Avoiding Traffic Black-Holes for Route Aggregation in IS-IS

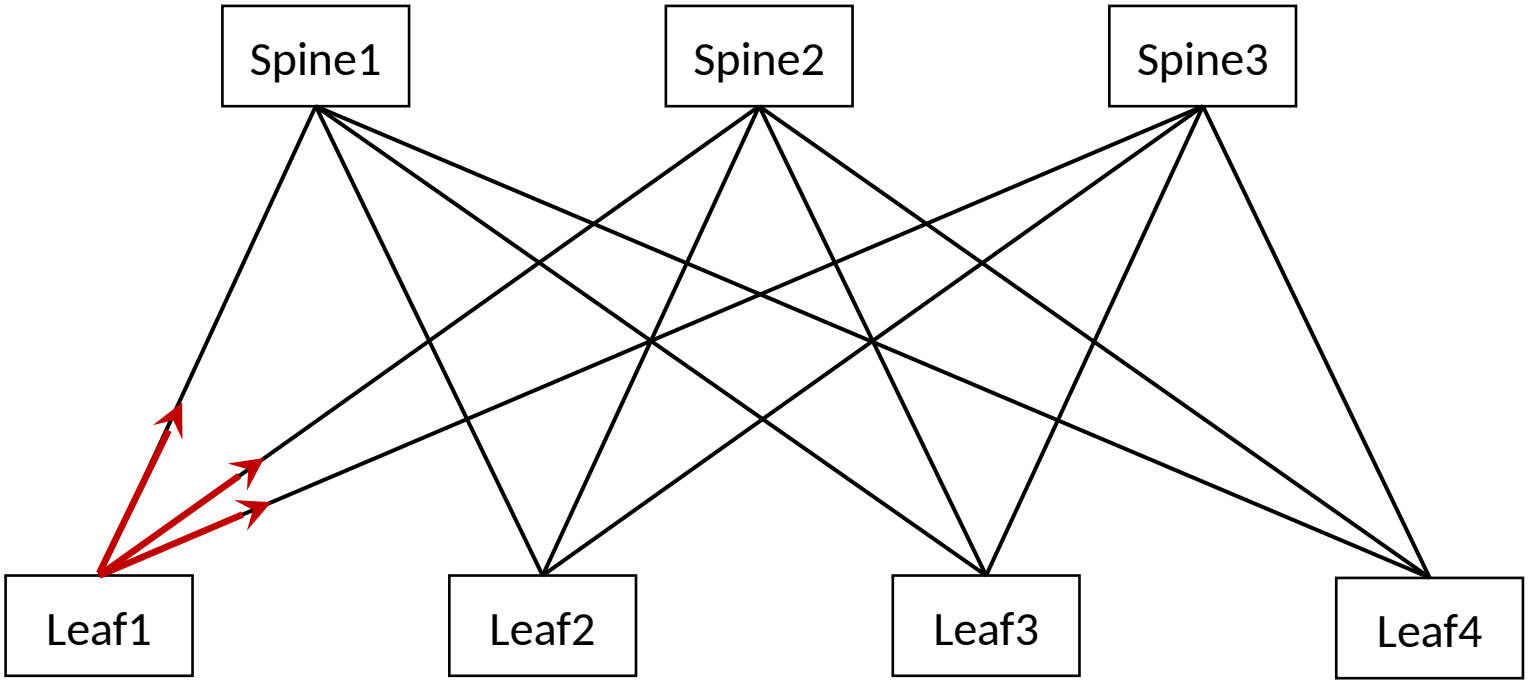
draft-chen-isis-black-hole-avoid-00

Zhe Chen, Xiaohu Xu

Huawei

2017.3.31

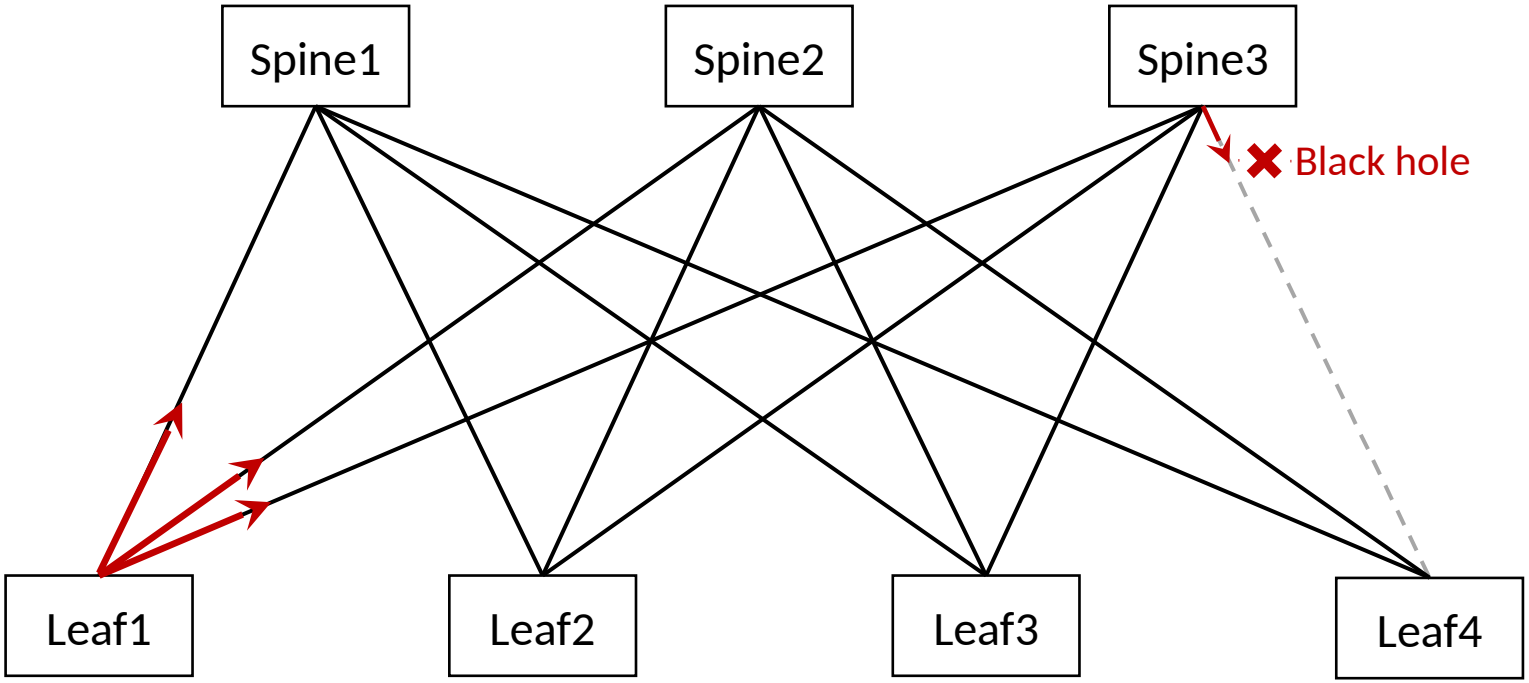
Problem Description



Dst	Nexthop
0.0.0.0/0	Spine1 Spine2 Spine3

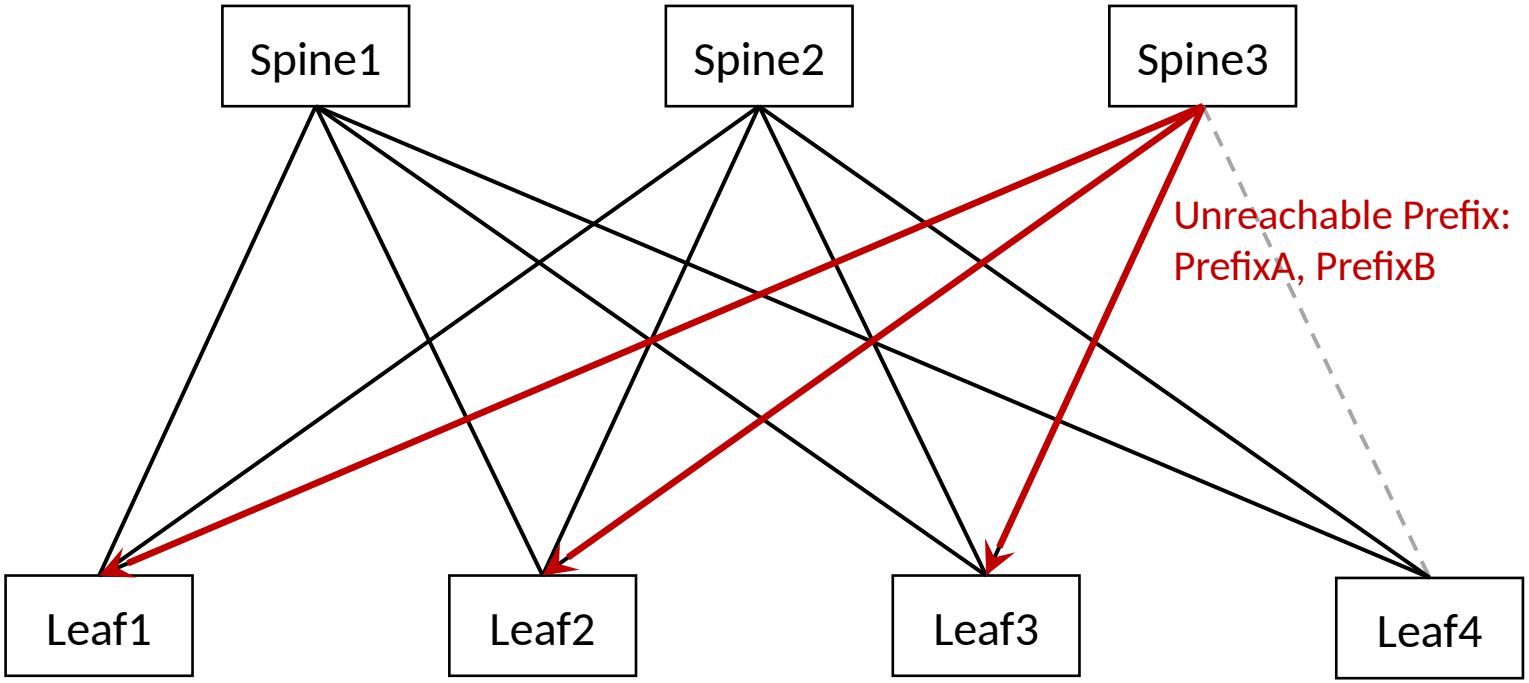
PrefixA
PrefixB

Problem Description

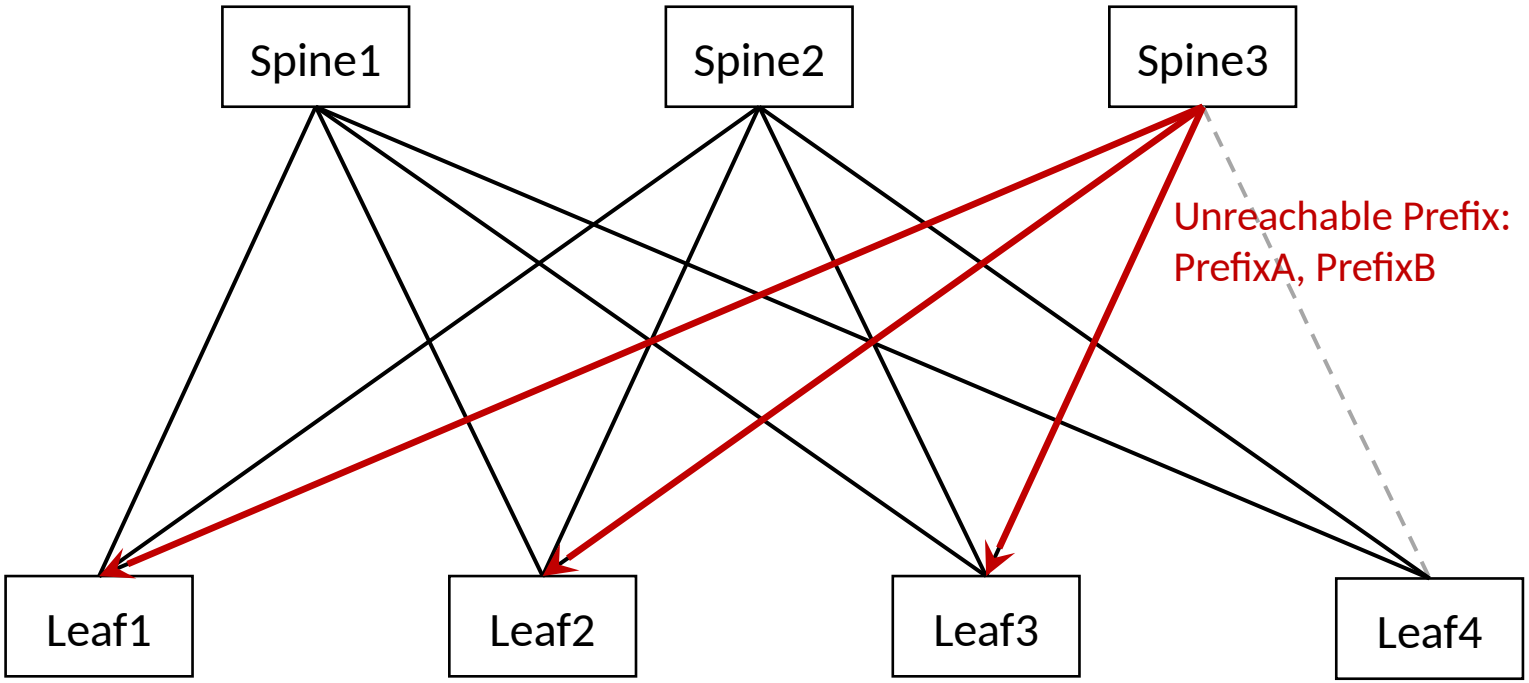


Dst	Nexthop
0.0.0.0/0	Spine1 Spine2 Spine3

Solution Overview



Dst	Nexthop
0.0.0.0/0	Spine1 Spine2 Spine3

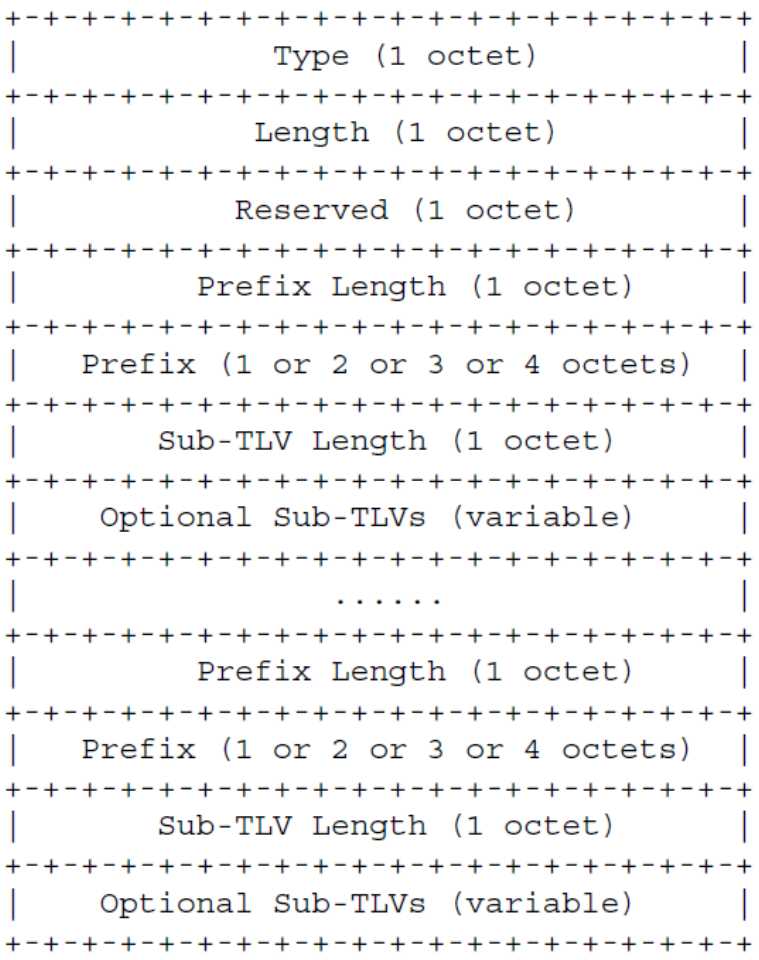


Dst	Nexthop
0.0.0.0/0	Spine1 Spine2 Spine3

Dst	Nexthop
0.0.0.0/0	Spine1 Spine2 Spine3
PrefixA	Spine1 Spine2
PrefixB	Spine1 Spine2

TLV Encoding

- IP Unreachability TLV:



Type: TBD.

Length: Length of the Value field of the TLV.

Reserved: Bits reserved for future usage.

Prefix Length: The value can be 0 to 32, indicating the number of effective bits in the Prefix field.

Prefix: Encoding the unreachable prefix in the minimal number of octets for the given number of effective bits (i.e., the Prefix Length field). The remaining bits of prefix SHOULD be set zero and ignored upon receipt.

Sub-TLV Length: Length of Sub-TLVs.

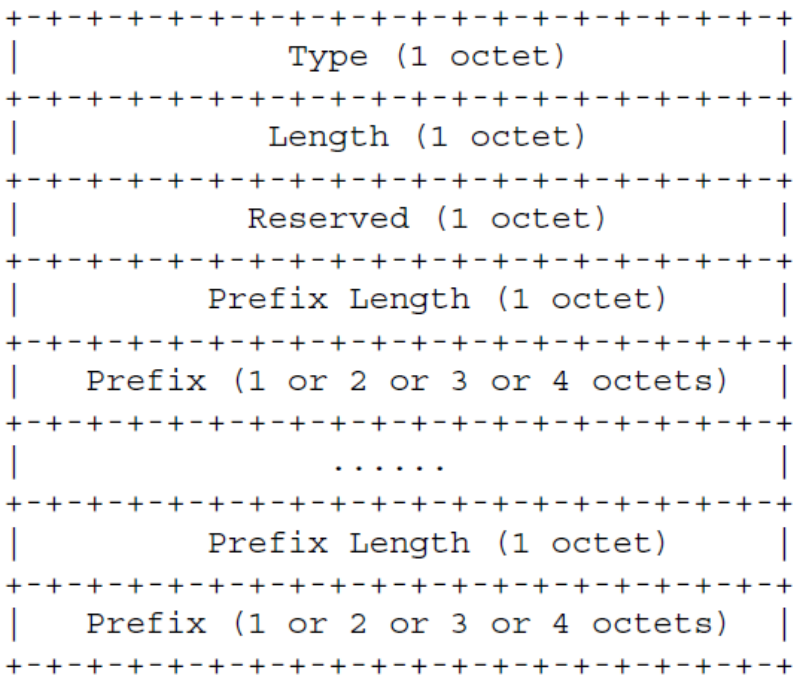
Sub-TLVs: Optional Sub-TLVs for future extension.

Details

- When link failure happens between a Spine node and a Le af node, the Spine node SHOULD:
  - 1) encode all prefixes attached to the Leaf node (i.e., the unreac hable prefixes) into the IP Unreachability TLV,
  - 2) append the IP Unreachability TLV to the IS-IS LSP, and
  - 3) send the LSP to every other Leaf node it connects to.
- The Leaf node SHOULD:
  - 1) install each of the unreachable prefixes into its routing table,
  - 2) set the next hop to be an ECMP group including all Spine nod es it connects to, except the one who advertises the unreachabl e prefix.

Alternative Solution

- Introduce a new Sub-TLV of the Extended IP Reachabili ty TLV ([RFC 5305]):



Type: TBD.

Length: Length of the Value field of the TLV.

Reserved: Bits reserved for future usage.

Prefix Length: The value can be 0 to 32, indicating the number of effective bits in the Prefix field.

Prefix: Encoding the unreachable prefix in the minimal number of octets for the given number of effective bits (i.e., the Prefix Length field). The remaining bits of prefix SHOULD be set zero and ignored upon receipt.

- When link failure happens, the Spine node SHOULD advertise a LSP to the Leaf node, which contains the Extended IP Reachability TLV (carrying the default route) with the new Sub-TLV (carrying the unreachable prefixes) appended.

Next Steps

- Add IPv6 Support.
- We need more reviews and comments.