Multi-domain Service Forwarding For NSH
draft-li-sfc-nsh-multi-domain-03

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on October 19, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include

Li                      Expires April 12, 2018                  [Page 1]
Abstract

This document describes the mechanism to achieve multi-domain service forwarding for NSH. The proposed mechanism adopts a horizontal deployment structure, which divides a multi-domain SFC into several segmental SFCs in the control plane and each domain creates its own SFP independently in data plane. A label is proposed to identify different cross-domain flows at the classifier by extending the metadata of the NSH encapsulation.

1. Introduction

Service Function Chaining (SFC) [RFC7665] is an architecture proposed to decouple traditional network service functions with corresponding physical resources. It is flexible and convenient for the network operator to deploy on-demand service functions and steer the traffic through them in sequence.

Network Service Header (NSH) [I-D.ietf-sfc-nsh] is defined as a data plane protocol to create dynamic service function chains. According to the NSH encapsulation, the flow can pass along pre-defined Service Function Path and exchange metadata among the Service Classifier, the
Service Function and the Service Function Forwarder for information sharing.

Network service forwarding in SFC is based on the combination of Service Path Identifier (SPI) and Service index (SI), which are defined in [I-D.ietf-sfc-nsh]. [I-D.kumar-sfc-nsh-forwarding] analyzes the NSH forwarding shortcomings and further discusses the separation of the service forwarding and the service delivery. However, it focuses on infrastructure service forwarding for NSH in a single domain. [I-D.ietf-sfc-hierarchical] proposes a hierarchical way to achieve multi-domain forwarding for SFC, which can be regarded as a vertical approach. Contrast to the vertical approach, a horizontal one is easy to scale in and out. Besides, the horizontal approach can decrease the overhead of the control plane, because it maintains more service traffic in the data plane.

Therefore, this document proposes a horizontal approach to achieve multi-domain service forwarding for NSH. The main idea is to divide a multi-domain SFC into several segmental SFCs according to domain partitions in the control plane and create corresponding SFPs in each domain by its classifier independently. A label is proposed to identify different cross-domain flows, which is encapsulated in the metadata.

1.1. Requirement Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Multi-domain Service Forwarding Mechanism

Figure 1 Multi-domain SFC Divide
As shown in Figure 1, according to requirements, it divides the whole SFC in the control plane and issues cross-domain forwarding tables to corresponding classifiers initially. Multi-domain service forwarding includes two aspects: the inter-domain service forwarding and the intra-domain service forwarding. The former is based on a unique label for each flow, and the latter is performed by the classifier. To achieve a unified service forwarding mechanism in multi-domain, this mechanism uses metadata in NSH encapsulation to carry the necessary forwarding information among different forwarding elements, such as the Service Classifier and the Service Function Forwarder.

3. Definition Of Terms

This document uses some terms defined in SFC architecture [RFC7665] and NSH [I-D.ietf-sfc-nsh] drafts for ease of understanding and the reader is advised to refer to those documents for up to date and additional information.

Segmental SFC: The cross-domain SFC is divided into several segmental SFCs according to the domain partition. Each segmental SFC is assigned to its corresponding domain.

Service Label: the label used to identify different flows can help the classifier create the SFP by its corresponding segmental SFC, which is issued from SFC control plane.

Cross-domain Forwarding Table: There are three columns in the table: Service Label, Next Classifier and Segmental SFC. The table matches a specific flow with its Service Label. The cross-domain forwarding of the flow is depended on the address of Next Classifier. The Cross-domain Forwarding Table is maintained by SFC control plane and issued to corresponding classifier according to the Segmental SFC partition.

3.1. Service Function Chaining Segmentation

At first, the control plane creates a SFC with a unique Service Label for the flow. Then, the SFC is divided into several segmental SFCs according to physical resource constraints. It is important to note that the control plane MUST NOT specify the SFP directly. The control plane is only responsible to indicate what service functions are required in each domain, and the corresponding SFP MUST be specified by the service classifier.

3.2. Inter-domain Service Forwarding

The Service Label is proposed to identify different flows when packets need to be cross-domain forwarded. When the service
classifier receives packets with NSH encapsulations, it checks the Service Label of the first packet to look up the address of the next classifier in its cross-domain forwarding table, which is issued by SFC control plane. Then the information of next classifier is written into the metadata and will be used by the last SFF in the Segmental SFC. Once the last SFF receives the packet from last SF in the domain which changes the SI to zero, it forwards the packets to next classifier directly without any modification of their NSH encapsulations.

The Service Label is only carried by the first packet of a certain flow with specific SPI. It’s beneficial to decrease header cost and improve forwarding efficiency.

3.3. Classification and Intra-domain Service Forwarding

The service classifier creates SFP for each flow according to the segmental SFC in its cross-domain forwarding table. After the control plane assigns segmental SFCs for different domains, the corresponding table is issued to the classifier in each domain. When the packet arrives at the classifier, its Service Label is used to find out the next segmental SFC. Then, the classifier creates a SFP for the flow according to that segmental SFC.

In this situation, the classifier in a segmental SFC SHOULD set the SI to the length of the segmental SFC.

4. Multi-domain Service Forwarding Encapsulation

In order to reduce overhead of metadata, the context header with MD Type = 0x2 is chosen to support multi-domain service forwarding. Figure 2 shows the allocation of the metadata.

<table>
<thead>
<tr>
<th>Ver</th>
<th>O</th>
<th>C</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>Length</th>
<th>MD-type=0x2</th>
<th>Next Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Path Identifier (SPI)</td>
<td>Service Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metadata Class=0x20</td>
<td>Type=0x1</td>
<td>R</td>
<td>Len=0x4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>Service Label</td>
<td>Next Classifier</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 Multi-domain Service Forwarding Encapsulation

Context Header Allocation Descriptions:
Metadata Class: It MUST be set to 0x20 as requested in Section 7.

C bit: It SHOULD be set indicating critical metadata exists.

Type: It MUST be set to 0x1 for multi-domain service forwarding.

Len: Because the length of the context header is constant, it MUST be set to 0x4.

D bit: The default value is zero, which means the metadata SHOULD be ignored. When set to 0x1 indicates that this packet need to be forwarded in multiple domains.

Service Label: Identifies different flows with different labels. The service classifier decides the service forwarding path of the packet according to its domain label with a forwarding table issued by control plane.

Next Classifier: Indicates the identifier of the classifier in next domain. Because of the SFC segmentation, each classifier only works in the SFC domain it belongs to. When the current segmental SFC is terminate, the last SFF will query the next segmental domain by this identifier.

All other flag fields are reserved for future use. Reserved bits MUST be set to zero and MUST be ignored upon receipt.

5. Multi-domain Service Forwarding Example

```
+----------+ +----------+              +----------+
|          | |          |              |          |
| Domain1  |   SF-a   | |   SF-b   |     Domain2  |   SF-c   |
|          | |          |              |          |
+---^--+---+ +---^--+---+              +---^--+---+
|  |         |  |                      |  |
+-------+ +---+--v---+ +---+--v---+    +-------+ +---+--v---+
|       | |          | |          |    |       | |          |
--->  CF1  +->  SFF1-1  +->  SFF1-2  +---->  CF2  +->  SFF2-1  +--->
|       | |          | |          |    |       | |          |
+-------+ +----------+ +----------+    +-------+ +----------+
```

Figure 3 Multi-domain Service Forwarding Example

This section describes the scenario shown in Figure 3, a packet flow pass through three service functions deployed in two domains. The Domain1 is consist of CF1, SFF1-1, SFF1-2, SF-a and SF-b; the Domain2 is consist of CF2, SFF2-1 and SF-c. The workflow is as follow:
1. SFC control plane creates the SFC with a Service Label for the specific flow and divides the whole SFC into several segmental SFCs.

2. The cross-domain forwarding table is issued to its corresponding classifier with specific Service Label, Next Classifier and Segmental SFC.

3. CF1 (the classifier in domain1) creates the SFP and the NSH encapsulation for the first packet of a certain flow according to its segmental SFC in domain1. It sets ‘D’ flag to 1 and fills in the Service Label and Next Classifier. The SI is set to the length of this segmental SFC.

4. CF1 forwards the packet to SFF1-1.

5. SFF1-1 determines SF-a as the next hop and forwards the packet.

6. After SF-a processes the packet, the packet is forwarded back to SFF with decremented SI.

7. SFF1-1 forwards the received packet to SFF1-2.

8. SFF1-2 determines SF-a as the next hop and forwards the packet.

9. Similar to SF-a, SF-b forwards the packet to SFF1-2 and decrements SI to zero.

10. When SFF1-2 receives the packet, it finds out that the segmental SFC is terminate in this domain because of the SI, and then SFF1-2 forwards the packet to next classifier without modification of the NSH encapsulation. The following packets of this flow (in same SPI in this domain) are forwarded with the same routing information.

11. When CF2 (the classifier in domain2) receives the first packet with an NSH encapsulation, CF2 will checks its Service Label.

12. According to the Service Label, CF2 looks up segmental SFC in the cross-domain forwarding table and creates the corresponding SFP and the NSH encapsulation.

13. CF2 sets the ‘D’ flag to zero and forwards the packet to SFF2-1. The action of its following packets are similar.

14. SFc processes the packet and forwards it back to SFF2-1 with decremented SI.
15. SFF2-1 finds out that the SFP is terminate and forwards the packet to the final Receiver.

6. Security Considerations

As with many other protocols, metadata of the NSH encapsulation can be spoofed or otherwise modified. It is important to protect the cross-domain packet from malicious modification, because the metadata contains sensitive information about the user and environment. Therefore, it is significant to ensure the integrity of the metadata and provide the protection of the user privacy.

7. IANA Considerations

IANA is requested to allocate a new class from the TLV Class defined in [I-D.ietf-sfc-nsh].

0x20 Multi-domain Forwarding Type

IANA is requested to set up a registry of "NSH Multi-domain Service Forwarding TLV Types". These are 7-bit values. Registry entries are assigned by using the "IETF Review" policy defined in [RFC5226].

IANA is requested to allocate two new types as follows:

- Type = 0x00 Reserved
- Type = 0x01 Multi-domain Service Forwarding

8. Conclusions

This document proposes a mechanism for multi-domain service forwarding based on a unique label. In order to relieve the pressure of the control plane, the multi-domain SFC is divided into segmental SFC according to the domain partitions, and the SFP in each domain is created independently.

9. References

9.1. Normative References


[I-D.ietf-sfc-nsh]
9.2. Informative References


[I-D.ietf-sfc-hierarchical]


[I-D.kumar-sfc-nsh-forwarding]


10. Acknowledgments

This work in this document was supported by National High Technology of China ("863 program") under Grant No.2015AA015702.
Authors’ Addresses

Guanwen Li
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: 14120079@bjtu.edu.cn

Guanglei Li
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: 15111035@bjtu.edu.cn

Taixin Li
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: 14111040@bjtu.edu.cn

Qi Xu
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: 15111046@bjtu.edu.cn

Bohao Feng
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: bohaofeng@bjtu.edu.cn

Huachun Zhou
Beijing Jiaotong University
Beijing 100044, P.R. China
Email: hchzhou@bjtu.edu.cn