



# 中国移动 G-SRv6 技术白皮书 (2020 年)

中国移动研究院

2020 年 9 月

## 前言

IPv6 作为我国新型基础设施建设的关键基础技术，加快推进 IPv6 规模部署是贯彻落实党中央、国务院关于建设网络强国的战略之一，可以预见网络将加快进入 IPv6 新时代。基于 IPv6 基础技术融合 SR（Segment Routing）实现 SRv6 支持网络协议简化、能力开放和可编程，从而灵活满足 5G、云网协同和泛在连接等新兴业务的需求。SRv6 采用 128bit IPv6 地址作为 SR SID（Segment ID），继承了 IPv6 地址灵活连接、全局路由和可编程的优点，以及 SR 源路由、网络简化和路径可回溯的优势，代表了基于 IPv6 技术为基础网络的重要发展方向，将成为下一代 IPv6 网络的核心技术。

但是，落地部署 SRv6 还需要解决承载效率、硬件兼容性、现网平滑升级等种种问题，中国移动联合业内合作伙伴创新提出 G-SRv6 头压缩优化方案并在 IETF 推动标准化进程，希望加快 SRv6 在国内的落地部署，拉开云网时代 IP 网络创新的大幕。

本白皮书旨在提出中国移动对于 G-SRv6 头压缩优化的技术方案、组网部署应用和互通实践规划。希望能够为产业在部署 SRv6 头压缩相关技术、产品和解决方案时提供参考和指引。

本白皮书的版权归中国移动所有，未经授权，任何单位或个人不得复制或拷贝本建议之部分或全部内容。

# 目录

1 技术背景 .....	4
2 概念和原理 .....	6
2.1 SRv6 头压缩概念 .....	6
2.2 SRv6 头压缩原理 .....	7
3 技术方案 .....	10
3.1 G-SRv6 数据面方案 .....	10
3.2 G-SRv6 控制面方案 .....	15
3.2.1 IS-IS 控制面 .....	16
3.2.2 BGP-LS 控制面 .....	19
3.2.3 BGP SRv6 Policy 控制面 .....	21
4 典型组网应用 .....	24
4.1 纯压缩场景 .....	25
4.2 混编压缩场景 .....	28
5 互通实践 .....	30
5.1 互通验证场景 .....	31
5.2 互通验证结果 .....	32
6 总结与展望 .....	33
缩略语列表 .....	35
参考文献 .....	37

## 1 技术背景

在云网融合时代大背景下，灵活敏捷的网络服务能力直接影响运营商的竞争力。SR（Segment Routing）是源路由技术的一种，SRv6 是 SR 技术在 IPv6 网络的应用。SRv6 的出现是一个巨大的创新，它结合 SDN 技术使能可编程的网络，这为云网时代的网络基础服务、增值网络服务提供了创新的土壤。

SRv6 继承了 SR-MPLS 的所有优点并进行了重大改进，通过把网络转发面统一到 IPv6，转发面只要 IPv6 路由可达就可建立跨域转发路径，不再需要专门的 MPLS 转发面；通过定义功能指令信息，使得在宏观网络的视角下实现报文转发处理的可编程。

SRv6 技术使用 128bit 的 IPv6 地址作为 Segment ID（SID），以其极简可编程的特性，实现了任意接入和任意联接，可灵活满足各种网络业务需求，同时提供更高的可靠性和可扩展性，将成为新一代 IP 承载网的核心技术。随着 SRv6 技术和标准逐渐成熟，主流的设备芯片厂商和开源平台都已经支持 SRv6，国内外运营商都开始了 SRv6 部署。

当前 SRv6 虽然具备上述诸多优点，但是缺点同样明显。主要的缺点在于当 SID 数目增多时，报文头开销过大。SRv6 报文头过长带来的挑战主要分为以下三个方面：

- 报文头开销大导致网络链路带宽利用率低，256byte 负荷包长在 8 层 SID 的情况下带宽利用率只有 60%左右；

- 新增报文头长度可能导致报文大小超过 MTU，导致分片或丢包，导致传输性能急剧下降；
- SRv6 报文处理对芯片要求高，老旧现网设备难以支持深度的报文头的复制操作。

这些问题也导致运营商现网无法平滑升级 SRv6。针对上述问题，业界提出了多种 128bit SID 压缩方案。这些方案虽然达到了压缩目标，但都或多或少存在一些缺陷：

- 不兼容标准 SRv6，影响实际部署

为支持存量演进和平滑升级，压缩方案需要支持和标准 SRv6 节点混合编排。此外，还需要兼容 SRH。压缩方案修改 SRH 格式，可能存在报文被标准 SRv6 节点丢弃的风险。

- 压缩长度连续可变，增大实现难度

压缩方案需要考虑网络的可扩展性和字节对齐等因素。连续可变的压缩方案加大了设计复杂度，实用性也有待商榷。

- 不支持原有地址规划，浪费地址空间

当前 SRv6 的部署大多基于已有的公网地址规划。压缩方案的地址规划如果与当前网络地址规划不兼容，会导致无法部署或者公网地址的浪费。

针对这些问题，中国移动联合业界合作伙伴提出新的 SRv6 头压缩技术方案 G-SRv6，该方案兼容标准 SRv6，使用固定的 SID 压缩长度，兼容原有地址规划，支持现网平滑升级是目前适用于运营商网络的业界领先方案。同时中国移动一方面在产业界推动产品实现，另一方面

在 IETF 推动标准化进程，以加快 SRv6 在国内的落地部署。

## 2 概念和原理

### 2.1 SRv6 头压缩概念

下面是标准 SRv6 扩展头 SRH 的封装格式：

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)			

图 1 标准 SRv6 SRH 封装

SRv6 数据包的扩展头封装长度为：40 Byte (IPv6 头) + 8 Byte (SRH 固定头) + 16\*N Byte (Segment List)。随着 SID 数量增加，SRv6 扩展头开销会增大。

根据 RFC8402、RFC8754 的定义，一个 SRv6 SID 是一个 128bit 的 IPv6 地址。在 SRv6 可编程网络中，它通常由如下三部分组成。

Locator	Function	Arguments
---------	----------	-----------

图 2 支持网络可编程标准 SRv6 SID 格式示意图

- Locator 是网络拓扑中分配给一个网络节点的标识，用于指导转发数据包到该节点。
- Function 用来指示网络节点要执行的转发动作。
- Arguments (Args) 是可选执行参数，可能包含流、服务或其

他任何相关信息。

Locator 可以细分为 B:N，其中 B 标识 SRv6 SID Block，一般由运营商分配给某个子网，通常用 Prefix 来表示；N 则是该子网内区分节点的标识。

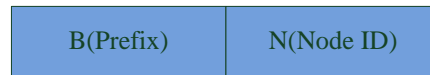


图 3 支持网络可编程典型 SRv6 Locator 示意图

从上面的 SRv6 报文头格式定义中，我们可以看出，报文头占比最大的部分就是 Segment List。在表项个数维持不变的情况下，SRv6 报文的头压缩，实际上就是针对单个 SID 的压缩。

下面是头压缩技术中常用概念，包括：

- 完整 SID，指 128bit 长度 SID；
- 压缩 SID，指缩短长度后的 SID；
- SID 压缩算法，指完整 SID 压缩成压缩 SID 的具体方法；
- 控制面协议扩展方案，包含节点支持压缩的能力发布，可压缩的 SID 的发布等；
- 数据面协议扩展方案，包含完整和压缩 SID 边界识别，Segment List 和 DA（Destination Address）的读写规则等；

## 2.2 SRv6 头压缩原理

实际上，一个 SRv6 网络中很多 SID 都具有共同前缀（Common Prefix），一个 SID List 中 SID 的共同前缀部分都是重复冗余的。所以，SID list 中去掉 SRv6 SID 中的 Common Prefix 部分和其他冗余部分，保

留 Node ID 和 Function ID 作为压缩 SID，可以很好的减少报文头开销。这也是本白皮书中描述的 SRv6 头压缩方案中的 G-SID（Generalized SID，通用段标识符）的基本压缩原理。

G-SID 是可压缩 SRv6 SID 的一部分，和 Common Prefix 以及 Argument/Padding 一起组合成完整的 SRv6 SID。G-SID 支持 32bit 或者 16bit 两种长度。实际从硬件处理效率、利旧和可扩展性的角度出发，32 bit 是一种更常用的压缩 SID 长度，因此本白皮书只重点介绍 32 bit 的压缩。

下面是典型的 32bit SRv6 压缩 G-SID 格式，该 G-SID 是由 128bit SID 中的 Node ID 和 Function ID 组成。128bit SID 格式是完整 SID，也即 SRv6 SID；32bit SID 为 G-SID，其是完整 SID 的变化部分。格式如下图所示：

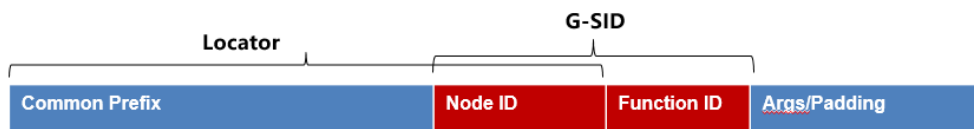


图 4 支持压缩的 SRv6 SID 格式图

完整 SID 和压缩 SID 的转换关系根据 SRv6 Locator 的规律可以采用如下方式：

G-SID 组成为 N（Node）：F（Function）

压缩方案首先引出的是网络规划问题，要保证规划的 Node 数量和 Function 数量的总位数不能超出压缩长度。实际需要根据压缩长度、Function 最大数量和 Node 最大数量综合分析，确定 Function ID 和 Node ID 位数。



为支持压缩，需要扩展控制面协议。将节点支持压缩的能力，可压缩的 SID 发布出去。

为支持压缩，需要扩展数据面协议。因为 SRH 中 Segment List 可以插入 G-SID,所以需要更新 Segment List 的访问复制逻辑和 DA 的替换逻辑，同时加入对完整 SID 和 G-SID 边界识别方法。

为方便理解 G-SID 和完整 SID 混编并确保 128bit 对齐,引入 G-SID Container 概念。

一个 G-SID Container 是准确的 128bit 长度，可以包含：

- 一个完整 SRv6 SID
- 多个 G-SID, 比如 1~4 个 32bit 的 G-SID 或者 1~8 个 16bit G-SID, 不满 128bit 则全零补齐

以 32bit G-SID 为例，G-SID Container 可能的格式如下图所示。

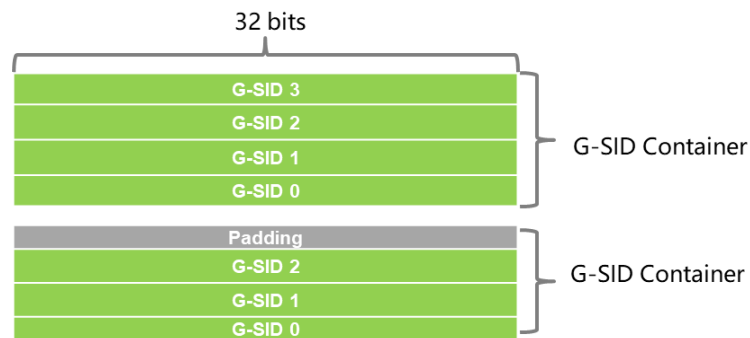


图 5 32-bit G-SID Container 格式图

由于 SRv6 SRH 中可以编码多种类型的 G-SID Container, 所以这个 SRH 我们称之为 Generalized SRH (G-SRH)。这种支持将通用的多种 Segment 编码到 SRH 中的方案是一种 SRv6 的升级，我们称之为 Generalized SRv6 (G-SRv6)。

### 3 技术方案

本节将介绍 G-SRv6 的整体技术方案，包括 G-SRv6 数据面方案和控制面方案两大部分。数据面部分主要描述压缩 G-SID 数据封装方式，控制面部分主要描述基于 IS-IS, BGP-LS 和 BGP SRv6 Policy 协议支持压缩 G-SID 的扩展。

G-SRv6 在数据封装上没有改变原生 SRH 格式和语义，天然兼容原生 SRv6 并可与原生 SRv6 混合部署。压缩 SID 推荐使用固定长度 32bit，现有芯片在支持 SR MPLS 的基础上就可以支持，大为降低实现难度。由于 G-SRv6 压缩原理是提取公共前缀，对于前缀的长度没有任何限制，可以灵活适应现有网络地址规划和路由规划，不会浪费运营商分配的地址空间。

#### 3.1 G-SRv6 数据面方案

原生 SRv6 数据封装主要体现在 IPv6 报文的 SRH 扩展头部，G-SRH 与 SRH[RFC8754]格式保持一致，没有对其格式和字段语义进行改动，兼容但支持将 128bit SRv6 SID 和 32bit G-SID 混合编程在 G-SRH 中，其格式如下图所示。

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)			

图 6 G-SRH 格式图

当 128 bit SRv6 SID 和 32 bit G-SID 混编在 G-SRH 中时，其编码示例如下图所示：

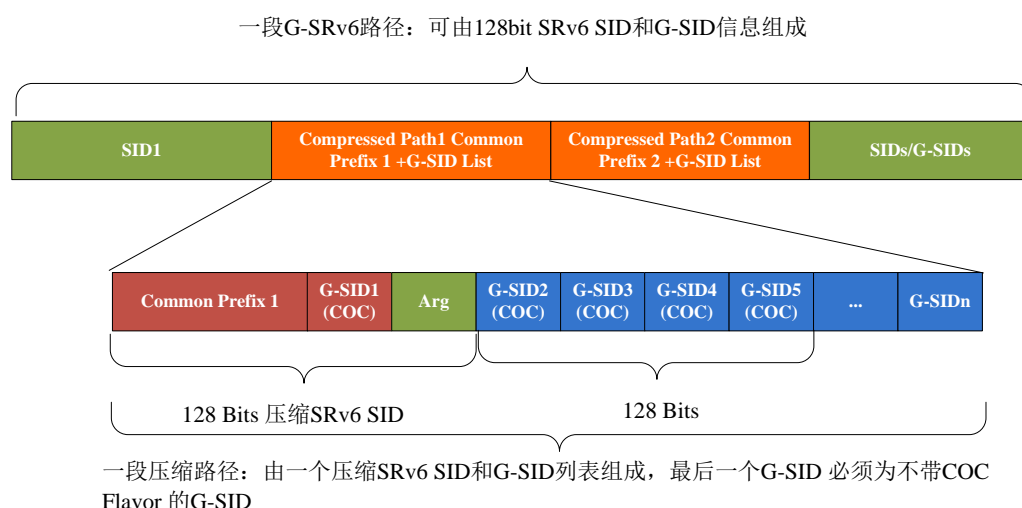


图 7 G-SRv6 混编 SID 图

一段 G-SRv6 路径可能由 SRv6 子路径和 SRv6 压缩子路径组成。SRv6 子路径由 SRv6 SID 编码。SRv6 压缩子路径由一个支持压缩的 128bit SRv6 SID 引导开始，并由随后的多个 G-SID 组成。在 G-SID List 结束时，需以 128bit 为界结束。G-SID List 结束的下一个 SID 为 128bit SRv6 SID，其可能是普通的 SRv6 SID，也可能是一个支持压缩的 SRv6 SID。

为标识 SID list 中 SRv6 压缩路径的起始和结束，也即 128bit SID

和 32bit 的 SID 之间的边界，需要再新增 Flavor,并发布对应 Flavor 的 SID。

表 1 新增 Flavor 表

SID Flavor 类型	功能描述
COC	标识更新 32 bit 的 G-SID 到目的地址。
COC16	标识更新 16 bit 的 G-SID 到目的地址。

当前 SRv6 的目的地址（Destination Address, DA）更新动作是将下一个 128bit SID 更新到目的地址,因此已有 SRv6 SID 自带更新 128bit SID 的 Flavor(如果该 SID 不是最后一个 SID),所以无需额外增加 Flavor 定义。

当 DA 中 SID 携带 COC Flavor 时，表示需要将下一个 32bit G-SID 更新到目的地址。因此当 DA 中 SID 为不携带 COC Flavor SID 时，表示将 SL--, 将下一个 128bit SID 更新到目的地址（SL>0 时）。

压缩 SRv6 在 SID List 中的编码规则为：

- 一段压缩 SRv6 路径的开始由一个 128bit 的 COC Flavor SID 指示，其包含 COC Flavor G-SID。此 SID 携带了完整的 SID 信息，包含 Common Prefix（CP）等信息，可用于与后续 G-SID 恢复完整的下一个 SID。
- 压缩路径中间的 G-SID 均为携带 COC 的 G-SID，指示下一个是 32bit G-SID。
- 压缩路径的最后一个 G-SID 必须为无 COC Flavor 的 32bit G-SID，用于指示压缩路径结束。其在 DA 中组合成的 SRv6 SID 按照 128bit SID 的处理规则处理。

最后一个 G-SID 为无 COC Flavor SID 的原因在于：其被更新到目的地址之后，DA 中的 SID 由于无 COC Flavor，因此其被节点按照普通 SRv6 SID 来处理，将下一个 128bit SID 更新到目的地址，可以起到标识这段压缩路径结束，从 32bit G-SID 切换到 128bit SID 处理的作用。

这种设计可以很好的支持两个场景：

- 混编场景中压缩路径结束，从 32bit 切换到 128bit 的场景
- 纯压缩场景中，最后一个 G-SID 为无 COC Flavor，因此最后一跳的处理与当前 SRv6 一致，更兼容。

此外，在压缩路径中，为了定位下一个 G-SID，需要新增 SI（Generalized SID Index）来定位其在 G-SID Container 中的位置。SI 放置在目的地址中的 G-SID 之后的 Arguments 的最低位。

本方案定义 SID 长度必须为 128bit，即没有 Padding。在 32bit G-SID 压缩中，SI 为 128bit SID 的最低 2bit。为方便硬件实现和未来扩展，本方案使用最后一个字节为 SI 值。而 Common Prefix 建议选取字节对齐的前缀，比如 64。

在生成压缩 SID 时，需预留空间给 SI，其值为 0。在转发的过程中，只有当活跃 SID（Active SID）为可压缩 SID 时，目的地址中的 SI 字段才有意义。此时，SL 指示了 G-SID Container 在 G-SRH 中的位置，而 SI 指示了 G-SID 在 G-SID Container 中的位置。按照 SID List 的倒序排列规则（即 Segment List[0]为最后一跳，Segment List[n]为第一跳），G-SID 的排列顺序也是倒序的，即 SI=3 时，定位到 G-SID Container 的最高 32bit 值，SI=0 时，定位到 G-SID Container 的最低 32bit 值。具体

示例见图 8。



图 8 SI（G-SID index）封装格式图

此方案基本思想为：通过 COC Flavor SID 指示更新下一个 32 bit G-SID 到目的地址，G-SID 位置由 SL 与 SI 定位。

以 32bit G-SID 为例，Endpoint 节点收到数据包的处理伪代码如下：

```
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
```

SRv6 压缩的处理流程由 COC Flavor SID 的触发，操作数据也局限在 COC Flavor SID 之内，对已有 SID 以及 SRH 的处理没有影响。其中 SI 在 DA 的 G-SID 之后。以纯压缩路径加 128bit VPN SID 为例，编码与 G-SID 更新示意图如下图所示：

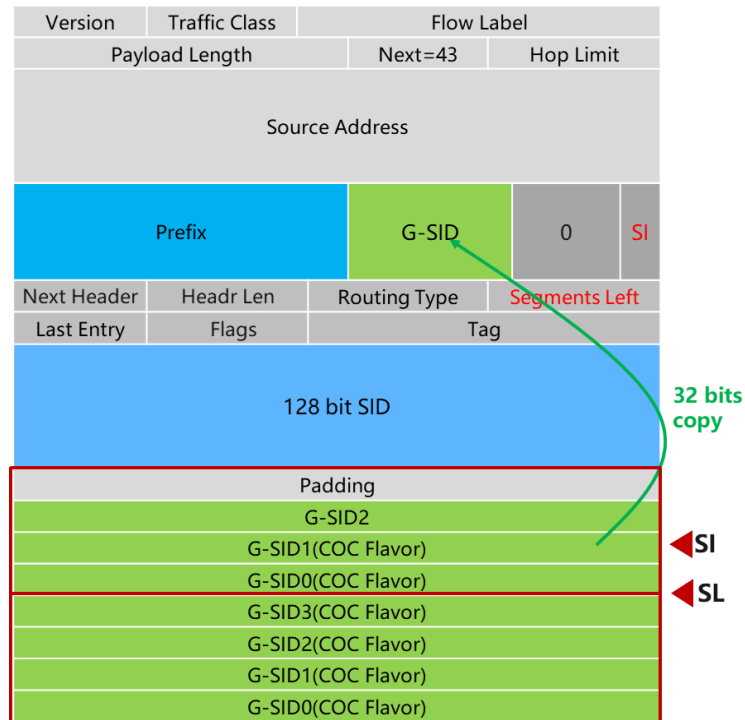


图 9 G-SRv6 编码示意图

### 3.2 G-SRv6 控制面方案

G-SRv6 方案的支持需要对控制面协议进行一定的扩展。原生 SRv6 对 ISIS 协议扩展以支持设备间收集 SID 属性信息 [draft-ietf-lsr-isis-srv6-extensions-09], 为了发布节点压缩能力和可压缩 SID 信息, G-SRv6 需要对 IS-IS 再进行扩展。原生 SRv6 为了收集 SID 属性信息到控制器已有 BGP-LS 的协议扩展 [draft-ietf-idr-bgppls-srv6-ext-03], 为了收集节点压缩能力和可压缩 SID 信息, G-SRv6 需要对 BGP-LS 再进行扩展。原生 SRv6 控制器完成算路后, 通过 BGP 扩展新的 NLRI [draft-ietf-idr-segment-routing-te-policy-08] 下发 SRv6 Policy 到头节点, 携带 SID 列表路径, 为了支持携带可压缩的 SID, G-SRv6 需要对 BGP TLV 进行扩展定义。

### 3.2.1 IS-IS 控制面

网络中的节点在完成 SRv6 SID Space 规划之后，需要根据要求完成 SRv6 SID 配置，并根据配置生成本地 SID 表，然后通过 IGP 协议将 SID 发布到网络中，因此需要对 IS-IS 等 IGP 协议进行对应扩展。以 IS-IS 协议为例，协议扩展如下：

- 扩展 SRv6 Capabilities sub-TLV，增加 C-flag，通告节点支持压缩的能力。

Type	Length		C	Flags
Optional sub-subs-TLVs				

图 10 SRv6 Capabilities sub-TLV 格式图

C-flag 由本方案定义，当 C-flag 设置为 1 时，为“当前节点具有 SRv6 压缩能力”，用来标志节点的 SRv6 压缩能力。

不支持压缩的 SDN 控制器或节点在收到扩展 C-flag 的 SRv6 Capabilities sub-TLV 时，应该忽略掉 C-flag。

节点的 SRv6 压缩能力可以通过 BGP-LS 上报给 SDN 控制器（参见 3.2.2）。SDN 控制器在计算 SRv6 TE 路径时，根据节点的压缩能力编排 SID list。C-flag 也可在节点算路时使用。

如果头节点没有上报 SRv6 压缩能力，那么 SDN 控制器不应该向该头节点下发包含 G-SID 的 SRv6 Policy。如果确实要下发，那么，只能包含无 COC Flavor 的 SID，且按照 128bit 编排。

- 扩展 SRv6 End SID sub-TLV, SRv6 End.X SID sub-TLV, SRv6 LAN



End.X SID sub-TLV 增加 C-flag 标识 SID 支持压缩。C-flag 位置如图 所示。

Type		Length			
Flags	C	Endpoint Behavior			
SID (128 bits)...					
SID (cont...)					
SID (cont...)					
SID (cont...)					
Sub-sub-tlv-len		Sub-sub-TLVs (variable)			

图 11 SRv6 End SID sub-TLV 格式图

Type		Length			
Flags	C	Routing Type	Segments Left		
Endpoint Behavior					
SID (128 bits)...					
SID (cont...)					
SID (cont...)					
SID (cont...)					
Sub-sub-tlv-len		Sub-sub-TLVs (variable)			

图 12 SRv6 End.X SID sub-TLV 格式图

Type	Length		System ID (6 octets)	
Flags	C	Algorithm	Weight	
Endpoint Behavior				
SID (128 bits)...				
SID (cont...)				
SID (cont...)				
SID (cont...)				
Sub-sub-tlv-len		Sub-sub-TLVs (variable)		

图 13 SRv6 LAN End.X SID sub-TLV 格式图

当 C-flag 置位时，Sub-sub-TLVs 中必须携带 SRv6 SID Structure Sub-sub-TLV。其中 LB Length 描述了 Common Prefix 的长度。G-SID 的长度由 LN Length（Node-ID 的长度）+ Fun. Length（Function ID）的长度决定，其可以为 32 或者 16。若 Node-ID 和 Function ID 长度之和不等于 32 或 16 则该 SID 被认为是非法的 SID，需丢弃该 SID。

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

图 14 SRv6 SID Structure sub-TLV 格式图

- 在发布 SID 时，需要扩展以上几个 SID sub-TLV 的 Endpoint Behavior 字段的 Code point，增加 COC/COC16 flavor 引入的 Behavior code point。

在 IGP 能力发布过程中，节点通过置位 SRv6 Capabilities TLV 的扩展 C-flag 来标识本节点支持压缩能力。

为支持 SRv6 压缩需要实例化对应的 SID 并将通过 IGP 等协议发

布到网络中或上送给控制器。节点在实例化 SID 时，将会生成对应的 SID 转发表项。

在发布 SID 时，需要通过置位 SID 对应的 TLV（如 SRv6 End SID sub-TLV, SRv6 End.X SID sub-TLV, SRv6 LAN End.X SID sub-TLV）中的 C-flag 来标识本 SID 的格式支持压缩。具体支持 32bit 压缩还是 16bit 压缩，由 LN Length + Fun. Length 的长度判断，若长度为 32 则为 32bit 压缩，若为 16 则为 16bit 压缩。

### 3.2.2 BGP-LS 控制面

BGP-LS 协议可以用于上送 SRv6 相关的信息。本文档基于 [draft-ietf-idr-bgpls-srv6-ext-02] 文稿扩展用于支持上报 SRv6 头压缩 SID 属性信息。

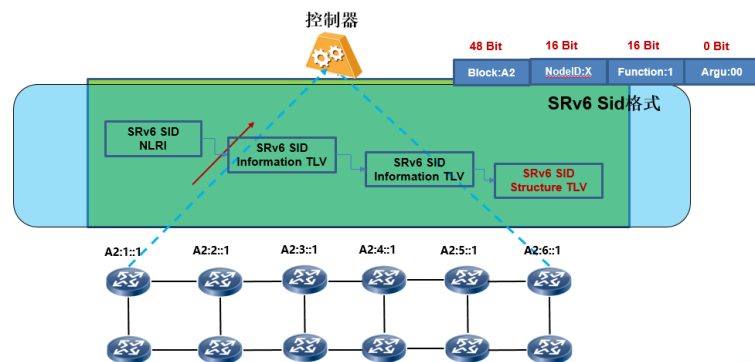


图 15 SRv6 BGP-LS 协议流程示意图

控制器（Controller）为了实现 TE 路径计算，需要了解全网拓扑信息、拓扑的 TE 属性以及 SRv6 属性等相关信息。控制器可以通过 BGP-LS 协议实现对网络拓扑信息的收集。为支持 SRv6 压缩，同样需要扩展三类信息。

- 扩展 SRv6 Capabilities sub-TLV，增加 C-flag，通告节点支持压缩的能力

当 C-flag 设置为 1 时，为“当前节点具有 SRv6 压缩能力”，用来标志节点的 SRv6 压缩能力。

- 扩展 SRv6 Link Attributes 下的, SRv6 End.X SID sub-TLV, SRv6 LAN End.X SID sub-TLV，增加 C-flag 标识 SID 支持压缩。

Type		Length	
Endpoint Behavior		Flags	Algorithm
Weight	Reserved	SID (16octets)	
SID (cont ...)			
SID (cont...)			
SID (cont...)			
SID (cont...)		Sub-TLVs (variable)...	

图 16 SRv6 End.X TLV 格式图

B	S	P	C	Rsvd
---	---	---	---	------

图 17 Flag 字段格式图

- 扩展 Node SID 对应的 End Behavior sub-TLV,增加 C-flag 标识支持压缩。

BGP-LS 信息来自 IGP，此处 C-flag 扩展逻辑类似 IGP 扩展。

Type	Length		
Endpoint Behavior	Flags	C	Algorithm

图 18 End Behavior sub-TLV 格式图

- 在发布 SID 时，需要扩展以上几个 SID sub-TLV 的 Endpoint

Behavior 字段的 Code point, 增加 COC/COC16 Flavor 引入的 Behavior code point。

建立 BGP-LS 连接时, 若节点支持 SRv6 压缩, 则需要将 SRv6 Capabilities sub-TLV 里面的 C-flag 置位。节点可以通过 BGP-LS 将支持压缩的 SID 信息上送控制器。在上送这些 SID 时, 需要将 SID 对应的 SRv6 End.X SID sub-TLV, SRv6 LAN End.X SID sub-TLV, 以及 End behavior sub-TLV 中置位 C-flag。当 C-flag 置位时, 必须携带 SID Structure TLV, 用于描述 SID 的格式。

### 3.2.3 BGP SRv6 Policy 控制面

控制器可以通过 BGP SR policy[draft-ietf-idr-segment-routing-te-policy-08]给首节点下发 SR policy。为支持压缩, 需要扩展 BGP SR Policy。BGP SR Policy 的格式如下:

```
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
        ...
```

为支持压缩，需要对协议做部分扩展：

节点是否支持压缩的能力无需扩展，其通过 BGP-LS 上送的信息获取。如果头节点支持压缩，则可以下发包含压缩 SID 的 SRv6 Policy，否则只应下发普通 SRv6 Policy。普通 SRv6 Policy 中可以包含 128bit 格式的可压缩 SID，其不携带 COC Flavor。

为描述 SID List 编码格式，需通过扩展定义 Segment List sub-TLV 下的 SID Encoding sub-TLV 来描述 G-SID 在 SID 中的具体位置，格式如下所示。此扩展可用于首节点感知 SID 信息的场景。若节点无需感知 SID 信息，则可以由控制器直接编码成 128bit 的 SRv6 SID，直接下发给首节点。本方案使用节点感知 SID 信息的解决方案。

Type	Length	Flag	Reserved
Block Length	G-SID Length	Reserved	

图 19 SID Encoding sub-TLV 格式图

- Type：类型。长度 1 字节。
- Length：TLV 长度，不包含 type 和 length 字段。长度 1 字节，取值 6。
- Flag：8bit， 暂无定义，发送时必须置零，接收时必须忽略。
- Reserved 字段：8bit 预留字段，发送时必须设置为 0，且接收时忽略。
- Block Length：8bit 值，当压缩 SID 为 32 时，理论取值范围可以为 0-94，本方案由于预留 8bit 做 SI 使用，所以规定 Block Length 长度不超过 88。

- G-SID Length: 8bit 值, 表示压缩 SID 长度。当前规范规定了 32bit G-SID 的压缩, 因此当节点收到 G-SID 为 128 和 32 之外的值, 均按照错误处理, 必须撤销该 NLRI[I-D.ietf-idr-segment-routing-te-policy-09]。
- Block Length = 0, G-SID Length = 128: 表示完整复制 128bit 到 SID list 中。
- Block Length = n, G-SID Length = 32: 表达复制 Block 之后的 32bit 到 SID list 中。
- Reserved 字段: 16bit 预留字段, 发送时必须设置为 0, 且接收时忽略。

节点在 G-SRH 编码时, 若 SID Encoding sub-TLV 描述的 SID 的 Block 长度为 0, G-SID 长度为 128, 则标识完整拷贝 128bit SID。否则按照 Block Length 值和 G-SID length 进行对应拷贝。从此 SID Encoding sub-TLV 之后一直到下一个 SID encoding sub-TLV 出现之前的所有的 SID 都按照当前 SID encoding sub-TLV 指示的 G-SID 位置, 将 G-SID 编码到 G-SRH 中。

修改后的 SID List 列表示例如下, 其中 SID Encoding sub-TLV1 定义了 Segment2 和 Segment3 的 G-SID 的位置信息, SID Encoding sub-TLV2 定义了 Segment4, Segment5 和 Segment6 的 G-SID 的位置信息。

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 典型组网应用

在严格路径 TE 场景中，逐跳的 SRv6 节点需要通过 SID 来指定，因此路径越长 SID 数目越多；在跨越多个 SRv6 域的场景，端到端的 SID 数目更多。使用 SRv6 头压缩技术可以有效减少 SRv6 报文开销，减轻设备压力。G-SRv6 对转发芯片的要求相比 SR MPLS 并没有提高，主流芯片现有转发封装能力就可以支持，可以在运营商网络中平滑升级。



如下图的企业分支互联场景,专线客户从A省分支访问B省分支,创建端到端 SRv6 头压缩通道。根据 SRv6 头压缩技术的部署情况,G-SRv6 场景可以细分为纯压缩及混编两种场景。

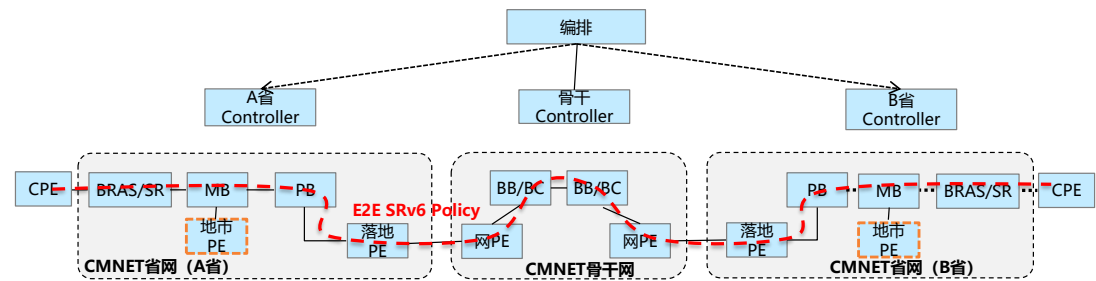


图 20 中国移动企业分支互联 SRv6 承载组网图

4.1 纯压缩场景

在纯压缩场景中，压缩域内的节点全部支持 SRv6 头压缩技术。节点的 common prefix 相同，也可以不同。下面以省内专网用户 MP-BGP L3VPN over SRv6 TE（压缩）业务为例，描述纯压缩场景：

- 专网在同一压缩域内, block 统一规划为 8000:A:B:C/64, node id 为 20bit, function id 为 12bit。
- 如第 3 章所述，压缩域内节点配置相应的节点压缩属性，为 END 和 END.X 分配具有 COC 属性和不具备 COC 属性的两类 SID，通过 IGP 和 BGP-LS 通告给其他节点及控制器。

如，PE1 为某接口分配两类 END.X 类型 SID

PE1 END.X      8000:A:B:C:1:1    COC Flavor  
                  8000:A:B:C:1:2    None COC Flavor

- 控制器根据配置的压缩 Policy 路径：

PE1->P1->P2->P3->P4->P5->PE2，根据 3.1 的编码原则，计算分配 SID 如下，并通过 BGP POLICY 下发给头节点 PE1。

表 2 纯压缩场景 SID 分配信息表

节点	SID 类型	SID(block 64bit)	SID 长度	SID 属性
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

- 头节点 PE1 将 VPN 流量映射到 Policy 路径中，根据控制器下发的 Segment List SID 信息，以及 PE2 分发的 VPN SID，构建完整的 G-SRH，封装好 IPV6 数据包后向 P1 发送，SRH 中第一个 SID 8000:A:B:C:1:1::是本地发布的 COC Flavor END.X，根据 3.1 节中的转发伪代码描述，执行 COC 属性 SID 转发操作：SL--，DA.SI=3,指向下一个 32bit G-SID 2:1，将 2:1 复制到 DA 相关位置，基于新的 DA 进行查表转发。此时 SRH 中 SL=2,目的地址为 8000:A:B:C:2:1::3。。
- 各节点 G-SRH 压缩 SID 示例如下

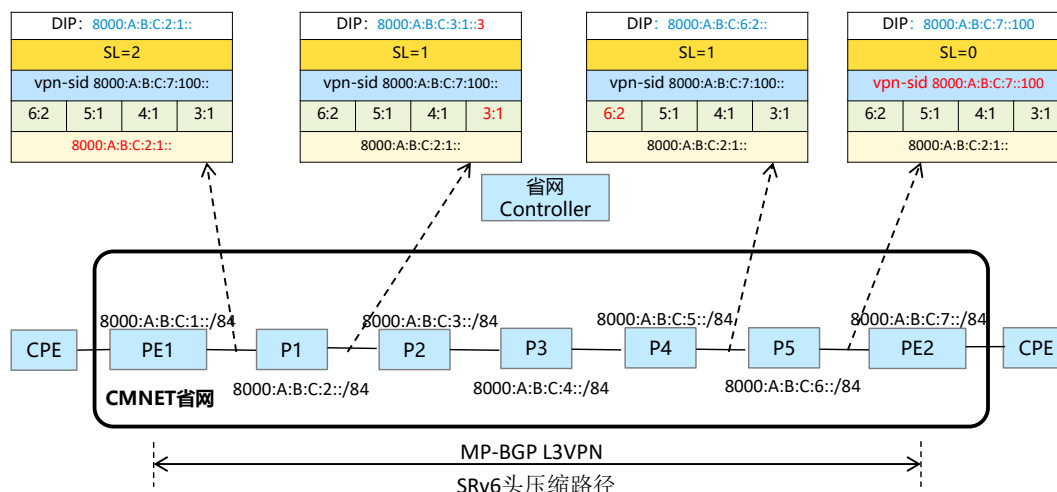


图 21 纯压缩场景转发过程示意图

- 节点 P1 收到数据包，目的地址为 8000:A:B:C:2:1::3 是本地发布的 COC Flavor END.X SID，根据 4.1 节中的转发伪代码描述，执行 COC 属性 SID 转发操作：SI--，指向下一个 32bit G-SID 3:1，将 3:1 复制到 DA 相关位置，基于新的 DA 进行查表转发。此时，SL=2，SI=2，目的地址变更为 8000:A:B:C:3:1::2。
- 后续节点 P2,P3,P4 根据收到报文的目的地址，匹配到本地发布的 COC Flavor 属性 END.X SID，根据 3.1 节中的转发伪代码描述，执行 COC 属性 SID 转发操作：SI--，基于 SL 和 SI，定位对应 G-SID 并据此更新 DA，查表转发。特别需要说明的是，节点 P4 执行 COC 属性 SID 转发操作由于 SI=0，表示读取到 G-SID container 边界，则执行 SL--，DA.SI=3，重新定位对应 G-SID 6:2，复制到 DA 相关位置，目的地址变为 8000:A:B:C:6:2::3。
- 节点 P5 收到数据包时，目的地址为 8000:A:B:C:6:2::3 是本地发布的 无 COC Flavor END.X SID，按照普通 SRv6 SID 转发处理流程，SL--，取下一个 128bit SID 8000:A:B:C:7::100（此处为 VPN

SID) 作为目的地址。

- 节点 PE2 收到数据包时, DA 为 8000:A:B:C:7::100,是不带 COC Flavor 的 VPN SID, 执行普通 SRv6 VPN SID 处理。

## 4.2 混编压缩场景

SRv6 头压缩部署有先后次序, 可能存在部分网络支持 SRv6 头压缩, 部分网络仅支持普通 SRv6 的情况。这导致跨网络端到端 SRv6 严格路径场景中, 存在部分节点支持 SRv6 头压缩, 而另一部分节点不支持头压缩。SRv6 头压缩技术需要考虑此种混编场景, 支持 32bit 压缩 G-SID 和普通 128bit SRv6 SID 共存于同一条 SRv6 TE 路径中。下面以跨省专网用户 MP-BGP L3VPN over SRv6 TE (压缩) 业务为例, 描述混编压缩场景。

- PE1、P1、P2、P3 为压缩域 1 节点, P4、P5、P6 为传统 SRv6 域节点, P7、P8、P9、PE2 为压缩域 2 节点。压缩域 1 block 统一规划为 8000:A:B:C/64, node id 为 20bit, function id 为 12bit; 压缩域 2 block 统一规划为 8000:A:B:E/64, node id 为 20bit, function id 为 12bit; 普通 SRv6 域 block 统一规划为 8000:A:B:D/64。
- 压缩域内节点配置相应的节点压缩属性, 为 END 和 END.X 分配具有 COC 属性和不具备 COC 属性的两类 SID, 通过 IGP 和 BGP-LS 通告给其他节点及控制器。普通 SRv6 节点只分配普通非压缩属性的 SID。

- 控制器根据配置的压缩 Policy 路径：

PE1->P1->P2->P3->P4->P5->P6->P7->P8->P9->PE2, 根据 3.1 中的编码原则, 计算分配 SID 如下, 并通过 BGP POLICY 下发给头节点 PE1。

表 3 混编压缩场景 SID 分配信息表

节点	SID 类型	SID(block 64bit)	SID 长度	SID 属性
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

- 头节点 PE1 将 VPN 流量映射到 Policy 路径中, 根据控制器下发的 Segment List SID 信息及 PE2 分发的 VPN SID, 构建完整的 G-SRH, 封装好 IPV6 数据包后向 P1 发送, 开始转发流程。SRH 中第一个 SID 8000:A:B:C:1:1:: 是本地发布的 COC Flavor END.X, 根据 3.1 节中的转发伪代码描述, 执行 COC 属性 SID 转发操作: SL--, DA.SI=3, 指向下一个 32bit G-SID 2:1, 将 2:1 复制到 DA 相关位置, 基于新的 DA 进行查表转发。此时, SL=6, SI=3, 目的地址变更为 8000:A:B:C:2:1::3。

各节点 G-SRH 压缩 SID 示例如下:

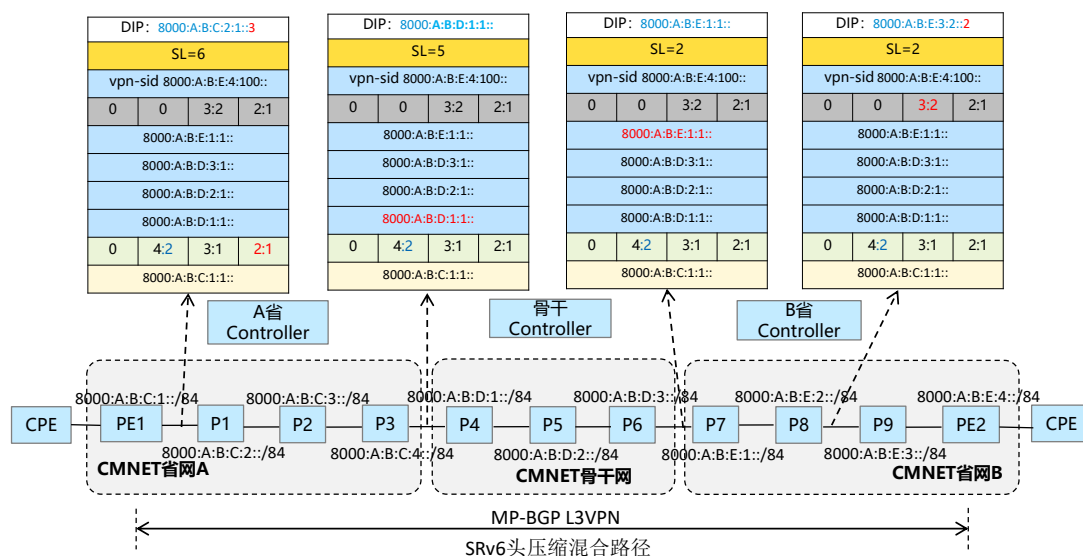


图 22 混编压缩场景转发过程示意图

- 后续节点 P1、P2、P7、P8 根据目的地址匹配到本地发布的 COC 属性 SID，执行 COC 属性 SID 转发操作，节点基于 SL 和 DA.SI 的值将对应 G-SID 更新到 DA，查表转发。
- 节点 P3、P9 根据目的地址，匹配到本地发布的 无 COC Flavor END.X SID，按照普通 SRv6 处理流程，SL--,取下一个 128bit SID 更新目的地址。
- 节点 P4、P5、P6 按照普通 SRv6 节点处理。
- 节点 PE2 收到数据包时，DA 为 8000:A:B:E:4::100,是不带 COC Flavor 的 VPN SID，执行普通 SRv6 VPN SID 处理。

## 5 互通实践

G-SRv6 头压缩优化方案以标准化为根基，支持异厂商互通组网，支持压缩 SRv6 与原生 SRv6 混编组网，支持控制器及设备的异厂商控

制和管理；设计立足高起点，可商用，配套 OAM、可靠性技术，以及 SDN VPN 业务的一键式开通等。

为推动 G-SRv6 压缩方案在中国移动现网部署，中国移动已制定完整的落地试点计划，分为实验室数据面异厂商原型机互通测试、实验室异厂商控制器及转发设备综合互通测试、省公司现网试点三个阶段。整体上，预计 2020 年 11 月份完成试点工作。目前，华为、中兴、新华三、锐捷等设备厂商和博通、盛科等芯片厂商已经实现了 G-SRv6 方案并在中国移动实验室完成了多厂商互通对接。

5.1 互通验证场景

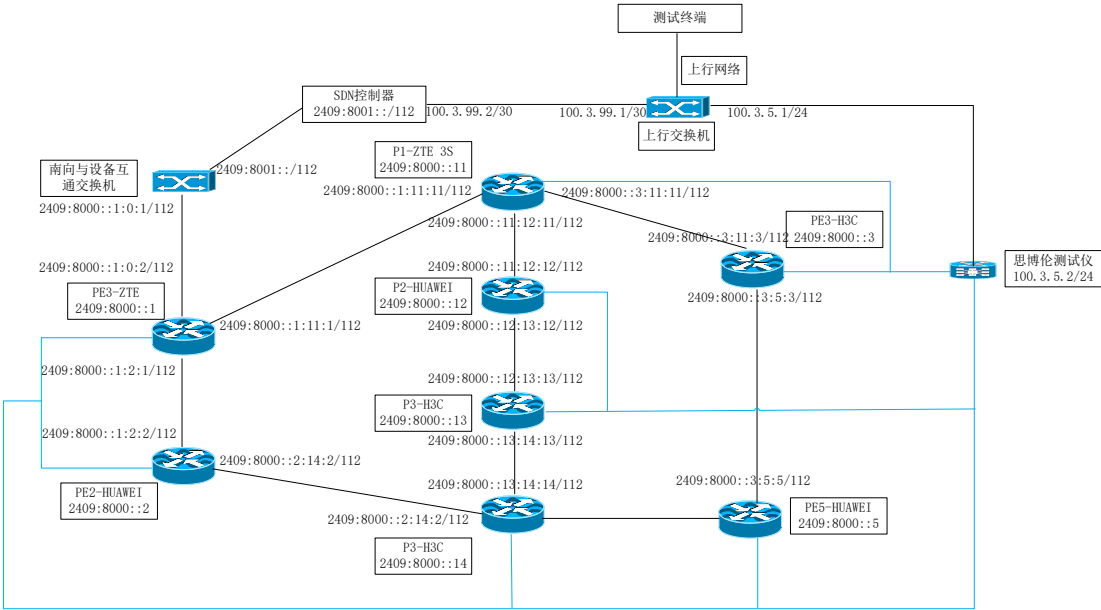


图 23 多厂商互通验证组网图

● 阶段一，实验室转发面互通测试场景

测试环境如图，中兴、华为、新华三多厂商路由器和博通、盛科的芯片样机以及锐捷的白盒交换机分别作为 PE 及 P 节点部署 SRv6 头压缩。

测试内容：多厂商设备转发面互通

测试涉及 SRv6 Policy 承载 BGP L3VPN 业务，纯压缩域（相同及不同网络前缀）、压缩域和非压缩域混编互通。

- 阶段二，实验室控制器、转发设备管理面和控制面互通场景

测试环境如图，中兴、华为、新华三多厂商路由器分别作为 PE 及 P 节点部署 SRv6 头压缩，控制器厂商包括中兴、华为、中盈。

测试内容：

控制器与各厂家转发设备实现管里面和控制面互通。

测试涉及 G-SRv6 压缩 SID 控制面互通、动态部署 SRv6 Policy 承载 BGP L3VPN 业务，纯压缩域（相同及不同网络前缀）、压缩域非压缩域混编互通，SBFD 等 OAM 及 HSB 可靠性验证。

- 阶段三，现网厂家试点场景

试点环境由现网业务类型决定，新增设备与现网设备组成压缩域与非压缩域混编组网。预计参与试点厂商为华为、中兴、新华三、中盈等，实现 BGP L3VPN over G-SRv6 一键式业务部署。

## 5.2 互通验证结果

2020 年 7 月初中国移动实验室完成设备厂商华为、中兴、新华三 G-SRv6 原型机数据面实验室互通测试，包括了单压缩域、跨压缩域、压缩域与非压缩域混编场景，P/PE 多类型路由器设备，并于 2020 年 8 月先后完成设备厂商锐捷白盒交换机和芯片厂商盛科、博通芯片样机 G-SRv6 数据面实验室互通测试。测试案例如 5.1 节所述，覆盖



G-SRv6 压缩优化的核心技术特征，实现单压缩域内、压缩域和非压缩域混编、跨多个压缩域的严格和松散路径互通。互通测试的顺利完成，验证了 G-SRv6 方案的可行性，并体现方案的高标准化水平，为下一阶段测试提供了坚实的基础。

接下来中国移动会如期开展阶段二和阶段三的实验室测试和现网试点工作。

## 6 总结与展望

随着 IPv6 在我国的加速部署，SRv6 的快速部署已成为大势所趋。但 SRv6 报文头开销大，尤其在多跳严格路径 TE、SFC 等场景下，SRv6 的部署效益受到较大限制。如何有效压缩 SRv6 报文头开销，已经成为当前国内外的研究热点。

中国移动充分认识到 SRv6 商用部署迫切性，同时也考虑到 SRv6 头压缩问题的重要性，携手华为、中兴、新华三、锐捷等设备厂商，博通、英特尔、美满、盛科等芯片厂商以及思博伦等测试仪厂商，共同推动 G-SRv6 压缩方案的发展。通过整合产业链各方力量，率先在国内形成产业生态，推动国内 SRv6 压缩标准的成熟。在国际标准合作上，也与思科等海外设备厂商和贝尔加拿大等海外运营商合作，稳步推进国际标准发展成熟，进一步提升 G-SRv6 标准在国际的影响力。

在 32bit 压缩中，G-SRv6 理论上可以将原有 SID 的封装效率提高到 75%左右，显著减少了报头开销，拓展了 SRv6 的应用部署范围。同时，G-SRv6 兼容 SRv6，支持压缩 SRv6 与原生 SRv6 混合编程，天

然支持与原生 SRv6 互联互通，可支持从普通 SRv6 存量演进，平滑升级到 G-SRv6。产品开发上，第一阶段多厂商互通测试已经在中国移动实验室完成，标志着 G-SRv6 技术方案的可行性已经得到验证，初步达成产业共识。

展望未来，未来网络将会是一个能力开放的网络，业务创新将基于全球网络和云基础设施的能力整合，进一步带来端、边、云、网的再分工。接下来，中国移动将稳步推动 G-SRv6 的方案试点落地，以商业部署为目标，加速推动产业生态成熟，形成完整产业链。让我们共同推动这一时代的到来。

## 缩略语列表

缩略语	英文全名	中文解释
IP	Internet Protocol	互联网协议
IPv6	Internet Protocol version 6	因特网协议版本 6
SR	Segment Routing	分段路由
SRv6	Segment Routing of IPv6	基于 IPv6 的分段路由
SID	Segment Identifier	段标识符
SRH	Segment Routing Header	分段路由扩展头
G-SID	Generalized SID	通用段标识符
G-SRv6	Generalized SRv6	通用 SRv6
G-SRH	Generalized SRH	通用分段路由扩展头
COC	Continue of Compression Flavor	指示本 SID 之后是压缩的 32 bits G-SID
SL	Segment Left	剩余 segment 数量
SI	Generalized SID Index	通用 SID 位置, 用于标识通用 SID 在 G-SID Container 中的位置
ISIS	Intermediate System-to-Intermediate System	中间系统到中间系统的域内路由信息交换协议
BGP	Border Gateway Protocol	边界网关协议
BGP-LS	BGP Link-State	BGP 链路状态
IGP	Interior Gateway Protocol	内部网关协议
SDN	Software Defined Network	软件定义网络
TE	Traffic Engineering	流量工程
LAN	Local Area Network	局域网
SAFI	Subsequent Address Family Identifier	子地址族标识
NLRI	Network Layer Reachable Information	网络层可达信息
MP-BGP	Multiprotocol extensions for BGP-4	BGP-4 的多协议扩展
VPN	Virtual Private Network	虚拟专用网
DA	Destination Address	目的地址
P	Provider	运营商骨干路由器
PE	Provider Edge	服务提供商网络边缘

OAM	Operation Administration and Maintenance	运行、管理和维护
SBFD	Seamless bidirectional forwarding detection	无缝双向转发检测
HSB	Hot standby	热备份
TI-LFA	Topology Independent Loop free Alternate	拓扑无关的无环路备份
SFC	Service Function Chain	服务链

## 参考文献

- [1] 程伟强,刘毅松,姜文颖,张庚. G-SRv6 头压缩优化技术研究与应用. 电信科学
- [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-bgppls-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