6MAN J. Hui
Internet-Draft JP. Vasseur
Intended status: Standards Track Cisco Systems, Inc
Expires: May 18, 2012 D. Culler
UC Berkeley
V. Manral
Hewlett Packard Co.
November 15, 2011

An IPv6 Routing Header for Source Routes with RPL
draft-ietf-6man-rpl-routing-header-05

Abstract

In Low power and Lossy Networks (LLNs), memory constraints on routers may limit them to maintaining at most a few routes. In some configurations, it is necessary to use these memory constrained routers to deliver datagrams to nodes within the LLN. The Routing for Low Power and Lossy Networks (RPL) protocol can be used in some deployments to store most, if not all, routes on one (e.g. the Directed Acyclic Graph (DAG) root) or few routers and forward the IPv6 datagram using a source routing technique to avoid large routing tables on memory constrained routers. This document specifies a new IPv6 Routing header type for delivering datagrams within a RPL Instance.

Table of Contents

1. Introduction

Routing for Low Power and Lossy Networks (RPL) is a distance vector IPv6 routing protocol designed for Low Power and Lossy networks (LLN) [I-D.ietf-roll-rpl]. Such networks are typically constrained in resources (limited communication data rate, processing power, energy capacity, memory). In particular, some LLN configurations may utilize LLN routers where memory constraints limit nodes to maintaining only a small number of default routes and no other destinations. However, it may be necessary to utilize such memory-constrained routers to forward datagrams and maintain reachability to destinations within the LLN.

To utilize paths that include memory-constrained routers, RPL relies on source routing. In one deployment model of RPL, necessary mechanisms are used to collect routing information at more capable routers and form paths from those routers to arbitrary destinations within a RPL Instance. However, a source routing mechanism supported by IPv6 is needed to deliver datagrams.

This document specifies the Source Routing Header (SRH) for use strictly between RPL routers in the same RPL Instance.

1.1. 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 [RFC2119].

2. Overview

The format of SRH draws from that of the Type 0 Routing header (RH0) [RFC2460]. However, SRH introduces mechanisms to compact the source route entries when all entries share the same prefix with the IPv6 Destination Address of a packet carrying a SRH, a typical scenario in LLNs using source routing. The compaction mechanism reduces consumption of scarce resources such as channel capacity.

SRH also differs from RH0 in the processing rules to alleviate security concerns that led to the deprecation of RH0 [RFC5095]. First, RPL routers implement a strict source route policy where each and every IPv6 hop is specified within the SRH. Second, a SRH is only used between RPL routers within a RPL Instance. RPL Border Routers, responsible for connecting other RPL Instances and IP domains that use other routing protocols, do not allow datagrams already carrying a SRH header to enter or exit a RPL Instance. Third, a RPL router drops datagrams that includes multiple addresses assigned to any interfaces on that router to avoid forwarding loops.

There are two cases that determine how to include a SRH when a RPL router requires the use of a SRH to deliver a datagram to its destination.

1. If the SRH specifies the complete path from source to destination, the router places the SRH directly in the datagram itself.

2. If the SRH only specifies a subset of the path from source to destination, the router uses IPv6-in-IPv6 tunneling [RFC2473] and places the SRH in the outer IPv6 header. Use of tunneling ensures that the datagram is delivered unmodified and that ICMP errors return to the source of the SRH rather than the source of the original datagram.

In a RPL network, Case 1 occurs when both source and destinations are within a RPL Instance and a single SRH is used to specify the entire path from source to destination, as shown in the following figure:

```
                   +--------------------+
                   |                    |
                   |  (S) -------> (D)  |
                   |                    |
                   +--------------------+
                          RPL Instance
```

In the above scenario, datagrams traveling from source, S, to destination, D, have the following packet structure:

```
                +--------+---------+-------------//-+
                | IPv6   | Source  | IPv6           |
                | Header | Routing | Payload        |
                |        | Header  |                |
                +--------+---------+-------------//-+
```

S's address is carried in the IPv6 Header's Source Address field. D's address is carried in the last entry of SRH for all but the last hop, when D's address is carried in the IPv6 Header's Destination Address field of the packet carrying the SRH.

In a RPL network, Case 2 occurs for all datagrams that have either source or destination outside the RPL Instance, as shown in the following diagram:

```
                  +-----------------+
                  |                 |
                  |  (S) -------> (BR1) -------> (D)
                  |                 |
                  +-----------------+
                       RPL Instance

                  +-----------------+
                  |                 |
                  |  (D) <------- (BR1) <------- (S)
                  |                 |
                  +-----------------+
                       RPL Instance
```

In the above scenario, datagrams that include the SRH in tunneled mode have the following packet structure when traveling within the RPL Instance:

```
           +--------+---------+--------+-------------//-+
           | Outer  | Source  | Inner  | IPv6           |
           | IPv6   | Routing | IPv6   | Payload        |
           | Header | Header  | Header |                |
           +--------+---------+--------+-------------//-+
                               <--- Original Packet --->
            <---           Tunneled Packet            --->
```

Note that the outer header (including the SRH) is added and removed by the RPL Border Router.

Case 2 also occurs whenever a RPL router needs to insert a source route when forwarding datagram. One such use case with RPL is to have all RPL traffic flow through a Border Router and have the Border Router use source routes to deliver datagrams to their final destination. When including the SRH using tunneled mode, the Border Router would encapsulate the received datagram unmodified using IPv6-in-IPv6 and include a SRH in the outer IPv6 header.

```
+-----------------+
|                 |
|  (S) -------\   |
|              \  |
|               (BR1)
|              /  |
|  (D) <------/   |
|                 |
+-----------------+
     RPL Domain
```

In the above scenario, datagrams travel from S to D through BR1. Between S and BR1, the datagrams are routed using the DAG built by RPL and do not contain a SRH. BR1 encapsulates received datagrams unmodified using IPv6-in-IPv6 and the SRH is included in the outer IPv6 header.

3. Format of the RPL Routing Header

The Source Routing Header has the following format:

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Next Header  |  Hdr Ext Len  | Routing Type  | Segments Left |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | CmprI | CmprE |  Pad  |               Reserved                |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    .                                                               .
    .                        Addresses[1..n]                        .
    .                                                               .
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Next Header — 8-bit selector. Identifies the type of header immediately following the Routing header. Uses the same values as the IPv4 Protocol field [RFC3232].

Hdr Ext Len — 8-bit unsigned integer. Length of the Routing header in 8-octet units, not including the first 8 octets. Note that when Addresses[1..n] are compressed (i.e. value of CmprI or CmprE is not 0), Hdr Ext Len does not equal twice the number of Addresses.

Routing Type — 8-bit selector. TBD by IANA.

Segments Left — 8-bit unsigned integer. Number of route segments remaining, i.e., number of explicitly listed intermediate nodes still to be visited before reaching the final destination.

CmprI — 4-bit unsigned integer. Number of prefix octets from each segment, except than the last segment, (i.e. segments 1 through n-1) that are elided. For example, a SRH carrying full IPv6 addresses in Addresses[1..n-1] sets CmprI to 0.

CmprE 4-bit unsigned integer. Number of prefix octets from the last segment (i.e. segment n) that are elided. For example, a SRH carrying a full IPv6 address in Addresses[n] sets CmprE to 0.

Pad 4-bit unsigned integer. Number of octets that are used for padding after Address[n] at the end of the SRH.

Address[1..n] Vector of addresses, numbered 1 to n. Each vector element in [1..n-1] has size (16 - CmprI) and element [n] has size (16-CmprE).

The SRH shares the same basic format as the Type 0 Routing header [RFC2460]. When carrying full IPv6 addresses, the CmprI, CmprE, and Pad fields are set to 0 and the only difference between the SRH and Type 0 encodings is the value of the Routing Type field.

A common network configuration for a RPL Instance is that all routers within a RPL Instance share a common prefix. The SRH introduces the CmprI, CmprE, and Pad fields to allow compaction of the Address[1..n] vector when all entries share the same prefix as the IPv6 Destination Address field of the packet carrying the SRH. The CmprI and CmprE field indicates the number of prefix octets that are shared with the IPv6 Destination Address of the packet carrying the SRH. The shared prefix octets are not carried within the Routing header and each entry in Address[1..n-1] has size (16 - CmprI) octets and Address[n] has size (16 - CmprE) octets. When CmprI or CmprE is non-zero, there may exist unused octets between the last entry, Address[n], and the end of the Routing header. The Pad field indicates the number of unused octets that are used for padding. Note that when CmprI and CmprE are both 0, Pad MUST carry a value of 0.

The SRH MUST NOT specify a path that visits a node more than once. When generating a SRH, the source may not know the mapping between IPv6 addresses and nodes. Minimally, the source MUST ensure that IPv6 Addresses do not appear more than once and the IPv6 Source and Destination addresses of the encapsulating datagram do not appear in the SRH.

Multicast addresses MUST NOT appear in a SRH, or in the IPv6 Destination Address field of a datagram carrying a SRH.

4. RPL Router Behavior

4.1. Generating Source Routing Headers

To deliver an IPv6 datagram to its destination, a router may need to generate a new SRH and specify a strict source route. When the router is the source of the original packet and the destination is known to be within the same RPL Instance, the router SHOULD include the SRH directly within the original packet. Otherwise, the router MUST use IPv6-in-IPv6 tunneling [RFC2473] and place the SRH in the tunnel header. Using IPv6-in-IPv6 tunneling ensures that the delivered datagram remains unmodified and that ICMPv6 errors generated by a SRH are sent back to the router that generated the routing header.

To avoid fragmentation, it is desirable to employ MTU sizes that allow for the header expansion (i.e. at least 1280 + 40 (outer IP header) + SRH_MAX_SIZE), where SRH_MAX_SIZE is the maximum path length for a given RPL network. To take advantage of this, however, the communicating endpoints need to be aware of the MTU along the path (i.e. through Path MTU Discovery). Unfortunately, the larger MTU size may not be available on all links (e.g. 1280 octets on 6LoWPAN links). However, it is expected that much of the traffic on these types of networks consists of much smaller messages than the MTU, so performance degradation through fragmentation would be limited.

4.2. Processing Source Routing Headers

As specified in [RFC2460], a routing header is not examined or processed until it reaches the node identified in the Destination Address field of the IPv6 header. In that node, dispatching on the Next Header field of the immediately preceding header causes the Routing header module to be invoked.

The function of SRH is intended to be very similar to IPv4's Strict Source and Record Route option [RFC0791]. After the routing header has been processed and the IPv6 datagram resubmitted to the IPv6 module for processing, the IPv6 Destination Address contains the next hop's address. When forwarding an IPv6 datagram that contains a SRH with a non-zero Segments Left value, if the IPv6 Destination Address is not on-link, a router SHOULD send an ICMP Destination Unreachable (ICMPv6 Type 1) message with ICMPv6 Code set to (TBD by IANA) to the packet's Source Address. This ICMPv6 Code indicates that the IPv6 Destination Address is not on-link and the router cannot satisfy the strict source route requirement. When generating ICMPv6 error messages, the rules in Section 2.4 of [RFC4443] must be observed.

To detect loops in the SRH, a router MUST determine if the SRH includes multiple addresses assigned to any interface on that router. If such addresses appear more than once and are separated by at least one address not assigned to that router, the router MUST drop the packet and SHOULD send an ICMP Parameter Problem, Code 0, to the Source Address. While this loop check does add significant per-packet processing overhead, it is required to mitigate traffic amplification attacks that led to the deprecation of RH0 [RFC5095].

The following describes the algorithm performed when processing a SRH:

```
    if Segments Left = 0 {
       proceed to process the next header in the packet, whose type is
       identified by the Next Header field in the Routing header
    }
    else {
       compute n, the number of addresses in the Routing header, by
       n = (((Hdr Ext Len * 8) - Pad - (16 - CmprE)) / (16 - CmprI)) + 1

       if Segments Left is greater than n {
          send an ICMP Parameter Problem, Code 0, message to the Source
          Address, pointing to the Segments Left field, and discard the
          packet
       }
       else {
          decrement Segments Left by 1

          compute i, the index of the next address to be visited in
          the address vector, by subtracting Segments Left from n

          if Address[i] or the IPv6 Destination Address is multicast {
             discard the packet
          }
          else if 2 or more entries in Address[1..n] are assigned to
                  local interface and are separated by at least one
                  address not assigned to local interface {
             send an ICMP Parameter Problem (Code 0) and discard the
             packet
          }
          else {
             swap the IPv6 Destination Address and Address[i]

             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
             }
             else {
                decrement the Hop Limit by 1

                resubmit the packet to the IPv6 module for transmission
                to the new destination
             }
          }
       }
    }
```

5. RPL Border Router Behavior

RPL Border Routers (referred to as LBRs in [I-D.ietf-roll-terminology]) are responsible for ensuring that a SRH is only used within the RPL Instance it was created.

For datagrams destined to the RPL Border Router, the router processes the packet in the usual way. For instance, if the SRH was included using tunneled mode and the RPL Border Router serves as the tunnel endpoint, the router removes the outer IPv6 header, at the same time removing the SRH as well.

Datagrams destined elsewhere within the same RPL Instance are forwarded to the correct interface.

Datagrams destined to nodes outside the RPL Instance are dropped if the outer-most IPv6 header contains a SRH.

6. Security Considerations

6.1. Source Routing Attacks

[RFC5095] deprecates the Type 0 Routing header due to a number of significant attacks that are referenced in that document. Such attacks include network discovery, bypassing filtering devices, denial-of-service, and defeating anycast.

Because this document specifies that SRH is only for use within a RPL Instance, such attacks cannot be mounted from outside a RPL Instance. As described in Section 5, RPL Border Routers MUST drop datagrams entering or exiting a RPL Instance that contain a SRH in the IPv6 Extension headers.

6.2. ICMPv6 Attacks

The generation of ICMPv6 error messages may be used to attempt denial-of-service attacks by sending error-causing SRH in back-to-back datagrams. An implementation that correctly follows Section 2.4 of [RFC4443] would be protected by the ICMPv6 rate limiting mechanism.

7. IANA Considerations

This document defines a new IPv6 Routing Type of TBD by IANA.

This document defines a new ICMPv6 Destination Unreachable Code of TBD by IANA to indicate that the router cannot satisfy the strict source-route requirement.

8. Acknowledgements

The authors thank Jari Arkko, Richard Kelsey, Suresh Krishnan, Erik Nordmark, Pascal Thubert, and Tim Winter for their comments and suggestions that helped shape this document.

9. Changes

(This section to be removed by the RFC editor.)

Draft 05:

- Address LC comments.

Draft 04:

- Updated text on recommendations for avoiding fragmentation.

- Clarify definition of CmprE where it is first mentioned.

- Change use of IPv6-in-IPv6 tunneling from SHOULD to MUST.

- Update packet processing pseudocode to match the text on sending back a parameter problem error.

- Recommend that non-RPL devices drop packets with SRH by default.

- Clarify packet structure figures.

- State that checking for cycles represents significant per-packet processing.

Draft 03:

- Removed any presumed values that are TBD by IANA.

Draft 02:

- Updated to send ICMP Destination Unreachable error only after the SRH has been processed.

- Updated pseudocode to reflect encoding changes in draft-01.

- Allow multiple addresses assigned to same node as long as they are not separated by other addresses.

Draft 01:

- Allow Addresses[1..n-1] and Addresses[n] to have a different number of bytes elided.

10. References

10.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[RFC4443] Conta, A., Deering, S., and M. Gupta, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, March 2006.

[RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation of Type 0 Routing Headers in IPv6", RFC 5095, December 2007.

10.2. Informative References

[I-D.ietf-roll-rpl] Winter, T., Thubert, P., Brandt, A., Clausen, T., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., and J. Vasseur, "RPL: IPv6 Routing Protocol for Low power and Lossy Networks", draft-ietf-roll-rpl-19 (work in progress), March 2011.

[I-D.ietf-roll-terminology] Vasseur, J., "Terminology in Low power And Lossy Networks", draft-ietf-roll-terminology-06 (work in progress), September 2011.

[RFC3232] Reynolds, J., "Assigned Numbers: RFC 1700 is Replaced by an On-line Database", RFC 3232, January 2002.

Authors' Addresses

Jonathan W. Hui
Cisco Systems, Inc
170 West Tasman Drive
San Jose, California 95134
USA

Phone: +408 424 1547
Email: jonhui@cisco.com

JP Vasseur
Cisco Systems, Inc
11, Rue Camille Desmoulins
Issy Les Moulineaux, 92782
France

Email: jpv@cisco.com

David E. Culler
UC Berkeley
465 Soda Hall
Berkeley, California 94720
USA

Phone: +510 643 7572
Email: culler@cs.berkeley.edu

Vishwas Manral
Hewlett Packard Co.
19111 Pruneridge Ave.
Cupertino, California 95014
USA

Email: vishwas.manral@hp.com