

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: 8 September 2022

Y. Liu  
F. Yang  
China Mobile  
A. Wang  
China Telecom  
X. Zhang  
China Unicom  
X. Geng  
Z. Li  
Huawei  
7 March 2022

MSR6(Multicast Source Routing over IPv6) Use Case  
draft-liu-msr6-use-cases-00

## Abstract

This document introduces the use cases for MSR6, including DCN and SD-WAN.

## Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

## Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on 8 September 2022.

## Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

Internet-Draft

[draft-liu-MSR6-use-cases-00](#)

March 2022

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Revised BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	MSR6 in DCN . . . . .	<a href="#">3</a>
<a href="#">3.</a>	MSR6 in SD-WAN . . . . .	<a href="#">6</a>
<a href="#">4.</a>	IANA Considerations . . . . .	<a href="#">8</a>
<a href="#">5.</a>	Security Considerations . . . . .	<a href="#">8</a>
<a href="#">6.</a>	Acknowledgements . . . . .	<a href="#">8</a>
<a href="#">7.</a>	Normative References . . . . .	<a href="#">8</a>
	Authors' Addresses . . . . .	<a href="#">9</a>

## [1.](#) Introduction

MSR6(multicast source routing over IPv6) ([\[I-D.cheng-spring-ipv6-msr-design-consideration\]](#)) defines multicast source routing over IPv6, just as SRv6 for unicast.

MSR6 encapsulation reuses IPv6 Header which could go through non-MSR6 IPv6 node and easily be used with other existing IPv6 extension headers.

Source routing brings no flow status in intermediate nodes and efficient encapsulation expense, which guarantees scalability. IPv6 extension header can be used for indicating multicast replication information.

MSR6 has two basic modes of forwarding: one is based on Shortest Path First(SPF), which is called MSR6 BE mode; the other is based on traffic engineered, which is called MSR6 TE mode.

When a multicast data packet enters an MSR6 domain, the ingress node encapsulates the packet with an IPv6 header with MSR6 function. The destination address of the IPv6 header steers the packet to the next replication node and the replication node replicates the packet based on MSR6 function, updates the destination address and iterates this process. Similar as SRv6 process. There are multiple methods to indicate multicast replication function and some of the potential solutions have been defined in individual drafts, independent with using bitstring or not. The existing multicast forwarding methods have been defined in IETF could be reused , e.g., BIER, P2MP Tree SID.

MSR6 could be used in the existing or potential multicast scenarios with the following benefits: 1. Native IPv6 based design, to enable multicast in an unified method with unicast. 2. Source Routing to achieve high scalability.

This document focuses on 2 of the use cases: DCN and SD-WAN, where MSR6 is more suitable compared to other existing multicast solutions defined in IETF.

## [2.](#) MSR6 in DCN

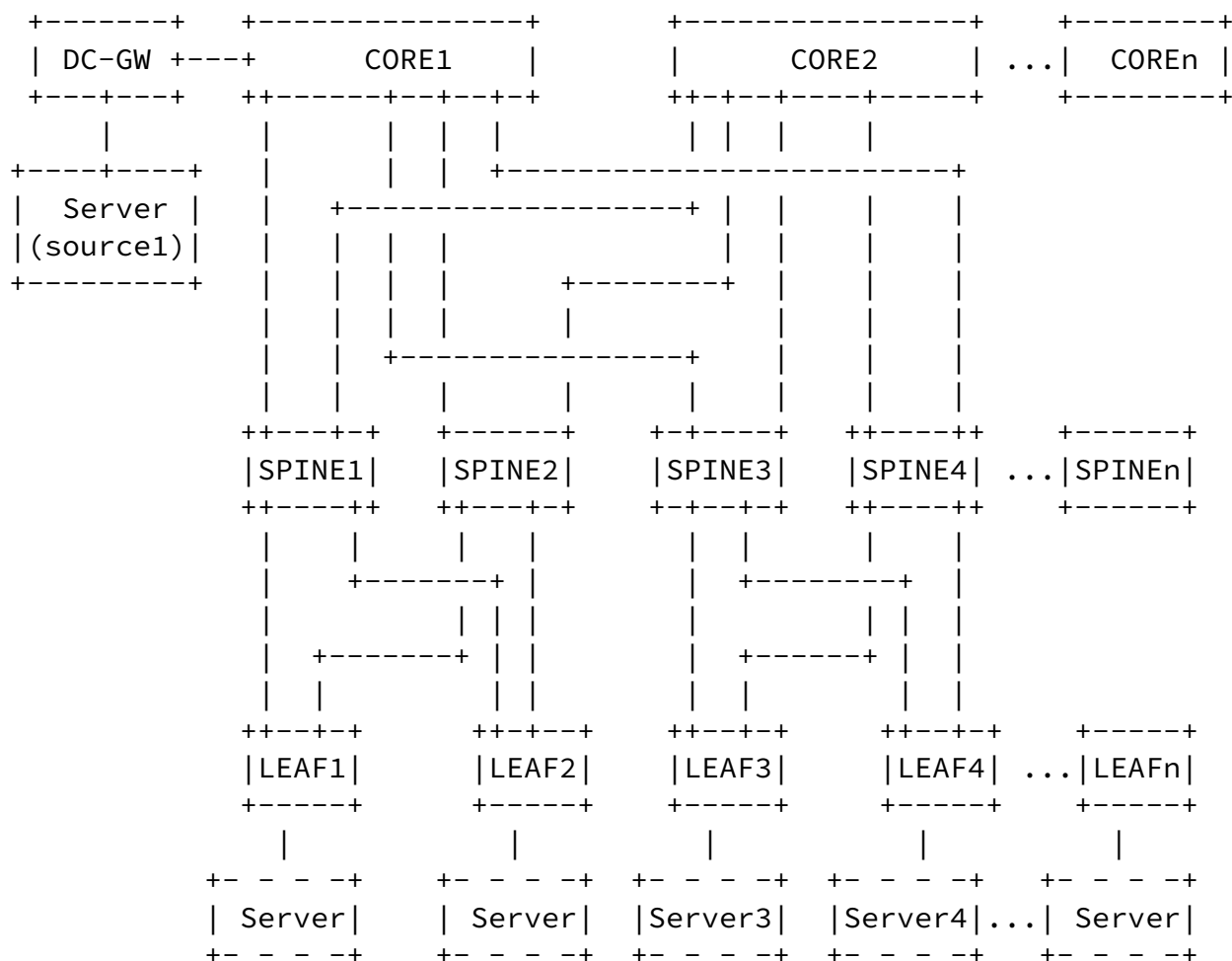
An increasing number of organizations operate dual-stack IPv4/IPv6 networks. Adoption of IPv6 has been delayed, partly due to network address translation (NAT) with IPv4, which takes private IP addresses and turns them into public IP addresses. An increasing number of organizations are adopting IPv6 in their clouds, driven by the public IPv4 space exhaustion, private IPv4 scarcity, especially within large-scale networks, and the need to provide connectivity to IPv6-only clients.

There are applications in data center with point-to-multipoint communication patterns that would benefit from native multicast. All these applications, when migrating to public clouds, will use host-based packet replication techniques. This leads to inefficiencies

for both tenants and providers, which inflates CPU load in datacenters, but also prevents tenants from sustaining high throughputs and low latencies for multicast workloads.

MSR6 could simplify multicast deployment in Data Center Network(DCN) with the capability of IPv6 based source routing.

The following figure shows an example of a data center network with dual-homes hosts for reliability. There are about 10k switches, 9k of which are leaves, and 100k adjacencies.



For multicast application in DCN, multicast source could be inside DCN or outside DCN.

For example, a multicast stream could be from source1 to service 3 and server 4. The multicast tree is from DC-Gateway to LEAF3 and LEAF4, which could be presented as:

```
DC-GW(ingress
node)-->CORE1---->SPINE3--[replicate]--->LEAF3(leaf)+LEAF4(leaf)
```

If BIER defined in [[RFC8279](#)] is used for P2MP tunnel in the network, bit position should be allocated for all egress nodes, i.e., 9k bit positions for all leaves a. Most of the bit positions are 0 and only 2 of them are set in the example. In this case, the BIER Header is inefficient and the encapsulation expense is unacceptable. Considering that the number of bit position also determines the the BIFT entry size, forwarding speed may also be affected.

There are some possible methods to improve the situation in BIER. For example "set" could be used to save the cost of bit position, but multiple packets are supposed to be sent when the BFR-ID of the

receivers belong to different set. And when the network size is large, the usefulness of set is not obvious. In the case showed above, even 10 Sets are planned, there needs about 9 hundreds bit positions for each packet and different set requests different BIFTs in each node.

In BIER-TE, BitStrings need to carry bits to indicate not only the receiving BFER but also the intermediate hops/links across which the packet must be sent. For the most common case, bit position should be allocated for all adjacencies. About 100k bit positions are requested. The bit position representing adjacencies that the multicast tree goes through are set and the rest of the bit positions are set to 0. In the example above, 7 bit positions are set in the bitstring. BIER-TE header is less efficient and the encapsulation expense is more significant, even compared to BIER. Also controller is supposed to allocate different BIFTs for 10k nodes;

Some methods introduced by [[I-D.ietf-bier-te-arch](#)] to improve the situation. "Set" could also be used, but not enough as the analysis above. There are some other methods for reducing the number of required bits, such as unicast (forward\_routed()), ECMP() or flood (DNC) over "uninteresting" sub- parts of the topology, which brings

different kinds of limitation for path planning.

In MSR6, only the nodes/adjacencies in the multicast tree are indicated in the packet just like the segment list in SRv6 used for unicast. Packet encapsulation overhead is proportional to the number of nodes or links through which the multicast tree passes (the encapsulation overhead of the path indication is 3\*segment length, which is 3\*128 bits in the example above). Local Bitstring (reduce the number of segments in the segment list, since it is no longer necessary to represent leaf nodes with segments, where the encapsulation overhead of the path indication is 1\*segment length, which is 1\*128 bits in the example above) and compression similar as SRv6 (reduce the length of each segment, where the encapsulation overhead of the path indication is 3\*segment length, which is 3\*32 bits, if the length of compressed segment is 32 bits, in the example above) could be used to reduce header expense further.

MSR6 enhances the scalability for multicast: when the multicast domain is large while the multicast tree is small, only the information of the multicast tree could be encapsulated in the packet. It means that the multicast packet encapsulation efficiency is not affected by the number of nodes/adjacencies in the network, which enhances the scalability of multicast domain scale.

### [3.](#) MSR6 in SD-WAN

[I-D.ietf-bess-bgp-sdwan-usage] discusses the usage and applicability of BGP as the control plane for multiple SDWAN scenarios.

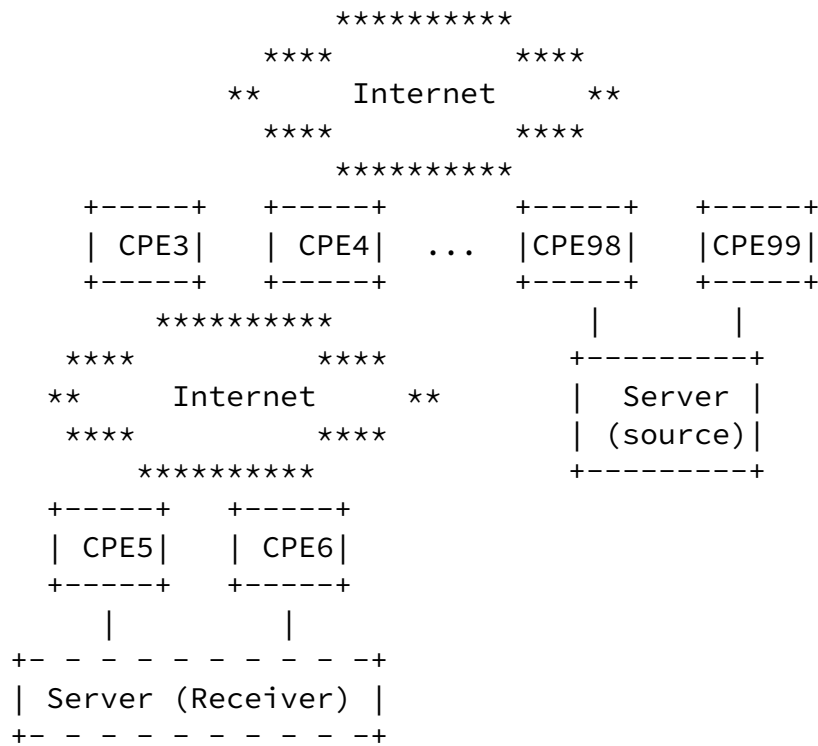
[[I-D.dukes-sr-for-sdwan](#)] and

[[I-D.dunbar-sr-sdwan-over-hybrid-networks](#)] describe how SR/SRv6 could be used in SD-WAN senario.

Security is one of the fundamental requiremnt in SD-WAN network.

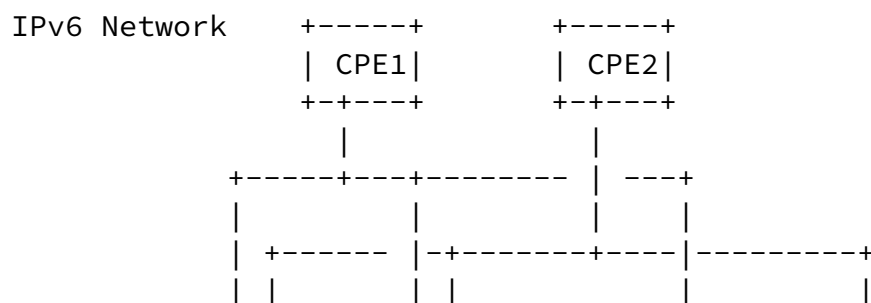
Multicast services for SD-WAN also request encryption. The following figure shows an example of SD-WAN multicast.

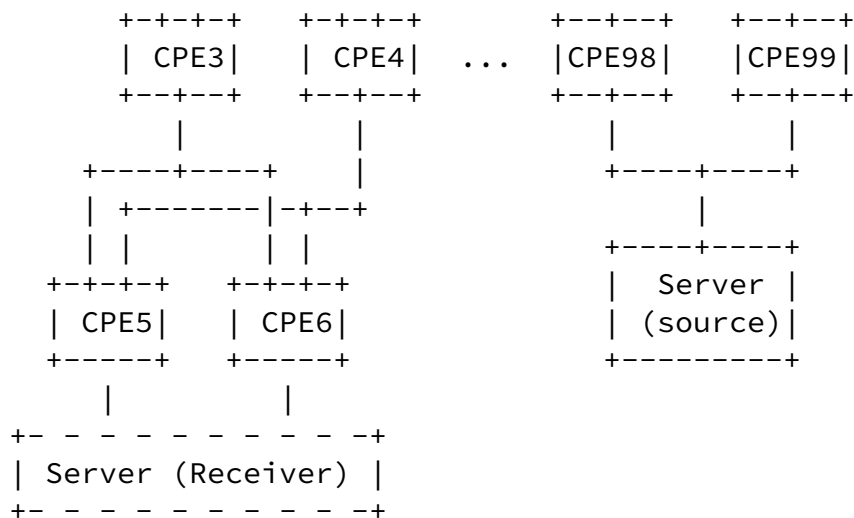
IPv6 Network	+-----+	+-----+
	CPE1	CPE2
	+-----+	+-----+



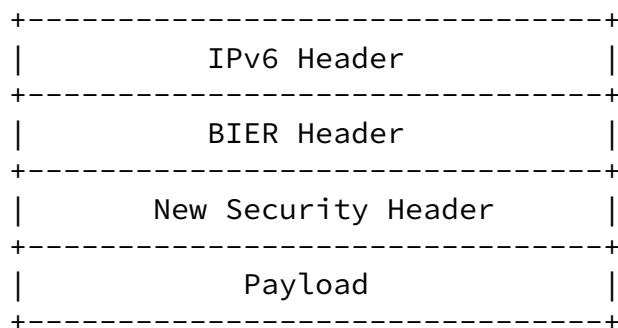
A multicast case in SD-WAN is from CE99 to CE3, CE5 and CE6. The multicast tree could presented as:

CE99(ingress node)-->CPE2--[replicate]-->CE3(leaf)+CE4--[replicate]-->CE5(leaf)+CE6(leaf)

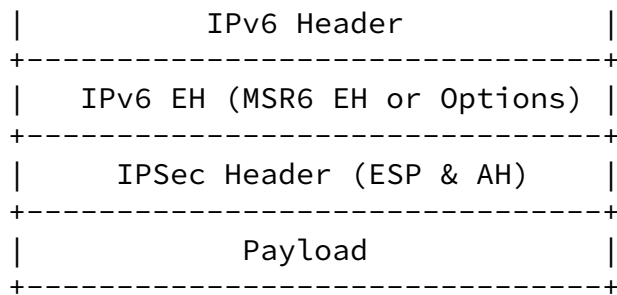




The independent layer design of BIER brings challenges to support authentication and security over Internet: a new Security Header, like IPsec, has to be defined in the BIER layer. If BIER is used in this case, the packet is supposed to encapsulated as the following to implement end to end multicast encryption:



For MSR6, which is designed based on native IPv6, it is allowed to reuse IPv6 Authentication header and Encapsulating Security Payload header. If MSR6 is used in this case, the packet is supposed to encapsulated as the following to implement end to end multicast security:



Just as IPsec, there are other existing functionalities that have been in IETF based on IPv6, for example fragmentation, network slicing, IOAM etc, which could all be reused in MSR6 which is based on IPv6 data plane. Comparingly, it has to be defined again if these functions/header are supposed to be used in BIER, which brings redundancy.

#### 4. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

#### 5. Security Considerations

#### 6. Acknowledgements

#### 7. Normative References

[I-D.cheng-spring-ipv6-msr-design-consideration]  
Cheng, W., Mishra, G., Li, Z., Wang, A., Qin, Z., and C. Fan, "Design Consideration of IPv6 Multicast Source Routing (MSR6)", Work in Progress, Internet-Draft, [draft-cheng-spring-ipv6-msr-design-consideration-01](https://www.ietf.org/archive/id/draft-cheng-spring-ipv6-msr-design-consideration-01), 25 October 2021, <<https://www.ietf.org/archive/id/draft-cheng-spring-ipv6-msr-design-consideration-01.txt>>.

[I-D.dukes-sr-for-sdwan]  
Dukes, D., Filsfils, C., Dawra, G., Garvia, P. C., Clad, F., and S. Salsano, "SR For SDWAN: VPN with Underlay SLA", Work in Progress, Internet-Draft, [draft-dukes-sr-for-sdwan-01](https://www.ietf.org/archive/id/draft-dukes-sr-for-sdwan-01), 27 April 2018, <<https://www.ietf.org/archive/id/draft-dukes-sr-for-sdwan-01.txt>>.

[I-D.dunbar-sr-sdwan-over-hybrid-networks]

Dunbar, L. and M. Toy, "SRv6 across SDWAN paths", Work in Progress, Internet-Draft, [draft-dunbar-sr-sdwan-over-hybrid-networks-07](#), 18 May 2021, <<https://www.ietf.org/archive/id/draft-dunbar-sr-sdwan-over-hybrid-networks-07.txt>>.

[I-D.ietf-bess-bgp-sdwan-usage]

Dunbar, L., Guichard, J., Sajassi, A., Drake, J., Najem, B., and D. Carrel, "BGP Usage for SDWAN Overlay Networks", Work in Progress, Internet-Draft, [draft-ietf-bess-bgp-sdwan-usage-04](#), 18 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-bess-bgp-sdwan-usage-04.txt>>.

[I-D.ietf-bier-te-arch]

Eckert, T., Menth, M., and G. Cauchie, "Tree Engineering for Bit Index Explicit Replication (BIER-TE)", Work in Progress, Internet-Draft, [draft-ietf-bier-te-arch-12](#), 28 January 2022, <<https://www.ietf.org/archive/id/draft-ietf-bier-te-arch-12.txt>>.

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

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", [RFC 8279](#), DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.

Authors' Addresses

Yisong Liu  
China Mobile  
Email: [liuyisong@chinamobile.com](mailto:liuyisong@chinamobile.com)

Feng Yang  
China Mobile  
Email: [yangfeng@chinamobile.com](mailto:yangfeng@chinamobile.com)

Aijun Wang

China Telecom  
Email: wangaj3@chinatelecom.cn

Liu, et al.

Expires 8 September 2022

[Page 9]

---

Internet-Draft

[draft-liu-MSR6-use-cases-00](#)

March 2022

Xueru Zhang  
China Unicom  
Email: zhangxr49@chinaunicom.cn

Xuesong Geng  
Huawei  
Email: gengxuesong@huawei.com

Zhenbin Li  
Huawei  
Email: lizhenbin@huawei.com

