

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: 24 October 2025

Y. Liu
China Mobile
X. Geng
Huawei Technologies
C. Lin
New H3C Technologies
24 April 2025

Encapsulation and Forwarding Process for IPv6 Multicast Source
Routing (MSR6) Data Plane
draft-liu-pim-msr6-encapsulation-and-forwarding-01

Abstract

This document specifies the mechanism of IPv6 Multicast Source Routing (MSR6), covering the encapsulation and forwarding processes in the data plane. It introduces the hierarchical bitstring structure (HBS) to organize replication information, and represent more information than flat bit strings. It indicates multicast forwarding behavior based on the local bitstring forwarding table, requires less BIFT state and improve scalability.

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 24, 2025.

Copyright Notice

Copyright (c) 2025 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 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

1. Introduction.....	2
2. Terminology.....	3
2.1. Terms.....	3
2.1.1. Hierarchical Bitstring Structure (HBS).....	3
2.2. Abbreviations.....	3
3. MSR6 Data Packet Format.....	3
3.1. MSR6 Forwarding Design Principles.....	3
3.2. MSR6 Header Encapsulation Format.....	4
3.3. Fixed-Length HBS Unit Encapsulation Format.....	5
4. MSR6 Data Forwarding Table.....	7
4.1. Network Node Forwarding Behavior.....	7
4.2. Local BIFT Forwarding Table.....	8
5. MSR6 Data Forwarding Process.....	10
5.1. Fixed-Length Mode Forwarding.....	10
5.2. BUD Node Forwarding.....	11
5.3. Multiple-SI Mode Forwarding.....	12
6. IANA Considerations.....	13
7. Security Considerations.....	13
8. References.....	13
8.1. Normative References.....	13
Authors' Addresses.....	14

1. Introduction

The document [draft-liu-msr6-problem-statement-01] addresses the challenges in existing multicast solutions for IPv6, such as the inefficiencies in scalability and state maintenance at network nodes. It highlights the need for MSR6 to improve multicast delivery by using a hierarchical bitstring structure (HBS) to optimize routing and reduce state burden. This approach aims to support modern applications like video streaming and large-scale content distribution, offering a more scalable and efficient multicast solution for complex network environments.

This document defines the mechanism of IPv6 Multicast Source Routing (MSR6) encapsulation and forwarding processes. It is intended to guide the design, development, and testing of data plane of the device supporting MSR6.

Multicast can reduce waste of network bandwidth and improve network resource utilization, in application scenarios such as multi-party video. The existing multicast solutions require complex control plane protocols and maintain a large number of multicast tree states. IPv6 Multicast Source Routing (MSR6) takes advantage of source routing on the IPv6 data plane to simplify multicast deployment, directing multicast forwarding at the ingress node and reducing the state of intermediate nodes.

The document addresses the challenges in existing multicast solutions for IPv6, such as the inefficiencies in scalability and state maintenance at network nodes. It highlights the need for MSR6 to improve multicast delivery by using a hierarchical bitstring structure (HBS) to optimize routing and reduce state burden. This approach aims to support modern applications like video streaming and large-scale content distribution, offering a more scalable and efficient multicast solution for complex network environments.

2. Terminology

2.1. Terms

2.1.1. Hierarchical Bitstring Structure (HBS)

A packet header structure formed by the root node according to the constructed multicast tree, where the multicast replication information of each network node in the multicast tree is identified by a bitstring structure and encapsulated in a certain order.

2.2. Abbreviations

BSL: Bitstring Length

HBS: Hierarchical Bitstring Structure

MSR6: IPv6 Multicast Source Routing

3. MSR6 Data Packet Format

3.1. MSR6 Forwarding Design Principles

MSR6 data plane forwarding is defined, adhering to the following principles:

Localized Bit Allocation: Uses local bit allocation and local BIFT tables for more compact bit allocation, reducing the bitstring length in packets. The local bit inherits the "bit segment" idea of [RFC9262]. In addition to indicating the destination of multicast

data, bits can also be given some special meanings, which represent the execution of specific functions at specific locations in the network. Function is defined locally by the node. It can specify a path for forwarding, or it can perform some user-defined behaviors.

Hierarchical Structure Expression: Uses HBS Unit to express MSR6 forwarding paths, with root nodes encapsulating the constructed multicast tree's HBS Units in order. Each structure carries a bitstring that specifies the local behavior.

Increased Payload Efficiency: Various methods are provided to reduce the MSR6 header length and increase payload efficiency, including removing invalid Units, multiple-SI mode.

3.2. MSR6 Header Encapsulation Format

The MSR6 header contains the basic parameters of MSR6 forwarding process and indicates how to access the replication information structure of the current node.

The MSR6 header follows [RFC8200] and is encapsulated in the IPv6 Routing header. The encapsulation format is shown in Figure 1.

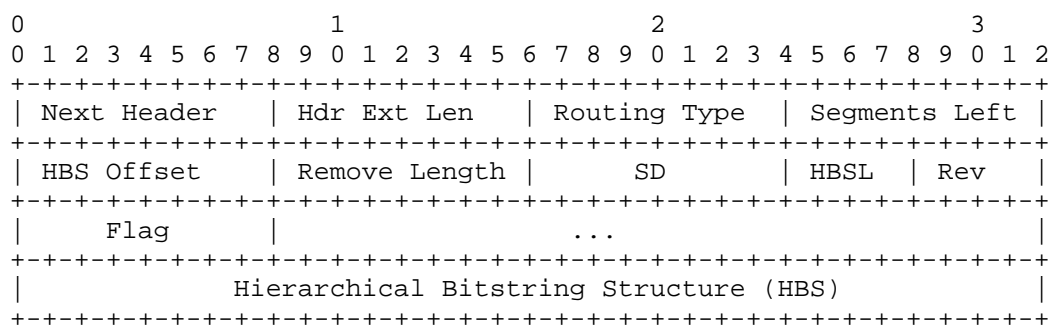


Figure 1: MSR6 Header Encapsulation Format

The fields are defined as follows:

- o Version: 4 bits, value is 6.
- o Traffic Class: 8 bits, similar to IPv4 TOS.
- o Flow Label: 20 bits, used to distinguish real-time traffic flows.
- o Payload Length: 16 bits, includes the length of the payload and extension headers.

- o Next Header: 8 bits, value 43 for Routing header.
- o Hop Limit: 8 bits, decremented by 1 at each hop.
- o Source Address: 128 bits, the sender's address.
- o Destination Address: 128 bits, the receiver's address.
- o Next Header: 8 bits, per [RFC8200].
- o Hdr Ext Len: 8 bits, per [RFC8200].
- o Routing Type: 8 bits, MSR6 type, temporarily 252.
- o Segments Left: 8 bits, set to 0 for MSR6.
- o HBS Offset: 8 bits, offset of the current HBS Unit.
- o Remove Length: 8 bits, length of removed HBS Units.
- o SD: 8 bits, sub-domain of the MSR6 packet.
- o HBSL: 4 bits, bitstring length of the current HBS Unit.
- o BHBSL: Base HBSL, 4 bits, base bitstring length of the packet.
- o Flag: 8 bits, includes the E, and OAM flags.
- o Hierarchical Bitstring Structure (HBS): Contains one or more HBS Units.

3.3. Fixed-Length HBS Unit Encapsulation Format

The HBS Unit is used to carry the bistring-based replication information of a node in the path, and also indicates the location information of multiple nodes at the next level (if any). Multiple HBS Units form a complete replication path, Each structure has the fixed-length BSL and the value is in the HBSL field of MSR6 Header.

The encapsulation format is shown in Figure 2.

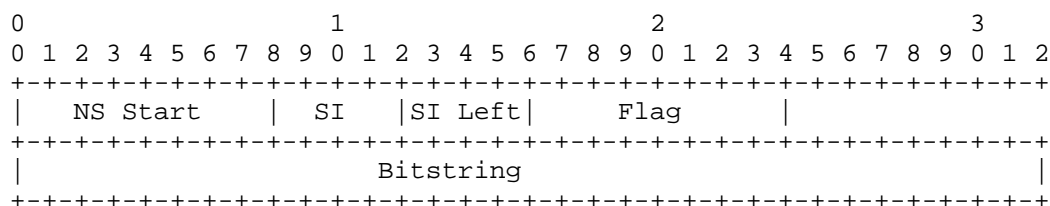


Figure 2: Fixed-Length HBS Unit Encapsulation Format

The fields are defined as follows:

- o NS Start: Next Segment Start, 8 bits, starting offset of the next HBS Unit.
- o SI: 4 bits, Set Identifier of the bitstring.
- o SI Left: 4 bits, remaining number of HBS Unit for other Set Identifier in current network node.
- o Flag: 8 bits, includes the NF, and DTP flags.
- o Bitstring: 16/32/64/128/256 bits, local bitstring for guiding packet replication.

4. MSR6 Data Forwarding Table

4.1. Network Node Forwarding Behavior

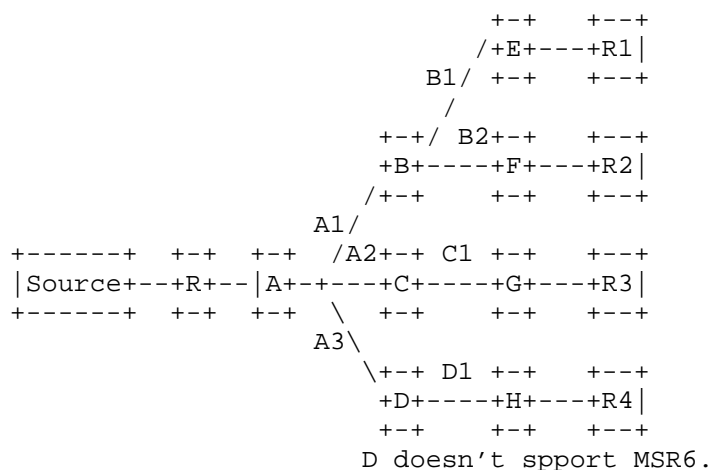


Figure 5: MSR6 Network Topology Diagram

In the network topology diagram shown in Figure 5, the nodes are defined as follows:

- o Source Node: This node connects to the multicast source, acting as the ingress node of the MSR6 network, such as node R in the diagram. It is responsible for constructing the multicast forwarding tree, encapsulating the multicast source traffic with the MSR6 header, and replicating the traffic for forwarding.
- o Intermediate Replication Node: These nodes, such as nodes A, B and C in the diagram, lie between the source node and the receiver nodes within the MSR6 network. They are responsible for table lookup and forwarding based on the MSR6 header.
- o Receiver Node: These nodes, such as nodes E, F, G and H in the diagram, connect to the multicast receivers and act as the egress nodes of the MSR6 network. They are responsible for decapsulating the MSR6 header and forwarding the packets to the multicast receivers.

4.2. Local BIFT Forwarding Table

Each network node maintains a local BIFT forwarding table containing the following information:

- o BSL: Bit String Length, one of 16 bit/32 bit/64 bit/128 bit/256 bit.
- o SD: Subdomain to which the MSR6 packet belongs.
- o SI: Set Identifier to which the bit position belongs.
- o Bit Position: Locally assigned bit position indicating a specific forwarding behavior.
- o Forwarding Attribute: Forwarding attributes associated with the bit position, such as:
 - Local_decap: Local decapsulation, where the MSR6 TE header is decapsulated locally and the packet is forwarded based on the inner protocol type.
 - Forward: The packet needs to be forwarded to the next HBS Unit and to the MSR6 neighbor according to the bit position.
- o Neighbor: The neighbor address to which the packet is forwarded (not present for Local_decap).
- o Output Interface: The output interface for forwarding the packet (not present for Local_decap).

Based on the network topology in Figure 6, the local BIFT tables for the nodes are as follows:

Source Node A's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Forward	B	A1
8	0	0	00000010	Forward	C	A2
8	0	0	00000100	Forward	G	A2

Intermediate Replication Node B's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Forward	E	B1
8	0	1	00000001	Forward	F	B2

Intermediate Replication Node C's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Forward	G	C1
8	0	0	00000100	Forward	G	C2

Receiver Node D's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Local_decap	None	None

Receiver Node E's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Local_decap	None	None

Receiver Node F's Local BIFT Table:

BSL	SD	SI	Bit Position	Forwarding Attribute	Neighbor	Output Interface
8	0	0	00000001	Local_decap	None	None

5. MSR6 Data Forwarding Process

5.1. Fixed-Length Mode Forwarding

The head node collects the necessary information and encapsulates the multicast tree information into the MSR6 TE packet header. It then offsets to the first HBS unit based on the HBS offset value in the current packet header.

o HBS Unit Position = HBS Offset Value * HBSL Byte Value.

After offsetting to the start of the HBS unit, the node uses the bit string, HBSL, and SI within the unit to find the corresponding BIFT forwarding table on the device. This table determines the neighbor nodes, output interfaces, and forwarding attributes for packet replication. The node will then perform local decapsulation (Local_decap), normal forwarding (Forward) based on the forwarding attributes.

o Local Decapsulation: The packet is decapsulated locally, and the MSR6 TE header is removed. The local multicast forwarding table is then checked based on the inner protocol type for forwarding.

o Normal Forwarding: The HBS offset value in the packet is updated to allow offsetting to the next HBS unit in subsequent processing. The packet is then forwarded to the MSR6 neighbor. The method to update the HBS offset value in the current packet is:

New HBS Offset = NS Start + bitposition in bitstring without local_decap.

5.2. BUD Node Forwarding

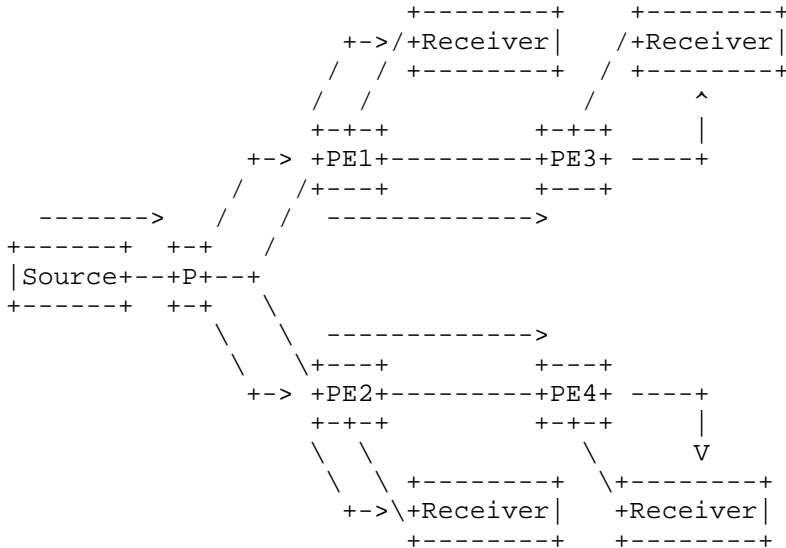


Figure 8: BUD Node Diagram

Nodes PE1 and PE2 play a BUD role, having both local receivers and lower-tier nodes like PE3. Node P cannot handle the packet by simply stripping the header, as some traffic on PE1 exits the MSR6 TE tunnel to local receivers, while other traffic continues to PE3 and PE4. When downstream BUD nodes are present, the node expects both normal forwarding and PHP forwarding simultaneously.

When a P device senses BUD nodes in the neighbor devices, it assigns two Bit Positions:

- o Type 1 Bit Position: Indicates that traffic leaving this interface has no further Segments, only local receivers.
- o Type 2 Bit Position: Indicates that traffic leaving this interface has further Segments, as shown in PE3 in the diagram.

Bit Positions for different interfaces should be as continuous as possible. For example, assuming the HBS Unit on node P identified by BSL can contain a 16-bit bit string:

- o P assigns Type 1 Bit Position 1 to PE1 and Type 1 Bit Position 2 to PE2.

- o P assigns Type 2 Bit Position 5 to PE1 and Type 2 Bit Position 6 to PE2.

Node P's HBS Unit composed of Type 1 BP is shown in Figure 9:

Next Segment Start	SI	Reserved	Bit String
FF	1	0	0000..00000011

Figure 9: HBS Unit Composed of Type 1 BP

- o Next Segment Start field is an invalid value 0xFF.

Node P's HBS Unit composed of Type 2 BP is shown in Figure 10:

Next Segment Start	SI	Reserved	Bit String
XX	1	0	0000..00110000

Figure 10: HBS Unit Composed of Type 2 BP

5.3. Multiple-SI Mode Forwarding

When the number of bits allocated by some nodes exceed the bit string length limit, same bit will be allocated in different Set Identifiers. HBS units for multiple SIs will be encapsulated, if the multicast packet needs to be replicated to multiple SIs in this node.

If SI Left field of HBS Unit greater than 0, then the node performs multiple-SI forwarding processing.

After offsetting to the start of the HBS unit for the first SI, the node performs fixed-length mode forwarding processing. The HBS offset value in the packet is updated to allow offsetting to the next SI HBS unit in subsequent processing. The packet is then performs fixed-length mode forwarding processing according to new

HBS unit. The method to update the HBS offset value in the current packet is:

Next SI HBS Offset = Next SI Start.

When an HBS Unit for one SI is processed, the SI Left field is reduced by 1. The multiple-SI forwarding process is completed until the SI Left field is reduced to 0.

6. IANA Considerations

TBD

7. Security Considerations

TBD

8. References

8.1. Normative References

- [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>>.
- [RFC9262] Eckert, T., Ed., Menth, M., and G. Cauchie, "Tree Engineering for Bit Index Explicit Replication (BIER-TE)", RFC 9262, DOI 10.17487/RFC9262, October 2022, <<https://www.rfc-editor.org/info/rfc9262>>.

Authors' Addresses

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com

Xuesong Geng
Huawei Technologies
Email: gengxuesong@huawei.com

Changwang Lin
New H3C Technologies
Email: linchangwang.04414@h3c.com

