SRv6 SPAN
draft-li-spring-srv6-span-00

Abstract

   As an important means for operation and maintenance (O&M), mirroring
   is the most direct and comprehensive technology for capturing data
   streams and forwarding information. Compared with other
   visualization technologies, it can not only obtain the content of an
   entire packet, but also add forwarding information of a network
   device to a mirror packet and send the packet to a remote analysis
   server.

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 .

Copyright Notice

Table of Contents

1. Introduction

   The mirroring technology is also required in network O&M and fault
   locating. Networking architectures and network scales vary with
   scenarios, and accordingly requirements for mirroring technologies
   are different. For example, port mirroring can meet mirroring
   requirements for some small and medium-sized networks. In terms of
   network architecture, on the physical link, a mirroring-enabled
   switch is directly connected to a server used to analyze mirror
   packets. Although it is simple to use the mirroring technology in
   this case, the deployment of the mirror server is greatly limited.

   With the expansion of the network scale, especially in a super-large
   data center, the overlay tunneling technology is also used for
   deploying the mirror server, so as to remove the limitation that the
   mirror server needs to be directly connected on the physical link
   during networking.

The conventional mirroring technology is designed and implemented based on the IPv4 network, and data stream-based remote mirroring is developed based on local port mirroring. Combined with the networking in the data center, overlay and underlay networks are connected using the GRE tunneling technology. As such, the analysis server of mirror data packets can be flexibly deployed, and requires only route reachability instead of direct connection on the physical link.

In an SRv6 network, to simplify the network architecture and ensure stable running, protocol types are reduced as much as possible. If GRE is selected as the basic forwarding protocol of the mirroring technology, more restrictions will be imposed on the application scope of the mirroring technology. In addition, SRv6 is capable of connecting overlay and underlay networks, and does not need any additional forwarding protocol.

To better use the mirroring technology in the SRv6 network, a native mirroring function needs to be designed based on the features of the SRv6 network. Moreover, the formats of the conventional mirror protocol packets and the existing device capabilities should be reused to the maximum extent, so that the software system of the existing mirror analysis server is compatible, thereby avoiding redeveloping the software of the mirror server due to the SRv6-based mirroring technology. This helps reduce the difficulty in implementing the SRv6 network.

[2]. Purposes for Proposing the SRv6 SPAN Technology

Based on the IPv4 network technology, local port mirroring and data stream-based remote mirroring are developed. In the networking of the data center, the analysis server for mirror data packets should have route reachability during deployment. Therefore, the conventional far-end mirroring technology is based on GRE. There are two reasons for using GRE: One is that packets encapsulated by GRE support route-based forwarding, and only route reachability is required for the deployment of the mirror analysis server in the data center. Second, the large-scale construction of data centers coincided with the introduction of the GRE technical standards in 2016 when Microsoft deployed GRE on a large scale in its data center.

Currently, two changes have taken place in the forwarding protocol technology of the data center: One is that VXLAN, as an overlay technology, exceeds GRE and is deployed in more data centers; the other is that the IPv6 development has reached a critical moment, and the SRv6 source routing technology based on the IPv6 data plane is expected to be widely deployed in the future. Therefore, the GRE-based mirroring technology deviates from the very-simple design

principle in a data center network architecture, and it is difficult
to use the GRE-based mirroring technology in a data center where
VXLAN has been deployed. Otherwise, the data plane and the control
plane will become more complex. In view of the above, this document
aims to design an SRv6 SPAN technology. On the one hand, it is born
to support the SRv6 network. SRv6 is capable of connecting overlay
and underlay networks and does not need IPv6 GRE, simplifying the
network architecture and protocols. On the other hand, the formats
of the conventional mirror protocol packets are reused to the maximum
extent, so that the software system of the existing mirror analysis
server is compatible, thereby avoiding redeveloping the software of
the mirror server due to the SRv6 SPAN technology. This helps reduce
the difficulty in implementing the SRv6 network.

2.1. Design and Implementation of the SRv6 SPAN Technology

This document aims to design an SRv6 SPAN technology, so as to: 1.
simplify the network architecture where the mirroring technology is
used and deployed; 2. be compatible with the packet formats of
conventional mirrors; 3. enhance the mirroring technology to meet new
requirements.

2.1.1. SRv6 SPAN Technology and Networking

SRv6 is used to connect overlay and underlay networks of mirror data
packets, without needing IPv6 GRE, thereby simplifying the network
architecture and protocols. However, the standard forwarding plane
of SRv6 may also be divided into two types: a standard SRv6 packet
carrying the SRH and a SRv6-BE packet carrying the SRH.

The SRv6 SPAN technology is designed to support SRv6-BE. This is
because IPv6 networks have been upgraded on a large scale, but the
SRv6 network is still under development. Different from standard
SRv6, SRv6-BE does not need to add the SRH but is directly borne on
the IPv6 tunnel, so that the SRv6/IPv6 network can use the mirroring
technology in this document to the maximum extent. Therefore, the
SRv6-BE mirroring technology only needs to support IPv6 forwarding,
which allows remote deployment of analysis servers for mirror
packets, while requires no SRv6 SRH processing capability. This
reduces the difficulty in deploying the SRv6 SPAN technology.

In addition, the SRv6 SPAN technology in this document further
supports the SRv6 network that carries the SRH. A network device
with SRv6 SRH processing capability can implement precise path
control by using a mirror packet encapsulated based on SRv6 SRH, and
can capture other forwarding information when the mirroring
technology is used.

   Therefore, the SRv6 SPAN technology is not only compatible with a
   device that supports only an IPv6 network, but also matches with a
   network device with SRv6 SRH processing capability. It can remotely
   deploy an analysis server used for mirror packets, implement path
   control, and enable extensible forwarding information capturing.

2.1.2. SRv6 SPAN Technology and Packet Formats

   The protocol packet formats of the SRv6 SPAN technology are
   compatible with the formats of the conventional ERSPAN protocol
   packet as far as possible. The formats of the conventional mirror
   protocol packets are reused to the maximum extent, so that the
   software system of the existing mirror analysis server is compatible,
   thereby avoiding redeveloping the software of the mirror server.

```
     0                   1                   2                   3
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | Next Hdr=144  |  Hdr Ext Len  | Routing Type  | Segments Left |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Last Entry   |     Flags     |               Tag             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |            Segment List[0] (128-bit IPv6 address)             |
    |                                                               |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |                                                               |
                                  ...
    |                                                               |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |            Segment List[n] (128-bit IPv6 address)             |
    |                                                               |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |            SRv6 SPAN Header                                    |
    |                                                               |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |             Origin  Packet                                    |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                       SRv6 SPAN Packet Format

Based on SRv6 packet formats, Next Header 144 is used to identify
SRv6 SPAN. In addition, the SRv6 SPAN Header format has been defined
in the first version of this document, and includes a 4-octet
sequence number and 12-octet portion forwarding information.

```
     0                   1                   2                   3
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |       Sequence Number (Increments per packet per session )    |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Ver  |           VLAN        | COS |BSO|T|     Session ID     |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          Timestamp                            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |           SGT            |P|    FT   |   Hw ID  |D|Gra|O|
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
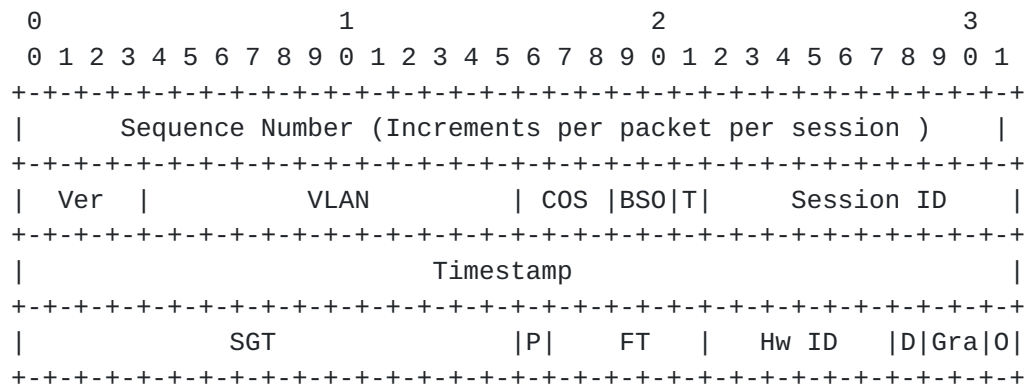
                      SRv6 SPAN Header Format

The various fields of the above SRv6 SPAN header are described in
this table:

Field	Length (bits)	Definition
Sequence number	32	For SRv6 SPAN packets sequence number，increased per packet per session，for detecting loss of mirror packet by analysis server.
Ver	4	SRv6 SPAN Encapsulation version. For SRv6 SPAN packets it is set to 0x6.
VLAN	12	VLAN of the frame monitored by an SRv6 SPAN source session: for ingress monitor this will be the original source VLAN whereas for egress monitor this will be the destination VLAN.
COS	3	Class of Service of the monitored frame. Ingress or egress CoS value is to be used depending on the monitor type/direction.
T	1	This bit indicates that the frame copy encapsulated in the SRv6 SPAN packet has been truncated. This occurs if the SRv6 SPAN encapsulated frame exceeds the configured MTU and hence has to be truncated.

   Session ID   10   Identification associated with each SRv6 SPAN
   (ERSPAN ID)   session. Must be unique between the source
                               and the receiver(s).

   BSO (Bad/Short/   2   A 2-bit value indicating the integrity of
   Oversized)   the payload carried by SRv6 SPAN:
                               00 --> Good frame with no error, or
                                      unknown integrity
                               11 --> Payload is a Bad Frame with CRC or
                                      Alignment Error
                               01 --> Payload is a Short Frame
                               10 --> Payload is an Oversized Frame

   Timestamp   32   The timestamp value needs to be derived
                               from a hardware clock which is
                               synchronized to the system-clock. This 32-
                               bit field should support at least a
                               timestamp granularity of 100 microseconds
                               (see the Timestamp Granularity field).

   SGT   16   Security Group Tag of the monitored frame.

   P   1   This bit indicates that the SRv6 SPAN payload
                               is an Ethernet protocol frame .

   FT (Frame Type)   5   This field can be used to reconstruct the
                               original frame's encapsulation if it is
                               supported by the receiver.
                               This field may also be used by SRv6 SPAN
                               engines to indicate that the mirrored
                               frame's L2 encapsulation header (or a
                               portion of it) was skipped and not
                               included in the SRv6 SPAN packet.
                               00000 --> Ethernet frame (802.3 frame)
                               00010 --> IP Packet
                               Other values --> Reserved for future use

   Hw (Hardware) ID 6   Unique identifier of an SRv6 SPAN engine
                               within a system.

   D (Direction)   1   Indicates whether the original frame was
                               SRv6 SPAN'ed in ingress or in egress.
                               Ingress (0) or Egress (1).

   Gra (Timestamp
   Granularity)   2   Time unit to be supported for time-
                               stamping:
                               00b --> granularity = 100 microseconds

```
                              01b --> granularity = 100 nanoseconds
                              10b --> granularity = IEEE 1588
                              TimeRepresentation format (see definition
                              below; with nanoseconds portion stored in
                              the Timestamp field and seconds portion
                              stored in the SRv6 SPAN platform-dependent
                              sub-header)

                              struct TimeRepresentation
                              {
                                 UInteger32 seconds;
                                 UInteger32 nanoseconds;
                              };
                              11b --> user configurable time unit
                              (platform dependent, for example specific
                              to an isolated non-synchronized system
                              with very high local accuracy)

     O (Optional
     Sub-header)       1      The O flag indicates whether or not the
                              optional platform-specific sub-header is
                              present. If it's present, the next octet
                              indicates the platform specific format
                              used (Platf ID). The SRv6 SPAN payload
                              starts after the O flag when O == 0b
                              or after 8 octets when O == 1b.

                      SRv6 SPAN Header Description

                                +---+
```

2.2. Future Considerations and Enhancements of the SRv6 SPAN Technology

In the first version of this document, the goal is to solve the
problem that the conventional mirroring technology is not compatible
with the existing SRv6 network.

In a later version of this document, the SRv6 SPAN technology will be
enhanced in three aspects:

* First, different requirements for mirroring capabilities in
multiple scenarios are met.

* Second, the capability of controlling a forwarding path and service
quality of a mirror data stream are enhanced;.

* Third, the capability of capturing required information is
enhanced.

SRv6 technology will be further developed, and enhancement points of
the SRv6 SPAN technology will be detailed in subsequent documents.
It is expected that the mirroring technology can be used as the
essential means for SRv6 network O&M and fault locating, so as to
facilitate the development of the SRv6 network.

3. Conclusion

The implementation of the SRv6 network mirroring function through
SRv6 SPAN technology is not only conducive to unifying the forwarding
data plane of the SRv6 network, avoiding the additional introduction
of GRE and other tunneling protocols, but also fully considering the
compatibility with the traditional mirroring message format in the
definition of SRv6 SPAN Header, Reduce the repeated development of
image analysis software to promote the development of SRv6 networks
and the deployment of SRv6 SPAN.

4. Security Considerations

TBD.

5. IANA Considerations

TBD.

Authors' Addresses

   Zhiqiang Li
   China Mobile
   Beijing 100053
   China

   Email: lizhiqiangyjy@chinamobile.com

   Tao Sun
   China Mobile
   Beijing 100053
   China

   Email: suntao@chinamobile.com

   Wei Cheng
   Centec Networks
   Suzhou 215000
   China

   Email: chengw@centecnetwors.com

   Junjie Wang
   Centec Networks
   Suzhou 21500
   China

   Email: wangjj@centecnetwors.com