Refresh Interval Independent facility FRR

draft-chandra-mpls-enhanced-frr-bypass-00

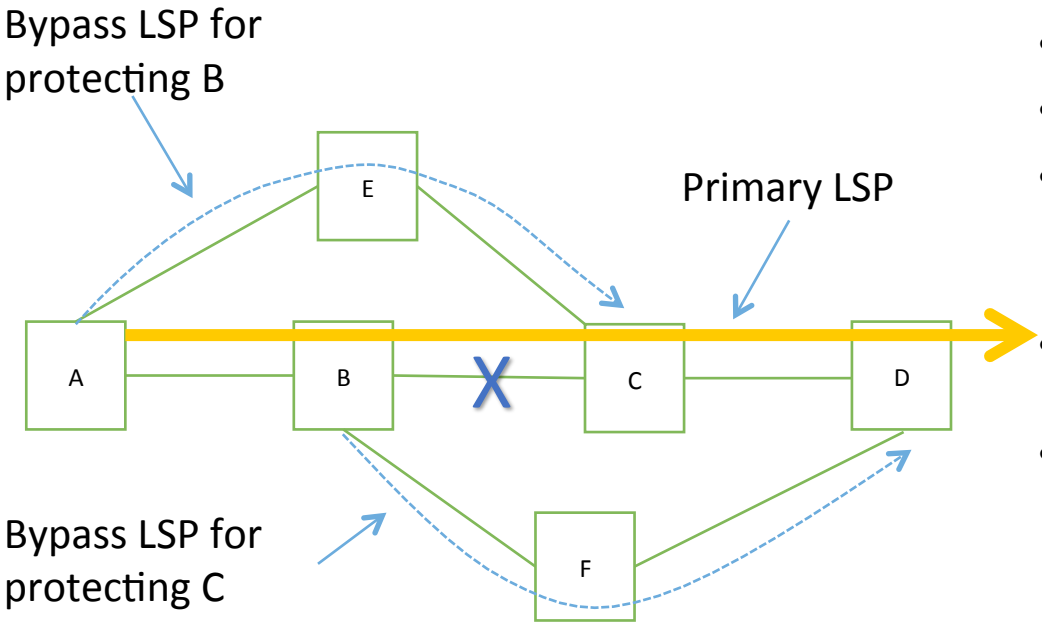
Chandra Ramachandran (csekar@juniper.net)

Yakov Rekhter (yakov@juniper.net)

Refresh independent state maintenance and stale state cleanup

- *"RSVP sends periodic refresh messages to maintain the state along the reserved path(s). In the absence of refresh messages, the state automatically times out and is deleted."* (RFC2205)
- Original RSVP-TE inherited this from RSVP
- 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 rest of this presentation covers fixes to RFC4090 to provide refresh independent stale state cleanup during facility-based FRR

Fast stale state cleanup during RSVP-TE Fast Reroute – fixing RFC4090 (1)

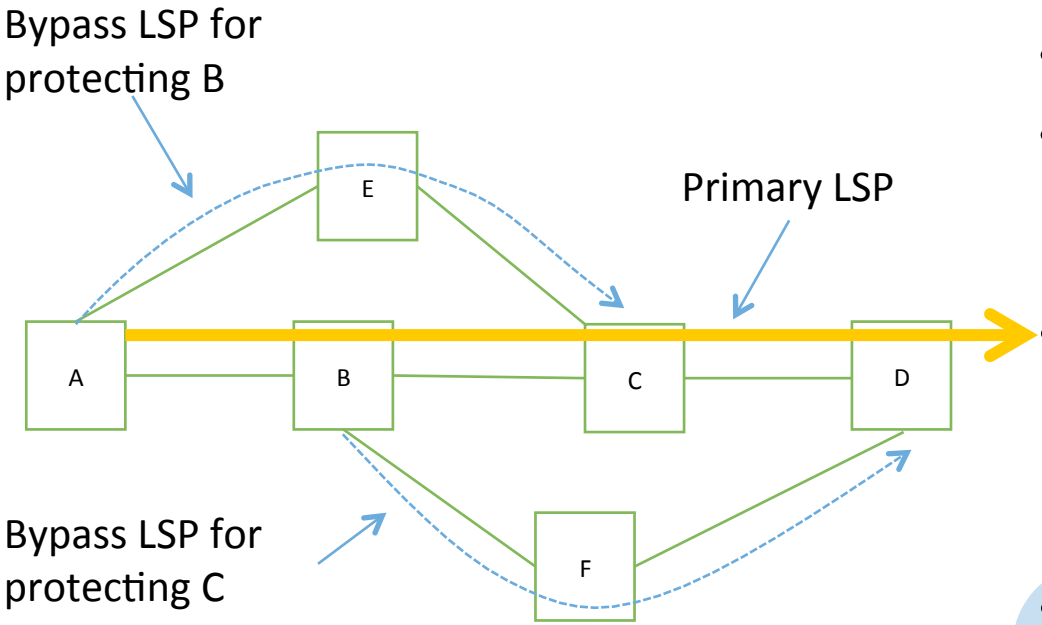


Protecting against B node failure:
A – Point of Local Repair (PLR)
C – Merge Point (MP)

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

- 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 refreshed*" (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 refresh timer expires*" (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
- Bottom line: 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

Fast stale state cleanup during RSVP-TE Fast Reroute – fixing RFC4090 (2)



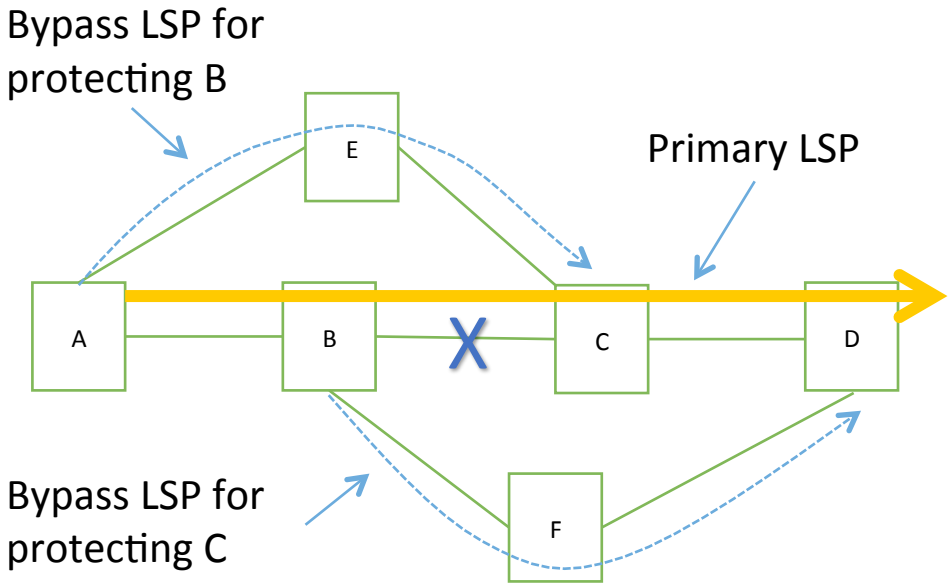
Protecting against B node failure:
A – Point of Local Repair (PLR)
C – Merge Point (MP)

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

- How can a node determine it is Merge Point for a protected LSP?
- "*Whenever the PLR has a backup path available, the PLR MUST set the local protection available flag*" (RFC 4090)
  - Specified only for RRO carried in RESV – do this for RRO carried in PATH too
- Initiate NodeID signaling adjacency between PLR and MP
  - "*A Hello session in which local and remote NodeIDs are used in source and destination fields of Hello packet*" (RFC 4558)
  - MP may delete LSP state if PLR goes down

- **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 and NodeID signaling adjacency with PLR is up.**

Fast stale state cleanup during RSVP-TE Fast Reroute – fixing RFC4090 (3)

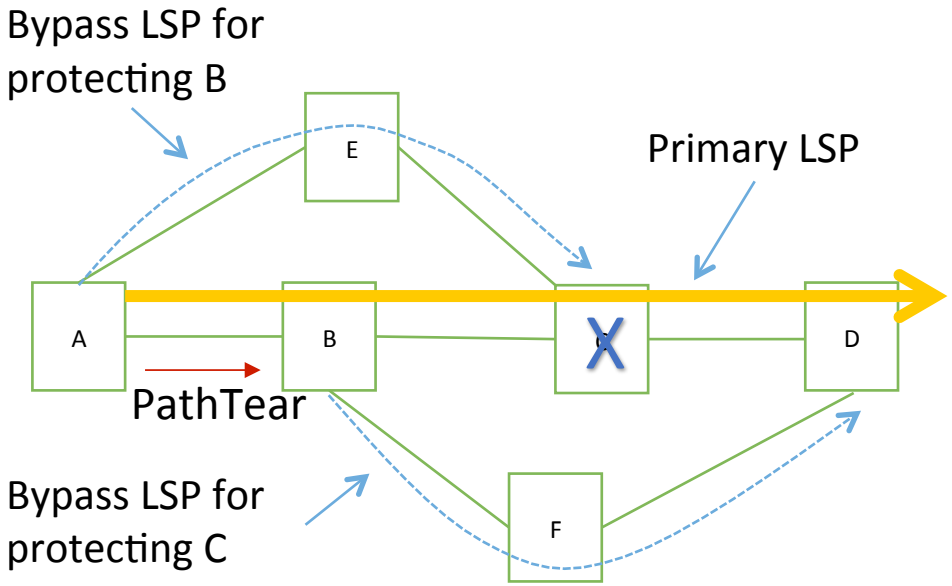


Protecting against B node failure:
A – Point of Local Repair (PLR)
C – Merge Point (MP)

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

- Link (B, C) goes down
- 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, C deletes RSVP-TE state for the primary LSP without waiting for refresh timeout
  - Fast stale state cleanup is accomplished !
- As C deletes RSVP-TE state, C sends "Conditional" PathTear to D
  - Instead of sending "vanilla" PathTear
- **"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)**
- When D receives "Conditional" PathTear, D does not delete the state for the primary LSP, as (a) D is a node protection Merge Point for that LSP, and (b) the LSP is currently being protected by the Point of Local Repair B, associated with this Merge Point
- **Problem fixed !!!**

Fast stale state cleanup during RSVP-TE Fast Reroute – fixing RFC4090 (4)



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)

- Router C goes down
- Ingress A initiates LSP tear down triggered by an administrative event
- B receives tear down message before B could initiate backup LSP signaling
- As B deletes state, B does not initiate backup signaling
  - *Receipt of PathTear (path teardown) message deletes matching path state* (RFC 2205)
- D would remove state only when refresh timeout expires
  - Again, RFC 4090 relies on refresh timeout for stale state cleanup!
- **"Remote" PathTear:**
  - **Originated by a PLR when it deletes the RSVP-TE LSP state before the PLR has completed backup LSP signaling**
  - **PLR sets its local NodeID address in HOP object**
  - **Receiver should accept PathTear when HOP object contains NodeID address of PLR and deletes the state**
- **All stale state issues are resolved.**

Backward compatibility

- RSVP-TE capability advertisement to provide seamless interoperability with the implementations that do not support the new extensions
  - Use RSVP Capability object (RFC 5063) to advertise support for new extensions
  - Requires IANA allocation in the Capability object value registry
- RSVP-TE Conditions object in “Conditional” PathTear belongs to Class-num 10bbbbbb (RFC 2205) so that implementations that do no support the new extensions:
  - *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-00