The analysis of SRv6 and potential improvement
draft-yan-rtgwg-srv6-constrain-analysis-00

Abstract
This document analyzes the constrains of Segment Routing (SR), especially in the aspect of the header consumption and multicast of SRv6. Some potential methods have been proposed to improve the performance and enable more functions.

Status of this Memo
This Internet-Draft is submitted to IETF 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/1id-abstracts.html

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

Copyright and License Notice
Copyright (c) 2018 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
Introduction

Segment Routing (SR) leverages the source routing paradigm. It allows a headend node to steer a packet flow along any path for specific objectives like Traffic Engineering (TE). A node steers a packet through an SR Policy instantiated as an ordered list of instructions. These instructions are stack-structural and Segment Routing can be directly instantiated on the IPv6 data plane through the use of the Segment Routing Header defined in [I-D.ietf-6man-segment-routing-header]. SRv6 refers to this SR instantiation on the IPv6 data plane. The current SRv6’s design is good however there are still some constrains in SID consumption and functional defects.

In this document, we analyze the constrains of SRv6’s design and try to propose the potential improvement methods.

Terminology

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].

Constrain Analysis

3.1 Segment Consumption

The SID of SRv6 is a LOC:FUNCT tuple where FUNCT is a locally defined label. If a router is required to forward the packet to a specific neighbor, or perform a specific function, a corresponding label/SID MUST be put into the header. It means that N IPv6 addresses are needed in the header if the packet is required to pass through N routers.

As shown in Figure 1, PE1 wants to transmit a flow to PE2. Each router in this path is required to execute the function of Rate Limit (RL). In this example, the total cost before the inner IP Packet will be 158 Bytes. This causes two problems. The first one is the low payload proportion. The average packet size on the Internet is around 500 bytes and this design will occupy near 30%. The second one is the low processing efficiency. The current network processor reads normally less than 100 bytes at one time. If the header is too long, it needs more time to process. Conversely, if the header can be compressed shorter, the processing efficiency will be improved a lot.
In this example, theoretically there are four redundant RL instructions. The potential improvements may come from two aspects. Firstly, the coupling of locator and function makes the segments cannot be reused. If the locator and function can be de-coupled by some methods, it potentially decreases the size of whole packet header especially when the number of routers is large and most of them execute the same function when forwarding the packet. Secondly, a simple path label may be employed instead of stacking multiple locators. A shorter path label can decrease the length of instruction list.

3.2 Multicast

SR-MPLS supports multicast but SRv6 can hardly do the same. As shown in Figure 2, CE1, CE2, CE3, CE4 and CE5, construct a Blue VPN. CE1 wants to send a broadcast frame to all the other CEs. Because of the localized semantics, different routers use different SIDs for the same instruction / metadata. Therefore, the multicast packet MUST be replicated at PE1 instead of the P-routers (P).

In other words, in current SRv6 scheme, the multicast packet MUST be replicated at the ingress PE because egress PE uses different SIDs for the same VPN, since the locator is contained in SIDs. This is not an unsolvable problems. If the locator and function can be de-coupled meanwhile all the P-routers perform same operation according to a uniform instruction suite, all the multicast packets will be same in...
the egress router so that the packet could be replicated at any P-routers.

```
 M2  M3
+-----------------+ +-----------------+  --M2-->
| IPv6 Header    | | IPv6 Header    |  +---|PE2+--+CE2| Blue VPN
| SID=C2::20    | | SID=C3::30    |  +---+  +---+ C2::20 for
+-----------------+ +-----------------+  +---|PE3++CE3| Blue VPN
| MAC PK     | | MAC PK     |  +---|PE4+--+CE4| Blue VPN
|  --M3-->
+-----------------+ +-----------------+  |   +---|PE5+--+CE5| Blue VPN
| IPv6 Header    | | IPv6 Header    |  +---+  +---+ C3::20 for
+-----------------+ +-----------------+  +---+  +---+ C3::20 for
```

Figure 2, Multicast with SRv6

### 4 Potential Improvement

This section presents potential solutions to address the constrains discussed above.

The core idea of the solutions is to de-couple instruction and locator carried in the data packet, by adopting globalized instructions. Globalized instruction is an instruction code that could be recognized by every node in a domain. Moreover, the packet processing logic associated with the instruction code SHOULD be same for all nodes in the domain. In this way, the packet header overhead could be significantly compressed, and multicast could be well supported.

#### 4.1 Decrease the Consumption

In particular, the example topology in Figure 1 could be used to
illustrate this idea. To limit packet rate for a specific flow, the packets sent from PE1 to PE2 SHOULD carry two globalized instructions. One has the semantic of "shortest-path forwarding the packet according to PE2's address", and the other has the semantic of "limiting packet rate based on flow identifier and the maximal packet rate". Each node along the forwarding path performs these two globalized instructions accordingly thus achieving the targeted network function (i.e., packet rate limitation) with minimal overhead in the packet header.

![Function diagram](image)

**Figure 3, Potential Solution for Speed Limitation**

### 4.2 Globalize Semantics

The example topology in Figure 2 could be used to illustrate how this idea benefits multicast scenarios. As shown in Figure 4, after receiving a broadcast Ethernet frame from CE1, the ingress node (i.e., PE1) only need to encapsulate the frame into a single packet, and the packet SHOULD carry two globalized instructions. One has the semantic of "forwarding and replicating (if needed) the packet according to the multicast address of group PE2-PE5", and the other has the semantic of "if there is no next-hop, striping the encapsulated instructions, looking up the VRF of blue VPN and forwarding the packet accordingly". Each transit node in the network forwards and replicates (if needed) the packet based on the multicast address, and the egress nodes perform corresponding VPN actions. Therefore, globalized instructions enable packet replication in transit nodes, thus eliminating bandwidth wasting.
Figure 4, Potential Solution for Multicast VPN

Note that the metadata (e.g., unicast and multicast IP addresses, flow identifier, maximal packet rate, and VPN identifier) associated with the globalized instructions should also be carried in the packets, and there should be an approach to efficiently organize those instructions and metadata into a reasonable packet format. Such an approach will be specified in separated documents.

5 Security Considerations

6 IANA Considerations

7 References

7.1 Normative References

7.2 Informative References

[SR-MPLS-MCAST] Dave Allan, etc., "A Framework for Computed Multicast Applied to SR-MPLS", Work in Progress, draft-allan-pim-sr-mpls-multicast-framework-00, June 1, 2018


Authors’ Addresses

Shen Yan
Huawei Technologies Co. Ltd
No. 156, Beiqing Rd,
Haidian Dist, Beijing,
China, 100095
EMail: yanshen@huawei.com

Zhe Chen
Huawei Technologies Co. Ltd
No. 156, Beiqing Rd,
Haidian Dist, Beijing,
China, 100095
EMail: chenzhe17@huawei.com