

6man
Internet-Draft
Intended status: Standards Track
Expires: November 7, 2019

R. Bonica
Juniper Networks
Y. Kamite
NTT Communications Corporation
T. Niwa
KDDI
A. Alston
D. Henriques
Liquid Telecom
N. So
F. Xu
Reliance Jio
G. Chen
Baidu
Y. Zhu
G. Yang
China Telecom
Y. Zhou
ByteDance
May 6, 2019

The IPv6 Compressed Routing Header (CRH)
draft-bonica-6man-comp-rtg-hdr-04

Abstract

This document defines the Compressed Routing Header (CRH). The CRH, like any other Routing header, contains a list of segment identifiers (SID). The CRH differs from other Routing headers in that its segment identifiers can be 8, 16 or 32 bits long.

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 November 7, 2019.

Copyright Notice

Copyright (c) 2019 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.	Requirements Language	4
3.	The Compressed Routing Header (CRH)	4
4.	The CRH Domain	6
5.	Segment Identifiers (SID)	6
6.	Processing Rules	7
6.1.	General	7
6.2.	CRH Specific	8
6.2.1.	Computing Minimum CRH Length	9
7.	Mutability	10
8.	Management Considerations	10
9.	Security Considerations	10
10.	IANA Considerations	11
11.	Acknowledgements	11
12.	References	11
12.1.	Normative References	11
12.2.	Informative References	12
Appendix A.	CRH Processing Examples	12
A.1.	Loose Source Routing	14
A.2.	Loose Source Routing Preserving The First SID	14
A.3.	Strict Source Routing	15
	Authors' Addresses	16

[1.](#) Introduction

An IPv6 [[RFC8200](#)] source node can steer a packet through a specific path to its destination. The source node defines the path as an ordered list of segments and encodes the path in an IPv6 Routing header. As per [[RFC8200](#)], all Routing headers includes the following fields:

- o Next Header - Identifies the header immediately following the Routing header.
- o Hdr Ext Len - Length of the Routing header.
- o Routing Type - Identifies the particular Routing header variant.
- o Segments Left - The number of segments still to be traversed before reaching the destination.
- o Type-specific Data - Syntax and semantics are defined by the Routing Type.

The following Routing types are currently defined:

- o Source Route (i.e., RH0) [[RFC5095](#)] (deprecated)
- o Type 2 Routing Header [[RFC6275](#)]
- o RPL Source Route Header [[RFC6554](#)]
- o Segment Routing Header (SRH)
[[I-D.ietf-6man-segment-routing-header](#)]

In each of the above-mentioned Routing Types, Type-specific Data contains a list of one or more segment identifiers (SID). Typically, a SID is an IPv6 address that identifies a segment endpoint. In the SRH, the SID may carry additional semantics.

In all cases, the SID is 128 bits long. Therefore, routing headers can be very large. For example, an 88-byte Source Route header is required to specify a path that contains six segments. The same can be said of the SRH.

Large Routing headers are undesirable for the following reasons:

- o Many ASIC-based forwarders copy the entire IPv6 extension header chain from buffer memory to on-chip memory. As the size of the IPv6 extension header chain increases, so does the cost of this copy.
- o Because Path MTU Discovery (PMTUD) [[RFC8201](#)] is not entirely reliable, many IPv6 hosts refrain from sending packets larger than the IPv6 minimum link MTU (i.e., 1280 bytes). When packets are small, the overhead imposed by large Routing headers becomes pronounced.

This document defines the Compressed Routing Header (CRH). The CRH, like any other Routing header, contains a list of SIDs. The CRH differs from other Routing headers in that its SIDs can be 8, 16, or 32 bits long.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

3. The Compressed Routing Header (CRH)

Figure 1 depicts the CRH.

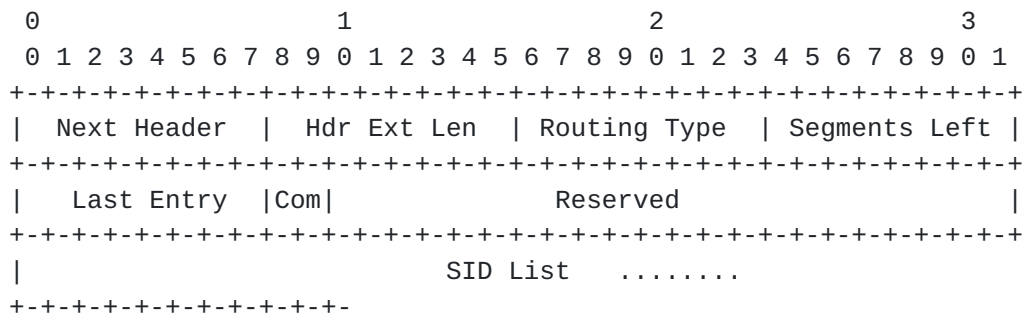


Figure 1: Compressed Routing Header (CRH)

The CRH contains the following fields:

- o Next Header - Defined in [[RFC8200](#)].
- o Hdr Ext Len - Defined in [[RFC8200](#)].
- o Routing Type - Defined in [[RFC8200](#)]. Value TBD by IANA. (Suggested value: 5)
- o Segments Left (SL) - Defined in [[RFC8200](#)].
- o Last Entry - 8 bits. Represents the index (zero based), in the Segment List, of the last element of the Segment List..
- o Com (Compression) - 2 bits. Represents the length of each entry in the SID List. Values are eight bits (0), sixteen bits (1), thirty-two bits (2), and reserved (3).

Figure 3: Sixteen-bit Encoding (Com equals 1)

- o A globally scoped IPv6 unicast address [[RFC4291](#)].
- o A Unique Local IPv6 Unicast Address (ULA) [[RFC4193](#)].
- o A Multicast address [[RFC4291](#)].

A strictly routed SID also maps to a link interface. Nodes send packets through that interface in order to access the segment endpoint.

Loosely routed SIDs have domain-wide significance. This means that within a CRH domain, a loosely routed SID MUST map to exactly one IPv6 address. By contrast, strictly routed SIDs have node-local significance. This means that within a CRH domain, one node can map a strictly routed SID to one address while another node maps the same strictly routed SID to a different address. See [Appendix A](#) for an example.

Forwarding nodes can learn the above-mentioned mappings from a central controller, from a distributed routing protocol or using any other means. The mechanisms that forwarding nodes use to learn the above-mentioned mappings are beyond the scope of this document.

6. Processing Rules

6.1. General

[RFC8200] defines rules that apply to IPv6 extension headers, in general, and IPv6 Routing headers, in particular. All of these rules apply to the CRH.

For example:

- o Extension headers (except for the Hop-by-Hop Options header) are not processed, inserted, or deleted by any node along a packet's delivery path, until the packet reaches the node (or each of the set of nodes, in the case of multicast) identified in the Destination Address field of the IPv6 header.
- o If, while processing a received packet, a node encounters a Routing header with an unrecognized Routing Type value, the required behavior of the node depends on the value of the Segments Left field. If Segments Left is zero, the node must ignore the Routing header and proceed to process the next header in the packet, whose type is identified by the Next Header field in the Routing header. If Segments Left is non-zero, the node must discard the packet and send an ICMP [\[RFC4443\]](#) Parameter Problem, Code 0, message to the packet's Source Address, pointing to the unrecognized Routing Type.
- o If, after processing a Routing header of a received packet, an intermediate node determines that the packet is to be forwarded onto a link whose link MTU is less than the size of the packet,

the node must discard the packet and send an ICMP Packet Too Big message to the packet's Source Address.

6.2. CRH Specific

When a node recognizes and processes a CRH, it executes the following procedure:

- o If the IPv6 Source Address is a link-local address, discard the packet.
- o If the IPv6 Source Address is a multicast address, discard the packet.
- o If Segments Left equal 0, skip over the CRH and process the next header in the packet.
- o If Segments Left is greater than Last Entry plus one, send an ICMP Parameter Problem, Code 0, message to the Source Address, pointing to the Segments Left field, and discard the packet.
- o If Com is equal to (3) Reserved, send an ICMP Parameter Problem, Code 0, message to the Source Address, pointing to the Com field, and discard the packet.
- o If the IPv6 Hop Limit is less than or equal to 1, send an ICMP Time Exceeded -- Hop Limit Exceeded in Transit message to the Source Address and discard the packet.
- o Compute L, the minimum CRH length (See [Section 6.2.1](#))
- o If L is greater than Hdr Ext Len, send an ICMP Parameter Problem, Code 0, message to the Source Address, pointing to the Last Entry field, and discard the packet.
- o Decrement Segments Left (i.e., SL).
- o Search for SID[SL] in the table that maps SID's to IPv6 addresses and interfaces. If SID[SL] cannot be found in that table, send an ICMP Parameter Problem, Code 0, message to the Source Address, pointing to SID[SL], and discard the packet.
- o Copy the address associated with SID[SL] to the IPv6 Destination Address.
- o If the IPv6 Destination address is a multicast address and SL is greater than zero, send an ICMP Parameter Problem, Code 0, message

to the Source Address, pointing to Segment List [SL], and discard the packet.

- o Decrement the IPv6 Hop Limit.
- o If SID[SL] is a loosely routed segment, resubmit the packet to the IPv6 module for transmission to the new destination.
- o If SID[SL] is a strictly routed segment, forward the packet through the interface that is associated with SID[SL].

The above stated rules are demonstrated in [Appendix A](#).

6.2.1. Computing Minimum CRH Length

The algorithm described in this section accepts the following CRH fields as its input parameters:

- o Compression (Com).
- o Last Entry.

It yields L, the minimum CRH length. The minimum CRH length is measured in 8-octet units, not including the first 8 octets.

<CODE BEGINS>

```
if (Com == 0 ) {          /* Eight bit encoding */
    L = ( ( Last Entry + 1 ) / 8 );
    if ( ( Last Entry + 1 ) % 8 )
        L++;
}
elseif (Com == 1 ) {      /* Sixteen bit encoding */
    L = ( ( Last Entry + 1 ) / 4 );
    if ( ( Last Entry + 1 ) % 4 )
        L++;
}
elseif (Com == 2 ) {      /* Thirty-two bit encoding */
    L = ( ( Last Entry + 1 ) / 2 );
    if ( ( Last Entry + 1 ) % 2 )
        L++;
}
else {                    /* Invalid Com */
    L = 0xFF
}

return(L)

<CODE ENDS>
```

7. Mutability

The Segments Left field is mutable and MAY be decremented by processing nodes. All remaining fields are immutable.

8. Management Considerations

PING and TRACEROUTE [[RFC2151](#)] both operate correctly in the presence of the CRH.

9. Security Considerations

The CRH can be used within trusted domains only. In order to enforce this requirement, domain edge routers MUST do one of the following:

- o Discard all inbound packets that contain a CRH
- o Authenticate [[RFC4302](#)] [[RFC4303](#)] all inbound packets that contain a CRH

10. IANA Considerations

This document makes the following registration in the Internet Protocol Version 6 (IPv6) Parameters "Routing Type" registry maintained by IANA:

Suggested Value	Description	Reference

5	Compressed Routing Header (CRH)	This document

11. Acknowledgements

Thanks to Joel Halpern, Tony Li, Gerald Schmidt, Nancy Shaw and Chandra Venkatraman for their comments.

12. References

12.1. Normative References

- [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>>.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", [RFC 4291](#), DOI 10.17487/RFC4291, February 2006, <<https://www.rfc-editor.org/info/rfc4291>>.
- [RFC4302] Kent, S., "IP Authentication Header", [RFC 4302](#), DOI 10.17487/RFC4302, December 2005, <<https://www.rfc-editor.org/info/rfc4302>>.
- [RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", [RFC 4303](#), DOI 10.17487/RFC4303, December 2005, <<https://www.rfc-editor.org/info/rfc4303>>.
- [RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, [RFC 4443](#), DOI 10.17487/RFC4443, March 2006, <<https://www.rfc-editor.org/info/rfc4443>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

- [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>>.
- [RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, [RFC 8201](#), DOI 10.17487/RFC8201, July 2017, <<https://www.rfc-editor.org/info/rfc8201>>.

12.2. Informative References

- [I-D.ietf-6man-segment-routing-header] Filsfils, C., Previdi, S., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", [draft-ietf-6man-segment-routing-header-18](#) (work in progress), April 2019.
- [RFC2151] Kessler, G. and S. Shepard, "A Primer On Internet and TCP/IP Tools and Utilities", FYI 30, [RFC 2151](#), DOI 10.17487/RFC2151, June 1997, <<https://www.rfc-editor.org/info/rfc2151>>.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", [RFC 4193](#), DOI 10.17487/RFC4193, October 2005, <<https://www.rfc-editor.org/info/rfc4193>>.
- [RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation of Type 0 Routing Headers in IPv6", [RFC 5095](#), DOI 10.17487/RFC5095, December 2007, <<https://www.rfc-editor.org/info/rfc5095>>.
- [RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", [RFC 6275](#), DOI 10.17487/RFC6275, July 2011, <<https://www.rfc-editor.org/info/rfc6275>>.
- [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>>.

Appendix A. CRH Processing Examples

This appendix provides examples of CRH processing in the following applications:

- o Loose source routing (Appendix A.1)

- o Loose source routing preserving the first SID (Appendix A.2)
- o Strict source routing (Appendix A.3)

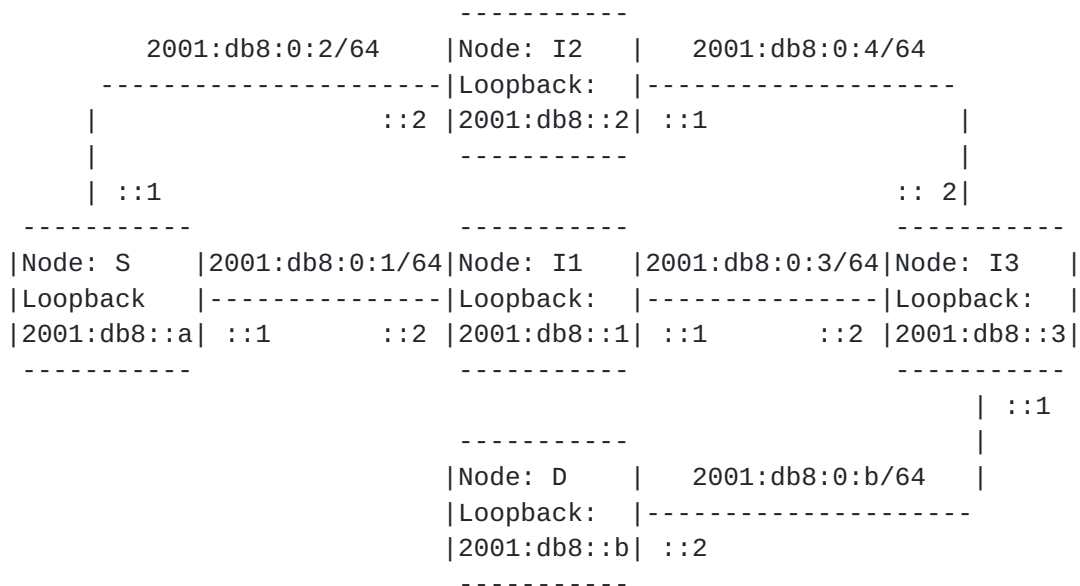


Figure 5: Reference Topology

Figure 5 provides a reference topology that is used in all examples.

Instantiating Node	SID	IPv6 Address
All	1	2001:db8::1
All	2	2001:db8::2
All	3	2001:db8::3
All	10	2001:db8::a
All	11	2001:db8::b

Table 1: Loosely Routed SIDs

Table 1 provides mappings for loosely routed SIDs. These mappings are instantiated on all nodes in the reference topology.

Instantiating Node	SID	IPv6 Address	Interface
S	129	2001:db8:0:1::2	S -> I1
S	130	2001:db8:0:2::2	S -> I2
I1	129	2001:db8:0:3::2	I1 -> I3
I2	129	2001:db8:0:4::2	I2 -> I3
I3	129	2001:db8:0:b::2	I3 -> D

Table 2: Strictly Routed SIDs

Table 2 provides mappings for strictly routed SIDs. These mappings are available on the instantiating node only.

[A.1.](#) Loose Source Routing

In this example, Node S sends a packet to Node D, specifying loose source route through Node I3. In this example, the first node in the path, I3, does not appear in the CRH segment list. Therefore, the destination node may not be able to send return traffic through the same path.

+-----+	
As the packet travels from S to I3:	
+-----+	
Source Address = 2001:db8::a	Last Entry = 0
Destination Address = 2001:db8::3	Segments Left = 1
	SID[0] = 11
+-----+	
+-----+	
+-----+	
As the packet travels from I3 to D:	
+-----+	
Source Address = 2001:db8::a	Last Entry = 0
Destination Address = 2001:db8::b	Segments Left = 0
	SID[0] = 11
+-----+	

[A.2.](#) Loose Source Routing Preserving The First SID

In this example, Node S sends a packet to Node D, specifying loose source route through Node I3. In this example, the first node in the path, I3, appears in the CRH segment list. Therefore, the destination node can send return traffic through the same path.

[illegible][illegible]

A.3. Strict Source Routing

In this example, Node S sends a packet to Node D, specifying the strict source route through I1 and I3.

+-----+-----+	
As the packet travels from S to I1:	
+-----+-----+	
Source Address = 2001:db8::a	Last Entry = 1
Destination Address = 2001:db8:0:1::2	Segments Left = 2
	SID[0] = 129
	SID[1] = 129
+-----+-----+	

+-----+	
As the packet travels from I1 to I3:	
+-----+	
Source Address = 2001:db8::a	Last Entry = 1
Destination Address = 2001:db8:0:3::2	Segments Left = 1
	SID[0] = 129
	SID[1] = 129
+-----+	


```
+-----+
| As the packet travels from I3 to D: |
+-----+
| Source Address = 2001:db8::a      | Last Entry = 1  |
| Destination Address = 2001:db8:0:b::2 | Segments Left = 0 |
|                                     | SID[0] = 129   |
|                                     | SID[1] = 129   |
+-----+
```

Authors' Addresses

Ron Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, Virginia 20171
USA

Email: rbonica@juniper.net

Yuji Kamite
NTT Communications Corporation
3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Email: : y.kamite@ntt.com

Tomonobu Niwa
KDDI
3-22-7, Yoyogi, Shibuya-ku
Tokyo 151-0053
Japan

Email: to-niwa@kddi.com

Andrew Alston
Liquid Telecom
Nairobi
Kenya

Email: Andrew.Alston@liquidtelecom.com

Daniam Henriques
Liquid Telecom
Johannesburg
South Africa

Email: daniam.henriques@liquidtelecom.com

Ning So
Reliance Jio
3010 Gaylord PKWY, Suite 150
Frisco, Texas 75034
USA

Email: Ning.So@ril.com

Fengman Xu
Reliance Jio
3010 Gaylord PKWY, Suite 150
Frisco, Texas 75034
USA

Email: Fengman.Xu@ril.com

Gang Chen
Baidu
No.10 Xibeiwang East Road Haidian District
Beijing 100193
P.R. China

Email: phdgang@gmail.com

Yongqing Zhu
China Telecom
109 West Zhongshan Ave, Tianhe District
Guangzhou
P.R. China

Email: zhuyq.gd@chinatelecom.cn

Guangming Yang
China Telecom
109 West Zhongshan Ave, Tianhe District
Guangzhou
P.R. China

Email: yanggm.gd@chinatelecom.cn

Yifeng Zhou
ByteDance
Building 1, AVIC Plaza, 43 N 3rd Ring W Rd Haidian District
Beijing 100000
P.R. China

Email: yifeng.zhou@bytedance.com

