TRILL-over-IP

draft-ietf-trill-over-ip-10.txt

IETF 99, Prague

Margaret Cullen margaret@painless-security.com
Mingui Zhang, Donald Eastlake, Dacheng Zhang.

Contents

- Basic Summary
  - What TRILL over IP is about
- Chronology
  - Status in IETF Process
- Outstanding Comments
  - From TSVART review and other commenters

Basic Summary

- "TRILL over IP" treats an IP network as a link connecting TRILL switch ports, thus providing a method to connect TRILL sites into a single TRILL campus.
- Two Scenarios are described in the draft
  - Remote Office Scenario
  - IP Backbone Scenario
- The draft specifies packet format and security. And it tries to cover transport considerations including congestion, MTU, fat flows, QoS, and middleboxes.

Basic Summary (continued)

- The existing TRILL-over-IP draft specifies three formats for TRILL over IP transport. It is extensible so additional formats can be specified in the future.
- A simplified version of the formats, on the wire, are as follows, when not secure. (Security adds an IPsec layer.)
  - Link(**UDP**(TRILL))
  - Link(UDP(**VXLAN**(Ethernet(TRILL))))
  - Link(**TCP**(TRILL))

Chronology

- Personal draft first posted 24 October 2011.
- WG draft first posted 24 March 2014.
- WG Last Call passed on 13 January 2017
- There are other comments but the best compendium is the TSVART review posted 15 June 2017:
  - https://www.ietf.org/mail-archive/web/trill/current/msg07801.html

Outstanding Comments

- Zero Checksum
- ECN
- Quality of Service
- TCP Transport
- Fragmentation
- Congestion Control
- NAT

Zero Checksum

- In order to run at line speed, it is likely that TRILL over IP needs to dispense with the IP header checksum and use zero. This is mentioned in the draft with a reference to RFC 6936 “Applicability Statement for the Use of IPv6 UDP Datagrams with Zero Checksums”.
- However the draft needs to be clearer about when TRILL over IP meets the conditions for use of zero checksum given in RFC 6936.

Congestion Control

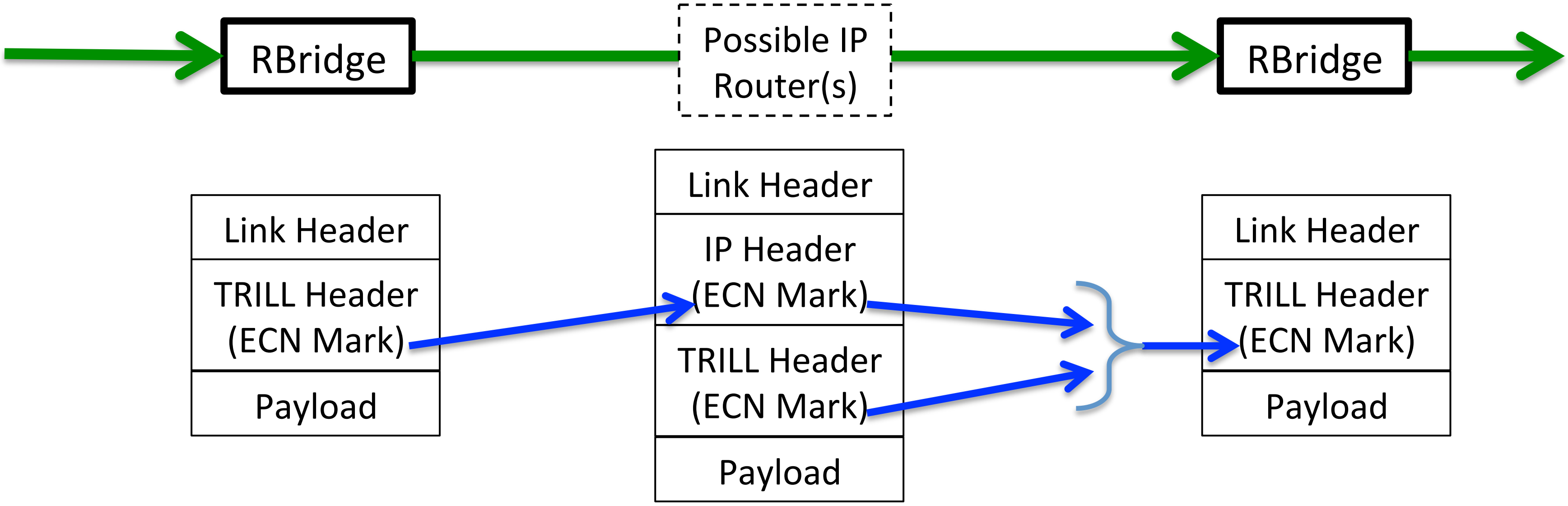
- Current draft is a good start on congestion control but needs a couple of additional things:
  - What is the effect of using serial unicast for multicast in some cases?
  - How can TRILL tell if the traffic it is carrying is congestion controlled?

ECN

- TRILL-over-IP currently says nothing about ECN (Explicit Congestion Notification, see https://tools.ietf.org/html/draft-ietf-trill-ecn-support-03).
- A TRILL-over-IP link is an IP tunnel and should support ECN as described in RFC 6040.
  – When the TRILL campus supports ECN, the TRILL-over-IP header should be ECN marked based on ECN marking in the TRILL Header of the packet being transported with IP.
  – If ECN is not supported in the particular TRILL campus, it does NOT seem worth while to dig past the TRILL layer to get ECN from/to inner IP traffic.

ECN (continued)

Should be supported:



RBridge

Possible IP Router(s)

RBridge

Link Header

TRILL Header (ECN Mark)

Payload

Link Header

IP Header (ECN Mark)

TRILL Header (ECN Mark)

Payload

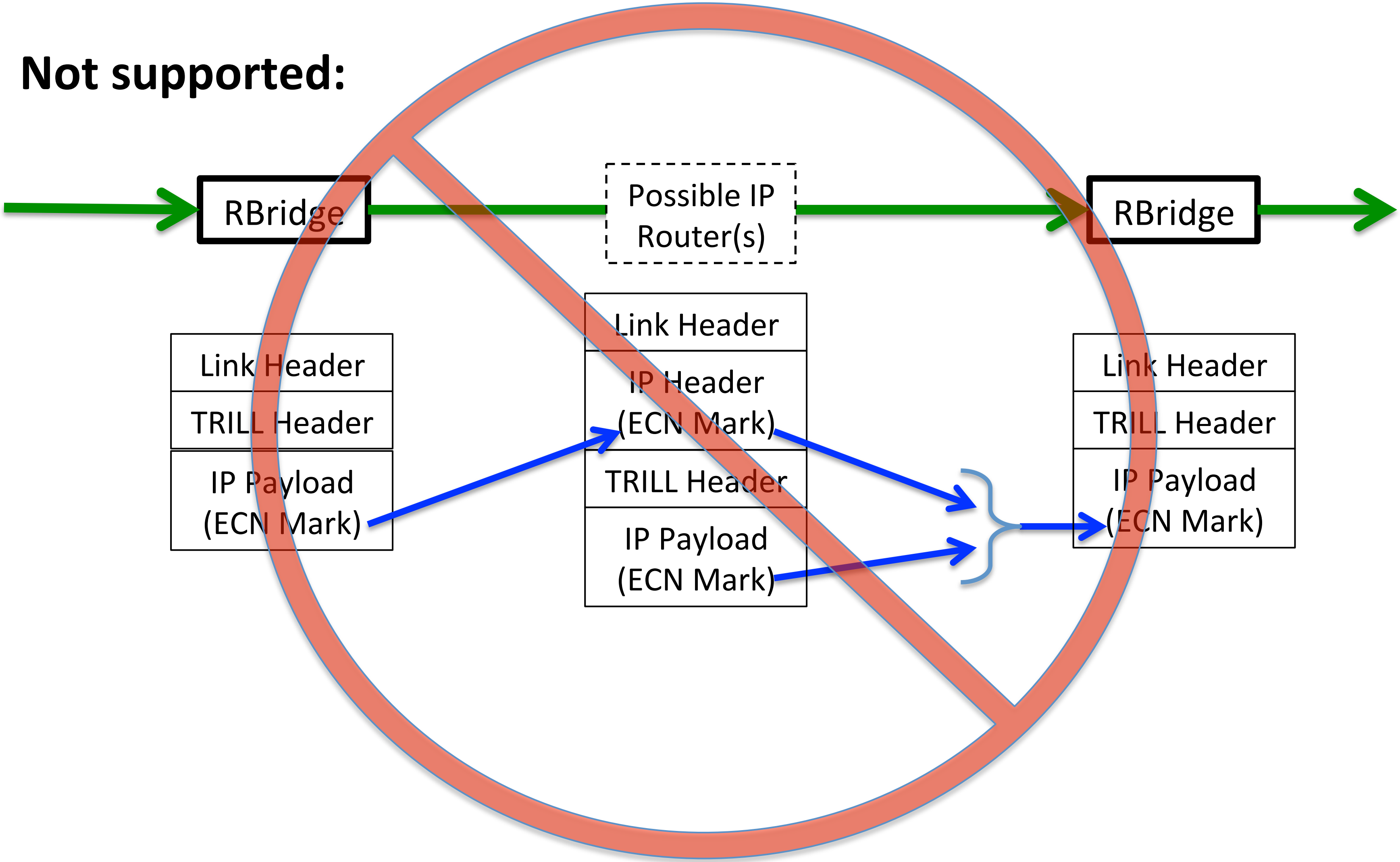
Link Header

TRILL Header (ECN Mark)

Payload

ECN (continued)

Not supported:



Quality of Service

- TRILL primarily uses a three bit priority field for QoS. There is also a one bit drop eligibility indicator. IP uses the six bit DSCP field.
- TRILL-over-IP provides a mapping from TRILL QoS to DSCP but
  - that mapping needs to be reviewed based on the latest RFCs
  - the draft needs to provide more warning that DSCP interpretation is variable particularly in the general Internet
- In the case of transport using TCP, a TCP connection per provided QoS level is needed. Commonly the TRILL QoS levels would be mapped to a smaller number of DSCP values.

TCP Transport

- TRILL-over-IP offers TCP as one format option.
- Call it “transport”, not “encapsulation”, since there is no 1-to-1 mapping between TRILL packets and TCP packets.
- Add framing with length field so that the incoming TCP packets can be parsed to extract/re-assemble TRILL packets.
- End point TCP tweaks for performance to achieve line speed such as disable NAGLE.

TCP Transport (continued)

- When TCP transport is used, a TRILL packet can be spread over multiple TCP packets. Thus MTUs less than the campus wide Sz “minimum” MTU are useful. MTU discovery needs to be extended down to a smaller configurable limit with a reasonable default minimum.

Fragmentation

- Fragmentation of UDP/TCP should be avoided. Fragmented packets are less reliable and fragments are usually blocked by NATs and firewalls.
- Within a Data Center or the like, use of UDP with fragmentation is consistent with the TRILL philosophy of mostly worrying about routing messages more than data messages.
- TCP format with restricted TCP packet size is the best way to avoid fragments to get through restricted MTU or NAT or firewall situations.

NAT

- Need to use static bindings.
- Neighbor RBridge addresses as reported in source IP addresses need to be mapped, using the static binding information, into actual remote IP addresses before being listed in the IS-IS Neighbor TLV.
- Generally need keep-alive messages.
- NATs commonly block fragments (see Fragmentation).
- Interaction with security: IPsec ESP packets probably need to be encapsulated in UDP to get through NATs.

Feedback? Questions?