SFC R. Penno
Internet-Draft C. Pignataro
Intended status: Standards Track C. Yen
Expires: March 24, 2016 E. Wang
K. Leung
Cisco Systems
September 21, 2015

Packet Generation in Service Function Chains
draft-penno-sfc-packet-00

Abstract

Service Functions (e.g., Firewall, NAT, Proxies and Intrusion Prevention Systems) generate packets in the reverse flow direction to the source of the current in-process packet/flow. In this document we discuss and propose how to support this required functionality within the SFC framework.

Requirements 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].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Service Functions (e.g., Firewall, NAT, Proxies and Intrusion Prevention Systems) generate packets in the reverse flow direction destined to the source of the current in-process packet/flow. This is a basic intrinsic functionality and therefore needs to be supported in a service function chaining deployment.

2. Problem Statement

The challenge of this functionality in service chain environments is that generated packets need to traverse in the reverse order the same Service Functions traversed by original packet that triggered the packet generation.

Although this might seem to be a straightforward problem, on further inspection there are a few interesting challenges that need to be solved. First and foremost a few requirements need to be met in order to allow a packet to make its way through back to its source through the service path:

- o A symmetric path ID needs to exist. Symmetric path is discussed in [SymmetricPaths]
- o The SF needs to be able encapsulate such error or proxy packets in a encapsulation transport such as VXLAN-GPE [I-D.ietf-nvo3-vxlan-gpe] + NSH header [I-D.ietf-sfc-nsh]
- o The SF needs to be able to determine, directly or indirectly, the symmetric path ID and associated next service-hop index or indicate reverse path for the service path ID in the original packet (TBD: verify or part)

3. Definitions and Acronyms

The reader should be familiar with the terms contained in [I-D.ietf-sfc-nsh] ,[I-D.ietf-sfc-architecture] and [I-D.ietf-nvo3-vxlan-gpe]

4. Service Function Behavior

When a Service Function wants to send packets to the reverse direction back to the source it needs to know the symmetric service path ID (if it exists) and associated service index. This information is not available to Service Functions since they do not need to perform a next-hop service lookup. There are four recommended approaches to solve this problem and we assume different implementations might make different implementation choices. (Editor note: "if it exists?", reverse service path ID is required on SF or SFF)

1. The SF can receive service path forwarding information in the same manner a SFF does. (Editor's note: not sure about this? SFF looks up next hop based on SPID/SI for normal SFF forwarding. In this case, SF needs to add the reverse SPID/SI which is similar to OAM support)

2. The SF can send the packet in the forward direction but set appropriate bits in the NSH header requesting a SFF to send the packet back to the source

3. The classifier can encode all information the SF needs to send a reverse packet in the metadata header

4. The controller uses a deterministic algorithm when creating the associated symmetric path ID and service index.

We will discuss the ramifications of these approaches in the next sections.

4.1. SF receives Reverse Forwarding Information

This solution is easy to understand but brings a change on how traditionally service functions operate. It requires SFs to receive and process a subset of the information a SFF does. When a SF wants to send a packet to the source, the SF uses information conveyed via the control plane to impose the correct NSH.

Advantages:

- Changes are restricted to SF and controller, no changes to SFF
- Incremental deployment possible
- No protocol between SF and SFF, which avoids interoperability issues
- No performance penalty on SFF due to in or out-of-band protocol

Disadvantages:

- SFs need to process and understand Rendered Service Path messages from controller

This solution can be characterized by putting the burden on the SF, but that brings the advantage of being self-contained (as well as providing a mechanism for other features). Also, many SFs have policy or classification function which in fact makes them a classifier and SF combination in practice.

4.2. SF requests SFF cooperation

These solutions can be characterized by distributing the burden between SF and SFF. In this section we discuss two possible in-band solutions: using OAM header and using a reserved bit 'R' in the NSH header.

4.2.1. OAM Header

When the SF needs to send a packet in the reverse direction it will set the OAM bit in the NSH header and use an OAM protocol [I-D.penno-sfc-trace] to request that the SFF impose a new, reverse-path NSH. Post imposition, the SFF forwards the packet correctly.

SF Reverse Packet Request

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+\
 |Ver|1|C|R|R|R|R|R|R|   Length  |  MD-type=0x1  |  OAM Protocol | |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
 |          Service Path ID                      | Service Index | |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
 |                Mandatory Context Header                       | |S
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |F
 |                Mandatory Context Header                       | |C
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
 |                Mandatory Context Header                       | |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
 |                Mandatory Context Header                       | |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ <
 |Rev. Pkt Req   |         Original NSH headers (optional)       | |O
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |A
                                                                  |M
                                                                 /
```

(postamble)

Ver: 1

OAM Bit: 1

Length: 6

MD-Type: 1

Next Protocol: OAM Protocol

Rev. Pkt Req: 1 Reverse packet request

Advantages:

- o SF does not need to process and understand control plane path messages.

- o Clear division of labor between SF and SFF.

- o Extensible

- o Original NSH header could be carried inside OAM protocol which leaves metadata headers available for SF-SFF communication.

Disadvantages:

- o SFFs need to process and understand a new OAM message type

- o Possible interoperability issues between SF-SFF

- o SFF Performance penalty

4.2.2. Service Function Forwarder Behavior

In the case where the SF has all the information to send the packet back to the origin then no changes are needed at the SFF. When an SF requests SFF cooperation the SFF MUST be able to process the OAM message used to signal reverse-path forwarding:

- o Process/decode OAM message

- o Examine and act on any metadata present in the NSH header

- o Examine its forwarding tables and find the symmetric path-id and index of the next service-hop

The symmetric path can be found in the Rendered Service Path Yang model [RSPYang] and is conveyed to the SFF when a path is constructed.

If a SFF does not understand the OAM message it just forwards the packet based on the original path-id and index. Since it is a special OAM packet, it tell other SFF and SF that they should process it differently. For example, a downstream intrusion detection SF might not associated flow state with this packet.

4.2.3. Reserved bit

In this solution the SF sets a reversed bit in the NSH that carries the semantic as the OAM solution discussed previously. This solution is simpler from a SF perspective but requires allocating one of the reversed bits. Another issue is that the metadata in the original packet might be overwritten by SFs or SFFs in the path.

When SFF received the reversed bit, it shall look up a preprogrammed table to map the Service Path ID and Index in the NSH into a new Service Path ID and Index. The SFF would then use the ID and Index pair to determine the SF/SFF which is in the reverse direction.

Advantages:

o No protocol header overhead

o Limited performance impact on SF

Disadvantages:

o Use of a reserved bit

o SFF Performance penalty

o Not extensible

4.3. Classifier Encodes Information

This solution allows the Service Function to send a reverse packet without interactions with the controller or SFF, therefore it is very attractive. Also, it does not need to have the OAM bit set or use a reserved bit. The penalty is that for a MD Type-1 packet a significant amount of information (48 bits) need to be encoded in the metadata section of the packet and this data can not be overwritten. Ideally this metadata would need to be added by the classifier.

The Rendered Service Path yang model [RSPYang] already provides all the necessary information that a classifier would need to add to the metadata header. An explanation of this method is better served with an examples.

4.3.1. Symmetric Service Paths

The figure below shows a simple SFC with symmetric service paths comprising three SFs.

```
   .....................SFP1 Forward........................>

     Forward SI  253          252          251

   +---+         .-.          .-.          .-.            +---+
   |   |        /   \        /   \        /   \           |   |
   | A +-------( SF1 )------( SF2 )------( SF3 )----------+ B |
   |   |        \   /        \   /        \   /           |   |
   +---+         `-'          `-'          `-'            +---+

     Reverse SI  253          254          255

    <....................SFP1 Reverse.........................


     SFP1 Forward ->  SF1 : SF2 : SF3
     SFP1 Reverse <-  SF1 : SF2 : SF3

     RSFP1 Forward -> SF1 : SF2 : SF3
     RSFP1 Reverse <- SF1 : SF2 : SF3
```

Figure 1: SFC example with symmetric path

Below we see the JSON objects of the two symmetric paths depicted above.


```
RENDERED_SERVICE_PATH_RESP_JSON = """
{
  "rendered-service-paths": {
    "rendered-service-path": [
      {
        "name": "SFC1-SFP1-Path-2-Reverse",
        "transport-type": "service-locator:vxlan-gpe",
        "parent-service-function-path": "SFC1-SFP1",
        "path-id": 3,
        "service-chain-name": "SFC1",
        "starting-index": 255,
        "rendered-service-path-hop": [
          {
            "hop-number": 0,
            "service-index": 255,
            "service-function-forwarder-locator": "eth0",
            "service-function-name": "SF3",
            "service-function-forwarder": "SFF3"
          },
          {
            "hop-number": 1,
            "service-index": 254,
```

```
              "service-function-forwarder-locator": "eth0",
              "service-function-name": "SF2",
              "service-function-forwarder": "SFF2"
            },
            {
              "hop-number": 2,
              "service-index": 253,
              "service-function-forwarder-locator": "eth0",
              "service-function-name": "SF1",
              "service-function-forwarder": "SFF1"
            }
          ],
          "symmetric-path-id": 2
        },
        {
          "name": "SFC1-SFP1-Path-2",
          "transport-type": "service-locator:vxlan-gpe",
          "parent-service-function-path": "SFC1-SFP1",
          "path-id": 2,
          "service-chain-name": "SFC1",
          "starting-index": 253,
          "rendered-service-path-hop": [
            {
              "hop-number": 0,
              "service-index": 253,
              "service-function-forwarder-locator": "eth0",
              "service-function-name": "SF1",
              "service-function-forwarder": "SFF1"
            },
            {
              "hop-number": 1,
              "service-index": 252,
              "service-function-forwarder-locator": "eth0",
              "service-function-name": "SF2",
              "service-function-forwarder": "SFF2"
            },
            {
              "hop-number": 2,
              "service-index": 251,
              "service-function-forwarder-locator": "eth0",
              "service-function-name": "SF3",
              "service-function-forwarder": "SFF3"
            }
          ],
          "symmetric-path-id": 3
        }
      ]
    }
```

}"""

We will assume the classifier will encode the following information in the metadata:

o symmetric path-id = 2 (24 bits)

o symmetric starting index = 253 (8 bits)

o symmetric number of hops = 3 (8 bits)

o starting index = 255 (8 bits)

In the method below we will assume SF will generate a reverse packet after decrementing the index of the current packet. We will call that current index.

If SF1 wants to generate a reverse packet it can find the appropriate index by applying the following algorithm:

```
current_index = 252

remaining_hops = symmetric_number_hops - starting_index - current_index
remaining_hops = 3 - (255 - 252) = 0
reverse_service_index = symmetric_starting_index - remaining_hops - 1
reverse_service_index = next_service_hop_index = 253 - 0 - 1 = 252
The "-1"  is necessary for the service index to point to the next service_hop.
```

If SF2 wants to send reverse packet:

current index = 253

```
remaining_hops = 3 - (255 - 253) = 1
reverse_service_index = next_service_hop_index = 253 - 1 - 1 = 251
```

IF SF3 wants to send reverse packet:

current index = 254

```
remaining_hops = 3 - (255 - 254) = 2
reverse_service_index = next_service_hop_index = 253 - 2 - 1 = 250
```

The following tables summarize the service indexes as calculated by each SF in the forward and reverse paths respectively.

Fwd SI = forward Service Index
Cur SI = Current Service Index
Gen SI = Service Index for Generated packets

RSFP1 Forward -
Number of Hops: 3
Forward Starting Index: 253
Reverse Starting Index: 255

SF	SF1	SF2	SF3
Fwd SI	253	252	251
Cur SI	252	251	250
Gen SI	252	253	254

RSFP1 Reverse -
Number of Hops: 3
Reverse Starting Index: 255
Forward Starting Index: 253

SF	SF1	SF2	SF3
Rev SI	253	254	255
Cur SI	252	253	254
Gen SI	252	251	250

Figure 2: Service indexes generated by each SF in the symmetric forward and reverse paths

4.3.2. Asymmetric Service Paths

In real world the forward and reverse paths can be asymmetric, comprising different set of SFs or SFs in different orders. The following figure illustrates an example. The forward path is composed of SF1, SF2, SF4 and SF5, while the reverse path skips SF5 and has SF3 in place of SF2.

```
                    ..........        .........
                   .          .      .         .
                  .     249    .    .    246    .
                 .              .  .             .
                .       .-.      ..      .-.      .
  .............        /   \            /   \      ....SFP1 Forward....>
                      ( SF2 )   247    ( SF5 )
  Forward SI   250   / \   / \        / \   /\
                    /   `-'   \      /   `-'  \
                   /           \    /          \
 +---+         .-./             `-./            \               +---+
 |   |        /   \            /   \             \              |   |
 | A +-------( SF1 )----------( SF4 )-------------+-------------+ B |
 |   |        \   /            \   /                            |   |
 +---+         `-'\             ,-'                             +---+
                   \           /
                    \   .-.   /
   Reverse SI  251   \ /   \ /  254
  <..........         ( SF3 )         .................SFP2 Reverse.....
              .        \   /         .
               .        `-'         .
                .                  .
                 .                .
                  .     253      .
                   .............
```

SFP1 Forward -> SF1 : SF2 : SF4 : SF5
SFP2 Reverse <- SF1 : SF3 : SF4

Figure 3: SFC example with asymmetric paths

An asymmetric SFC can have completely independent forward and reverse paths. An SF's location in the forward path can be different from that in the reverse path. An SF may appear in one direction of path only. In order to use the same algorithm to calculate the service index generated by an SF, one design option is to insert special NOP SFs in the rendered service paths so that each SF is positioned symmetrically in the forward and reverse rendered paths. The SFP corresponding to the example above is:

SFP1 Forward -> SF1 : SF2 : NOP : SF4 : SF5

SFP2 Reverse <- SF1 : NOP : SF3 : SF4 : NOP

The NOP SF is assigned with a sequential service index the same way as a regular SF. The SFF receiving a packet with the service path ID and service index corresponding to a NOP SF should advance the

service index till the service index points to a regular SF. Implementation can use a loopback interface or other methods on the SFF to skip the NOP SFs.

Once the NOP SF is inserted in the rendered service paths, the forward and reverse paths become symmetric. The same algorithm can be applied by the SFs to generate service indexes in the opposite directional path. The following tables list the service indexes corresponding to the example above.

Fwd SI = forward Service Index
Cur SI = Current Service Index
Gen SI = Service Index for Generated packets

RSP1 Forward -
  Number of hops: 5
  Forward Starting Index: 250
  Reverse Starting Index: 255

SF	SF1	SF2	NOP	SF4	SF5
Fwd SI	250	249	248	247	246
Cur SI	249	248	247	246	245
Gen SI	250	251	N/A	253	254

RSP1 Reverse -
  Number of hops: 5
  Reverse Starting Index: 255
  Forward Starting Index: 250

SF	SF1	NOP	SF3	SF4	NOP
Rev SI	251	252	253	254	255
Cur SI	250	251	252	253	254
Gen SI	249	N/A	247	246	N/A

4.3.3. Analysis

Advantages:

- o SF does not need to request SFF cooperation or contact controller
- o No SFF performance impact

Disadvantages:

- o Metadata overhead in case MD-Type 2 is used
- o Relies on classifier or SFF to encode metadata information
- o If classifier will encode information it needs to receive and process rendered service path information
- o SFF needs to decrement NOP associated indexes

4.4. Reversed Path derived using Forward Path ID and Index Method

In this simplified model, no extra storage is required from the NSH and SFF does not need to know how to handle the reversed packet nor does it know about it. Reverse Path is programmed by Orchestrator and used by SF having the need to send upstream traffic.

Instead of defining a new Service Path ID, the same Service Path ID is used. The Orchestrator must define the reverse chain of service using a different range of Service Path Index. It is also assumed that the reverse packet must go through the same number of Services as its forward path. It is proposed that Service Path Index (SPI) 1..127 and 255..129 are the exact mirror of each other.

Here is an example: SF1, SF2, and SF3 are identified using Service Path Index (SPI) 8, 7 and 6 respectively.

Path 100 Index 8 - SF1

Path 100 Index 7 - SF2

Path 100 Index 6 - SF3

Path 100 Index 5 - Terminate

At the same time, Orchestrator programs SPI 248, 249 and 250 as SF1, SF2 and SF3. Orchestrator also programs SPI 247 as "terminate". Reverse-SPI = 256 - SPI.

Path 100 Index 247 - Terminate

Path 100 Index 248 (256 - 8) - SF1

Path 100 Index 249 (256 - 7) - SF2

Path 100 Index 250 (256 - 6) - SF3

If SF3 needs to send the packet in reverse direction, it calculates the new SPI as 256 - 6 (6 is the SPI of the packet) and obtained 250. It then subtract the SPI by 1 and send the packet back to SFF

Subsequently, SFF received the packet and sees the SPI 249. It then diverts the packet to SF2, etc. Eventually, the packet SPI will drop to 247 and the SFF will strip off the NSH and deliver the packet.

The same mechanism works even if SF1 later decided to send back another upstream packet. The packet can ping-pong between SF1 and SF3 using existing mechanism.

Advantages:

- No precious NSH area is consumed

- SF self-contained solution

- No SFF performance impact and no cooperation needed

- No Special Classification required

Disadvantages:

- SPI range is reduced and may become incompatible with existing topology

- Assumption that the reverse path Service Functions are the same as forward path, only in reverse

- Reverse paths need to use Service Index = 128 for loop detection instead of SI = 0.

An alternative to reducing Service Path Index range is to make use of a different Service Path ID, e.g. the most significant bit. The bit can be flipped when the SF needs to send packet in reverse. However, the negation of the SPI is still required, e.g. SPI 6 becomes SPI 250

In either case, the SF must have the knowledge through Orchestrator that the reverse path has been programmed and the method (SPI only or SPI + SPID bit) to use.

The mechanism to insert NOP to keep reverse path symmetric as described in section 4.3.2 can be applied in this method as well.

5. Other solutions

We explored other solution that we deemed to complex or that would bring a severe performance penalty:

o An out-of-band request-response protocol between SF-SFF. Given that some service functions need to be able to generate packets quite often this will would create a considerable performance penalty. Specially given the fact that path-ids (and their symmetric counterpart) might change and SF would not be notified, therefore caching benefits will be limited.

o An out-of-band request-response protocol between SF-Controller. Given that admin or network conditions can trigger service path creation, update or deletions a SF would not be aware of new path attributes. The controller should be able to push new information as it becomes available to the interested parties.

o SF (or SFF) punts the packet back to the controller. This solution obviously has severe scaling limitations.

6. Implementation

The solutions "Reversed Path derived using Forward Path ID and Index Method" and "SF receives Reverse Forwarding Information" were implemented in Opendaylight

7. IANA Considerations

TBD

8. Security Considerations

9. Acknowledgements

Paul Quinn, Jim Guichard

10. Changes

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, DOI 10.17487/RFC2616, June 1999, <http://www.rfc-editor.org/info/rfc2616>.

11.2. Informative References

[I-D.ietf-nvo3-vxlan-gpe] Quinn, P., Manur, R., Kreeger, L., Lewis, D., Maino, F., Smith, M., Agarwal, P., Yong, L., Xu, X., Elzur, U., Garg, P., and D. Melman, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-00 (work in progress), May 2015.

[I-D.ietf-sfc-architecture] Halpern, J. and C. Pignataro, "Service Function Chaining (SFC) Architecture", draft-ietf-sfc-architecture-11 (work in progress), July 2015.

[I-D.ietf-sfc-nsh] Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-01 (work in progress), July 2015.

[I-D.penno-sfc-trace] Penno, R., Quinn, P., Pignataro, C., and D. Zhou, "Services Function Chaining Traceroute", draft-penno-sfc-trace-02 (work in progress), March 2015.

[I-D.penno-sfc-yang] Penno, R., Quinn, P., Zhou, D., and J. Li, "Yang Data Model for Service Function Chaining", draft-penno-sfc-yang-13 (work in progress), March 2015.

[RSPYang] Opendaylight, , "Rendered Service Path Yang Model", February 2011, <https://github.com/opendaylight/sfc/blob/master/sfc-model/src/main/yang/rendered-service-path.yang>.

[SymmetricPaths]
IETF, , "Symmetric Paths", February 2011,
<https://tools.ietf.org/html/draft-ietf-sfc-architecture-11#section-2.2>.

Authors' Addresses

Reinaldo Penno
Cisco Systems
170 West Tasman Dr
San Jose CA
USA

Email: repenno@cisco.com

Carlos Pignataro
Cisco Systems
170 West Tasman Dr
San Jose CA
USA

Email: cpignata@cisco.com

Chui-Tin Yen
Cisco Systems
170 West Tasman Dr
San Jose CA
USA

Email: tin@cisco.com

Eric Wang
Cisco Systems
170 West Tasman Dr
San Jose CA
USA

Email: ejwang@cisco.com

Kent Leung
Cisco Systems
170 West Tasman Dr
San Jose CA
USA

Email: kleung@cisco.com