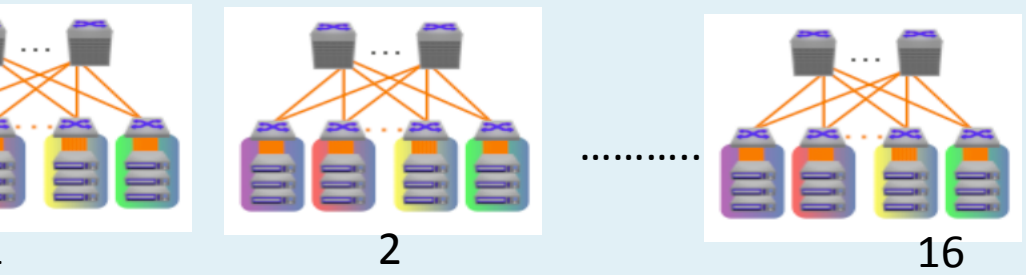
draft-zhang-6man-scale-large-datacenter-00.txt
Improving Scalability of Switching System in Large Data Center

Current Implementations for 2 Million IPv6 End-hosts

Option 1 – Multiple Clusters

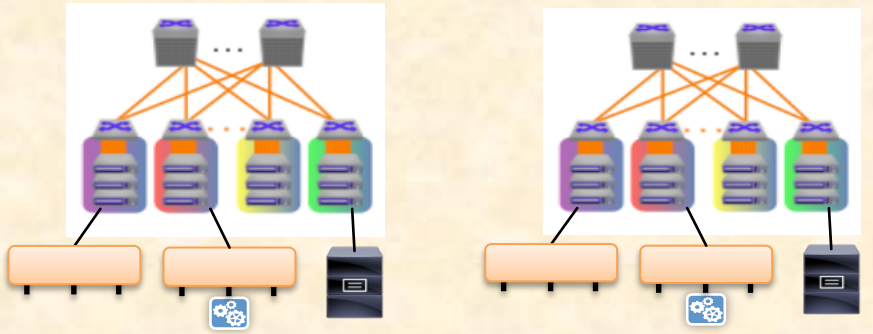
oadCom based (T2) implementation
8K IPv6 host routes on leaf
clusters
-64 spines



rge number of spine switches
LAN/subnets cannot span across clusters

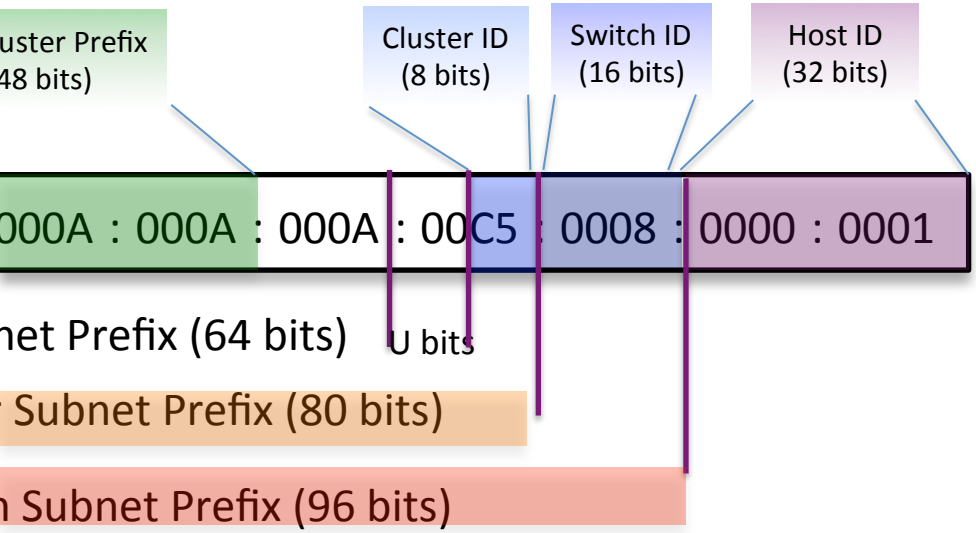
Option 2 – Large FIB Table on Sp

1 million IPv6 host routes on spine
128K IPv6 host routes on leaf
2 cluster
4-8 spines



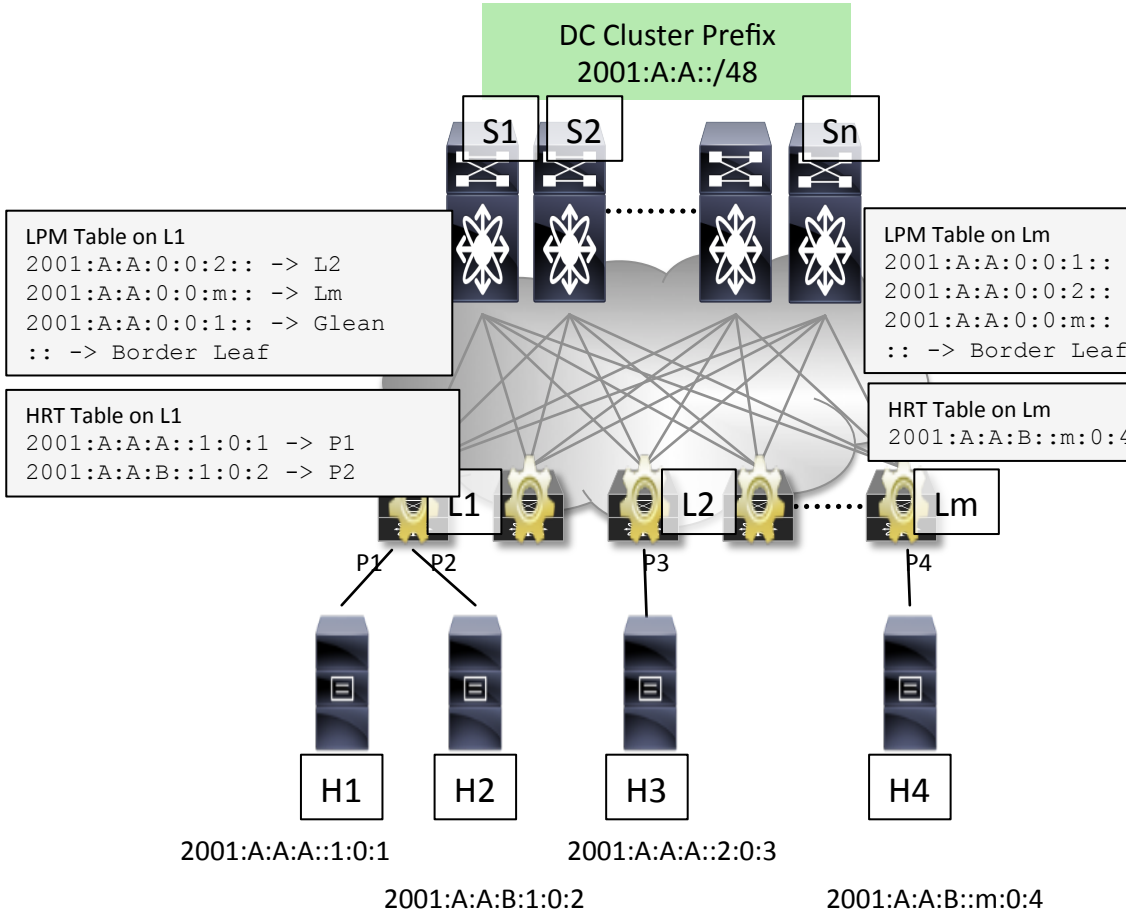
- Cost more on spine switches
- Hard to troubleshoot
- VLAN/subnets cannot span across cl

Scale Through Aggregation



ster Prefix	2001:A:A::/48
t Prefix	2001:A:A:A::/64
Subnet Prefix (SSP)	2001:A:A:**A**:C5:8::/96
gated SSP (ASSP)	2001:A:A:**0**:C5:8::/96
r Subnet Prefix (CSP)	2001:A:A:A:C5::/89
Address	2001:A:A:A:C5:8:0:1/128

Optimizing FIB table in hardware
Distributed Neighbor Discovery



Advantages – Simple, Fast & Scalable

1. Scale to multi-millions hosts/VMs in single DC cluster with low cost ASICs
2. Better scalability in hardware
   - Small FIB table size on all switches
3. Better scalability in software
   - Distributed Neighbor Discovery
   - No multicast ND messages between access switches
   - No host routes advertisement
5. Easy to manage and troubleshoot