BFD in Segment Routing Networks Using MPLS Dataplane

draft-mirsky-spring-bfd

Greg Mirsky
Jeff Tantsura
Mach Chen
Ilya Varlashkin

BFD over MPLS dataplane

- RFC 5884 has defined use of BFD Asynchronous mode over MPLS LSP
- BFD over SR-MPLS SHOULD use LSP Ping to bootstrap BFD session
- In addition to requirements stated in draft-ietf-mpls-spring-lsp-ping:
  - Initiator MUST include FEC(s) corresponding to the destination segment.
  - Initiator, i.e. ingress LSR, MAY include FECs corresponding to some or all of segments imposed in the label stack by the ingress LSR to communicate the segments traversed.

  add:
  - When LSP Ping is used to bootstrap a BFD session the FEC corresponding to the destination segment to be associated with the BFD session MUST be as the very last sub-TLV in the Target FEC TLV.

- BFD control packet encapsulation:
  - with IP/UDP header MUST:
    - destination IP address 128/8 for IPv4 address or 0:0:0:0:FFFF:7F00/104 for IPv6 address;
    - use UDP destination port 3784
  - ACH encapsulation use GAL and G-ACh type 0x0007
BFD Reverse Path

- Ingress LER A periodically transmits BFD control messages over MPLS LSP
- Egress LER B periodically transmits BFD control messages, per RFC 5884, over path selected based on local policy:
  - IP network using UDP destination port 4784
  - Reverse path segment route with IP/UDP encapsulation (UDP destination port 3784) or ACH encapsulation
- Failure in the reverse path of the BFD session may be interpreted as LSP failure
Control BFD Reverse Path

- New optional BFD Reverse Path TLV
- Used with BFD Discriminator TLV
- Instructs egress BFD to transmit BFD control packets over the specified MPLS LSP
- Re-use sub-TLVs defined in draft-ietf-mpls-spring-lsp-ping
- BFD Reverse Path TLV may contain none, one or more sub-TLVs
- If none sub-TLV has been found in the BFD Reverse Path TLV, then the egress BFD MUST transmit BFD control packets over IP network
New Segment Routing Static MPLS Tunnel sub-TLV

- Ordered list of Label Stack Elements with the top of the stack label as Label Entry 1 and the bottom of the stack label – Label Entry N
- BFD Reverse TLV MAY include zero or one SR Static MPLS Tunnel sub-TLV
- If no sub-TLVs present in the BFD Reverse Path TLV – the egress MUST switch the reverse BFD session to be transmitted over IP network
- If more than one SR Static MPLS Tunnel sub-TLVs present in the BFD Reverse Path TLV, the remote peer MUST send MPLS LSP Echo Reply with Return Code value set to “Too Many TLVs Detected” (new code)

<table>
<thead>
<tr>
<th>SegRouting MPLS sub-TLV Type</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label Entry 1</td>
<td></td>
</tr>
<tr>
<td>Label Entry N</td>
<td></td>
</tr>
</tbody>
</table>
BFD Demand mode
(based on draft-mirsky-bfd-mpls-demand)

- RFC 5880 defined BFD Demand mode
- BFD node controls mode of its peer, i.e. the BFD node MAY switch its peer into and out of the Demand mode
- To verify bi-directional continuity the node in Demand mode MAY initiate Poll sequence by simply setting Poll (P) bit in BFD control messages sent periodically to its peer
- BFD node in the Demand mode MAY send BFD control messages with Poll (P) bit set if any of its parameters have changed
Theory of operation I

- Ingress LER A bootstraps the BFD session to LER B using LSP Ping
- BFD session between A and B in Async mode reaches Up state
- BFD node A switches mode to Demand using Poll sequence
- Node B ceases transmission of periodic BFD packets
Ingress LER A bootstraps the BFD session to LER B using LSP Ping
• BFD session between A and B in Async mode reaches Up state
• BFD node A switches mode to Demand using Poll sequence
• Node B ceases transmission of periodic BFD packets
• Node B detects failure
  − Node B initiates Poll sequence with Diagnostic code set to Control Detection Time Expired
  − Because Node A have received failure notification from the node B (RDI) it sends BFD control packet with Final bit set over IP network as following:
    • destination IP address MUST be set to the destination IP address of the LSP Ping Echo request message
    • destination UDP port set to 4784
    • Final (F) flag in BFD control packet MUST be set
    • Demand (D) flag in BFD control packet MUST be cleared
  − Node A moves BFD session state to Down
  − Node A switches the BFD session to Async mode
  − Node A transmits BFD control packets periodically at slow rate
Next steps

• Your comments, suggestions, questions always welcome and greatly appreciated
• Which WG to anchor – MPLS or SPRING?