TRILL Resilient Distribution Trees

draft-zhang-trill-resilient-trees-04.txt
Mingui, Tissa, Jana, Ayan, Anoop, Donald
When a distribution tree fails

• When a link on distribution tree (DT) fails, it is recovered through campus wide re-convergence.

• It may lead to considerable long disruption to ongoing multicast traffic.

• Protection mechanisms should be designed to mitigate this disruption.
Multicast FRR using backup DT

• TRILL switches can calculate multiple trees.
• Akin to the IGP multicast Fast ReRouting (FRR) mechanisms
  – TRILL can install a backup DT in advance.
  – If a link on the primary DT fails, use the backup DT directly without DT calculation and installation.
Usage of Affinity Sub-TLV

• RFC6326bis defines Affinity Sub-TLV. It explicitly assigns a link on a DT.
  – It is called “Affinity Link” in this doc

• The Affinity Link is not necessarily on the shortest path trees.

• It’s utilized to manipulate the backup DT calculation.
DT calculation with the Affinity Link

- Suppose the Affinity Link is RB4->RB5, tree root is RB1
  - {Nickname=RB5, Num of Trees=1, Tree-num of roots=RB1}
- Delete all incoming links of RB5 except the affinity link RB4->RB5
- Compute the DT according to SPF calculation on the sub topology
- Link RB4->RB5 will surely appear on the DT
Protecting a link on primary DT

• Link RB1-RB2 on the primary tree is protected by the backup tree.
Maximally edge disjoint DTs

- The more Affinity Links are intentionally assigned, the more links of the backup DT can be pinpointed. Maximally disjoint primary & backup DTs can be set up in this way.

*root*

- Full topology
- tree 1
- tree 2

Red and blue edges are disjoint.
Backup DT pruning

1. **Primary Tree (Tree 1)**
   - Root
   - RB1
   - RB3
   - RB5
   - RB6
   - RB7
   - RB9
   - RB10
   - Receivers

2. **Backup Tree (Tree 2)**
   - Root
   - RB1
   - RB2
   - RB3
   - RB5
   - RB6
   - RB7
   - RB9
   - RB10
   - Receivers

3. **Pruning**
   - Primary tree 1 is pruned at RB7, RB9, and RB10.
   - Backup tree 2 is pruned at RB7, RB9, and RB10.
Suppose RB7 is the multicast source while RB9 and RB10 are the receivers. When RB1-RB5 fails, RB7 will switch the multicast traffic from tree1 to tree2.
Global 1+1 protection

- RB7 replicates multicast packets and send them along both trees. Receivers RB9 & RB10 accept only one copy from the primary tree using Reverse Path Forwarding Check (RPFC).
- When RB1-RB5 fails, RB9&RB10 change their RPFC and accept the other copy from the backup tree.
Local protection

- When RB1-RB5 fails, RB1 locally switches to the backup tree.
Thanks!