TRILL Working Group
Internet Draft
Intended Category: Proposed Standard

M. Zhang
D. Eastlake 3rd
Huawei
R. Perlman
EMC
M. Cullen
Painless Security
H. Zhai
JIT
D. Liu
China Telcom Co., Ltd

Expires: May 27, 2017
November 23, 2016

TRILL Multilevel Using Unique Nicknames
draft-ietf-trill-multilevel-unique-nickname-01.txt

Abstract

TRILL routing can be extended to support multiple levels by building on the multilevel feature of IS-IS routing. Depending on how nicknames are managed, there are two primary alternatives to realize TRILL multilevel: the unique nickname approach and the aggregated nickname approach as discussed in [MultiL]. This document specifies the unique nickname approach. This approach gives unique nicknames to all TRILL switches across the multilevel TRILL campus.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html

Copyright and License Notice

Copyright (c) 2016 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 (http://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. Acronyms and Terminology 3
3. Data Routing 4
3.1. Unicast Routing 4
3.2. Multicast Routing 5
3.2.1. Local Distribution Trees 5
3.2.2. Global Distribution Trees 5
4. Protocol Basics and Extensions 8
4.1. Multilevel TRILL Basics 8
4.2. Nickname Allocation 8
4.3. Nickname Announcements 9
4.4. Capability Indication 11
5. Mix with Aggregated nickname Areas 11
6. Security Considerations 11
7. IANA Considerations 12
8. References 12
8.1. Normative References 12
8.2. Informative References 13
Author's Addresses 15

1. Introduction

The multiple level feature of [IS-IS] can increase the scalability of TRILL as discussed in [MultiL]. However, multilevel IS-IS needs some extensions to support the TRILL multilevel feature. The two most significant extensions are how TRILL switch nicknames are managed and how distribution trees are handled [MultiL].

There are two primary alternatives to realize TRILL multilevel

[MultiL]. One approach, which is referred as the "aggregated nickname" approach, involves assigning nicknames to the areas, and allowing nicknames to be reused in different areas, by having the border TRILL switches rewrite nickname fields when entering or leaving an area. For more description about the aggregated nickname approach, one can refer to [MultiL] and [SingleN]. The other approach, which is referred as the "unique nickname" approach, is specified in this document. Unique nickname approach gives unique nicknames to all the TRILL switches in the multilevel campus, by having the Level-1/Level-2 border TRILL switches advertise into the Level 1 area which nicknames are not available for assignment in the area, and insert into Level 2 area which nicknames are used by this area so that other areas cannot use them anymore, as well as informing the rest of the campus how to reach the nicknames residing in this area. In the document, protocol extensions that support such advertisement are specified.

Each RBridge in a unique nickname area calculates two types of trees: local distribution trees and global distributions trees. For multi-destination traffic that is limited to an area, the packets will be flooded on the local distribution tree. Otherwise, the multi-destination packets will be flooded along the global distribution tree.

In the unique nickname approach, nicknames are globally valid so that border RBridges do not rewrite the nickname field of TRILL data packets that are transitions between Level 1 and Level 2, as border RBrides do in the aggregated nickname approach. If a border RBridge is a transit node on a forwarding path, it does not learn MAC addresses of the TRILL data packets forwarded along this path. Testing and maintenance operations that originate in one area and terminate in a different area are also simplified [MultiL]. For these reasons, unique nickname approach might realize simpler border RBridges than the aggregated nickname approach. However, the unique nickname approach is less scalable and may be less well suited for very large campuses.

2. Acronyms and Terminology

Data Label: VLAN or FGL

IS-IS: Intermediate System to Intermediate System [IS-IS]

RBridge: A device implementing the TRILL protocol.

TRILL: TRansparent Interconnection of Lots of Links or Tunneled Routing in the Link Layer [RFC6325].

TRILL switch: An alternative name for an RBridge.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Data Routing

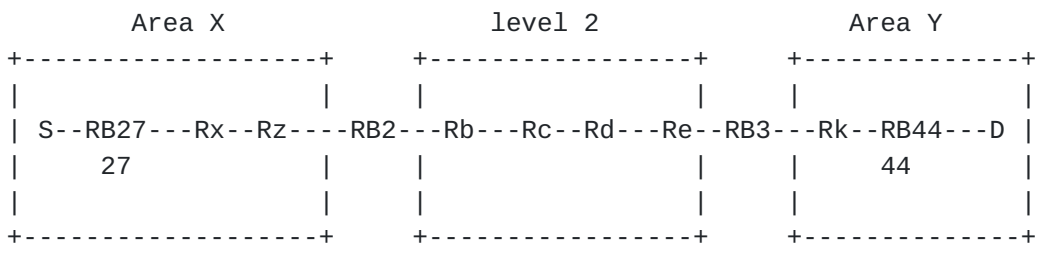


Figure 3.1: An example topology for TRILL multilevel

Figure 3.1 is adapted from the example topology of [MultiL].

The routing processes are described in the following two subsections.

3.1. Unicast Routing

The plain RBridge RB27 has a different view of the topology of the TRILL campus than its border RBridge RB2. For an outward path that reaches an RBridge not in the same area (say RB44), RB27 calculates the segment of the path in Area X, the border RBridge RB2 calculates the segment in Level 2, while the border RBridge to the destination area, RBridge RB3, calculates the segment from itself to RB44.

Let's say that S transmits a frame to destination D and let's say that D's location is learned by the relevant TRILL switches already. These relevant switches have learned the following:

1) RB27 has learned that D is connected to nickname 44.

The following sequence of events will occur:

- S transmits an Ethernet frame with source MAC = S and destination MAC = D.

- RB27 encapsulates with a TRILL header with ingress RBridge = 27, and egress RBridge = 44 producing a TRILL Data packet.

- RB2 has announced in the Level 1 IS-IS instance in Area X, that it owns all nicknames of other areas, including 44. Therefore, IS-IS routes the packet to RB2.

- The packet is forwarded through Level 2, from RB2 to RB3, which has advertised, in Level 2, it owns the nickname 44.

- RB3, when forwarding into Area Y, does not change the ingress nickname 27 or the egress nickname 44.

- RB44, when decapsulating, learns that S is attached to nickname 27.

3.2. Multicast Routing

The scope of multicast routing is defined by the tree root nickname. A tree with a Level 2 tree root nickname is global and a tree with Level 1 tree root nickname is local. See Section 4.2 for the Level 1 and Level 2 nickname allocation.

Border RBridges announce the global trees to be calculated only for those Data Labels that span across areas. APPsub-TLVs as specified in Section 3.2 of [TreeSel] will be advertised for this purpose. Based on the Data Label, an ingress RBridge can determine whether a global tree or a local tree is to be used for a TRILL multi-destination Data packet.

If there are legacy TRILL switches that do not understand the APPsub-TLVs for tree selection, configuration MUST guarantee that global Data Labels are disabled on these legacy TRILL switches (Otherwise, the legacy TRILL switches might use local trees for multi-destination traffic with a global scope.). These legacy TRILL switches may use global trees to flood multi-destination packets with a scope of the local area. Those global trees MUST be pruned at the border TRILL switches based on Data Labels.

3.2.1. Local Distribution Trees

The root RBridge RB1 of a local distribution tree resides in the area. RBridges in this area calculate this local tree based on the link state information of this area, using RB1's nickname as the root. Protocol behaviors for local distribution trees have been specified in 4.5 of [RFC6325]. The only different is that the local distribution tree spans this area only. A multi-destination packet with an egress nickname of the root RBridge of a local tree MUST NOT be leaked into Level 2 at the border RBridge.

3.2.2. Global Distribution Trees

Within Level 2, the RBridge with the highest tree root priority advertises the set of global trees by providing a list of Level 2 RBridge nicknames just as defined in Section 4.5 of [RFC6325].

According to [RFC6325], the RBridge with the highest root priority advertises the tree roots for a Level 1 area. There has to be a border RBridge with the highest root tree priority in each area so that it can advertises the global tree root nicknames into the area. Also, this border RBridge needs to advertise the set of local distribution trees by providing another set of nicknames. Since nicknames of global tree roots and local tree roots indicate different flooding scopes, these two set MUST NOT overlap. If a border RBridge has been assigned both as a global tree root and a local tree root, it has to acquire both a global tree root nickname(s) and local tree root nickname(s). However, non-border RBridges in an area do not differentiate between a global tree root nickname and a local tree root nickname.

Suppose RB3 is the RBridge with the highest tree root priority within Level 2, and RB2 is the highest tree root priority in Area X. RB2 advertises in Area X that nickname RB3 is the root of a distribution tree. Figure 3.2 through Figure 3.5 illustrate how different RBridges view the global distribution tree.

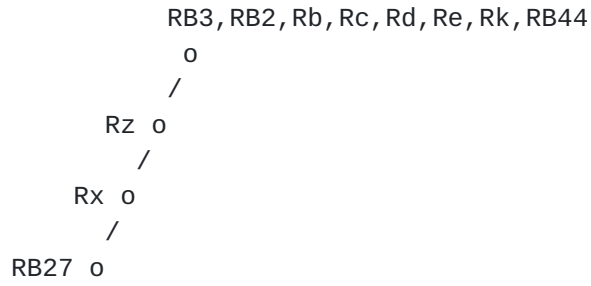


Figure 3.2: RB27's view of the global distribution tree

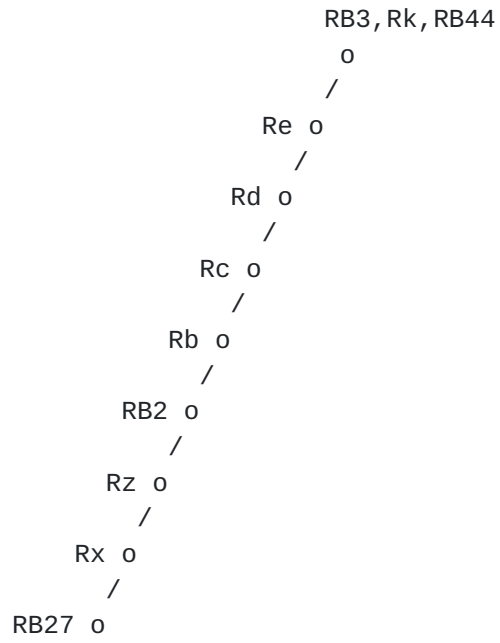


Figure 3.3: RB2's view of the global distribution tree

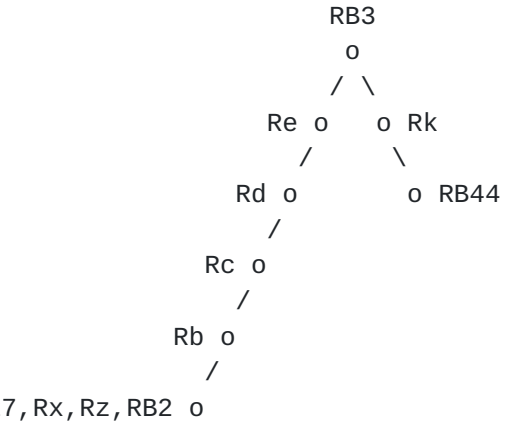


Figure 3.4: RB3's view of the global distribution tree

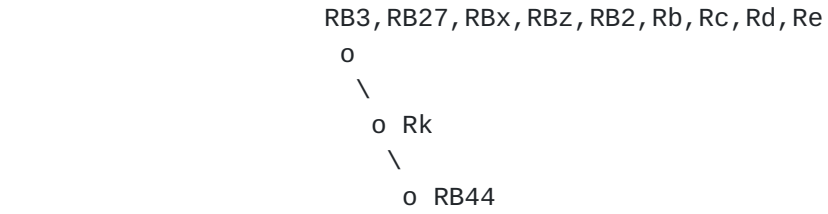


Figure 3.5: RB44's view of the global distribution tree

The following sequence of events will occur when a multi-destination TRILL Data packet is forwarded using the global distribution tree:

- RB27 produces a multi-destination (M bit is one) TRILL Data packet with ingress RBridge = 27 and egress RBridge = 3. RB27 floods this packet using the segment of the global distribution tree that resides in Area X.

- RB2, when flooding the packet in Level 2, uses the segment of the global distribution tree that resides in Level 2.

- RB3, when flooding the packet into Area Y, uses the segment of the global distribution tree that resides in Area Y.

- The multicast listener RB44, when decapsulating the received packet, learns that S is attached to nickname 27.

4. Protocol Basics and Extensions

4.1. Multilevel TRILL Basics

Multilevel TRILL builds on the multilevel feature of [IS-IS]. Border RBridges are in both a Level 1 area and in Level 2. They establish adjacency with Level 1 RBridges as specified in [RFC7177] and [RFC6325]. They establish adjacency with Level 2 RBridges in exactly the same way except that (1) for a LAN link the IS-IS Hellos used are Level 2 Hello PDUs [IS-IS] and (2) for a point-to-point link the Level is configured and indicated in flags in the point-to-point Hello. The state machines for Level 1 and Level 2 adjacency are independent and two RBridges on the same LAN link can have any adjacency state for Level 1 and, separately, any adjacency state for Level 2. Level 1 and Level 2 link state flooding are independent using Level 1 and Level 2 versions of the relevant IS-IS PDUs (LSP, CSNP, PSNP, FS-LSP, FS-CSNP and FS-PSNP). Thus Level 1 link state information stays within a Level 1 area and Level 2 link state information stays in Level 2 unless there are specific provisions for leaking (copying) information between levels. This is why multilevel can address the TRILL scalability issues as specified in Section 2 of [MultiL].

The former "campus wide" minimum acceptable link size Sz is calculated as before by Level 1 RBridges (including border RBridges) using the originatingLSPBufferSize advertised in Level 1 LSP so it is area local in multilevel TRILL. A minimum acceptable link size in Level 2, called Sz2, is calculated by the RBridges participating in Level 2 in the same way as Sz is calculated but using the originatingLSPBufferSize distributed in Level 2 LSPs.

4.2. Nickname Allocation

Level 2 RBridges contend for nicknames in the range from 0xF000 through 0xFBFF the same way as specified in [RFC6325], using Level 2 LSPs. The highest priority border router for a Level 1 area should contend with others in Level 2 for smallish blocks of nicknames for the range from 0x0001 to 0xEFFF. Blocks of 64 aligned on multiple of 64 boundaries are RECOMMENDED in this document.

The nickname contention in Level 2 will figure out which blocks of nicknames are available for an area and which blocks of nicknames are used else where. The NickBlockFlags APPsub-TLV as specified in Section 4.3 will be used by the border RBridge(s) to announce the nickname availability.

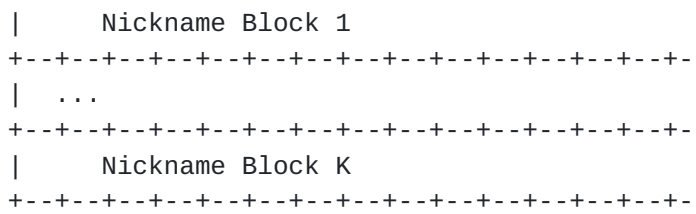
4.3. Nickname Announcements

Border RBridges need to exchange nickname information between Level 1 and Level 2, otherwise forwarding paths inward/outward will not be calculated. For this purpose, border RBridges need to fabricate nickname announcements. Sub-TLVs used for such artificial announcements are specified as follows.

Besides its own nickname(s), a border RBridge needs to announce, in its area, the ownership of all external nicknames that are reachable from this border RBridge. These external nicknames include nicknames used in other unique nickname areas and nicknames in Level 2. Non-border RBridge nicknames within aggregated nickname areas are excluded. Also, a border RBridge needs to announce, in Level 2, the ownership of all nicknames within its area. From listening to these Level 2 announcements, border RBridges can figure out the nicknames used by other areas.

RBridges in the TRILL base protocol use the Nickname Sub-TLV as specified in Section 2.3.2 of [RFC7176] to announce the ownership of nicknames. However, it becomes uneconomic to use this Sub-TLV to announce a mass of internal/external nicknames. To address this issue, border RBridges should make use of the NickBlockFlags APPsub-TLV to advertise into the Level 1 area the inclusive range of nicknames that are available or not for self allocation by the Level 1 RBridges in that area. Its structure is as follows:





- Type: tbd2 (TRILL NickBlockFlags)

- Length: 2 + 2*K where K is the number of nickname blocks.

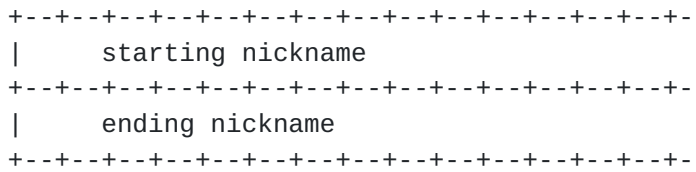
- OK:

  - When this bit is set to 1, the blocks of nicknames in this APPsub-TLV are available for Level 1 use of the area. The APPsub-TLV will be advertised in both Level 1 and Level 2. For nicknames that fall in the ranges or the nickname blocks, RBridges of Level 2 always route to the originating border RBridge, just as if this border RBridge owns these nicknames.

  - When this bit is set to 0, it indicates that the nicknames covered by the nickname blocks are being used in Level 2 or other areas so that they are not available for Level 1 use of the area. The APPsub-TLV will be advertised into Level 1 only. For nicknames that fall in the ranges of the nickname blocks, RBridges of the area always route to the originating border RBridge, just as if this border RBridge owns these nicknames.

- RESV: reserved for future flag allocation. MUST be sent as zero and ignored on receipt.

- Nickname Block: a starting and ending nickname as follows:



For nicknames in these ranges, other RBridges will deem that they are owned by the originating border RBridge. The paths to nicknames that fall in these ranges will be calculated to reach the originating border RBridge. TRILL Data packets with egress nicknames that are neither in these ranges nor announced by any RBridge in the area MUST be discarded.

There might be multiple border RBridges connected to the same area. Each border RBridges may advertise a subset of the entire internal/external nickname space in order to realize load balance. However, optimization of such load balance is an implementation issue and is out the scope of this document.

As specified in Section 4.2.6 of [RFC6325], multiple border RBridges may claim the same nicknames outward and/or inward. Other RBridges add those nicknames as if they are attached to all of those border RBridges.

4.4. Capability Indication

All border RBridge MUST understand the NickBlockFlags APPsub-TLV. Non border RBridges in an area SHOULD understand the NickBlockFlags APPsub-TLV. If an RBridge within an area understands the NickBlockFlags APPsub-TLV, it MUST indicate this capability by announcing it in its TRILL-VER Sub-TLV. (See Section 7).

If there are RBridges that do not understand the NickBlockFlags APPsub-TLV, border RBridges of the area will also use the traditional Nickname Sub-TLV [RFC7176] to announce into the area those nicknames covered by the nickname blocks of the NickBlockFlags APPsub-TLV whose OK is 0. The available range of nicknames for this area should be configure on these traditional RBridges.

5. Mix with Aggregated nickname Areas

The design of TRILL multilevel allows a mixture of unique nickname areas and aggregated nickname areas (see Section 1.2 of [MultiL]). Usage of nickname space must be planed so that nicknames used in any one unique nickname area and Level 2 are never used in any other areas which includes unique nickname areas as well as aggregated nickname areas. In other words, nickname re-usage is merely allowed among aggregated nickname areas.

Border RBridges of an aggregated area need to announce nicknames heard from Level 2 into their area like just like an unique nickname border RBridge. But these RBridges do not announce nicknames of their area into Level 2.

Each border RBridge of the aggregated areas will appear on the global tree, as specified in Section 4.1, as a single node. The global trees for unique nickname areas span unique nickname areas and Level 2 but never reach the inside of aggregated areas.

6. Security Considerations

With TRILL multilevel, flooding of control traffic for link state information of Level 1 and Level 2 is separated. This addresses the TRILL scalability issues as specified in Section 2 of [MultiL] and also confines the effective scope of possible malicious events. However, due to the nature that unique nickname areas share a unique nickname space, border RBridges still have to leak nickname information between levels. For this purpose, border RBridges need to fabricate the nickname announcements as specified in Section 4.3. Malicious devices may also fake the NickBlockFlags APPsub-TLV to announce a range of nicknames. By doing this, the attacker can attract TRILL data packets that are originally to reach a bunch of other RBridges. For this reason, RBridges SHOULD be configured to include the IS-IS Authentication TLV (10) in the IS-IS PDUs that contains the NickBlockFlags APPsub-TLV, so that IS-IS security ([RFC5304] [RFC5310]) can be used to secure the network.

If border RBridges do not prune multi-destination distribution tree traffic in Data Labels that are configured to be area local, then traffic that should have been contained within an area might be wrongly delivered to end stations in that Data Label in other areas. This would generally violate security constraints.

For general TRILL Security Considerations, see [RFC6325].

7. IANA Considerations

IANA is requested to register a new flag bit with mnemonic "B" (Block of Nicknames) under the TRILL-VER Sub-TLV Capabilities registry.

Bit	Mnemonic	Description	Reference
tbd1	B	Able to handle the NickBlockFlags APPsub-TLV	[This document]

IANA is requested to assign a new type for the NickBlockFlags APPsub-TLV from the range available below 256 and add the following entry to the "TRILL APPsub-TLV Types under IS-IS TLV 251 Application Identifier 1" registry as follows:

Type	Name	Reference
tbd2	NickBlockFlags	[This document]

8. References

8.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, DOI 10.17487/RFC6325, July 2011, <http://www.rfc-editor.org/info/rfc6325>.

[TreeSel] Li, Y., Eastlake, D., et al, "TRILL: Data Label based Tree Selection for Multi-destination Data", draft-ietf-trill-tree-selection, Work in Progress.

[RFC7176] Eastlake 3rd, D., Senevirathne, T., Ghanwani, A., Dutt, D., and A. Banerjee, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 7176, DOI 10.17487/RFC7176, May 2014, <http://www.rfc-editor.org/info/rfc7176>.

[RFC7177] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Yang, H., and V. Manral, "Transparent Interconnection of Lots of Links (TRILL): Adjacency", RFC 7177, DOI 10.17487/RFC7177, May 2014, <http://www.rfc-editor.org/info/rfc7177>.

[IS-IS] International Organization for Standardization, "Information technology -- Telecommunications and information exchange between systems -- Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, November 2002.

8.2. Informative References

[MultiL] Perlman, R., Eastlake, D., et al, "Alternatives for Multilevel TRILL (Transparent Interconnection of Lots of Links)", draft-ietf-trill-rbridge-multilevel, Work in Progress.

[SingleN] Zhang, M., Eastlake, D., et al, "Single Area Border RBridge Nickname for TRILL Multilevel", draft-ietf-trill-multilevel-single-nickname, Work in Progress.

[RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <http://www.rfc-editor.org/info/rfc5304>.

[RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <http://www.rfc-editor.org/info/rfc5310>.

Author's Addresses

Mingui Zhang
Huawei Technologies
No. 156 Beiqing Rd., Haidian District
Beijing 100095
China

Phone: +86-13810702575
Email: zhangmingui@huawei.com

Donald E. Eastlake 3rd
Huawei Technologies
155 Beaver Street
Milford, MA 01757
United States

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Radia Perlman
EMC
2010 256th Avenue NE, #200
Bellevue, WA 98007
United States

Email: radia@alum.mit.edu

Margaret Cullen
Painless Security
14 Summer St. Suite 202
Malden, MA 02148
United States

Email: margaret@painless-security.com

Hongjun Zhai
Jinling Institute of Technology
99 Hongjing Avenue, Jiangning District
Nanjing, Jiangsu 211169
China

Email: honjun.zhai@tom.com

Dongxin Liu
China Telcom Co., Ltd
109 West Zhongshan Ave, Tianhe District
Guangzhou 510630
P.R. China

Email: liudx@gsta.com