China Mobile Technical White Paper on G-SRv6

(2020)

China Mobile Research Institute

2020.09

Preface

IPv6 is a key fundamental technology to build new infrastructure in China. Acceleration of massive IPv6 deployment is one of the strategies to build the nation into a strong cyber-power. Obviously the network will accelerate in evolution into the new era of IPv6. SRv6, by combining IPv6 and SR (Segment Routing), enables to simplify network protocols, open network capabilities and support network programmability. Thus it could flexibly meet the requirements of emerging services such as 5G, cloud network collaboration and ubiquitous connectivity. SRv6 makes uses a 128-bit IPv6 address as the SR SID (Segment ID), inheriting the advantages of flexible connection, global routing, and programmable IPv6 addresses, as well as the advantages of SR source routing, network simplification, and path traceability. It is an important evolution direction of IPv6-based network, and it will become the key technology of the next generation IPv6 network.

However, there are also various challenges such as carrying efficiency, hardware compatibility, and smooth upgrade of existing deployed network, to be solved for successful SRv6 deployment. China Mobile comes up with innovative SRv6 head compression solution with industry partner and drives standardization of that in IETF, which would further speed up SRv6 deployment in China and open the curtain of IP network innovation in the era of cloud network.

This white paper aims to elaborate China Mobile's technical proposal, network deployment application and inter-operability practice planning for G-SRv6 header compression solution. It targets to provide reference and guidance for the industry during deployment of SRv6 head compression-related technologies, products and solutions.

Contents

1 Technical Background

Under the background of the convergence of cloud and network, flexible and agile network service capabilities will directly determine operator's competitiveness. Segment Routing (SR) is a kind of source routing technology. SRv6 is the application of SR technology in IPv6 network. The emergence of SRv6 is a huge innovation. It combines SDN technology to enable programmable networks, which provides innovative soil for basic network services and value-added network services in the cloud network era. SRv6 inherits all the benefits of SR-MPLS and makes significant improvements. By unifying the network forwarding surface to IPv6, the forwarding surface establishes a cross-domain forwarding path as long as the IPv6 routing is accessible and a dedicated MPLS forwarding surface is no longer required. By defining functional instruction information, the message forwarding process can be programmed from the perspective of macro network.

SRv6 uses 128-bit IPv6 address as Segment ID (SID). With its extremely simple and programmable characteristic, it realizes arbitrary access and arbitrary connection, which can flexibly meet various network service requirements and provide higher reliability and extensibility, and will become the core technology of the next generation IP network.

While SRv6 technologies and standards become mature, SRv6 has been widely supported by major OEMs, silicon vendors and open-source platform. And SRv6 gradually starts deployment in the network of domestic and international operators.

Despite of lots of great advantages of SRv6 discussed as above, there are also fatal disadvantages. The key one comes from excessive SRH header overhead, especially when the number of SID increases. The drawbacks of long SRv6 header could be analyzed from 3 angles:

- Low bandwidth utilization of network link by big overhead of SRH headers.
  In the case of a 256byte packet with 8 SIDs, the bandwidth utilization ratio is only about 60%;

- Packet size beyond MTU if big SRH headers are added, which may result in packet fragmentation or packet loss, and thus a sharp decrease in transmission performance;
- Tight requirements on chipset to support existing SRv6 SRH processing. Existing deployed network nodes and new equipment nodes would have challenges in supporting in-depth SRH manipulations.

These problems also lead to the inability of operators to smoothly upgrade SRv6 on existing networks. To solve this block issue, there are multiple proposals developed in the industry to compress 128-bit SID into smaller one. Although these proposals achieve the target of SRv6 compression, all of them have some limitations, more or less:

- Incompatible with classic SRv6, which blocks the deployment.

To support evolution and smooth upgrade of existing network, the compression proposal needs to support co-existence with standard SRv6 compliant network nodes. In addition, it needs to be compatible with standard SRv6 SRH. Otherwise any new SRv6 header introduced by the compression proposal may result in the risk of the packets being discarded by standard SRv6 compliant network nodes.

- Arbitrary variable compressed SRv6 length, which increases the burden of implementation.

The compression proposal needs to consider network scalability and byte alignment. Arbitrary variable SRv6 compressed proposals increase the design complexity, and there is also doubt whether it is practicable.

- Waste of IPv6 address space due to no support of existing IP address plan.

Current SRv6 deployment is mostly based on the existing plan of public network address. If the address plan of compression proposals is incompatible with the current one, it will result in failure of deployment or waste of public network addresses.

China Mobile together with the industry partner propose the SRv6 head compression, which is fully compatible with standard SRv6. It makes use of fixed SID compression size, which is compatible with current network address plan and support smooth upgrade of the existing network. It is the leading solution for operator networks. Currently China Mobile is promoting production of this proposal in the industry while China Mobile is working with industry to promote standardization of this proposal in IETF. With all of above efforts, China Mobile hopes to accelerate the deployment of SRv6 in China.

2 Theory of Operations

2.1 Concept of SRv6 Head Compression

The encapsulation format of the standard classic SRv6 header, SRH, is listed as below:

Fig 2-1: Standard SRv6 SRH package

Next Header	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)			
Optional Type Length Value objects(variable)			

A SRv6 packet consists of: 40 Bytes (IPv6 header) + 8 Bytes (SRH fixed header) + 16*N Bytes (Segment List)。As the number of SIDs increases, the SRv6 header overhead increases.

According to the definition in RFC8402 and RFC8754, a SRv6 SID is an IPv6 address of 128 bits. In a SRv6 programmable network, it usually consists of the following three parts.

Fig 2-2: Support network programmable standard SRv6 SID format schematic

Locator	Function	Arguments

- Locator is the identifier assigned to a network node in the network topology to direct the forwarding of packets to this node.
- Function is used to indicate the forwarding action to be performed by the network node.
- Arguments (Args) is optional execution parameters that may contain flow, service, or any other relevant information. Arguments (Args).

The Locator can be further divided into B:N, where B identifies the SRv6 SID Block, which is normally assigned to a subnet by the operator, and is usually denoted by Prefix. N is the identity of the distinguishing node in the subnet.

Fig 2-3: Support schematic of typical NETWORK programmable SRv6 Locator

B(Prefix)	N(Node ID)

From the above definition of the SRv6 header format, we can see that the largest proportion of the header is the Segment List. When the number of entries remains unchanged, compression of a SRv6 header actually means to compress a single SID.

Below are common concepts used in head compression techniques, including:

- A Full SID: 128-bit length SID；
- A Compressed SID: Shortened SID；
- A SID compression algorithm: A specific method to compress a full SID into a compressed SID；
- Control plane protocol extension proposal, including the propagation of SRv6 compression capability of each network node, the propagation of compressed SIDs, etc；
- Data plane protocol extension proposal, including identification of the full SID and compressed SID, Segment List and rules for DA (Destination Address) read and write, etc；

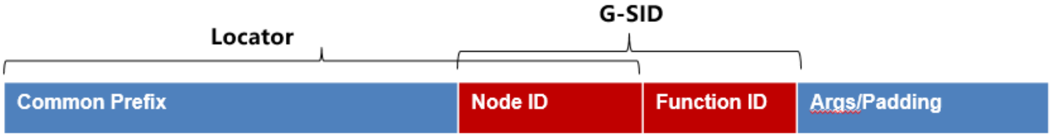
2.2 SRv6 Head Compression Principles

Actually, many SIDs have a common prefix in a SRv6 network. The common prefix part of SIDs in the SID list is redundant. Therefore, in the SID list, if Common Prefix and other redundant parts of the SRv6 SIDs are removed, while keeping the Node ID and Function ID for SID compression, the header overhead can be well reduced. This is also the basic compression principle for the G-SID (Generalized SID) in the SRv6 header compression proposal described in this white paper.

The G-SID is part of the compressed SRv6 SID, combined with Common Prefix and Argument/Padding to form the full SRv6 SID. G-SID supports both 32-bit and 16-bit lengths. In fact, from the perspective of hardware processing efficiency, reusability and scalability, 32-bit is more common for SID compression. This white paper will only focus on 32-bit compression.

A typical 32-bit SRv6 compressed G-SID format is defined as below, which consists of the Node ID and Function ID in the 128-bit SID. The 128-bit SID format is the standard classic SID, also known as SRv6 SID. The 32-bit SID is a G-SID, which is the variable part of the standard classic SID. The format is shown in the figure as below:

Fig 2-4: SRv6 SID format that support compression



According to principle of the SRv6 Locator, the transformation relationship between the standard classic SID and the compressed SID can be described in the following ways:

G-SID is composed of N (Node) : F (Function)

Network planning shall be considered first to choose a perfect SRv6 compression proposal. The total number of nodes and functions in the planning should not exceed the compression length. In reality, the bits allocated for Function ID and Node ID should be determined based on the comprehensive analysis of the compression length, the maximum number of functions and the maximum number of nodes.

In order to support compression, the control plane protocols need to be extended. Each network node should propagate and publish their capabilities in supporting SRH compression and the compressed SIDs to the whole network.

In order to support compression, the data plane needs to be extended as well. Because the Segment List in SRH can be inserted into G-SID, it is necessary to update how Segment List is replicated and how DA is replaced. We also need to add how to identify the boundary between the standard classic SID and G-SID.

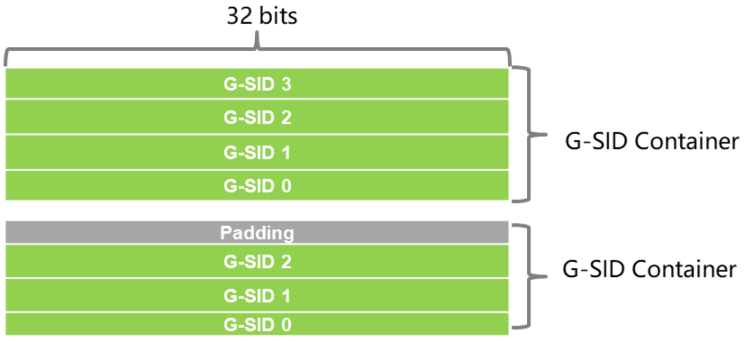
In order to facilitate the understanding how G-SID and standard classic SID could co-exist with 128-bit alignment, the concept of G-SID Container is introduced.

A G-SID Container is exactly of 128-bit length and it could contain：

- A standard classic SRv6 SID
- Or multiple G-SID, such as 1-4 of 32-bit G-SID or 1-8 of 16-bit G-SID. Zero padding if less than 128-bit.

Taking 32-bit G-SID as an example, the possible format of a G-SID container is shown in the figure below.

Fig 2-5: 32-bit G-SID Container format



As many types of G-SID Container can be encoded in SRv6 SRH, we call this SRH as Generalized SRH (G-SRH). The solution to support common multiple segments encoded in SRH is a kind of upgrade from current SRv6, which we call

Generalized SRv6 (G-SRv6).

3 Key Technologies

This section will introduce the overall technical solution of G-SRv6, including G-SRv6 data plane solution and control plane solution. The data plane mainly describes the compressed G-SID data encapsulation method. The control plane section will address the extensions of IS-IS, BGP-LS and BGP SRv6 Policy to support G-SRv6.

G-SRv6 does not change the native SRH format and semantics in terms of data encapsulation. It is naturally compatible with native SRv6 and can be deployed in combination with native SRv6. The compressed SID is recommended to use a fixed length of 32 bits, and the existing chips can support SR MPLS, which greatly reduces the difficulty of implementation. Since the principle of G-SRv6 compression is to extract the common prefix, there is no restriction on the length of the prefix, which can flexibly adapt to the existing network address planning and routing planning, and will not waste the address space allocated by the operator.

3.1 Data Plane Solution for G-SRv6

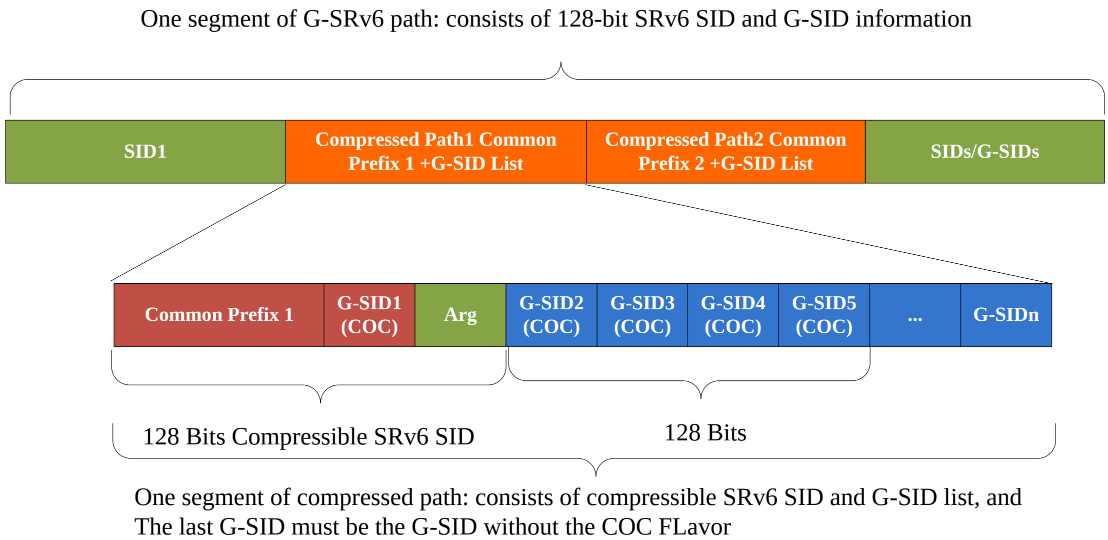
Native SRv6 data encapsulation is mainly reflected in the SRH extension header of IPv6 packets. G-SRH format is consistent with that of SRH [RFC8754], and there are no changes for its format and field semantics. G-SRH supports mixed encoding of 128-bit SRv6 SID and 32-bit G-SID. The format is shown in the figure below. When the 128-bit SRv6 SID and 32-bit G-SID co-exists in G-SRH, an encoding example is shown as below.

Fig 3-1: G-SRH format

Next Header	Hdr Ext Len	Routing Type	Segments Left
Last Entry	Flags	Tag	
G-SID Container[0](128 bits value)			
. . .			
G-SID Container[n](128 bits value)			
Optional Type Length Value objects(variable)			

A G-SRv6 path may consist of SRv6 path segments and SRv6 compression path segments. A SRv6 path segment is encoded by the SRv6 SID. The SRv6 compression path segment starts with a 128-bit SRv6 SID that supports compression and is composed of multiple following G-SIDs. It must end with 128-bit as the boundary. The next SID at the end of the G-SID list is a 128-bit SRv6 SID, which may be a normal SRv6 SID or a G-SRv6 SID with compression supports.

Fig 3-2: The G-SRV6 hybrid SID example



To identify the start point and end point of a SRv6 compression path in the SID list, in other words, the boundary between the 128-bit SID and 32-bit SID, additional flavors should be defined, and the SID for the corresponding flavor shall be published via control plane as well.

Table 3-1: New Flavor Table

SID Flavor types	Functional Description
COC	Identify the 32-bit G-SID to be updated to the destination address
COC16	Identify the 16-bit G-SID to be updated to the destination address

Current update action for SRv6 Destination Address (DA) is to update the Destination Address with the next 128-bit SID, so standard classic SRv6 SID carries the Flavor to update 128-bit SID (if it is not the last SID), and thus there is no need to introduce additional Flavor definitions.

When the SID in the DA carries the COC Flavor, it indicates that the destination address shall be updated with the next 32-bit G-SID. Therefore, when the SID in the DA does not carry the COC Flavor SID, it means that the SL--, and update the destination address with the next 128-bit SID (when SL>0).

The encoding rules for compressed SRv6 in SID List are as follows:

- The start of a compressed SRv6 path is indicated by a 128-bit SID of COC Flavor, which contains the G-SID of the COC Flavor. This SID carries the full SID information, including Common Prefix (CP) and other information, and it can be used to restore the full next SID with the subsequent G-SID.
- The G-SIDs in the middle of a compression path are the ones carrying COC, indicating that the next one is a 32-bit G-SID.
- The last G-SID of a compression path must be a 32-bit G-SID without COC Flavor, indicating the end of the compression path. The SRv6 SID combined in DA will be processed according to the 128-bit SID processing rule.

The rationale for the last G-SID as Non-COC Flavor is that, after it is used to update the destination address, the SID in DA has no COC flavor, and it will be processed by the node as normal SRv6 SID. The destination address will be updated by the next 128 bit SID. This can be used to mark the end of a compression path, and to switch from 32-bit G-SID to 128-bit SID processing.

This design works well for two scenarios:

- Hybrid scenario, where the compression path ends, and switches from 32bit to 128bit.
- Pure compression scenario, where the last G-SID is without COC flavor, and the processing of the last hop is consistent with current SRv6 with full compatibility.

In addition, in order to locate the next G-SID in the compressed path, SI (Generalized SID Index) needs to be added to locate its position in the G-SID Container. SI is placed in the lowest bits of Arguments after G-SID in the destination address.

In this proposal, the SID length shall be 128-bit, and there is no padding. In 32-bit G-SID compression, SI is the last 2 bits of 128-bit SID. To facilitate hardware implementation and future proof, the last byte is used as the SI value in this proposal. It is recommended to use byte-aligned prefix for common prefix, such as 64.

When generating a compressed SID, some space needs to be reserved for SI, and its value is 0. During the forwarding process, the SI field in the destination address is meaningful only when the active SID is a compressed SID. At this time, the SL indicates the location of the G-SID Container in the G-SRH, and the SI indicates the location of the G-SID in the G-SID Container. According to the reverse ordering rules used in SID List (Segment List[0] is the last hop, Segment List[n] is the first hop), the order of G-SID is also reversed. For example, for SI=3, it locates to the highest 32bit value of the G-SID Container; when SI=0, the lowest 32bit value of the G-SID Container is located. A specific example is shown in Figure 3-4.

Fig 3-3: SI (G-SID Index) encapsulation format

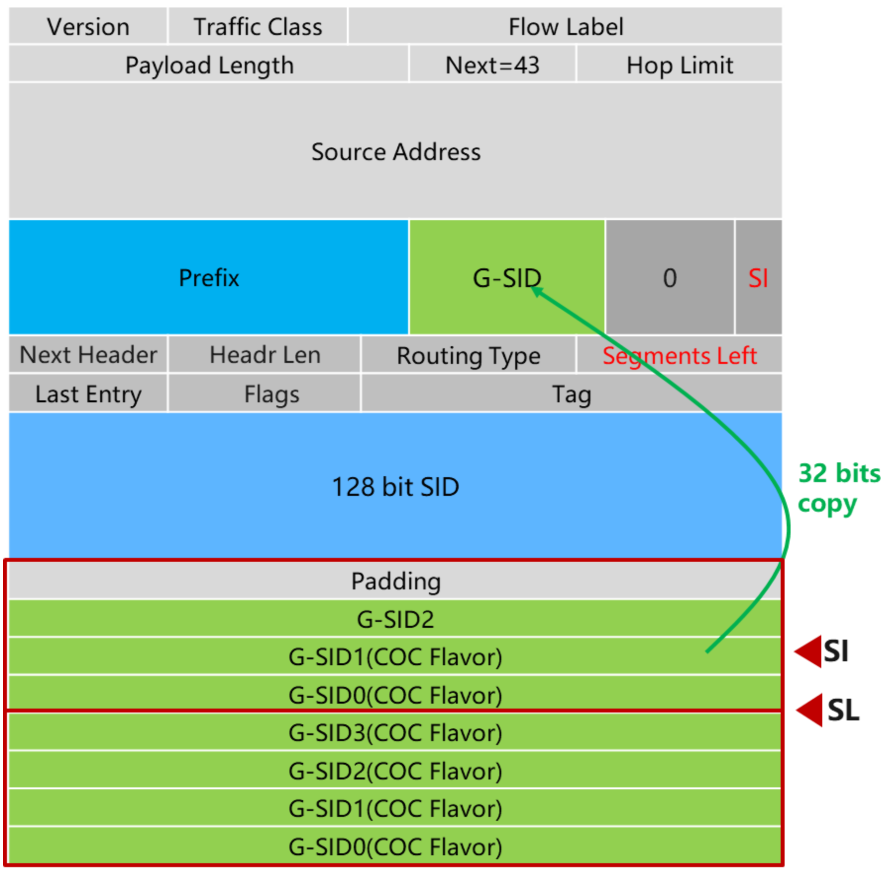
The basic idea of this scheme is that, to update the destination address with the next 32 bits G-SID through the COC Flavor SID instruction, and the G-SID position is identified by SL and SI.

Taking 32-bit G-SID as an example, when packets are received on the Endpoint node, the pseudo code for packet processing is as follows:

```
if IPv6 hits a COC Flavor SID    //32-bit SID process
    if DA.SI == 0:               //the first SID of the next G-SID Container
        SL--;
        DA.SA = 3;
    else                         // next G-SID in the current G-SID Container
        DA.SI--;
    DA[CP..CP + 31] = SRH[SL][DA.SI];
    Forward the packet based on new DA;
else                             //the normal SRv6 branch is not affected
    SRv6 processing
```

The SRv6 compression processing flow would only be triggered by the COC Flavor SID, and the operating data would also be limited to the COC Flavor SID, which has no impacts on the processing of the existing standard classic SID and SRH. The SI is after the G-SID in DA. Taking a pure compression path and 128-bit VPN SID as an example, the illustration diagram for encoding and G-SID update is shown below.

Fig 3-4: G-SRv6 encoding



3.2 Control Plane Solution for G-SRv6

The solution of G-SRv6 requests some extension of standard classic SRv6 control plane protocols to enable the use of compressed SID in G-SRv6. The native classic SRv6 already extends ISIS protocol [draft-ietf-lsr-isis-srv6-extensions-08] to support advertising SID attributes to all the SRv6 nodes in the network . In order to further advertise each node’s capability in compression and compressed G-SID information, G-SRv6 needs to further extend the IS-IS protocol. The classic SRv6 already extends the BGP-LS protocol [draft-ietf-idr-bgpls-srv6-ext-02] for SDN controller to collect SID attributes. For the similar reason to enable SDN controller to collect compression capability and compressed SID information, G-SRv6 needs to further extend the BGP-LS protocol as well. Once path computation done, the SDN controller uses BGP to advertises the set of SRv6 policies to head-end node, and the SID list carried in the newly defined SRv6 policy NLRI[draft-ietf-idr-segment-routing-te-policy-08]. In order to carry the compressed SID in SID list, G-SRv6 needs to extend the BGP TLVs as well.

3.2.1 IS-IS Control Plane

After the SRv6 SID space planning is done, the nodes in the network need to complete the SRv6 SID configuration as required, generate the local SID table according to the configuration, and then advertise the SID information to the network through IGP. Therefore, it is necessary to extend IGP protocols such as IS-IS.

Taking the IS-IS protocol as example:

(1) Extend SRv6 Capabilities sub-TLV to add C-flag to advertise the compression capability of the node.

Fig 3-5: SRv6 Capabilities sub-TLV

<table>
<tr><td>Type</td><td>Length</td><td></td><td>C</td><td>Flags</td></tr>
<tr><td colspan="5">Optional sub-subs-TLVs</td></tr>
</table>

C-flag is used to indicate the SRv6 compression capability of the node. If set, it means that the current node has SRv6 SID compression capability.

The SDN controller or node without the support of SRv6 SID compression should ignore the C-flag upon receipt of SRv6 Capabilities sub-TLV with extended C-flag.

The SRv6 compression capability of a node can be reported to the SDN controller through BGP-LS (see Section 3.2.2). When computing the SRv6 TE path, the SDN controller orchestrates the SID list according to the compression capability of the node. C-flag can also be used in node path computation.

If the headend does not support the SRv6 compression capability, the SDN controller should not advertise the SRv6 policies containing the G-SID to the headend. If the G-SID really needs to be advertised, the SRv6 policies can only contain the SID without COC Flavor, and can only be orchestrated as 128-bit.

(2) To extend SRv6 End SID sub-TLV, SRv6 End.X SID sub-TLV, and SRv6 LAN End.X SID sub-TLV to add C-flag to indicate if this SID supports compression or not. The position of C-flag is shown in the figures as below.

Fig 3-6: The SRv6 End SID sub-TLV

<table>
<tr><td colspan="2">Type</td><td>Length</td><td colspan="2"></td></tr>
<tr><td>Flags</td><td>C</td><td colspan="2">Endpoint Behavior</td><td></td></tr>
<tr><td colspan="5">SID (128 bits)...</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="2">Sub-sub-tlv-len</td><td colspan="3">Sub-sub-TLVs (variable)</td></tr>
</table>

Fig 3-7: End.X SID sun-TLV

<table>
<tr><td colspan="2">Type</td><td>Length</td><td colspan="2"></td></tr>
<tr><td>Flags</td><td>C</td><td>Routing Type</td><td>Segments Left</td><td></td></tr>
<tr><td colspan="3">Endpoint Behavior</td><td colspan="2"></td></tr>
<tr><td colspan="5">SID (128 bits)...</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="2">Sub-sub-tlv-len</td><td colspan="3">Sub-sub-TLVs (variable)</td></tr>
</table>

Fig 3-8: LAN End.X SID sub-TLV

<table>
<tr><td colspan="2">Type</td><td>Length</td><td colspan="2">System ID (6 octets)</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td>Flags</td><td>C</td><td>Algorithm</td><td>Weight</td><td></td></tr>
<tr><td colspan="3">Endpoint Behavior</td><td colspan="2"></td></tr>
<tr><td colspan="5">SID (128 bits)...</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="5">SID (cont...)</td></tr>
<tr><td colspan="2">Sub-sub-tlv-len</td><td colspan="3">Sub-sub-TLVs (variable)</td></tr>
</table>

If C-flag is set, the SRv6 SID Structure Sub-sub-TLV must be carried in the Sub-sub-TLVs . Wherein, the LB Length represents the length of Common Prefix. The length of G-SID is determined by LN Length (length of Node-ID) and Fun. Length (length of Function ID), which can be 32 or 16. If the length of Node-ID + Function ID is other values than 32 or 16, this SID is regarded as an illegal SID and must be discarded.

Fig 3-9: SRv6 SID Structure sub-TLV format

Type	Length		
LB Length	LN Length	Fun.Length	Arg.Length

(3) To extend the Code points in the Endpoint Behavior field of the above SID sub-TLVs to add Behavior code point introduced by COC/COC16 flavor upon advertising SID.

During IGP capability advertising process, , the node will advertise its compression capability by setting the extended C-flag of SRv6 Capabilities TLV.

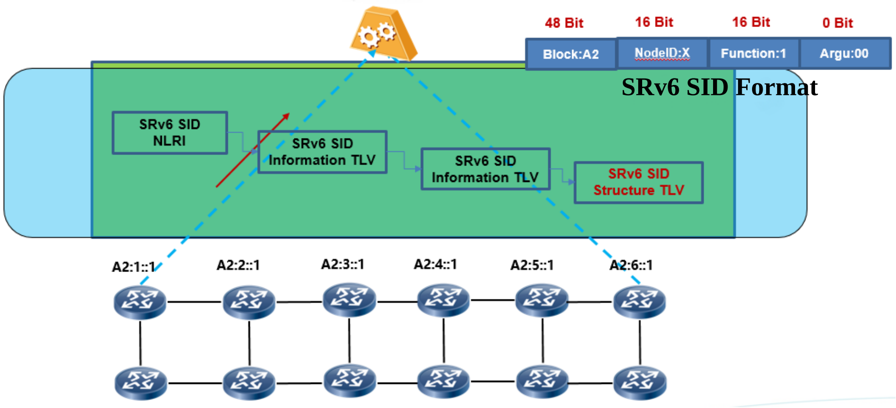
Then in order to support SRv6 compression, the corresponding SID needs to be instantiated and distributed to the network through IGP or reported to the controller. During the process of corresponding SID instantiation, the forwarding table entries will be generated for this SID on the node.

For specific SID advertisement, the C-flag in the TLV (for example, SRv6 End SID sub-TLV, SRv6 End.X SID sub-TLV, SRv6 LAN End.X SID sub-TLV) indicates if this SID supports compression or not. The G-SID compression could be either 32-bit or 16-bit. If the value of (LN Length + Fun. Length) is 32, it means this G-SID supports 32-bit compression. If the value of (LN Length + Fun. Length) is 16, it means this G-SID supports 16-bit compression.

3.2.2 BGP-LS Control Plane

The BGP-LS protocol can be used to advertise the SRv6 information pertaining to one node and reported the SRv6-related information to SDN controller. In this solution, the extension based on [draft-ietf-idr-bgpls-srv6-ext-02] to report the SID attribute of SRv6 header compression information.

Fig 3-10: SRv6 BGP-LS protocol Process



In order to compute TE paths, the SDN controller needs to know the topology information of the whole network as well as TE attributes and SRv6 attributes of the topology. BGP-LS provides the functionality for the SDN controller to collect the SRv6 capability of each node and instantiated SIDs of each node as well. To make G-SRv6 support SID compression, three types of TLVs also need to be extended.

(1) Extend SRv6 Capabilities sub-TLV to add C-flag to advertise the compression capability of the node.

C-flag is used to indicate the SRv6 compression capability of the node. If set, the node supports G-SRv6 compression

(2) Extend SRv6 End.X SID sub-TLV and SRv6 LAN End.X SID sub-TLV under SRv6 Link Attributes to add C-flag to indicate whether SID compression is supported.

Fig 3-11: SRv6 End.X TLV Format

<table>
<tr><td colspan="2">Type</td><td colspan="2">Length</td></tr>
<tr><td colspan="2">Endpoint Behavior</td><td>Flags</td><td>Algorithm</td></tr>
<tr><td>Weight</td><td>Reserved</td><td colspan="2">SID (16octets)</td></tr>
<tr><td colspan="4">SID (cont ...)</td></tr>
<tr><td colspan="4">SID (cont...)</td></tr>
<tr><td colspan="4">SID (cont...)</td></tr>
<tr><td colspan="2">SID (cont...)</td><td colspan="2">Sub-TLVs (variable)...</td></tr>
</table>

Fig 3-12: Flag Format

B	S	P	C	Rsvd

(3) Extend End Behavior sub-TLV of the Node SID to add C-flag to indicate the compression capability.

The BGP-LS information comes from IGP. The C-flag extension here is same as that of IGP extension.

Fig 3-13: End Behavior sub-TLV Format

<table>
<tr><td>Type</td><td colspan="3">Length</td></tr>
<tr><td>Endpoint Behavior</td><td>Flags</td><td>C</td><td>Algorithm</td></tr>
</table>

(4) Extend the Code points in the Endpoint Behavior field of the above SID sub-TLVs to add Behavior code point introduced by COC/COC16 flavor.

During BGP-LS connection setup stage, if a node supports SRv6 compression, the C-flag in SRv6 Capabilities sub-TLV needs to be set to 1. The node advertises the SID information including compression capability to the controller via BGP-LS. For the SID information encapsulated in advertising message, the C-flag in the SRv6 End.X SID sub-TLV, SRv6 LAN End.X SID sub-TLV and End behavior sub-TLV corresponding to the SIDs should be set. If C-flag is set to 1, the SID Structure TLV must be carried to describe the SID format.

3.2.3 BGP SRv6 Policy Control Plane

Typically, a SDN controller will define the set of SRv6 policies and advertise them to the headend using BGP SR policy [draft-ietf-idr-segment-routing-te-policy-08]. The BGP SR Policy needs to be extended to support SRv6 compression. BGP SR Policy is defined as below:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
  Tunnel Encaps Attribute (23)
    Tunnel Type: SR Policy
       Binding SID
       Preference
       Priority
       Policy Name
       Explicit NULL Label Policy (ENLP)
       Segment List
          Weight
          Segment
          Segment
                ...
```

The protocol needs to be partially extended to support compression.

There is no need the extension of SR Policy. Instead it could be obtained through the information sent by the BGP-LS on node's capability to support SRv6 compression.

If the headend node supports SRv6 compression, the SRv6 Policy containing the compressed SID can be propagated. Otherwise, only the common SRv6 Policy should be propagated. The common SRv6 Policy can contain the compressible SID in the format of 128-bit but does not carry COC Flavor.

The following SID Encoding sub-TLV under Segment List sub-TLV needs to be extended to indicate the specific position of G-SID in SID. The extended format is as follows. This extension is applicable to the scenarios like headend which request to decode SID information. For other nodes without the need to decode SID information, the controller can directly encode it into 128-bit SRv6 SID and advertise it to the headend.

Fig 3-14: SID Encoding sub-TLV Format

Type	Length	Flag	Reserved
Block Length	G-SID Length	Reserved	

- Type: Type of the TLV, 1 byte.
- Length: TLV length, not including type and length fields. Length: 1 byte, value: 6.
- Flag: 8-bit, to be defined in the future. It must be set to zero at the sending node and must be ignored when received.
- Reserved field: 8-bit reserved field. It must be set to zero at sending node and must be ignored when received.
- Block Length: 8-bit. If the SID is compressed to 32-bit, the value can be 0~94. In this solution, 8-bit are reserved for SI, so the maximum value of Block Length field is 88.
- G-SID Length: 8-bit, it represents the length of ccompressed SID. In current proposal only 32-bit G-SID is supported. If a node receives G-SID Length other than 128 of classical SID and 32 of G-SID, it should treat them as errors and the NLRI[I-D.ietf-idr-segment-routing-te-policy-09] must be cancelled.
    - Block Length = 0, G-SID Length = 128: Copy total 128-bit SID to the SID list.
    - Block Length = n, G-SID Length = 32: Copy 32-bit G-SID after Block to the SID list.
- Reserved field: 16-bit reserved field. It must be set to zero at the sending node and must be ignored when received.

In G-SRH encoding, if the Block length of the SID in SID Encoding sub-TLV is 0, and the G-SID length is 128, it indicates the G-SID is exactly copied from classic 128-bit SID. Otherwise, the G-SID is partial copied from classic 128-bit SID according to Block Length and G-SID length. All SIDs after this SID Encoding sub-TLV are encapsulated into G-SRH according to the G-SID position indicated by the current SID encoding sub-TLV till the appearance of the next SID encoding sub-TLV.

An example of the extended SID List showed as follows, the G-SID of Segment2 and Segment3 defined in SID Encoding sub-TLV1 field, and the SID Encoding sub-TLV2 field defines the G-SID of Segment4, Segment5 and Segment6.

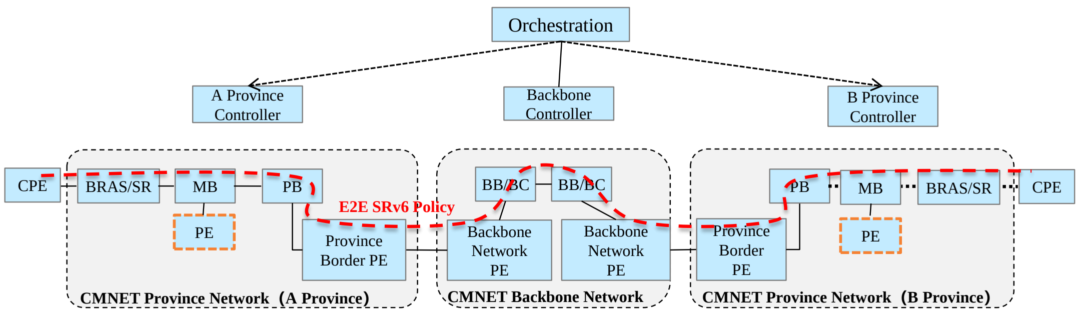
```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
    Tunnel Encaps Attribute (23)
      Tunnel Type: SR Policy
           Binding SID
           Preference
           Priority
           Policy Name
           Explicit NULL Label Policy (ENLP)
           Segment List
           Weight
           Segment1
           SID Encoding sub-TLV1
           Segment2
           Segment3
           SID Encoding sub-TLV2
           Segment4
           Segment5
           Segment6
            ...
```

4 Typical Network Applications

In the strict path TE scenario, the hop-by-hop SRv6 node needs to be specified by SID, so the longer the path is, the more the SIDs are. In scenarios that cross multiple SRv6 domains, there are more end-to-end SIDs. The SRv6 header compression technology can effectively reduce the SRv6 packet overhead and alleviate challenges on the equipment nodes. G-SRv6 has no new requirements on forwarding chips than SR MPLS. The forwarding and encapsulation capabilities of the current mainstream chips can support G-SRV6, so it can be upgraded smoothly in operator networks.

In the enterprise branch interconnection scenario shown in the figure below, the customer that order private line service could access the branch of province B from that of province A to create an end-to-end compressed SRv6 path. According to the deployment of the SRv6 header compression technology, the G-SRv6 deployment scenarios can be divided into pure compression scenarios and hybrid compression scenarios.

Fig 4-1: SRv6 bearing networking of China Mobile Enterprise Branch Interconnection



4.1 Pure Compression Scenario

In the pure compression scenario, all nodes in the compression domain support SRv6 header compression capabilities. The nodes may have the same common prefix or different common prefixes. We take the MP-BGP L3VPN over SRv6 TE (compression) service of the intra-provincial private network users as an example to describe the pure compression scenario.

- The private networks are in the same compression domain. The block is planned as 8000: A:B:C/64, the Node ID is 20 bits and the Function ID is 12 bits.
- As shown in Chapter 3, the nodes in the compression domain are configured

with the corresponding node compression attributes. The END and END.X SIDs with and without COC Flavor are allocated, which are advertised to other nodes and controllers through IGP and BGP-LS.

For example, PE1 allocates two types of END.X SIDs to an interface.

PE1 END.X 8000:A:B: C:1:1 COC Flavor

8000:A:B:C:1:2 None COC Flavor

- The controller allocates the SIDs as follows according to the configured compression Policy path: PE1->P1->P2->P3->P4->P5->PE2 and the coding principle in Section 3.1:

Table 4-1: SID Allocation in Pure Compression Scenario

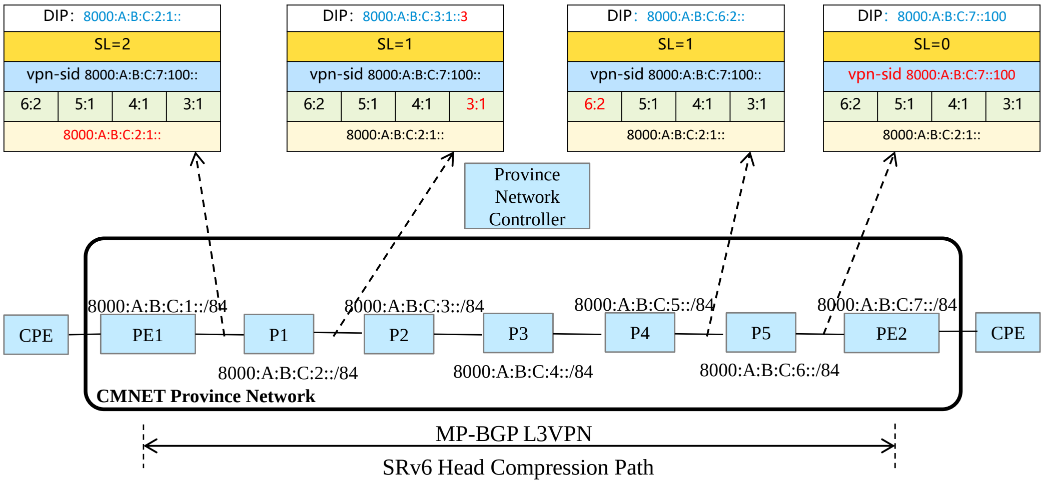
Node	SID type	SID(block 64bit)	SID length	SID attribute
PE1	END.X	8000:A:B:C:1:1::	128	COC
P1	END.X	8000:A:B:C:2:1::	32	COC
P2	END.X	8000:A:B:C:3:1::	32	COC
P3	END.X	8000:A:B:C:4:1::	32	COC
P4	END.X	8000:A:B:C:5:1::	32	COC
P5	END.X	8000:A:B:C:6:2::	32	None COC

The SIDs are delivered to the head node PE1 through BGP POLICY.

- The head node PE1 maps the VPN traffic to the Policy path, and constructs a complete G-SRH according to the Segment List SID information sent by the controller and the VPN SID distributed by the PE2. And then it encapsulates the IPV6 packet and sends it to the P1 and starts the forwarding process. The first SID in SRH 8000:A:B:C:1:1:: is the COC Flavor END.X published locally. According to the description of the forwarding pseudo-code in Section 3.1, the COC attribute SID forwarding operation is performed: SL--, DA.SI=3, point to the next 32bit G-SID 2:1, copy 2:1 to the DA relative position, and perform table lookup forwarding based on the new DA. In this case, SL=2 in SRH, and the destination address is 8000: A:B:C:2:1::3.

An example of the compressed SID of each node in G-SRH is as follows:

Fig 4-2: Forwarding Process in Pure Compression Scenario



- Node P1 receives the data packet with the destination address as 8000: A:B:C:2:1::3, which is the local COC Flavored END.X SID. According to the description of the forwarding pseudocode in Section 3.1, it performs COC Flavored SID forwarding process: SI-- directed to the next 32 bits G-SID 3:1. The compressed SIDs 3:1 will be copied to the related position of DA, and the packet will be forwarded after table lookup based on the new DA. At this time, SL=2, SI=2, the destination address is changed to 8000: A:B:C:3:1::2.
- Subsequently, nodes P2, P3 and P4 match the local COC Flavored END.X SIDs according to the destination addresses of the received packets. According to the description of the forwarding pseudocode in Section 3.1, they perform COC Flavor SID forwarding process: SI--. Based on SL and SI, and locate the corresponding G-SID and update DA accordingly, and then lookup the forwarding table and forward the packets accordingly. In particular, the node P4 performs the COC attribute SID forwarding operation. Since SI=0, it means that the boundary of the G-SID container is read. SL--, DA.SI=3 is executed, and the relocation corresponds to G-SID 6:2 Copy to the relative position of DA, the destination address becomes 8000:A:B:C:6:2::3.

- When the node P5 receives the packet with DA as 8000: A:B:C:6:2::3, which is the local END.X SID without COC Flavor. According to the SRv6 SID forwarding process: SL--, it will use the next 128 bits SID 8000: A:B:C:7:: 100 (VPN SID here) as the destination address.
- When the node PE2 receives the data packet with DA as 8000: A:B:C:7:: 100, which is VPN SID without COC Flavor. It will perform the standard SRv6 VPN SID processing.

4.2 Hybrid Compression Scenario

SRv6 header compression technology will be in a phased deployment. There may be some cases where some part of network support SRv6 header compression, but other parts of network only support standard classic SRv6. As a result, on an end-to-end SRv6 strict path across compressed SRv6 domain and standard classic SRv6 domain, some nodes support SRv6 header compression, while others do not. The SRv6 header compression technology needs to consider such hybrid scenarios and support coexistence of 32 bits compression G-SID and standard classic 128 bits SRv6 SID on the same SRv6 TE path. Now we use the MP-BGP L3VPN over SRv6 TE (compressed) service from an inter-provincial private network user as an example to describe the hybrid compression scenario.

- PE1, P1, P2 and P3 are nodes in compression domain 1, P4, P5 and P6 are nodes in standard classic SRv6 domain, and P7, P8, P9 and PE2 are nodes in compression domain 2. The block in compression domain 1 is planned as 8000: A:B:C/64, the Node ID is 20 bits, and the Function ID is 12 bits. The block in compression domain 2 is planned as 8000: A:B:E/64, the Node ID is 20 bits, and the Function ID is 12 bits. The block in standard SRv6 domain is planned as 8000: A:B:D/64.
- The nodes in the compression domain are configured with the corresponding node compression attributes. Two type of SIDs are allocated for the END and END.X SIDs with and without COC Flavor, which are advertised to other

nodes and controllers through IGP and BGP-LS. Standard classic SRv6 nodes are only allocated with standard SIDs without compression attribute.

- The controller allocates the SIDs as follows according to the configured compression Policy path:
  PE1->P1->P2->P3->P4->P5->P6->P7->P8->P9->PE2 and the coding principle in Section 3.1.

Tab 4-2: SID Allocation in Hybrid Compression Scenario

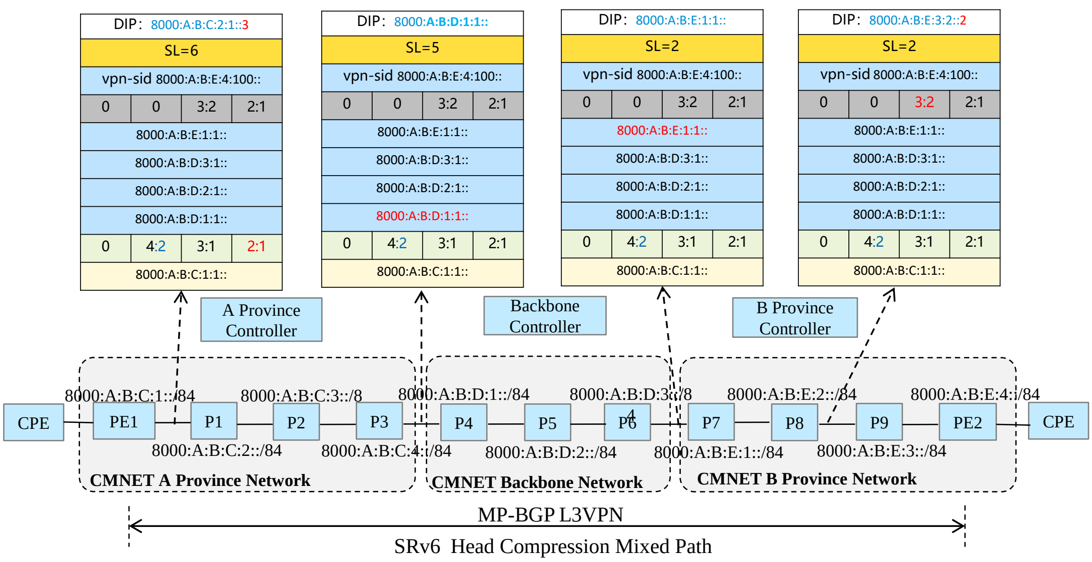
Node	SID type	SIDSID (block 64bit)	SID length	SID attribute
PE1	END.X	8000:A:B:C:1:1::	128	COC
P1	END.X	8000:A:B:C:2:1::	32	COC
P2	END.X	8000:A:B:C:3:1::	32	COC
P3	END.X	8000:A:B:C:4:2::	32	None COC
P4	END.X	8000:A:B:D:1:1::	128	None COC
P5	END.X	8000:A:B:D:2:1::	128	None COC
P6	END.X	8000:A:B:D:3:1::	128	None COC
P7	END.X	8000:A:B:E:1:1::	128	COC
P8	END.X	8000:A:B:E:2:1::	32	COC
P9	END.X	8000:A:B:E:3:2::	32	None COC

The SIDs are delivered to the head node PE1 through BGP POLICY.

- The head node PE1 maps the VPN traffic to the Policy path, and constructs a complete G-SRH according to the Segment List SID information sent by the controller and the VPN SID distributed by the PE2. It then encapsulates the IPV6 packet and sends it to P1 by starting the forwarding process. The first SID in SRH is 8000:A:B:C:1:1::, which is locally COC Flavored END.X SID. According to the description of the forwarding pseudocode in Section 3.1, it performs COC Flavored SID forwarding process: SL--, DA.SI=3, and then directed to the next 32 bits G-SID 2:1. The compressed SIDs 2:1 is copied to the related position of DA, and the packet will be forwarded after forwarding table lookup based on the new DA. At this time, SL=6, SI=3, the destination address is changed to 8000: A:B:C:2:1::3.

An example of the compressed SID of each node in G-SRH is follow：

Fig 4-3: The forwarding process of mixing compressed scene



- Subsequently, the nodes, P1, P2, P7 and P8, match the local COC Flavored SIDs according to the destination address, then perform the COC Favored SID forwarding process. The nodes update the corresponding G-SID to DA according to the SL and DA.SI values, and then lookup the forwarding table and forward the packet accordingly.
- Node P3 and P9 match local COC Flavored END.X SID according to the destination address. According to the standard SRv6 processing flow: SL--, they take the next 128 bits SID to update the destination address.
- Node P4, P5 and P6 process as standard SRv6 nodes.
- When the node PE2 receives the packet with DA as 8000: A:B:E:4:: 100, which is a VPN SID without COC Flavor. It is processed as standard SRv6 VPN SID.

5 Best Practices of Interoperability Test

Based on CMCC G-SRv6 technical specification, the G-SRv6 header compression solution has enabled multi-vendor inter-operability and network in a single compressed domain and hybrid domains across compressed domains and classic SRv6 domain. It has been successfully proved to work with inter-operability, networking, management and controls with equipment nodes and SDN controllers from multiple different vendors.

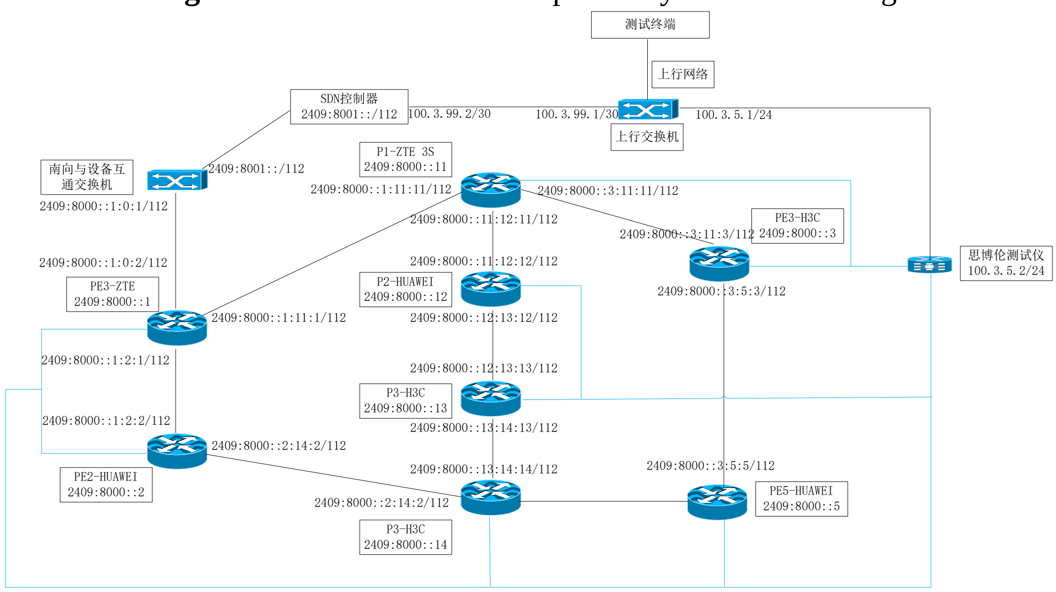
To promote the deployment of the G-SRv6 compression solution in the operator network, China Mobile has launched a thorough pilot plan for that, which is divided into three phases:

- Data plane interoperability tests in China Mobile Lab
- Management plane & control plane interoperability tests in China Mobile Lab
- Field trial in existing provincial network

In general, the field trial would be completed by November 2020. Till now, equipment vendors including Huawei, ZTE, H3C and Ruijie and chip vendors including Broadcom and Centec have successfully implemented the G-SRv6 solution in their products and successfully conducted multi-vendor inter-operability tests in the China Mobile Lab.

5.1 Interoperability Test Scenario

Fig 5-1: Multi-Vendor Interoperability Test Networking



- Phase 1: Data plane interoperability tests in China Mobile Lab

The test setup is as shown in the Figure 5-1. routers from ZTE, Huawei and H3C, silicon SVK from Broadcom(JR2) and Centec and white-box from Ruijie deploy their equipment nodes as PE and P nodes respectively for SRv6 header compression interoperability test.

Test scope: Multi-vendor data plane interoperability.

The test includes BGP L3VPN service over a single pure compression domain (with same and different prefix) and over hybrid of compression domains and standard classic domains.

- Phase 2: Management plane and control plane interoperability tests in China Mobile Lab

The test setup is as shown in the figure 5-1. ZTE, Huawei, H3C and other vendors deploy their routers as PE and P nodes respectively for SRv6 header. Controller vendors include ZTE, Huawei and China Unitechs.

Test Scope:

Interoperability among the controllers and equipment nodes from multiple vendors.

The tests include control plane interoperability for G-SRv6 SID compression, dynamic deployment of SRv6 Policy to carry BGP L3VPN services over a single compression domain (with same and different network prefix) and over hybrid of compression domain and uncompressed SRv6 classic domains, SBFD OAM and HSB reliability verification.

- Phase 3: Field trial in existing provincial network

The tests will depend on existing network and current service offering. New equipment nodes with G-SRv6 capability will be added into that network. It is expected that the participating vendors will be Huawei, ZTE, H3C, China Unitechs, etc., to implement BGP L3VPN over G-SRv6 one-click service deployment.

5.2 Interoperability Test Result

At the beginning of July 2020, China Mobile Lab completed the lab G-SRv6 forwarding plane interoperability tests, covering multiple vendors Huawei, ZTE and H3C. The tests covers single compression domain, multi-compression domains, hybrid domains of compressed and classic SRv6 with P/PE router equipment nodes. In August 2020, additional data plane interoperability tests have been successfully conducted with Ruijie’s white-box switch and chip prototype of silicon vendors’ SVKs from Broadcom and Centec. The test cases, as described in Section 5.1, cover the key technical features of G-SRv6 optimized compression proposal, and enable the interoperability of strict and loose paths within a single compression domain, hybrid domains of compressed G-SRv6 and classic SRv6, and multiple compression domains. The success of the interoperability tests proves the feasibility of the G-SRv6 solution, which builds a solid foundation for the next phase.

6 Conclusion and Outlook

With the accelerated deployment of IPv6 in China, the rapid deployment of SRv6 has become the next trend. However, the packet header overhead of SRv6 is an outstanding challenge, especially in the multi-hop strict path for TE and SFC use scenarios, the efficiency of SRv6 deployment are greatly impacted. How to effectively compress the packet header overhead of SRv6 has become a focus of research over the world.

Given the urgency of commercial deployment of SRv6 and also the importance of SRv6 header compression, China Mobile has been leading a multiple-vendor team in the industry, including OEMs from Huawei, ZTE, H3C and Ruijue, sillion suppliers from Broadcom, Intel, Marvel and Centec, and test equipment provider such as Spirent to promote the G-SRv6 solution. Through tightly cooperating with the whole industry, China Mobile is committed to build an industrial ecosystem, and promote the maturity of G-SRv6 standardization in China. Meanwhile for international

G-SRv6 standardization, China Mobile has also been cooperating with Cisco and Bell Canada to broaden G-SRv6 influence in international standard bodies.

With 32-bit compression, G-SRv6 can increase the SID encapsulation efficiency up to about 75% in theory with significant reduction of SRv6 overhead. At the same time, G-SRv6 is backward compatible with SRv6. It supports hybrid programming of compression SRv6 and standard classic SRv6. In nature G-SRv6 can be interworking with standard SRv6 with smooth evolution from standard SRv6. Now the first phase of multi-vendor interoperability tests has been completed in the lab of China Mobile, which proves that the technical feasibility of G-SRv6 and the preliminary industrial consensus has been achieved.

Looking forward, the future network will provide the open capabilities. Service innovations will be based on the synergy of the global networks and cloud infrastructure, which results in further function repartition among terminals, network edge, clouds, and networks. In order to achieve that goal, China Mobile will steadily promote the pilot of the G-SRv6 and accelerate the maturity of the industrial ecosystem with commercial deployment as our targets. Let's work together to promote the arrival of this new era.

Abbreviations

IP	Internet Protocol
IPv6	Internet Protocol version 6
SR	Segment Routing
SRv6	Segment Routing of IPv6
SID	Segment Identifier
SRH	Segment Routing Header
G-SID	Generalized SID
G-SRv6	Generalized SRv6
G-SRH	Generalized SRH
COC	Continue of Compression Flavor
SL	Segment Left
SI	Generalized SID Index
ISIS	Intermediate System-to-Intermediate System
BGP	Border Gateway Protocol
BGP-LS	BGP Link-State
IGP	Interior Gateway Protocol
SDN	Software Defined Network
TE	Traffic Engineering
LAN	Local Area Network
SAFI	Subsequent Address Family Identifier
NLRI	Network Layer Reachable Information
MP-BGP	Multiprotocol extensions for BGP-4
VPN	Virtual Private Network
DA	Destination Address
P	Provider
PE	Provider Edge
OAM	Operation Administration and Maintenance
SBFD	Seamless bidirectional forwarding detection
HSB	Hot standby
TI-LFA	Topology Independent Loop free Alternate
SFC	Service Function Chain

Reference

[1] Weiqiang Cheng, Yisong Liu, Wenying Jiang, ZhangGeng. Research and application of G-SRv6 head compressions optimization technology, Telecommunications Science.

[2] draft-cl-spring-generalized-srv6-for-cmpr-01, IETF, 2020

[3] RFC8754, IETF, 2020

[4] draft-ietf-spring-srv6-network-programming-19, IETF, 2020

[5] draft-ietf-lsr-isis-srv6-extensions-09, IETF, 2020

[6] draft-ietf-idr-bgpls-srv6-ext-03, IETF, 2020

[7] draft-ietf-idr-segment-routing-te-policy-09, IETF, 2020

[8] draft-cheng-spring-shorter-srv6-sid-requirement-02, IETF, 2020

[9] draft-filsfilscheng-spring-srv6-srh-comp-sl-enc-01, IETF, 2020