Refresh Interval Independent facility FRR

draft-chandra-mpls-enhanced-frr-bypass-01

Chandrasekar Ramachandran (csekar@juniper.net)

Markus Jork (mjork@juniper.net)

Vishnu Pavan Beeram (vbeeram@juniper.net)

Ina Minei (inaminei@google.com)

Dante Pacella (dante.j.pacella@verizon.com)

Ebben Aries (exa@fb.com)

Slide #1

Refresh independent state maintenance and stale state cleanup

- Reliance on periodic refreshes and refresh timeouts for RSVP-TE state maintenance and stale state cleanup is problematic from the scalability point of view
  - Due to the need to provide fast state synchronization between routers AND
  - Due to the need to limit the amount of stale state that a router has to maintain AND
  - Due to the need to limit the rate of RSVP-TE control plane traffic that a router has to handle
  - See Section 1 of RFC 2961 for more...
- There are existing mechanisms that allow to eliminate reliance on periodic refreshes and refresh timeouts for RSVP-TE state maintenance and stale state cleanup
  - Reliable exchange of *all* RSVP messages using refresh reduction (rfc2961)
  - Coupling state of individual LSPs with the state of RSVP signaling adjacency
- The existing mechanisms do not cover stale state cleanup during facility-based FRR (RFC4090)
- The presentation covers fixes proposed to RFC 4090 to provide refresh independent stale state cleanup during facility-based FRR
  - And also the updates to those fixes after IETF91 (changes from 00 to 01 of the draft)

Slide #2

Fast stale state cleanup during RSVP-TE Fast Reroute – fixing RFC4090
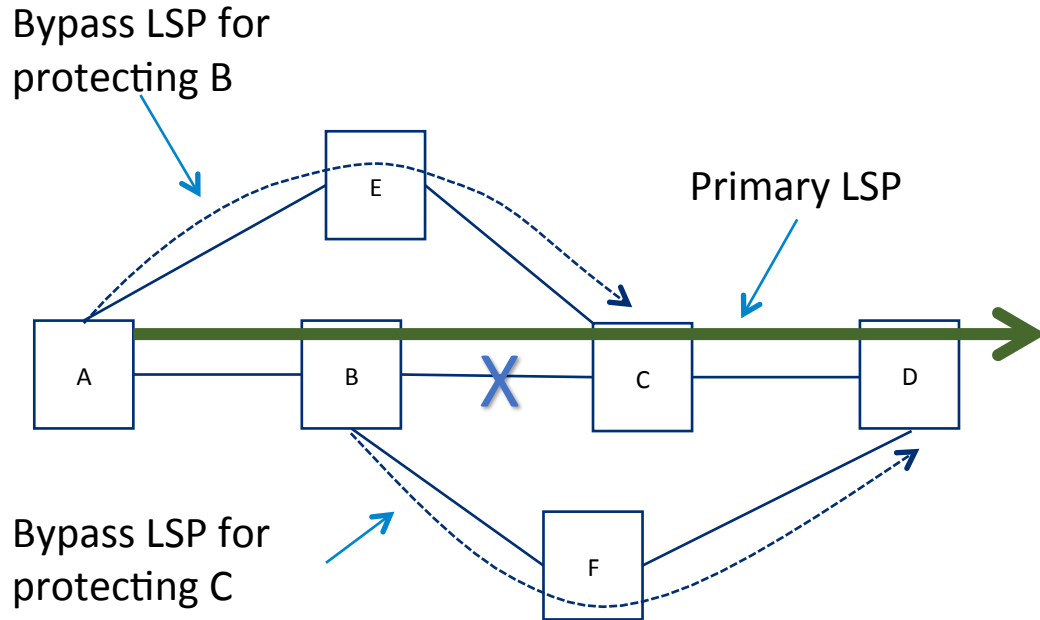
Bypass LSP for protecting B

Primary LSP

Bypass LSP for protecting C

Protecting against B node failure:
   A – Point of Local Repair (PLR)
   C – Merge Point (MP)
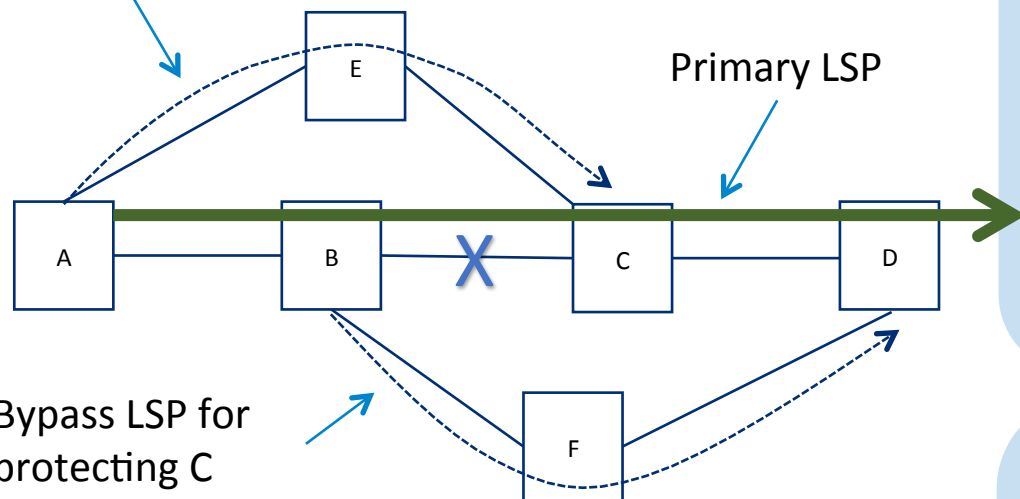
Protecting against C node failure:
   B – Point of Local Repair (PLR)
   D – Merge Point (MP)

Slide #3

- Link (B, C) goes down

- Router A does not initiate node protection FRR (as B is still up)

- When B detects link failure, B initiates node protection FRR with D as Merge Point

  - As part of FRR, B initiates signaling of the backup LSP

- When C detects link failure, "*PathTear and ResvErr messages MUST NOT be sent immediately*" (RFC4090)

- Furthermore, C "*SHOULD reset the refresh timers … as if they had just been <u>refreshed</u>*" (RFC4090)
  - To give B time "to begin refreshing state via the bypass LSP"
    - To give B time to signal the backup LSP

- C removes the state only "*if it has not been refreshed before the <u>refresh timer expires</u>*" (RFC4090)
  - When C removes the state, C sends PathTear to D, but **hopefully** by that time D receives from B signaling for the backup LSP

- <u>Bottom line</u>: RFC4090 relies on refresh timeout for stale state cleanup during RSVP-TE Fast Reroute (FRR), BUT

- **Stale state cleanup should not depend on refresh timeout !**
  - For the reasons explained in the previous slide

Refresh Independent FRR – MP determination & Conditional PathTear

Bypass LSP for protecting B

E

Primary LSP

A   B   ✗   C   D

F

Bypass LSP for protecting C

Protecting against B node failure:
  A – Point of Local Repair (PLR)
  C – Merge Point (MP)

Protecting against C node failure:
  B – Point of Local Repair (PLR)
  D – Merge Point (MP)

Slide #4

- **MP determination:**
  - **Whenever PLR has a backup path available, the PLR sets "Local protection available" flag in RRO carried in PATH**
    - **If PLR has a node protecting backup path, the PLR also sets "node protection" flag**
  - **PLR initiates NodeID Hello session to MP**
  - **A node concludes it is MP if PLR has set protection flags in PATH RRO <u>and</u> NodeID signaling adjacency with PLR is up.**

- **"Conditional" PathTear:**
  - **Originated when a router deletes the RSVP-TE state associated with a particular primary LSP (similar to "vanilla" PathTear)**
  - **Receiver should retain the state for that LSP *on the conditions* that (a) the receiver is a node protection Merge Point, and (b) the LSP is currently being protected by the Point of Local Repair associated with this Merge Point**
  - **Otherwise the receiver deletes the state (just like "vanilla" PathTear)**

Refresh Independent FRR – Remote PathTear
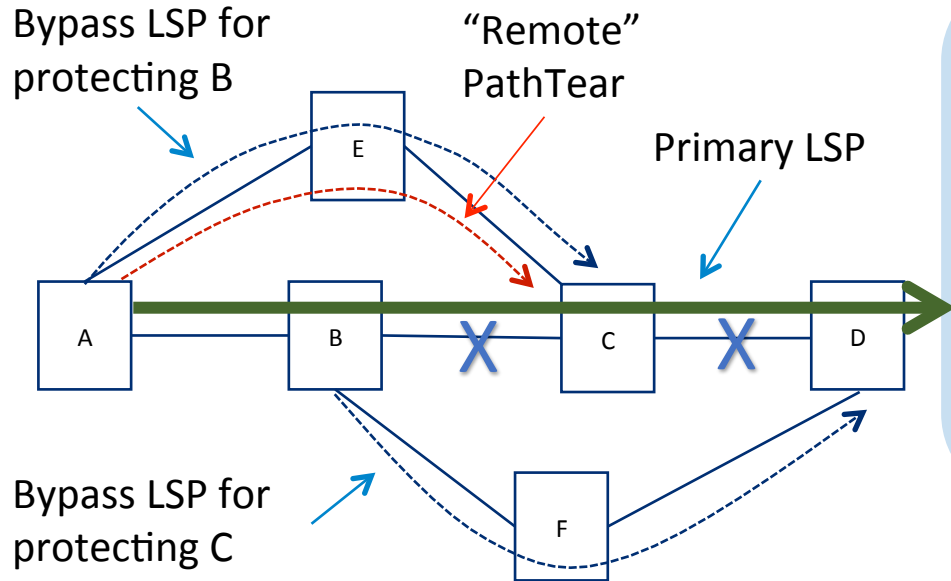
Bypass LSP for protecting B

"Remote" PathTear

Primary LSP

Bypass LSP for protecting C

Protecting against B node failure:
  A – Point of Local Repair (PLR)
  C – Merge Point (MP)

Protecting against C node failure:
  B – Point of Local Repair (PLR)
  D – Merge Point (MP)

Slide #5

- **"Remote" PathTear:**
  - **Originated by a PLR when:**
    - **PLR deletes the RSVP-TE LSP state before the PLR has completed backup LSP signaling**
    - **PLR detects change in RRO carried in Resv message indicating NP-MP is no longer in LSP path**
  - **PLR sets its local NodeID address in HOP object**
  - **Receiver should accept PathTear when HOP object contains NodeID address of PLR and delete the state**

- How does "Remote" PathTear work in back-back link failures?

- Link C-D & B-C fail in succession (in any order), Routers C and B initiate FRR

- Router B updates RRO in Resv message sent to Router A
  - Router C is no longer present in Resv RRO

- On processing RRO carried in Resv, Router A originates "Remote" PathTear to C

- Router C deletes LSP state

- RSVP-TE capability advertisement to provide seamless
  - RSVP-TE capability advertisement to provide seamless interoperability with the implementations that do not support the

- RSVP-TE Conditions object in "Conditional" num
  10bbbbbb (RFC 2205) so that implementations that do no support the new extensions:
  - process the message as "vanilla"
  - *ignore the object, neither forwarding nor sending an error message* (RFC 2205)
  - process the message as "vanilla" PathTear

For more details…

https://tools.ietf.org/html/draft-chandra-mpls-enhanced-frr-bypass-01