IS-IS Spine-Leaf

RTGWG IETF 98

draft-shen-isis-spine-leaf-ext-03

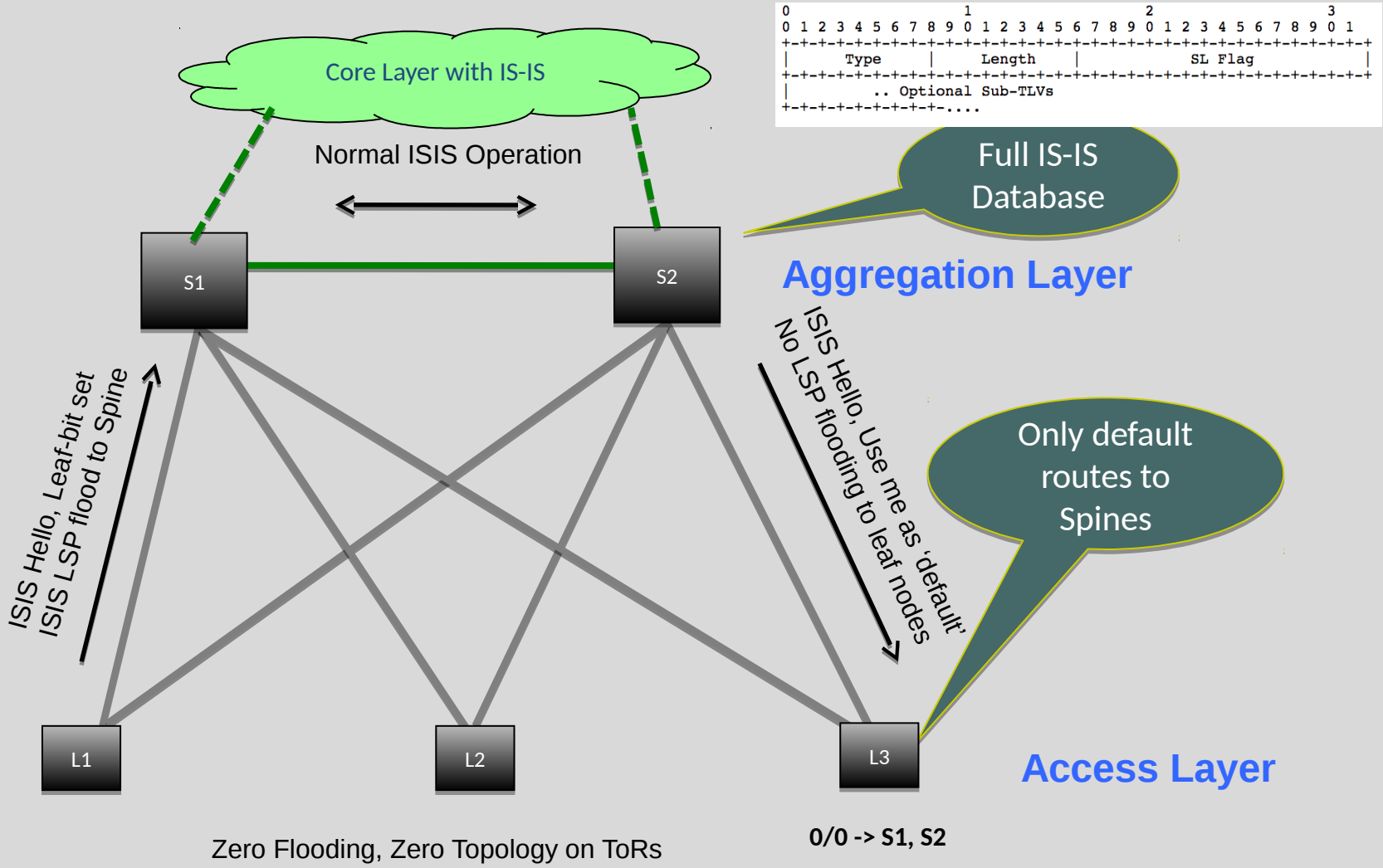
March 30, 2017

Naiming S., Les G., Sanjay T.

Motivations

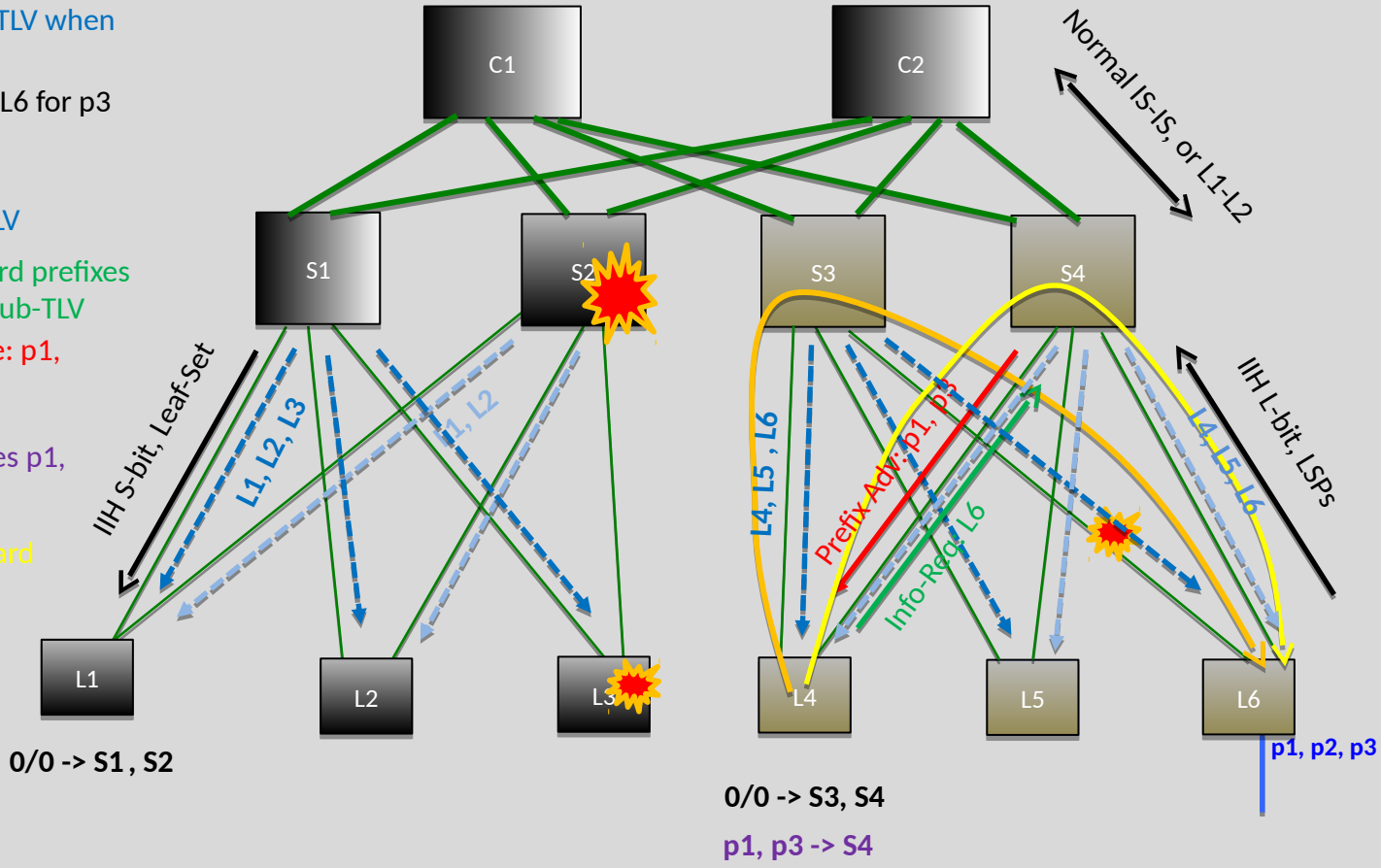
- Spine-Leaf popular in Data Centers
- Normally leaf-to-leaf traffic goes through one of the spine nodes
- Basically ECMP load-sharing from leaf to spine nodes
- Rich mesh of spine-leaf IGP topology generates LSP flooding issues, in particular in the events of link/node down
- This draft enables zero-flooding and zero topology for ToR's, allows IS-IS fitting better in DC environment

Extension Basics



Link/Node Down (CLOS)

- S1-S4 include Leaf-Set sub-TLV when sending Spine-Leaf TLV
- L4 picks S3 0/0, forward to L6 for p3
- **S3-L6 link down**
- S3 Leaf-Set lost L6 in sub-TLV
- L4 picks S4, sending "forward prefixes behind node L6" Info-Req sub-TLV
- S4 replies with "Prefixes are: p1, and p3 for L6" with IP/IPv6 Reachability
- L4 adds more specific entries p1, p3 with nexthop to S4
- L4 picks S4 lookup p3, forward to L6 for p3
- **Leaf L3 Node down. Nothing special to do**
- **Spine S2 Node down. Nothing special to do**



Node Down

Link Down

Spine-Leaf Discussion

- IS-IS with Spine-Leaf extension can be the only protocol or can work together with BGP in DC
- Other networks vs DC networks (this draft helps to meet the DC special requirements)
- Why one protocol is better than two working together, in both software and management? (we also have ARP, ND, BFD, LLDP, etc.)
- Other rich features (past 20 years) using e.g. BGP-EVPN or other overlay protocols, multicast, TE, SR, etc.
- Topology-less on leaf nodes can also do TE in DC. Among spine nodes of PODs or DCIs; PCE between IS-IS levels; Controller injects explicit paths or SIDs using topology from spine nodes