

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 28, 2022

W. Cheng
China Mobile
G. Mishra
Verizon Inc.
Z. Li
Huawei Technologies
A. Wang
China Telecom
Z. Qin
China Unicom
C. Fan
New H3C Technologies
October 25, 2021

Design Consideration of IPv6 Multicast Source Routing (MSR6)
draft-cheng-spring-ipv6-msr-design-consideration-01

Abstract

This document discusses the design consideration of IPv6 source routing multicast solution.

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 April 28, 2022.

Internet-Draft

Design Consideration of MSR6

October 2021

Copyright Notice

Copyright (c) 2021 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 (<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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Terminology	3
3.	Reference Mode	3
4.	Deployment Modes	4
4.1.	Conventional Multicast Deployment	4
4.2.	Host-Initiated Multicast Deployment	5
4.3.	Multicast Overlay Network	6
5.	Design Consideration	7
6.	Requirements	9
6.1.	Path Programming	9
6.2.	Resource Assurance	9
6.3.	Deterministic Delay	9
6.4.	Performance Measurement	10
6.5.	Reliability	10
6.6.	Forwarding Efficiency	10
7.	Use Cases	11
7.1.	MVPN or GTM service	11
7.2.	File Delivery	11
7.3.	WebRTC Relaying	12
8.	IANA Considerations	12
9.	Security Considerations	12
10.	Acknowledgements	12
11.	Normative References	12
	Authors' Addresses	15

[1.](#) Introduction

Multicast could provide efficient P2MP service without bandwidth waste. The increasing amount of live video traffic in the network bring new requirements for multicast solutions. Traditional multicast solutions request multicast tree-building on control plane

and maintaining end-to-end tree state per flow, which impacts router state capacity and network convergence time.

There has been a lot of work in IETF to simplify service deployment, in which Source Routing is a very important technology. Source routing is able to reduce the state of intermediate nodes and indicate unicast forwarding in the ingress nodes, which could simplify unicast deployment. Source routing requires sufficient flexibility on the forwarding plane and IPv6 has the advantage with good scalability. Particularly, based on the IPv6 data plane, SRv6 could be deployed not only in a controlled domain where routers and hosts are running SRv6 cooperatively.

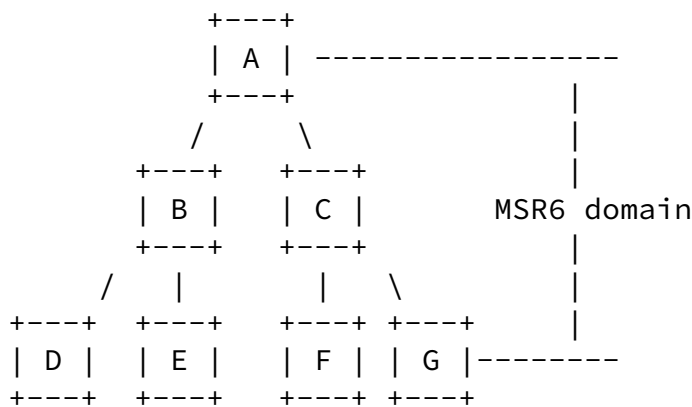
With the same stateless design, there is another work called Bit Index Explicit Replication (BIER) [[RFC8279](#)] introduced in IETF. The BIER solution is different than the traditional multicast and the benefits have been understood by the community. However, the BIER architecture is designed to be Layer-2.6 independent layer, and aims to be used in a domain comprised of routers only, where the Ingress router encapsulate a received multicast packet with BIER header and traverse through the domain, wherein the Egress router decapsulate the multicast packet from the BIER encapsulation and forward the multicast packet further as traditional multicast does.

With the deployment of SRv6, there is arising requirements to build solutions where hosts and routers are cooperatively deployed in a controlled domain, and stateless multicast solution start from hosts. Therefore, it is important to also have a layer-3 stateless multicast solution based on IPv6 data plane, which we called multicast source routing (MSR6).

This document discusses the design consideration of IPv6 multicast source routing (MSR6) solution. It combines the SRv6 Network programming concept and BIER concept to build a new Layer-3 stateless multicast solution.

[2. Terminology](#)

[3. Reference Mode](#)



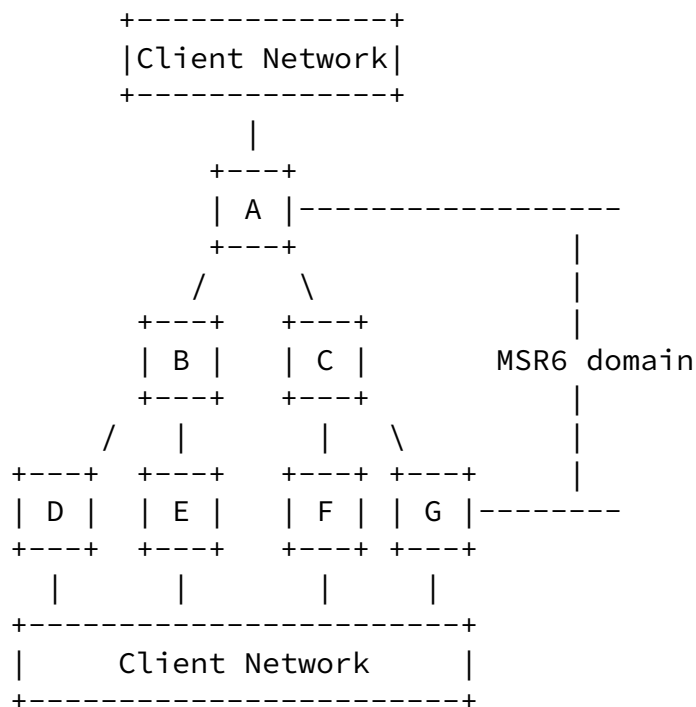
In the reference topology illustrated in Figure 1, it is assumed:

- o The MSR6 domain is comprised by node A, B, C, D, E, F and G.
- o Node A is root node. Node B and C are transit nodes. Node D, E, F and G are leaf nodes.
- o The root and leaf nodes can be hosts or routers.
- o The transit nodes are routers functioning the MSR6 routing and forwarding procedures.

[4. Deployment Modes](#)

[4.1. Conventional Multicast Deployment](#)

The first deployment mode is depicted in below figure.

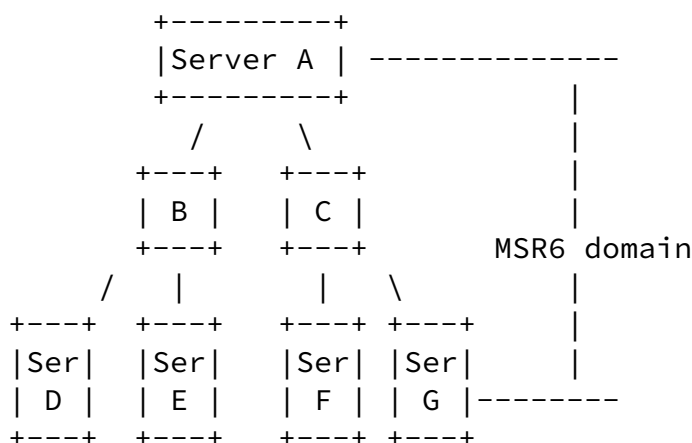


In this case, MSR6 domain is comprised of routers only. The Ingress router A encapsulates a multicast packet received from client Network

with a tunnel header and the encapsulated packet will traverse through the domain, wherein the Egress router decapsulate the multicast packet from the BIER encapsulation and forward the multicast packet further to the client network. This is similar to the BIER architecture with a difference that the the MSR6 uses IPv6 based dataplane.

4.2. Host-Initiated Multicast Deployment

The second deployment mode is dipicted in below figure.

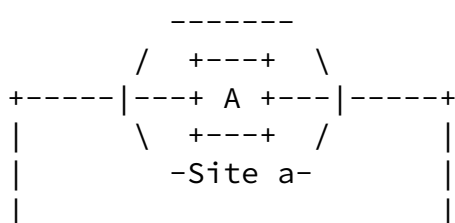


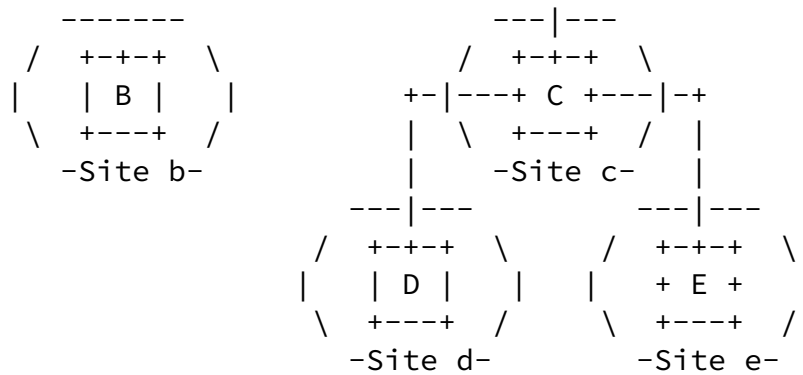
In this case, MSR6 domain is comprised of hosts and routers, where node A/D/E/F/G are hosts which is usually a server instead of the

user-equipment(UE). MSR6 packet is originated at Server A, and destined to Server D, E, F and G.

4.3. Multicast Overlay Network

In the Multicast Overlay Network, there is a multicast proxy node in each site, and the contents allocation from one site to serveral sites could be implemented by the multicast proxy nodes.





An example of the Multicast Overlay Network could be as follows:

- o Each client of the sites collect the information of the multicast proxy nodes and their connection relationship;
- o Several clients, from site b,d,e, request the same contents in one client X of site a at the same time;
- o Client X generates a P2MP path from Site a to site b,d,e and the path is encoded into an MRH A;
- o Client X sends out the requested contents by an IPv6 packet with MRH A, indicating the P2MP path explicitly;
- o The packet is transported to the clients in site b,d,e through proxy A in site a and Proxy C in site c;

5. Design Consideration

Firstly, MSR6 needs to support the basic multicast functionalities, including:

- o P2MP Forwarding: replicate and forward multicast packet to the next replication nodes;

- o Multicast Flow Overlay: multicast service, such as MVPN
- o P2MP OAM functions: Ping/Traceroute/BFD

In addition to this, it is necessary for MSR6 to meet the need of high quality service with high reliability, including:

- o Traffic Engineering: explicit path specification to satisfy different kinds of requirements
- o FRR
- o E2E Protection
- o Advanced network measurement functions, including: performance measurement and In-situ Flow Information Telemetry, which is the basis for traffic engineering and high quality transport service.

The IPv6 multicast source routing should take use of the advantages of source routing to reduce the state of the network as much as possible. That is, it should satisfy the above requirements with high scalability.

However, MSR6 is not about starting from scratch. The existing IETF work should be reused as much as possible:

- o BIER [[RFC8279](#)] and [[RFC8296](#)]

Bit Index Explicit Replication (BIER) defined in [[RFC8279](#)] is an architecture providing optimal multicast forwarding without requiring intermediate routers to maintain any per-flow state by using a multicast-specific BIER header. BIER use bitstring in the BIER header to indicate leaf nodes which gives an efficient solution for stateless multicast solution. flow without the requirement of Traffic Engineering.

- o SRv6([RFC8986](#))

SRv6 has advantages in indicating explicit paths, which brings convenience for unicast TE and FRR. MSR6 TE should refer to the

experience of SRv6. In addition, SRv6 provides flexible path

programming capability with the definition of different type of segments. MSR6 could make use of existing segments in the design of TE/FRR. For example, path segment ([\[I-D.ietf-srv6-path-segment\]](#)) could help to enhance the performance measurement capability. In the meantime, SRv6 compression ([\[I-D.srcompdt-spring-compression-requirement\]](#)) is under discussion to reduce encapsulation overhead, which could also be reused by MSR6.

- o The existing and ongoing IPv6 extensions

- 1) Existing functionalities including fragmentation and security

Multicast packets need to be fragmented and secured as they pass through the IPv6 network. This can be done using IPv6 Fragmentation and ESP(Encapsulating Security Payload) defined in [\[RFC8200\]](#). Work about Path MTU [\[I-D.ietf-idr-sr-policy-path-mtu\]](#) which supports fragmentation, is also under discussion. All these existing work should be reused in the MSR6.

- 2) New network functionalities based on the ongoing IPv6 Extensions, including Network Slicing, Deterministic Networking(DetNet), IOAM.etc.

Network slicing ([\[I-D.ietf-teas-ietf-network-slices\]](#)) and DetNet ([\[RFC8655\]](#)) are being introduced to satisfy the quality service requirements of critical services. IOAM ([\[I-D.ietf-ippm-ioam-data\]](#)) is also introduced to implement in-situ network measurement. IPv6 data plane is being used to support network slicing ([\[I-D.li-6man-e2e-ietf-network-slicing\]](#) and [\[I-D.dong-6man-enhanced-vpn-vtn-id\]](#)), Detnet ([\[I-D.geng-spring-srv6-for-detnet\]](#) and [\[I-D.geng-spring-sr-redundancy-protection\]](#)), IOAM ([\[I-D.ietf-ippm-ioam-data\]](#)), etc. Multicast service can also benefit from these new network functionalities to improve quality of service. MSR6 could reuse the ongoing work based on IPv6 extensions to implement the functionalities for multicast services.

- 3) Future possible work based on IPv6 extensions, including Application Aware Network (APN)

APN ([\[I-D.li-apn-framework\]](#)) is used to provide more granular services, which can use IPv6 extension header to carry APN information for the purpose of steering traffic, etc. MSR6 can combine with APN to map the traffic to different Network-based multicast services/functionalities according to the APN information in the IPv6 data plane.

4) MSR6 is also supposed to be started at the Host based on IPv6

In [[RFC8754](#)], it is supposed that a host can originate the IPv6 source routing packet. MSR6 should take use of the native IPv6 design and support originating the IPv6 packet by the host.

[6.](#) Requirements

This section describes the overall requirements which need to be fulfilled by MSR6.

[6.1.](#) Path Programming

When the multicast root and leafs are determined, the multicast service may be forwarded along different multicast paths/trees. Each multicast tree has different performance attribute, such as bandwidth, delay, etc. For instance, tree A is the shortest path from root 1 to leaf 1-100, which has the lowest delay, while the other tree B from the same root to same leafs can provide more bandwidth than tree A. Therefore, the delay-sensitive traffic like cloud AR/VR traffic SHOULD be forwarded along with tree A, and the traffic that is delay-insensitive but requiring large bandwidth like 8k HD Video be forwarded along with tree B.

In order to meet the different SLA requirements, IPv6-based MSR6 MUST support path programming.

[6.2.](#) Resource Assurance

Network slicing is another method to provide SLA differentiated services, because it is able to provide separate and independent logical network over the physical network infrastructure. Each Network Slicing has its own allocated resource, which can also meet the specific SLA requirements for different multicast service. Or an independent network slice could be allocated to all the multicast service, which could provide multicast service with better transport performance than normal unicast service.

So in order to provide SLA guaranteed services, MSR6 SHOULD support network slicing.

[6.3.](#) Deterministic Delay

Some delay-sensitive multicast traffic may have the strict requirements for network delay, especially in some scenario of IoT or enterprise. Deterministic Networking (DetNet) defined in [[RFC8655](#)]

could provide service with E2E latency upper bound for both unicast and multicast. Therefore, MSR6 should support DetNet.

[6.4.](#) Performance Measurement

Many OAM mechanisms are used to support network operation. Performance Measurement (PM) is one of the most important part of OAM. With PM, the real-time QoS of the forwarding path, like delay, packet loss ratio and throughput, can be measured.

PM can be implemented in one of three ways: active, passive, or hybrid defined in [\[RFC7799\]](#), differing in whether OAM packets need to be proactively sent.

On-path telemetry, defined in [\[I-D.song-opsawg-ifit-framework\]](#), is an hybrid mode OAM/PM mechanism, which provides better accuracy than active PM.

Some multicast scenarios have strong requirement for PM, therefore on-path Performance Measurement SHOULD be supported in MSR6.

[6.5.](#) Reliability

In scenarios where multicast is used to carry video streams, packet loss can lead to impaired video quality. So high reliability is required in these scenarios. E2E protection and local protection of node and links MUST be supported in MSR6. Furthermore, root and leaf node protection SHOULD be supported.

[6.6.](#) Forwarding Efficiency

Multicast is used for saving bandwidth cost with the capability of replication and forwarding.

Path Maximum Transmission Unit indicates the maximum size of a packet that it can be forwarded along a path. Setting an appropriate PMTU for packets can avoid fragmentation or dropping, so that the forwarding efficiency can be raised.

Also, the overhead of packets MUST be added very carefully since it affects the forwarding efficiency directly. For example, when many SIDs are inserted in an SRv6 packet, the overhead of the SID list is

too large.

Therefore, the MSR6 MUST support PMTU probing and configuration. In addition, it SHOULD support compression when it is necessary.

[7.](#) Use Cases

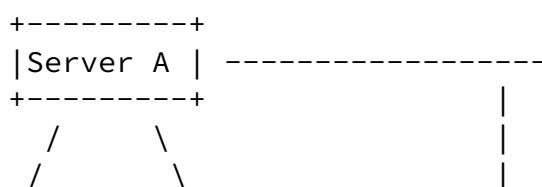
[7.1.](#) MVPN or GTM service

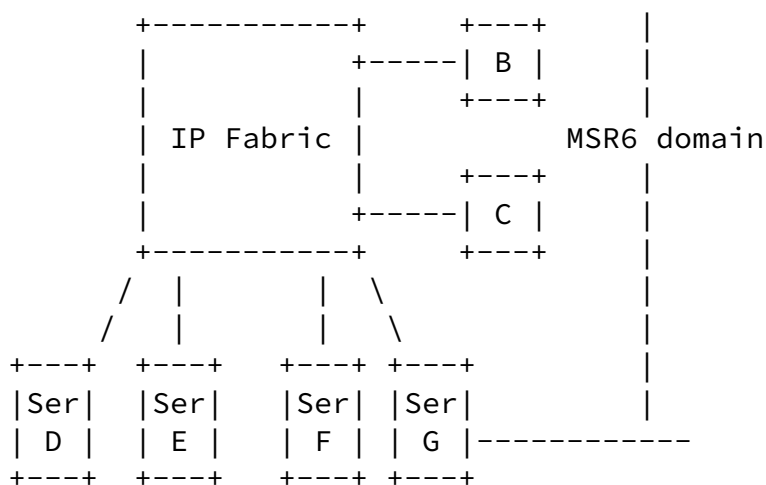
MVPN(multicast virtual private network, [RFC6513](#)) and GTM(Global Table Multicast, [RFC7716](#)) are usual cases for conventional deployment mode, where MSR6 acts as P-tunnel of c-multicast packets.

[7.2.](#) File Delivery

File Delivery over Unidirectional Transport (FLUTE, [RFC6726](#)) is a common use case for multicast. The traditional multicast or MSR6 conventional deployment mode are both valid candidates for FLUTE, where the application in host could Join a multicast Group address, and receive multicast packet carrying file content.

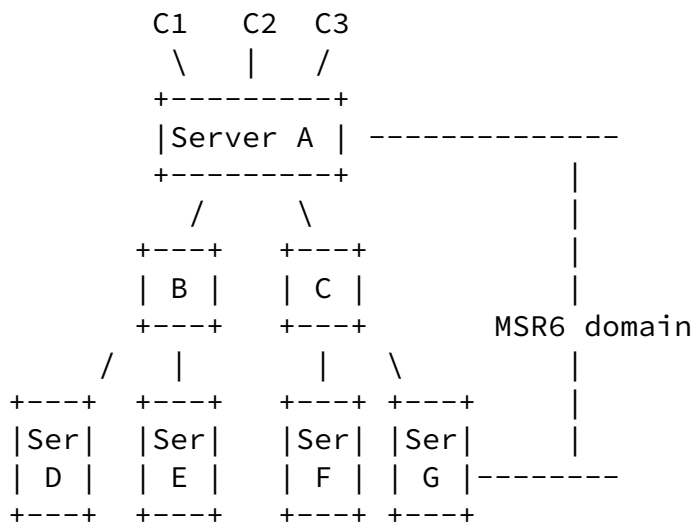
In some other cases, File delivery to multiple hosts efficiently (for example, video editing) may need within a data center network, but the data center may not support underlay multicast routing of any form in the Fabric infrastructure. In such cases, Host-initialized MSR6 could be used as figure depicted below. Server A originates a MSR6 packet carrying file content bytes and sends to B or C. Router B or C receives the MSR6 packet and replicate to Server D/E/F/G. Server D/E/F/G receives the MSR6 packet and gets the file content bytes within it. Router B or C acts as overlay Multicast Service Node [[RFC8293](#)] in this case.





[7.3.](#) WebRTC Relaying

Web Real-Time Communications (WebRTC) is now an official World Wide Web Consortium (W3C) and Internet Engineering Task Force (IETF) standard. In supporting of multiparty conference, WebRTC Gateway is usually deployed to switch audio and video packets between multiple clients. The figure depicts the case, where C1 to C7 are WebRTC client (e.g., browser), and Server A, D, E, F and G are WebRTC gateways. Within the MSR6 domain, MSR6 is used in both hosts and routers.



| | | |
C4 C5 C6 C7

[8.](#) IANA Considerations

This document makes no request of IANA.

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

[9.](#) Security Considerations

[10.](#) Acknowledgements

[11.](#) Normative References

[I-D.dong-6man-enhanced-vpn-vtn-id]
Dong, J., Li, Z., Xie, C., Ma, C., and G. Mishra,
"Carrying Virtual Transport Network (VTN) Identifier in
IPv6 Extension Header", [draft-dong-6man-enhanced-vpn-vtn-id-06](#) (work in progress), October 2021.

Cheng, et al.

Expires April 28, 2022

[Page 12]

Internet-Draft

Design Consideration of MSR6

October 2021

[I-D.geng-spring-sr-redundancy-protection]
Geng, X., Chen, M., Yang, F., Garvia, P. C., and G.
Mishra, "SRv6 for Redundancy Protection", [draft-geng-spring-sr-redundancy-protection-05](#) (work in progress),
August 2021.

[I-D.geng-spring-srv6-for-detnet]
Geng, X., Li, Z., and M. Chen, "SRv6 for Deterministic
Networking (DetNet)", [draft-geng-spring-srv6-for-detnet-01](#)
(work in progress), July 2020.

[I-D.ietf-idr-sr-policy-path-mtu]
Li, C., Zhu, Y., Sawaf, A. E., and Z. Li, "Segment Routing
Path MTU in BGP", [draft-ietf-idr-sr-policy-path-mtu-04](#)
(work in progress), October 2021.

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields

for In-situ OAM", [draft-ietf-ippm-ioam-data-15](#) (work in progress), October 2021.

[I-D.ietf-spring-srv6-path-segment]

Li, C., Cheng, W., Chen, M., Dhody, D., Gandhi, R., and Y. Zhu, "Path Segment for SRv6 (Segment Routing in IPv6)", [draft-ietf-spring-srv6-path-segment-02](#) (work in progress), May 2021.

[I-D.ietf-teas-ietf-network-slices]

Farrel, A., Gray, E., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "Framework for IETF Network Slices", [draft-ietf-teas-ietf-network-slices-04](#) (work in progress), August 2021.

[I-D.li-6man-e2e-ietf-network-slicing]

Li, Z. and J. Dong, "Encapsulation of End-to-End IETF Network Slice Information in IPv6", [draft-li-6man-e2e-ietf-network-slicing-00](#) (work in progress), April 2021.

[I-D.li-apn-framework]

Li, Z., Peng, S., Voyer, D., Li, C., Liu, P., Cao, C., Mishra, G., Ebisawa, K., Previdi, S., and J. N. Guichard, "Application-aware Networking (APN) Framework", [draft-li-apn-framework-03](#) (work in progress), May 2021.

[I-D.song-opsawg-ifit-framework]

Song, H., Qin, F., Chen, H., Jin, J., and J. Shin, "In-situ Flow Information Telemetry", [draft-song-opsawg-ifit-framework-16](#) (work in progress), October 2021.

Cheng, et al.

Expires April 28, 2022

[Page 13]

Internet-Draft

Design Consideration of MSR6

October 2021

[I-D.srcompdt-spring-compression-requirement]

Cheng, W., Xie, C., Bonica, R., Dukes, D., Li, C., Shaofu, P., and W. Henderickx, "Compressed SRv6 SID List Requirements", [draft-srcompdt-spring-compression-requirement-07](#) (work in progress), July 2021.

[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>>.

- [RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", [RFC 7799](#), DOI 10.17487/RFC7799, May 2016, <<https://www.rfc-editor.org/info/rfc7799>>.
- [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, [RFC 8200](#), DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.
- [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>>.
- [RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", [RFC 8296](#), DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/info/rfc8296>>.
- [RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", [RFC 8655](#), DOI 10.17487/RFC8655, October 2019, <<https://www.rfc-editor.org/info/rfc8655>>.
- [RFC8663] Xu, X., Bryant, S., Farrel, A., Hassan, S., Henderickx, W., and Z. Li, "MPLS Segment Routing over IP", [RFC 8663](#), DOI 10.17487/RFC8663, December 2019, <<https://www.rfc-editor.org/info/rfc8663>>.
- [RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", [RFC 8754](#), DOI 10.17487/RFC8754, March 2020, <<https://www.rfc-editor.org/info/rfc8754>>.

- [RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", [RFC 8986](#), DOI 10.17487/RFC8986, February 2021,

<<https://www.rfc-editor.org/info/rfc8986>>.

Authors' Addresses

Weiqiang Cheng
China Mobile

Email: chengweiqiang@chinamobile.com

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Zhenbin Li
Huawei Technologies

Email: lizhenbin@huawei.com

Aijun Wang
China Telecom

Email: wangaj3@chinatelecom.cn

Zhuangzhuang Qin
China Unicom

Email: qinzhuangzhuang@chinaunicom.cn

Chi Fan
New H3C Technologies

Email: fanchi@h3c.com