An MPLS-Based Forwarding Plane for Service Function Chaining

draft-farrel-mpls-sfc-02
Stewart Bryant <stewart.bryant@gmail.com>
Adrian Farrel <afarrel@juniper.net>
John Drake <jdrake@juniper.net>

IETF-100, Singapore, November 2017
Where To Have This Discussion?

• SFC WG has developed problem statement and architecture for SFC
  – We re-use these
• SFC chartered to work on “generic encapsulation” that is “agnostic to the
  layer at which it is applied”
  – Has developed the NSH
• This work is specific to an MPLS forwarding plane and uses an
  MPLS encapsulation
  – In this meeting we present in MPLS and SFC
    • Need review from experts
      – Want to be sure MPLS parts work
      – Want to be sure SFC parts work
    • Some functions need specific MPLS extensions
  – Let the chairs and ADs work out where the work belongs
Recall the SFC Architecture

- Packets flow from source to destination
- Packets are classified onto a Service Function Path (SFP)
- SFP traverses a series of Service Function Forwarders (SFFs)
- Each SFF delivers packets on the SFP to a specific Service Function Instance (SFI)
- SFC Proxy may be placed between SFF and SFI
Objectives / Non-objectives

1. Not trying to replace or obsolete NSH
2. Looking at a specific environment where deployed MPLS routers can serve as SFFs
   - No change to forwarding plane
   - Able to forward SFC packets “at line speed”
3. Support both modes of MPLS forwarding
   - Label swapping (also known as traditional MPLS)
   - Label stacking (also known as MPLS Segment Routing)
   - (Advanced function: allow “mix” of swapping and stacking)
4. Aim to get high level of SFC functionality
   - Possible that some features will be sacrificed in compromise with desire to achieve points 2 and 3
   - Must support SFC architecture (RFC 7665)
   - Should support metadata
   - Try to integrate with control plane solutions
     • draft-ietf-bess-nsh-bgp-control-plane
Overview of Solution

• Basic building block is a two-label unit
  • Labels included as Label Stack Entries
  • Neither of the labels can be in the range 0..15
    – I.e., must not overlap with Special Purpose Label values
  • This two-label unit is used differently for label swapping and label stacking
MPLS Label Swapping

- Tunnels between SFFs “as normal”
  - Of course, we are interested in MPLS as the transport
- SPI and SI used “as normal” for NSH
  - Some limitation as SPI is constrained here to 20 bits

### MPLS-SFC processing...

<table>
<thead>
<tr>
<th>Tunnel Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFC Context Label = SPI</td>
</tr>
</tbody>
</table>

| Service Function Label = SI |

| Payload |

- Labels are looked up and acted on by SFF to determine next hop
  - Maybe forward to SFI or SFC proxy
  - Maybe forward to next SFF
- In some cases action can be achieved simply through SPI
- In other cases need the two label context
  - SI is updated before further forwarding (it’s a swap)
  - SPI and SI set during classification
    - Potentially also during re-classification
Label Stacking (MPLS-SR)

• Tunnel labels might not be needed
  – SFC Context Label can be the SID of the SFF

• MPLS-SR processing
  • Stack of two-label units
  • SFC Context Label
    • Tells you how to get to next SFF (tunnel of SR)
  • SF Label identifies the SF given the SFC context
  • Pop the two-label unit before forwarding to next SFF
What About Metadata?

- MPLS encapsulation not well suited for carrying "arbitrary" metadata
- We define an Extended Special Purpose Label
  - This three-label sequence can be included at the bottom of the label stack
  - Metadata label is an index into a store of metadata
    - Must also not use 0..15
    - Store may be populated though management plane, control plane, or in-band (next slide)
      - This approach is not good for "per-packet metadata" (e.g., hashes)
      - Works fine for per-SFP or per-flow metadata
In-Band Metadata Distribution

• Consider draft-farrel-sfc-convent
  – Defines use of NSH with Next Protocol == None
  – Can be used to send NSH packets along an SFP without carrying payload (but still carrying metadata)
  – This draft defines how to do this in MPLS

<table>
<thead>
<tr>
<th>15 = Extended Special Purpose Label Follows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata Present Indicator (MPI)</td>
</tr>
<tr>
<td>Metadata Label</td>
</tr>
<tr>
<td>Length</td>
</tr>
<tr>
<td>Metadata</td>
</tr>
</tbody>
</table>

• Use an Extended Special Purpose Label
  • Hence, a three label sequence
  • Placed at the bottom of the label stack
  • Rest of stack exactly as for SFP
  • Metadata carried as payload
    • Formatted as TLV
    • Type field defined by SFC WG for NSH
    • Metadata as defined by SFC WG
Next Steps

• There are always things to polish, but...
  – This is now relative stable
  – Support for swapping and stacking in a common way took some effort, but has good benefits

• Fits with BESS control plane work

• To the authors it seems “obvious”
  – What do other people think?

• The authors think this is in charter for MPLS WG
  – Use of special purpose labels belongs in MPLS
  – But obviously it needs review by SFC WG

• Actions for chairs
  – Decide where this belongs
  – Think about adoption