribution and distributed SPF calculation.**
  - **All OSPF routers within the CLOS network are connected with controllers via a management LAN.**
  - **OSPF routers within the CLOS network just need to exchange OSPF Hello packets among them so as to discover OSPF neighbors.**
  - **The link-state information is only required to be exchanged between OSPF routers and controllers which are elected as OSPF DR/BDR for the management LAN.**



# Solution Details

- **After the exchange of OSPF Hello packets among OSPF routers, the OSPF neighbor relationship among them would transitions to and remains in the TWO-WAY state.**
- **OSPF routers would originate LSAs accordingly. However, these LSAs just need to be advertised to the controllers being elected as OSPF DR/BDR for the management LAN.**
  - **OSPF routers SHOULD NOT be elected as OSPF DR/BDR for the management LAN (This is done by setting the DR Priority of those OSPF routers to zero).**

# Solution Details (con't)

- **To further reduce the flood of multicast OSPF packets over the management LAN, OSPF routers SHOULD send OSPF packets as unicasts.**
  - **OSPF routers SHOULD send unicast OSPF Hello packets periodically to the controllers being elected as OSPF DR/BDR. In other words, OSPF routers would not send any OSPF Hello packet over the management LAN until they have found OSPF DR/BDR for the management LAN.**
  - **OSPF routers SHOULD send other types of OSPF packets to the controllers being elected as OSPF DR/BDR as unicasts as well.**

# Solution Details (con't)

- To avoid the data traffic from being forwarded across the management LAN, the cost of all OSPF routers' interfaces to the management LAN SHOULD be set to the maximum value.
- When a given OSPF router lost its connection to the management LAN, it SHOULD actively establish FULL adjacency with at least one of its OSPF neighbors within the CLOS network which has full adjacency with the controllers.
- The controllers being elected as OSPF DR would send OSPF packets as multicasts or unicasts as normal.
- Link State Acknowledgment packets are RECOMMENDED to be sent as unicasts rather than multicasts.

# Next Steps

- **Comments?**