

Workgroup: MSR6  
Internet-Draft:  
draft-eckert-msr6-problem-statement-00  
Published: 24 October 2022  
Intended Status: Informational  
Expires: 27 April 2023  
A T. Eckert, Ed.  
uFuturewei Technologies USA  
t  
h  
o  
r  
s  
:

## **Yet another Problem Statement for IPv6 Multicast Source Routing (MSR6)**

### **Abstract**

This draft summarizes additional personal opinions of the author for why existing stateless multicast solutions BIER/BIER-TE are not sufficient to support the operator, architecture and use-case goals that MSR6 proposes to solve.

This document is complementary to problems outlined in I-D.liu-msr6-problem-statement and in no way affects any of them. Instead, it attempts to look more at lower-layer functional challenges of current stateless source routing multicast solutions (BIER/BIER-TE), as well as architectural, network and protocol design ecosystem requirements of operators.

### **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 27 April 2023.

### **Copyright Notice**

Copyright (c) 2022 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. Problems](#)
  - [2.1. P1: BIER: Efficiency of delivery for sparse groups in large networks](#)
  - [2.2. P2: BIER-TE: Packet header efficiency of delivery to sparse groups in large networks](#)
  - [2.3. P3: BIER/BIER-TE: Efficient amount of per-router state \(BIFT\)](#)
  - [2.4. P4: More complex, multi-protocol host stack requirements with BIER](#)
  - [2.5. P5: Support for single, IPv6/RFC8200 compliant network forwarding plane](#)
  - [2.6. P6: No Support for SRv6 architecture](#)
  - [2.7. P7: Inability to perform per-hop IPv6 forwarding plane features](#)
  - [2.8. P8: Routing convergence speed and complexity](#)
- [3. Explanations](#)
  - [3.1. E1: BIER: Efficiency of delivery to sparse groups in large networks](#)
  - [3.2. E2: BIER-TE: Packet header efficiency of delivery to sparse groups in large networks](#)
  - [3.3. E3: BIER/BIER-TE: Efficient amount of per-router state \(BIFT\)](#)
  - [3.4. E4: More complex, multi-protocol host stack requirements with BIER](#)
  - [3.5. E5: Support for single, IPv6/RFC8200 compliant network forwarding plane](#)
    - [3.5.1. Example 1: IPv4 multicast in MPLS/VPN solutions vs. native MPLS multicast](#)
    - [3.5.2. Example 2: Native IPv6 multicast vs. IPv4 multicast](#)
    - [3.5.3. Example 3: BIERin6](#)
    - [3.5.4. Example 4: Routing for multicast](#)
  - [3.6. E6: No Support for SRv6 architecture](#)
  - [3.7. E7: Inability to perform per-hop IPv6 forwarding plane features](#)
  - [3.8. E8: Routing convergence speed and complexity in large networks](#)
- [4. Candidate MSR6 solution discussion](#)
  - [4.1. S1: BIER: Efficiency of delivery for sparse groups in large networks](#)
  - [4.2. S2: BIER-TE: Efficiency of delivery for sparse groups in large networks](#)
  - [4.3. S3: BIER/BIER-TE: Large amount of per-BFR state \(BIFT\)](#)
  - [4.4. S4: More complex, multi-protocol host stack requirements with BIER](#)
  - [4.5. S5: Support for single, IPv6/RFC8200 compliant network forwarding plane](#)
  - [4.6. S6: No Support for SRv6 architecture](#)
  - [4.7. S7: Inability to perform per-hop IPv6 forwarding plane features](#)

- [4.8. S8: Routing convergence speed and complexity](#)
- [5. Open issues](#)
- [6. Changelog](#)
- [7. Informative References](#)
- [Appendix A. Other considerations towards MSR6](#)
  - [A.1. Design separation between tree building and other functions](#)
  - [A.2. Reuse of BIER principles and technology components as much as possible](#)
  - [A.3. MSR6 as experimental \(not standards track\) work](#)
- [Author's Address](#)

## 1. Introduction

This draft summarizes additional personal opinions of the author for why existing stateless multicast solutions BIER/BIER-TE are not sufficient to support the operator, architecture and use-case goals that MSR6 proposes to solve.

This document is complementary to problems outlined in [[I-D.liu-msr6-problem-statement](#)] and in no way affects any of them. Instead, it attempts to look more at lower-layer functional challenges of current stateless source routing multicast solutions (BIER/BIER-TE), as well as architectural, network and protocol design ecosystem requirements of operators.

This draft is solely kept separate at this time, because the author felt it would be impossible for him to merge the text without doing disservice to that pre-existing draft due to time constraints and the inability to have reviewed and agreed upon this documents amongst the larger authorship of [[I-D.liu-msr6-problem-statement](#)].

This document is structured as follows.

The Problems section ([Section 2](#)) summaries problems with existing solutions and numbers them with Pi. The author found that it is very difficult trying to find the right degree of detail, but hopes that the presented list is a good compromise between too much detail and to abstract.

The Explanations section ([Section 3](#)) provides for each problem additional explanation to avoid that readers have to understand the problems only after reading up on a multitude of RFC and/or having similar deployment experiences.

On some points, the explanation section also presents what the author considers to be evidence as to how the IETF has established case precedents over the past decades how these problems have been solved in (in the authors opinion) similar cases.

These prior cases have often been controversial, but in the opinion of the author, they finally resulted in best serving the network operators and their customers, even at the expense that this produced more IETF standards than if some entity could have been able to persuade all customers and operators that there is a single one-fits-all approach to standards.

Even though the IESG did ask to only produce a problem statement, the author found that even the problems where best understood by

participants of the IETF that he discussed them with if he also provided an outline of the proposed solutions, instead of them having try to guess how hey could be solved. This is what the [Section 4](#) section discusses, and of course it is not meant to taint the opinions of readers as to how any of the problem should be solved.

Explanations for problem  $P_i$  are numbered  $E_i$ , Solution considerations for  $P_i$  are numbered  $S_i$ .

Finally, the document lists open issues and ends with with an appendix on other ([Appendix A](#)) considerations about how MSR6 might be operated adjacent to BIER in the IETF and why creation of such a WG is beneficial beyond the mere solving of a specific set of problems.

## **2. Problems**

### **2.1. P1: BIER: Efficiency of delivery for sparse groups in large networks**

In networks with (a) a large number of BFER (e.g.: 60k) and (b) a source-routing header size as favored by products (e.g. 256 bit), BIER will need to send in a stochastic manner as many packets as ingress replication (no BIER) instead only one packet.

### **2.2. P2: BIER-TE: Packet header efficiency of delivery to sparse groups in large networks**

Problem [Section 2.1](#) is larger in BIER-TE because significant number of bits in BIER-TE bitstrings need to be allowed to steering of the traffic ("tree engineering") making bitstrings effectively even smaller compared to BIER, increasing this Bitstring fragmentation problem.

### **2.3. P3: BIER/BIER-TE: Efficient amount of per-router state (BIFT)**

BIER/BIER-TE requires in every BIER router forwarding state for every egress router. In large networks, this can be as much as e.g.: 60,000 entries (or more) or maybe as much as twice as many with BIER-TE. This is unacceptable for many networks with this number of end-routers or hosts and mostly sparse multicast trees.

### **2.4. P4: More complex, multi-protocol host stack requirements with BIER**

Controlled networks such as Data-Centers with VM to VM IP multicast, or IoT networks do use IPV6 multicast. To use end-to-end stateless multicast with BIER, they would need to implement in addition to IPV6 host stacks BIER BFIR/BFER stacks as well as any BIER specific mechanisms for the routing underlay to use IPV6 as a flow overlay. This duplication of host stacks is undesirable for cost of development, operations code-space as well as hardware implementation (Smart NICs etc.).

## **2.5. P5: Support for single, IPv6/RFC8200 compliant network forwarding plane**

BIER is a separate (layer 2.5?) forwarding plane but not an extensions to IPv6/RFC8200. This is operationally undesirable for IPv6-only networks.

## **2.6. P6: No Support for SRv6 architecture**

BIER/BIER-TE including BIERin6 do not support the SRv6 architecture, making it impossible for networks that utilize a common source routing architecture across unicast and multicast to as easy as possible share any applicable or future SRv6 functionality between unicast and multicast and likely also simplify OAM and design.

## **2.7. P7: Inability to perform per-hop IPv6 forwarding plane features**

Operators depend on a wide variety of collateral forwarding plane features, even hop-by-hop or on dedicated intermediate hops which are not the ingress or egress router of the packet. Policies and/or Telemetry of these features are tied to IPv6 source and/or destination address. Important examples include, but are not limited to:

1. IPfix and related services for traffic monitoring
2. Traffic filtering (including (S,G))
3. QoS ( (S,G) policies)
4. IPsec/ESP ((S,G) SA policies)

## **2.8. P8: Routing convergence speed and complexity**

BIER/BIER-TE do not allow to build lowest outage, fast-converging network services because of their reliance on a) IGP re-routing and b) large BIFT forwarding tables.

## **3. Explanations**

### **3.1. E1: BIER: Efficiency of delivery to sparse groups in large networks**

A single BIER packet with a 256 bitstring can only create copies to the 256 BFER addressed via this bitstring. In a network with 60k BFER, at least 235 different Bitstrings need to be defined ( $256 * 235 > 60k$ ). A message to even a small number of BFER such as 10 could statistically easily require to send 10 packets, because each of the BFER is in a different bitstring.

Efficient replication to sparse multicast groups in large networks was at the core of the original IP multicast protocol PIM Sparse Mode in the early 1990th. The "Sparse" in this name exactly refers to such a group with few members in a large network. It can thus be seen as a day 1 requirement against any scalable IP multicast protocol, stateful or stateless.

### **3.2. E2: BIER-TE: Packet header efficiency of delivery to sparse groups in large networks**

Note that BIER-TE was designed only shortly after BIER and hence one primary goal was to re-use as much as possible of the BIER forwarding plane to make adoption as easy as possible. This excluded a lot of the more forward looking options being proposed for MSR6 now ([Section 4.1](#) / [Section 4.2](#)).

### **3.3. E3: BIER/BIER-TE: Efficient amount of per-router state (BIFT)**

With "flat bitstrings" as in BIER/BIER-TE each router needs a forwarding entry similar to unicast forwarding entries. In large networks, even enterprises, this is easy in the thousands (for edge routers) and another possible factor ten when running multicast source-routing end-to-end into hosts - which is feasible with newer proposed technologies but not flat bitstrings of BIER/BIER-TE (see [Section 4.1](#) / [Section 4.2](#)).

### **3.4. E4: More complex, multi-protocol host stack requirements with BIER**

Introducing a new host stack for a new network protocol like BIER without UDP tunneling requires most commonly OS upgrades. This is not a feasible requirement for reasonable fast adoption in many markets such as industrial, IoT or hosts in regulated environments such as aviation, finance industry, defense and the like. See [Section 4.4](#) how MSR can avoid this issue.

Whether in the OS or application (when using UDP tunneling), additional network stacks such as BIER typically incur more code-space overhead than extensions to an existing stack. This can be an issue for constrained code space hosts.

### **3.5. E5: Support for single, IPv6/RFC8200 compliant network forwarding plane**

The IETF has a long and successful history of extending different type of unicast network designs with multicast support that best matches and extends those underlying unicast mechanisms - even when vendors initially promoted less well aligned multicast technologies that was more readily available and solutions/workarounds could be built to support initial use-cases of customers that way.

While specific multicast technical detail benefits of adopting to the new unicast network design can be developed, this often only happens as a result of deployment experience with the new network technology and multicast later in the process. Instead, the most core reason for the IETF to build such new technologies was the operator/market preference and the logic that multicast functions should be available as a core architectural extension of any unicast network protocol architecture. This is why IP Multicast was added to IPv6 and MPLS and why MSR6 now wants to be chartered for adding source-routed IPv6 multicast based on expanding on RFC8200 principles.

Making multicast follow architectural frameworks of unicast and expanding it has a large set of benefits for operators, and can best

be described as the overall ecosystem complexity of operating (OAM) a network, but also in the often met expectation that it simplifies the adoption of future extensions/enhancements across unicast and multicast without having to re-define them, because the multicast solution is built from a different architecture.

The following lists a few mayor instances where the author thinks this principle can be observed. Of course it is always possible to create a detailed technical difference between any prior examples and the case for MSR6, but the author thinks that that is missing the larger point of the requirement:

Operator preferences may or may not be justifiable on selective detail technical points versus a pre-existing solution, but instead the IETF has repeatedly shown that it serves operator preferences when the proposed solutions are technically solid (working), require interoperability, and have a sufficient critical mass of supporters and active participants that can bring standardization to a successful finish.

Existence of prior, competing but alternating solutions have never been in the way of such additional technologies in the IETF, or else IETF would help first-comers to establish monopolies over market spaces where their solutions themselves may look sufficient but is not preferred by customers for the variety of reasons mentioned here.

### **3.5.1. Example 1: IPv4 multicast in MPLS/VPN solutions vs. native MPLS multicast**

In the early 2000th, (large) service providers where increasingly adopting MPLS (with IPv4) as their single network forwarding plane, primarily to be able to introduce L3VPN services. The only form of multicast that high-speed hardware forwarding could support in available products back then was IPv4 multicast (no MPLS), which is why vendors implemented and standardized in the IETF so-called "draft-rosen" that provided Multicast VPN (MVPN) for MPLS networks with L3VPN services.

Nevertheless, operators that where successfully deploying this IPv4 multicast based MVPN service continued to ask vendors to give them a "native MPLS" forwarding plane for their MVPN services. Some aspects of that solution, such as RSVP-TE/P2MP did provide easy to understand new functionality, such as traffic engineering, but the most commonly now used MPLS multicast solution, mLDP, (multicast LDP) provides significantly the same functionality that PIM can provide. Only the encapsulation of packets in the forwarding plane is different, but not the services that are supported across those two variants.

### **3.5.2. Example 2: Native IPv6 multicast vs. IPv4 multicast**

When IPv6 was introduced, it was not even a question to include it in the standardization of IPv6, even though in practice, most business critical use cases where equally feasible by sticking to IPv4: Most customers where using dual-plane networks and many of them continued to use IPv4 multicast because neither the

applications nor hardware forwarding in routers did support IPv6 multicast.

The difference over today was, that at that time, no vendor stood up and claimed IPv6 multicast should not be developed in the IETF because IPv4 multicast was sufficient, and if necessary with end-to-end tunneling of IPv6 multicast packets across IPv4 multicast it as well as hop-by-hop tunneling of IPv4 multicast over IPv6 only hops. But this is exactly what is done today with BIERin6:

### 3.5.3. Example 3: BIERin6

The BIER-WG is working on [[I-D.ietf-bier-bierin6](#)] as the proposed solution for how to use BIER in IPv6 networks. It does not change the deployment model of using BIER forwarding on every replicating hop (ideally for every router), but instead it recommends to:

1. Encapsulate IPv6 multicast packets end-to-end over BIER
2. Encapsulate BIER packets hop-by-hop across IPv6 when there are one or more intervening routers not supporting BIER

This approach is very similar to how in Example 1 and 2 networks that wanted to operationally standardize on only MPLS (Example 1) on every hop or only IPv6 (Example 2) on every hop, where asked by vendors to instead use IPv4 multicast because it existed and was hence easier to sell with existing vendor products. Instead of adding Source-routing to IPv6 multicast, BIERin6 suggests two layers of tunneling to make BIER fit into an otherwise IPv6 only network, effectively turning the network into a dual-plane IPv6-unicast-only + BIER-multicast-only network.

### 3.5.4. Example 4: Routing for multicast

The re-use of unicast technologies for multicast also expands into the control plane. Initial multicast designs where proposing multicast specific routing protocols like CBT, MOSPF and DVMRP, one of the core selling points that made PIM successful in the IETF was that it was re-using any deployed IP unicast routing protocol. Even when functionally for IP multicast itself, each of the prior protocols had arguably benefits (some of which where later also done with PIM, such as CBT->Bidir-PIM).

While re-use of existing unicast routing protocols may also be seen as a functional benefit in itself, these functional advantages where far less when doing the control plane for MPLS multicast.

mLDP was picked instead of PIM, because it allows customers to troubleshoot signaling from a single (LDP/TCP) basis for both unicast and multicast forwarding plane label binding and have more comparable behavior in a variety of corner situations.

Other MVPN signaling where initially continuing to use PIM in draft-rosen, but where later on demand of customers re-specified into BGP to suit the operator preference for a single routing protocol BGP for all services in the network (BESS model). The functionality itself was not the differentiator over PIM, instead the original ask was to do all that works well with PIM equally with BGP.



### 3.6. E6: No Support for SRv6 architecture

BIER/BIER-TE is a separate source-routing multicast architecture that is based on re-using the single BIER forwarding mechanisms in any type of network, whatever its preferred unicast architecture is.

Because there was and is no well-established mechanism for MPLS forwarding planes how to extend MPLS forwarding planes with extension headers, BIER did have to come up with its own header for MPLS networks. This makes it fundamentally different from IPv6, where [\[RFC8200\]](#) defines the overall source-routing mechanism, and SRv6 is the one instance of it. [\[RFC6554\]](#) is another instance.

Note though that an IETF MPLS design team is currently (2021/2022) working on future common extension header design for MPLS networks, so it is but a theoretical but still interesting question if in a time when MPLS had such established standards, MPLS network operators would have accepted a BIER header not specifically following all rules for source routing and header construction established by such pre-existing MPLS rules. This nevertheless is exactly one of the problems IPv6 network operators would like to see solved by MSR6 because those definitions do exist in [\[RFC8200\]](#) for IPv6 networks.

### 3.7. E7: Inability to perform per-hop IPv6 forwarding plane features

IPFIX [\[RFC7011\]](#) is a protocol to export traffic flow telemetry. It is widely implemented and deployed for IP/IPv6 multicast, allowing to capture per multicast (S,G) traffic flow statistics for network OAM services.

With BIER and BIERin6, there is no knowledge at the BIER forwarding layer about multicast (S,G) flows. This only exists on the flow-overlay layer, which is not present on non-ingress/egress BFRs.

The underlying forwarding functionality to capture traffic statistics for IPfix flow export has also been used in other common industry functions, such as MediaTrace mechanisms.

Policies for IPv6 multicast traffic, such as expressed through ACLs or other router constructs, are used for a variety of functions.

1. Filtering of traffic based on multicast destination group ranges, includes administrative scoping. For example, large service providers often have two tiers of service elements, where the PE are not the last point of policy control, but instead the first P hop could be. This avoid more complex CsC style additional layers.
2. Destination address group bits of IPv6 addresses have been used to indicate target QoS classes for traffic more flexibly and easier inter-domain than DSCP.
3. IPsec encryption/decryption can be inserted on any multi-hop path segments of an IP/IPv6 network using the IP multicast (S,G) address information as part of setting up the IPsec SA policies (this is not specific to IP multicast but comes for

free from IPsec being designed against IP and thus the address information that IPv6 multicast also shares).

Applying policies at arbitrary points along the network path is also more important for MSR6 than BIER, because MSR6 is not solely focused on only P/PE network designs, and the overall design of the BIER architecture with flow overlay, BIER layer and routing underlay is very much expecting such a network design: BIER removes the knowledge about the actual BIER payload (IPv6 in our case) from the BFR midpoints and tunnels it over BIER between BFIR and BFER. This is not sufficient for all MSR6 deployment options.

### **3.8. E8: Routing convergence speed and complexity in large networks**

Multicast services have commonly higher service level objectives than so-called Internet Best-Effort traffic, yet customers are looking for equally easy to operationalise solutions for multicast in the network.

With BIER and a large set of destinations (BFER), the BIFT hardware forwarding tables on every hop in the network need to be updated with BIFT state for all destinations. For tenths of thousands of destinations, this becomes as slow as updating the same number of unicast FIB entries. (it is also a waste of BIFT state when potentially at any point in time only a small percentage of destinations actually does need to receive traffic).

This ultimate hardware forwarding goal is preceded with all the well understood convergence issues of IGPs (micro-loops etc), as well as the time IGPs take to converge.

The known solution to this problem is to utilize strict explicit trees to deliver traffic. This takes the IGP in the network out of the picture for hop-by-hop forwarding, and when senders do utilize e.g. link-state information to calculate paths for trees, they can constrain their calculations to only the destination addresses of interest, but not the whole set of destinations in the network.

Even though BIER-TE can express strict trees, it can not overcome these issue well, because of the flat bitstring design (derived from BIER), it is not possible to build strict-steered multicast trees and simultaneously reduce BIFT state. Instead, attempting to solve this issue with BIER-TE makes BIFT larger.

## **4. Candidate MSR6 solution discussion**

### **4.1. S1: BIER: Efficiency of delivery for sparse groups in large networks**

[\[I-D.eckert-msr6-rbs\]](#) and various other MSR6 proposal introduce different stateless multicast source routing encodings than the BIER/BIER-TE "flat" bitstrings. These all overcome the problem by more intelligent encoding. [\[I-D.eckert-bier-cgm2-rbs\]](#) also contains preliminary scalability simulation results comparing the performance against BIER. See also `{#exp}`.

#### **4.2. S2: BIER-TE: Efficiency of delivery for sparse groups in large networks**

Solutions to this problem such as [[I-D.eckert-msr6-rbs](#)] are even stronger solutions to [Section 2.2](#) than to [Section 2.1](#) because BIER-TE suffers more from the problem than BIER. Equally, see also {#exp}.

#### **4.3. S3: BIER/BIER-TE: Large amount of per-BFR state (BIFT)**

In MSR6 proposals such as [[I-D.eckert-msr6-rbs](#)], each router only needs as much per-BFR state (BIFT) as that router has neighbors, which typically is 10..200, therefore order or magnitudes less than required in BIER/BIER-TE. This applies equally to using such proposals instead of BIER or BIER-TE.

#### **4.4. S4: More complex, multi-protocol host stack requirements with BIER**

IPv6 socket APIs exist today on every host stack for every OS, especially a large variety of IoT host stacks, but also host stacks that are old and will not be able to receive OS-level updates, or that for regulatory and other business reasons can not receive available OS-level updates (because of avoid need for re-validation).

IPv6 host stacks also have the socket API option to insert from application level arbitrary IPv6 extension headers. This allows to implement MSR6 solely in user-land without additional (UDP) tunneling required.

#### **4.5. S5: Support for single, IPv6/RFC8200 compliant network forwarding plane**

MSR6 proposed to use [[RFC8200](#)] derived IPv6 extension headers for source routing, similar to unicast source routing headers for IPv6 in [[RFC6554](#)] or [[RFC8754](#)] with IPv6 destination address rewrite on every source routing hop. When the source routing header information indicates only the destinations, this is called by MSR6 Best Effort (BE) mode. When the source-routing information indicates the delivery tree, this is called by MSR6 "Tree Engineered" (TE) mode.

To support IPv6 multicast without multiple extension header, MSR6 options such as [[I-D.eckert-msr6-rbs](#)] propose to include not only the source-routing information to perform BE/TE functions (as in BIER/BIER-TE), but also to include the IPv6 multicast destination address. This is necessary because the IPv6 destination address is not transparent end-to-end in RFC8200 compliant source-routing.

It is also another significant difference over BIER, where IPv6 multicast can only be supported as a flow overlay. And it is the reason why at the layer of IPv6 with MSR all IPv6 multicast address information is directly available, even when the final IPv6 destination address is not in the base but and IPv6 extension header.

#### **4.6. S6: No Support for SRv6 architecture**

As described in [Section 4.5](#), MSR6 proposes extension header similar in spirit and goals to SRH. It should equally create standards document that allow to apply the SRv6 terminology and concept of SIDs (Segment IDentifiers) to MSR6 when this is desired by the network operator.

By mean of using IPv6 addressing in MSR6, it may also be possible to utilize existing SRv6/SID benefits with multicast (such as inclusion of programability aspects in the MSR6 IPv6 addresses). This is subject to further study.

Note that IPv6 multicast did equally explore and benefit from special encodings into the IPv6 address space, arguably similar to SRv6 adding additional functionality via IPv6 address bits. This includes specifically "Unicast-Prefix-based IPv6 Multicast Addresses", [\[RFC3306\]](#) and even more so "Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address", [\[RFC3956\]](#).

#### **4.7. S7: Inability to perform per-hop IPv6 forwarding plane features**

MSR6 enables all collateral IPv6 forwarding plane features along the path by explicitly exposing intentionally the same source and destination address as non-source-routed IPv6 multicast packets.

This is unlike BIERin6, where this information is not in the possible outer IPv6 header (on loose hops), nor the BIER header, but only in the BIER end-to-end encapsulated payload header if it is IPv6. On BFR, the IPv6 context (such as VRF) in which this IPv6 header exists is not known, so any capturing of information via such "deep packet inspection" mechanism comes at a lot more problems to solve and would hardly be something that could be used for standardized IETF mechanisms such as IPfix.

The most fundamental privacy rule for network protocol designs is to not leverage packet data that is not explicitly meant to be exposed at a particular layer/point in the network, and on an MSR6 hop, the MSR6 extension header is explicitly exposed, representing the multicast (S,G) flow information, whereas in BIER this is never the case, but only on BFIR/BFER in the flow-overlay layer.

With MSR6, there is an additional overhead for routers to examine the destination IPv6 address from an extension header, but that is the cost of doing IPv6 business, because with the way [\[RFC8200\]](#) expects source routing to operate, this can not be further optimized. Of course, this means that it is highly important to ensure that exactly this IPv6 multicast destination address has ideally not multiple options of extension header or positions in it, but that MSR6 will develop one-fits-all for this component of the solution.

#### **4.8. S8: Routing convergence speed and complexity**

MSR6 proposal that attempt to solve this problem are those that allow to encode in the source-routing header sparse trees for large networks such that no IGP routing, but only hop-by-hop forwarding is required. These solutions do likewise also require only small

amounts of BIFT forwarding state per router that only changes when link-local MSR6 neighbors change {I-D.eckert-msr6-rbs}} is an example of such proposals.

## 5. Open issues

Probably many...

This version of the document does likely not well enough highlight in all places when the solution to a problem could also be had by enhancements to BIER, but it does so in hopefully the most important places.

## 6. Changelog

00 - initial version.

## 7. Informative References

[I-D.eckert-bier-cgm2-rbs] Eckert, T. T. and B. Xu, "Carrier Grade Minimalist Multicast (CGM2) using Bit Index Explicit Replication (BIER) with Recursive BitString Structure (RBS) Addresses", Work in Progress, Internet-Draft, draft-eckert-bier-cgm2-rbs-01, 9 February 2022, <<https://www.ietf.org/archive/id/draft-eckert-bier-cgm2-rbs-01.txt>>.

[I-D.eckert-msr6-rbs] Eckert, T. T., Geng, X., Zheng, X., Meng, R., and F. Li, "Recursive Bitstring Structure (RBS) for Multicast Source Routing over IPv6 (MSR6)", Work in Progress, Internet-Draft, draft-eckert-msr6-rbs-00, 11 July 2022, <<https://www.ietf.org/archive/id/draft-eckert-msr6-rbs-00.txt>>.

[I-D.ietf-bier-bierin6] Zhang, Z., Zhang, Z. J., Wijnands, I., Mishra, M. P., Bidgoli, H., and G. S. Mishra, "Supporting BIER in IPv6 Networks (BIERin6)", Work in Progress, Internet-Draft, draft-ietf-bier-bierin6-05, 4 September 2022, <<https://www.ietf.org/archive/id/draft-ietf-bier-bierin6-05.txt>>.

[I-D.liu-msr6-problem-statement] Liu, Y., Jiang, T., Eckert, T. T., Li, Z., Mishra, G. S., Qin, Z., Lin, C., and X. Geng, "Problem Statement of IPv6 Multicast Source Routing (MSR6)", Work in Progress, Internet-Draft, draft-liu-msr6-problem-statement-01, 21 October 2022, <<https://www.ietf.org/archive/id/draft-liu-msr6-problem-statement-01.txt>>.

[RFC3306] Haberman, B. and D. Thaler, "Unicast-Prefix-based IPv6 Multicast Addresses", RFC 3306, DOI 10.17487/RFC3306, August 2002, <<https://www.rfc-editor.org/info/rfc3306>>.

[RFC3956] Savola, P. and B. Haberman, "Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address", RFC 3956, DOI 10.17487/RFC3956, November 2004, <<https://www.rfc-editor.org/info/rfc3956>>.

**[RFC5015]**

Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano, "Bidirectional Protocol Independent Multicast (BIDIR-PIM)", RFC 5015, DOI 10.17487/RFC5015, October 2007, <<https://www.rfc-editor.org/info/rfc5015>>.

**[RFC6554]**

Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, DOI 10.17487/RFC6554, March 2012, <<https://www.rfc-editor.org/info/rfc6554>>.

**[RFC7011]**

Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <<https://www.rfc-editor.org/info/rfc7011>>.

**[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>>.

**[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>>.

**[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>>.

## **Appendix A. Other considerations towards MSR6**

### **A.1. Design separation between tree building and other functions**

BIER is designed as a common stateless multicast forwarding layer meant to be applicable to any network (or ever higher layer) protocol that can be tunneled/encapsulated across it ("flow-overlay").

This incurs the price of designing a flow overlay layer on BFIR/BFER (ingress/egress BIER edge routers), and not having any functionality of that encapsulated flow overlay traffic available on the BIER midpoints (BFR) - unless any such functionality is meticulously mapped into the BIER packet header, such as MPLS TC and IP/IPv6 DSCP as specified in [RFC8296].

This design maps very well the common P/PE design in service providers, but it does not necessarily match arbitrary IPv6 networks in industry, IoT, data-center (interconnect), or even service provider networks (aggregation, multi-AS,...). Instead, when such technologies are used, network designs across all features have to be built to enable P routers to be mostly feature-less.

IPv6 MSR6 instead is intended to allow re-using by default any pre-source-routing IPv6 feature as much as possible on any hop of the network. And minimize the requirements of introducing network protocol P/PE separation designs across all functions in the network to enable this.

One example for this principle is an early stage in reducing IP multicast tree state in data centers.

Originally, OAM for IP multicast traffic in DC was performed on PIM-SM (S,G) state (per state packet/byte counters), but diagnosis of PIM-SM state is challenging in itself and scales linearly with the amount of applications running - operators did not like it. Even worse, the amount of PIM-SM state deployed in the target data centers was about to exceed performance of routers. This led to the development of Bidir-PIM [[RFC5015](#)] which reduced PIM state to (\*,G), for example in typical data-centers from 20,000 states to 200 states at the time.

But Bidir-PIM made it impossible to have per (S,G) counters anymore on every hop through the multicast tree state. Instead, IPfix for IP multicast was developed, which could be configured for arbitrary subsets of multicast flows (based on (S,G) ACL policies), thus also not being tied to the scaling problems of the IP multicast tree building, and by virtue of operating completely independent of IP multicast it also had no tie-in with IP multicast protocols and can therefore also help to diagnose them.

## **A.2. Reuse of BIER principles and technology components as much as possible**

MSR6 efforts considers BIER-WG to be one of its most important adjacencies in the IETF and does not want to re-invent anything unnecessarily that BIER already solves well if applicable to IPv6 only source routing mechanisms.

Likewise, MSR6 efforts would welcome if new technology ideas originally experimented and used in MSR, but equally applicable to BIER can also be shared across MSR6 and BIER (such as new representations of the source-routing information).

## **A.3. MSR6 as experimental (not standards track) work**

In the opinion of the author, the ability to experiment with new and better technology components in MSR6 is easier when an MSR6 WG would initially be targeted for experimental status - in the same way as BIER initially was. BIER WG has for some time now been "upgraded" to standards track work, and also its RFC where upgraded to standard track, even though this does not show in the documents themselves (for historical reasons of the RFC publishing process).

With BIER (WG and RFCs) being standards track now, it still can still adopt experimental work, but when BIER establishes itself as a deployed solution in the market, it often becomes less easy to publish new, experimental work with the same name, because this raises the risk of customers not being able to well distinguish between more proven, older BIER options and newer, potentially experimental ones.

Therefore, experimental new technology options may be better first proven in the context of MSR6 work even when equally applicable to MSR6 and BIER. In the opinion of the author this specifically includes the new tree encoding structures proposed by different MSR6 drafts (including the authors [[I-D.eckert-msr6-rbs](#)]).

The core tenets of MSR6 such as its use of RFC8200 source routing mechanisms for end-to-end IPv6 multicast could easily be standards track, but may simply be best targeted as experimental too, solely to be sure that sufficient practical/operational experience work is gained - and tracked by appropriate IETF work - before they are upgraded to standards level.

#### **Author's Address**

Toerless Eckert (editor)  
Futurewei Technologies USA  
2220 Central Expressway  
Santa Clara, CA 95050  
United States of America

Email: [tte@cs.fau.de](mailto:tte@cs.fau.de)