

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: April 24, 2014

Mingui Zhang
Huawei
Russ White
Verisign
Hongjun Zhai
ZTE
October 21, 2013

Control Plane Requirements for TRILL Active/Active Edge
draft-zhang-trill-active-active-cp-req-00.txt

Abstract

TRILL Active/Active Edge enables a Multi-Chassis Link Aggregation Group to connect to multiple RBridges which can ingress and egress data traffic for the same VLAN at the same time. The purpose of introducing the TRILL Active/Active Edge is to increase the access bandwidth and resilience. This new access type puts forward new requirements for TRILL control plane. Current TRILL control plane need be extended in order to make the TRILL Active/Active Edge operational. Requirements are developed in this document as the guidelines for designing those specific control plane functions.

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/lid-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) 2013 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	3
2. Acronyms and Terminology	3
2.1. Acronyms	3
2.2. Terminology	3
3. Control Plane Requirements	3
3.1. Discovery	4
3.2. Election	4
3.3. Forwarding Information Synchronization	4
3.4. Failure Detection and Notification	5
3.5. Communicating Configuration Information	5
4. Security Considerations	5
5. IANA Considerations	6
6. References	6
6.1. Normative References	6
6.2. Informative References	6
Author's Addresses	8

1. Introduction

TRILL makes use of IS-IS link state routing to provide least-cost forwarding between TRILL switches. At the edge, [RFC6349] already defines an active-standby access for end-stations. An active-active method is to be added in TRILL so that end stations are able to increase the bandwidth and resilience of their access to a TRILL campus using Multi-Chassis Link Aggregation Group (MC-LAG) [PS]. Unlike a LAN link, MC-LAG does not exchange TRILL IS-IS PDUs. The TRILL switch ports attached to the MC-LAG demarcate the edge of TRILL and no adjacency can be formed on top of the MC-LAG.

From the point of view of the end stations, the MC-LAG is treated as if it were a single link and those RBridges connected to the other end of the MC-LAG need operate as if they were a single TRILL switch. Thus an Active/Active Edge (AAE) is set up. However, it doesn't mean that RBridges in the AAE can be connected to only one MC-LAG. It's possible that one port of an RBridge is connected to one MC-LAG and the other port is connected to another.

To achieve the TRILL Active/Active Edge, some functions should be added to the current TRILL control plane. There are several possible places to host these functions. For example, the TRILL IS-IS, TRILL BFD, ESADI protocol and TRILL Channel are potential choices [BFD] [ESADI] [Channel]. This document describes the high-level requirements for these new control plane functions. When specific protocols are designed, these requirements should be followed.

2. Acronyms and Terminology

2.1. Acronyms

MC-LAG: Multi-Chassis Link Aggregation Group
IS-IS: Intermediate System to Intermediate System
TRILL: TRAnsparent Interconnection of Lots of Links
AAE: Active/Active Edge

2.2. Terminology

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].

Familiarity with [RFC6325], [RFC6327], [6327bis] and [RFC6439] is assumed in this document.

3. Control Plane Requirements

This section specifies the requirements on the functions that the TRILL control plane need to provide for the AAE.

3.1. Discovery

When an RBridge is attached to an MC-LAG, it need to recognize other RBridges attached to this MC-LAG as in the same AAE. It also need to notify other RBridges the fact that it joins or leaves the AAE.

The identification of the AAE is REQUIRED during the discovery. The System Identifier of the MC-LAG is a choice. RBridges in an AAE MUST include this ID in the Protocol Data Units of the control plane protocol to achieve the discovery.

3.2. Election

RBridges per [RFC6325] run a protocol on the link to elect the Designated RBridge (DRB). The DRB performs some common tasks for the link. For example, the DRB gives the link a pseudonode nickname.

RBridges in an AAE need elect a master node to carry on common tasks for the AAE as well. Since MC-LAG disables the delivery of TRILL IS-IS PDUs. The TRILL IS-IS election protocol defined in Section 2.1 of [RFC6325] is not applicable here.

A substitute election protocol SHOULD be set up. Such protocol should reuse the decision process algorithm defined in [ISIS] to avoid introducing too much complexity into TRILL.

3.3. Forwarding Information Synchronization

As an example, AAE members may learn MAC addresses through data plane learning and ESADI protocol. MAC addresses learnt by one member should be shared to other members in the same AAE in order to reduce the unknown unicast traffic [PS]. As another example, the IGMP (Internet Group Management Protocol) [RFC3376] snooping on those ports attached to a MC-LAG has to be synchronized. A protocol is needed to synchronize these forwarding information among AAE members and keep the information updated in a timely manner.

MAC address may be frequently updated due to data plane learning. It is REQUIRED that the CPU is not overloaded due to the control plane MAC address update. Therefore the protocol should define a minimum updating interval.

Due to the overhead that may be produced, this protocol SHOULD be confined in the scope of the AAE members. Obviously, it SHOULD NOT be realized though the extension of TRILL IS-IS, which may otherwise

introduce a heavy burden to current TRILL's control plane. TRILL ESADI or TRILL Channel are proper candidates to realize such protocol.

3.4. Failure Detection and Notification

When a link of the MC-LAG to an AAE member RBridge or this RBridge itself is failed, this failure should be detected and notified to other AAE member RBridges through the control plane as soon as possible. RBridges other than AAE member RBridges may need be notified as well so that these RBridges can change their forwarding paths to avoid the failure.

The failed AAE member RBridge leaves the AAE. It's possible that there is less than two RBridges in the AAE due to the failure. Then the AAE should be destroyed. If the AAE is not destroyed, the failure notification will trigger a re-convergence of the AAE. It is optional to establish a dedicated session such as BFD to detect the failure in order to enable a fast convergence. All AAE members MUST run into a consensus converge state just like the convergence in IGP routing protocols. The control plane protocols need take the design of the re-convergence algorithm into consideration.

3.5. Communicating Configuration Information

Configuration on RBridges to enable the operation of an AAE should be minimized. Some of the configuration is local while the other is of global sense and need be conveyed through the control plane to other RBridges. An example is the Affinity Sub-TLV defined in [6326bis] and used in [CMT].

The communication of the configuration information through the control plane helps to settle mis-configuration. For example, enabled VLANs of every port attached to the same MC-LAG MUST be the same. An RBridge in an AAE can report the enabled VLANs to others through the control plane so that a mis-configured VLAN can be rectified or trigger an alarm to the management plane.

4. Security Considerations

Security issue should be considered when a specific extension is made to the existing TRILL control plane.

Authenticity for contents transported in IS-IS PDUs is enforced using regular IS-IS security mechanism [ISIS][RFC5310].

For security considerations pertain to extensions hosted by TRILL BFD and ESADI, corresponding documents should refer to the Security

Considerations in [BFD], [ESADI] and [Channel].

5. IANA Considerations

This document requires no IANA actions. RFC Editor: please remove this section before publication.

6. References

6.1. Normative References

- [RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.
- [RFC6327] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Dutt, D., and V. Manral, "Routing Bridges (RBridges): Adjacency", RFC 6327, July 2011.
- [6327bis] D. Eastlake, R. Perlman, et al, "TRILL: Adjacency", draft-ietf-trill-rfc6327bis-01.txt, July 2013, working in progress.
- [RFC6439] Perlman, R., Eastlake, D., Li, Y., Banerjee, A., and F. Hu, "Routing Bridges (RBridges): Appointed Forwarders", RFC 6439, November 2011.
- [BFD] V. Manral, D. Eastlake, et al, "TRILL (Transparent Interconnection of Lots of Links): Bidirectional Forwarding Detection (BFD) Support", draft-ietf-trill-rbridge-bfd-07.txt, July 2012, working in progress.
- [ESADI] H. Zhai, F. Hu, et al, "TRILL (Transparent Interconnection of Lots of Links): ESADI (End Station Address Distribution Information) Protocol", draft-ietf-trill-esadi-03.txt, July 2013, working in progress.
- [6326bis] D. Eastlake, T. Senevirathne, et al, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", draft-ietf-isis-rfc6326bis-01.txt, April 2013, working in progress.
- [Channel] D. Eastlake, V Manral, et al, "TRILL: RBridge Channel Support", draft-ietf-trill-rbridge-channel-08.txt, July 2012, working in progress.

6.2. Informative References

- [PS] M. Zhang and D. Eastlake, "Problem Statement: TRILL Active/Active Edge", draft-zhang-trill-aggregation-04.txt, August 2013, working in progress.

- [ISIS] ISO, "Intermediate system to Intermediate system routing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002.

- [CMT] T. Senevirathne, J. Pathangi, et al, "Coordinated Multicast Trees (CMT)for TRILL", draft-ietf-trill-cmt-02.txt, November 2012, working in progress.

- [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, February 2009.

Author's Addresses

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

Email: zhangmingui@huawei.com

Russ White
Verisign
12061 Bluemont Way
Reston, VA 20190
USA

Email: riwhite@verisign.com

Hongjun Zhai
ZTE
68 Zijinghua Road, Yuhuatai District
Nanjing, Jiangsu 210012
China

Phone: +86 25 52877345
Email: zhai.hongjun@zte.com.cn