TRILL Working Group	H. Zhai
Internet-Draft	F. Hu
Intended status: Standards Track	ZTE
Expires: May 17, 2012	R. Perlman
	Intel Labs
	D. Eastlake 3rd
	Huawei
	November 14, 2011

RBridge: Pseudonode Nickname
draft-hu-trill-pseudonode-nickname-01

Abstract

The Appointed Forwarder on a link for VLAN-x is the RBridge that ingresses native frames from the link and egresses native frames to the link in VLAN-x. If the appointed forwarder for an end station is changed, the remote data traffic to the end station could fail. This document is proposed to assign a nickname for pseudonode identifying a multi-access link to solve the issue. When any appointed forwarder encapsulates a packet, it uses the pseudonode nickname as "ingress nickname" rather than its own nickname. If it does, then if the appointed forwarder changes, or the DRB changes, and the pseudonode still uses the same nickname, then the remote RBridge caches won't need to change, and the data traffic to the end station would reach the link uninterruptedly.

Status of this Memo

Copyright Notice

Table of Contents

1. Problem Statement

The IETF TRILL protocol [RFC6325] provides optimal pair-wise data frame forwarding without configuration, safe forwarding even during periods of temporary loops, and support for multipathing of both unicast and multicast traffic. TRILL accomplishes this by using [IS-IS] [RFC1195] link state routing and encapsulating traffic using a header that includes a hop count. The design supports VLANs and optimization of the distribution of multi-destination frames based on VLANs and IP derived multicast groups. Devices that implement TRILL are called RBridges.

The AF (Appointed Forwarder) on a link for VLAN-x is the RBridge that ingresses native frames from the link and egresses native frames to the link in VLAN-x. If the appointed forwarder for an end station goes down and a different RBridge is appointed as appointed forwarder on the link, the end station will not perceive the changes. Therefore, the cache in remote RBridge could not be correct until it receives the data traffic from the end station, and the traffic from the remote RBridge to the end station could fail for a while. It is even worse for the Swap Nickname Field approach in multi-level TRILL network, for the egress RBridge of remote level 1 area cannot update the correspondence of MAC/VLAN-x and the pair of {ingress nickname, swap ingress nickname} until it receives the data traffic from end station [Mltrill].

Pseudonode nickname is proposed in this document to solve the above issue. Pseudonode nickname is assigned by DRB and used to identify a multi-access link. With pseudonode nickname, the data traffic to the end station can reach the destination link uninterruptedly and be forwarded to the end station by other RBridge even if the appointed forwarder for the VLAN on the link is changed.

This document is organized as following: Section 2 is the concept of pseudonode nickname. Section 3 introduces the LSP announcement mechanism for the pseudonode nickname. Section 4 describes the RBridges processing of transit frame traffic when considering pseudonode nickname. Section 6 specifies pseudonode nickname capability TLV and pseudonode nickname TLV format.

Familiarity with [RFC6325] is assumed in this document.

1.1. Terminology and Acronyms

This document uses the acronyms defined in [RFC6325] and the following additional acronym:
AF - Appointed Forwarder
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].
When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in [RFC2119].

2. Pseudonode Nickname

Pseudonode nickname is used to identify a link and SHOULD be reused after DRB changed. It is assigned by DRB on the link. When the RBridge becomes DRB and it doesn't find the pseudonode nickname from TRILL Hello of other RBridges, DRB assigns and announces a pseudonode nickname in its TRILL Hello on the link. If the new DRB obtains the pseudonode nickname from the TRILL Hellos of adjacent RBridges on the link, it reuses this nickname. The nickname for the pseudonode SHOULD keep unchanged even if the DRB or AF changed.
All the RBridges on the link should support pseudonode nickname, otherwise the RBridges that don't understand pseudonode nickname on the link cannot forward the encapsulated TRILL frame with pseudonode nickname. Each RBridge on the link announces its pseudonode nickname capability in its TRILL Hello. Only if DRB checks that all the adjacencies in Report state support and enable the pseudonode nickname capability, DRB assigns pseudonode nickname on the link. If not, DRB MUST NOT announce the pseudonode nickname in its pseudonode LSP in the TRILL campus network, otherwise, the remote data traffic may be forwarded to the RBridge without pseudonode nickname capability, and be discarded in the RBridge.
The bypass pseudonode bit is used to determine whether DRB should generate the pseudonode LSP. When bypass pseudonode bit is reset, the DRB should support pseudonode function and generate the pseudonode LSP [RFC6327]. So if DRB assigns pseudonode nickname on the link, the bypass pseudonode bit MUST be reset in its TRILL Hello.
For an acess port, no TRILL frames, except TRILL-Hello frames, can be transmitted (see Section 4.9.1 in [RFC6325]). However, some TRILL data frames may be transmitted among RBridges on a pseudonode nickname enabled link (see Section 4). Therefore, if an RBridge supports pseudonode nickname function on a link, its port(or ports) to this link MUST NOT be access port (or ports), i.e., the TRILL traffic disable (access port) bit MUST NOT be set on the port (or ports).

3. LSP Announcement

Pseudonode nickname is only announced in the DRB's pseudonode LSP in TRILL campus. If one of the RBridges on the link is disabled of the pseudonode nickname function, that is, DRB receives a TRILL Hello without pseudonode nickname capability enabled from the port on the link, the pseudonode nickname function should be disabled on the link, and then DRB updates its pseudonode LSP which doesn't include pseudonode nickname TLV in the TRILL campus campus.

While if an RBridge (not DRB) supporting pseudonode nickname joins into or exits from the link, it has no influence to the pseudonode nickname LSP originated by DRB. If an RBridge is selected as new DRB and the pseudonode nickname capability on the link is confirmed, it will generate and flood pseudonode LSP including the pseudonode nickname TLV in the TRILL campus. If DRB finds that the pseudonode nickname function is disabled on the link, it will updates its pseudonode LSP which doesn't include pseudonode nickname TLV in the TRILL campus network.

The pseudonode nickname is participated in path computing. The procedure of path computing of pseudonode nickname is same as the routing computing of IPv4 or IPv6 address in layer 3 IS-IS network [RFC1195].

Furthermore, pseudonode nickname can also be used to designate a distribution tree rooted at the pseudonode it identifies. If it is desired for a pseudonode to be a tree root, the DRB MUST advertise in its pseudonode LSP a "tree root" priority for this pseudonode nickname.

4. Influece on Transit Frame Processing

In TRILL protocol, Layer 2 frames are divided into five categories: native frames, TRILL Data frames, Layer 2 control frames, TRILL control frames and TRILL other frames(see Section 1.4 of [RFC6325]). Only the first two categories of frames are forwarded by RBridges, so they are called transit frames. Pseudonode nickname are only used for transit frames in this specification.

From a forwarding standpoint, transit frames may be classified into two categories: unicast and multi-destination. Section 4.1 covers the changes in processing unicast frames on RBridges that participates in a pseudonode nickname group. Section 4.2 describes the special processing of multi-destination frames on RBridges with pseudonode nickname capability enabled.

4.1. Unicast

The processing of unicast frames on both ingress and egress RBridges will be influenced by pseudonode nickname. However, the processing on transit RBridges remains unchanged.

Section 4.1.1 covers the changes of processing native frames on a pseudonode nickname participated ingress RBridge. Section 4.1.2 describes two methods to process TRILL data frames on egress RBridge.

4.1.1. Ingress Processing

When a VLAN-x tagged native frame is sent onto a multi-access link, only the appointed forwarder for that VLAN-x can ingress this frame into TRILL campus. If the pseudonode nickname capability is enabled on the link, the default is the forwarder uses the link's pseudonode nickname rather than the RBridge's nickname as ingress nickname when it converts a native frame received from this link, to TRILL data frames. This forwarder MAY use other nickname(such as the RBridge's nickname) than the pseudonode nickname as ingress nickname when it does TRILL-encapsulation to native frames received from this link if it has been configured to do so. The encapsulation of the native frame is as same as Section 4.1 of [RFC6325] except for the ingress nickname in TRILL header.

4.1.2. Egress Processing

On receiving a unicast TRILL data frame, the egress nickname in the TRILL header is examined, and if it is unknown or reserved, the frame is discarded. Then the Inner.VLAN ID, i.e., VLAN-x, is checked. If it is 0x0 or 0xFFF, the frame is discarded.
This RBridge will be the egress RBridge for the TRILL data frame, if the egress nickname is one of the RBridge's nicknames or one of the pseudonode nicknames of the connected links. If the egress RBridge is the VLAN-x forwarder on the destination link for this TRILL data frame, the frame is processed and the original self-learning is performed by this RBridge as described in [RFC6325]. Otherwise, the frame will be re-encapsulated and transmitted on the link by the egress RBridge. Only the VLAN-x forwarder can decapsulate the TRILL data frame to native form and forward it to the end station on the link, which is consistent with the principle of ingressing and egressing native frame into and out of TRILL campus, i.e., there is only a single RBridge on each link that is in charge of ingressing and egressing native frames from and to that link [RFC6327].
There are two methods for the egress to transmit the re-encapsulated TRILL data frame to VLAN-x forwarder on the link. In Section 4.1.2.1, the egress unicasts the re-encapsulated TRILL data frame to the VLAN-x forwarder, and in Section 4.1.2.2, the egress multicasts the TRILL data frame on the link.

4.1.2.1. Unicasting to VLAN-x Forwarder

To make the final hop, i.e., the egress RBridge (not VLAN-x forwarder), work for a frame addressed to the pseudonode, the forwarding table has to be based on {nickname, VLAN}, instead of {nickname} currently. In the couple of {nickname, VLAN}, nickname is the pseudonode nickname, and VLAN is the VLAN Id of VLAN-x forwarder on this link. If there are several appointed forwarders (each for a VLAN) on this link, several entries (each for a forwarder) exist in the forwarding table. In the couple of {nickname, VLAN}, the VLAN will be ignored if the nickname is not a pseudonode nickname on one of local links, and will be set to invalid value(such as 0x0 or 0xFFF). In other words, if the VLAN in an entry is invalid, the nickname is not a pseudonode nickname.

If the RBridge is not VLAN-x forwarder on the link, it goes to its forwarding table that says, based on the pseudonode nickname and VLAN-x Id, which of its RBridge neighbors, i.e., VLAN-x forwarder on this link, to forward to. The forwarder is identified by the next hop MAC address in the found entry from the above table, which is one of the unicast MAC addresses on one of its ports connected directly on this link. The TRILL data frame is discarded if no entry is found. Otherwise, the outer frame header of the TRILL data frame is stripped, the TRILL header remains unchanged, and a new outer frame header is prepended before the frame is forwarded to the VLAN-x forwarder on the link. For the forwarded frame, the Outer.MacSA is the MAC address of the transmitting port on the destination link, the Ouer.MacDA is the next hop MAC address in the found entry and the Outer.VLAN is the designated VLAN on the destination link.

If the above re-encapsulated TRILL data frame is received by a stale VLAN-x forwarder on the destination link, it will be dropped by the RBridge. Otherwise, the re-encasulated frame is processed as [RFC6325], and the Inner.MacSA and Inner.VLAN ID are, by default,learned as associated with the ingress nickname unless that nickname is unknown.

4.1.2.2. Multicasting to VLAN-x Forwarder

Alternatively, a special multicast MAC address, named "AF RBridges on this link", can be introduced for the final hop to forward such a TRILL data frame. The scope of the above MAC is limited to local link, just as the MAC for IS-IS hello PDUs. If a TRILL data frame is addressed to this special MAC and transmitted on a link, all the Appointed Forwarder (AF) RBridges on the link will process it to some extent.

With "AF RBridges on this link" MAC address, the forwarding table can remain unchanged in form, i.e., still based {nickname}. For an entry, the next hop MAC address will be "AF RBridges on this link", if the nickname is the pseudonode nickname on one of local links. In other words, if the nickname is a pseudonode nickname, the next hop MAC MUST be "AF RBridges on this link".

If not VLAN-x forwarder, the final hop RBridge, RBn, looks up its forwarding table, based on the egress nickname in TRILL header of the received frame. The frame will be discarded if no entry is found. Otherwise, RBn re-encapsulates the frame just like what a transit Rridge does, except that the TRILL header remains unchanged and the Ouer.MacDA is "AF RBridges on this link". If the egress nickname is pseudonode nickname, the re-encapsulated TRILL data frame is multicasted onto the link.

The TRILL data frame with "AF RBridges on this link" as Ouer.MacDA is discarded by other RBridges, which are not AF RBridges, on the link. Otherwise, the Inner.VLAN ID, i.e., VLAN-x, is checked. If the VLAN ID is not valid or the receiving RBridge, RBi, is not VLAN-x forwarder on this link, the frame is also discarded. Else, the TRILL data frame is decapsulated into native form and forwarded to the destination end station, and the Inner.MacSA and Inner.VLAN ID are also, by default, learned as associated with the ingress nickname unless that nickname is unknown by RBi.

4.1.2.3. Comparison

With the Unicasting method described in Section 4.1.2.1, the re-encapsulated TRILL data frame by the final hop RBridge is only processed by the VLAN-x forwarder on the link, which can reduce the burden of other RBridges as much as possible. But the forwarding table on ingress/egress SHOULD be changed to be based on {nickname, VLAN}, instead of {nickname}, where each AF Rbridge on a local link is identified by the pseudonode nickname and the vlan id of the AF on the link.

With Multicasting method described in Section 4.1.2.2, although all the AF RBridges, except for the final hop RBridge, on the link are required to process, to some extent, the re-encapsulated TRILL data frame, only the VLAN-x forwarder decapsulates the frame to native form and forwards it to the destination end station. However, the forwarding table can remain the same as current table in form, i.e., only based on {nickname}.

4.2. Multi-Destination

If pseudondoe nickname function is enabled on a link, a forwarder SHOULD use the link's pseudonode nickname as ingress nickname, except that it has been configured not to do so, when it does TRILL-encapsulation for a native frame received from this link. In TRILL campus, multi-destination TRILL data frames are propagated along the distribution trees chosen by ingress RBridges. To limit the amount of state necessary to perform the RPF (Reverse Path Forwarding) check, for a forwarder on the link where the pseudonode nickname function is enabled, it MUST select a tree that the DRB has announced (in its psedonode LSP) to be one of those that (the RBridges on) this link might use, when it uses this nickname as ingress nickname in doing TRILL-encapsulation to a native frame.

RBridges use SPF (Shortest Path First) algorithm, instead of spanning tree, to calculate distribution tree based on link state information. So the pseudonode, standing for a link, exists in distribution trees if the DRB advertises pseudonode LSP for this link. If a forwarder is not attached directly to the pseudonode in the chosen tree, the use of the link's pseudonode nickname as ingress nickname by this forwarder may mess up the RPF check along this tree.

For example, a simple topology is given in Figure 1, where RB1 through RB4 are RBridges, H1 and H2 are end stations, S1 and S2 are serial links, and E1 is a multi-access link. The numbers at the ends of each link are the metrics of RBridges' ports to the link.

```
                    +------+
                    |  RB1 |
             H1     +------+
              O      /1   \2
              |     /S1    \S2
              |   5/        \4
+------+   2 +------+      +------+
|  RB5 |-----|  RB2 |      |  RB3 |
+------+ 1   +------+      +------+
                |3            |1
                |             |
       ------------------------------- E1
              |                 |
              |2                |
           +------+             O
           |  RB4 |            H2
           +------+
```

Based on the topology, a distribution tree rooted at RB1 can be calculated by each RBridge, which is given in Figure 2, where the node represented by a triangle is the pseudonode for E1. H1 and H2 are not RBridges, so they are not on this tree.

```
        +------+
        |  RB1 |
        +------+
         /1   \2
        /      \
       /        \
 +------+      +------+
 |  RB2 |      |  RB3 |
 +------+      +------+
     |2           |1
     |            |
 +------+         +
 |  RB5 |        / \
 +------+       / E1\
               +-----+
                  |0
                  |
               +------+
               |  RB4 |
               +------+
```

Let's assume that the pseudonode nickname function is enabled on link E1 and the pseudonode nickname is PseNickE1. Then, on this tree, RB1 expects to receive a multi-desitnation TRILL data frame with PseNickE1 as ingress nickname only from the right port. When RB2 ingresses a native frames from link E1 and propogates it along this tree, it uses PseNickE1 as ingress nickname. However, when this TRILL data frame arrives at RB1 from the left port, it will be dropped by RB1 for the failure of RPF check.

Two solutions are given to fix the above problem in this document. Section 4.2.1 covers special distribution trees identified by pseudonode nicknames. Section 4.2.2 gives an optional method, which modifies the multi-destination frame processing behavior of RBridges, to some extent, to solve this problem.

4.2.1. Distribution Tree Rooted at Pseudonode

In TRILL protocol, a distribution tree is designated by a root nickname which identifies an RBridge or a pseudonode. So it is possible to use a pseudonode nickname as root nickname to calculate a tree. For examble, based on the topology in Figure 1, a distribution tree rooted at E1 is given in Figure 3.

```
          +  0    +-------+
         / \------|  RB4 |
        / E1\     +-------+
       +-----+
        /0  \0
       /     \
      /       \
 +-------+   +-------+
 |  RB2 |   |  RB3 |
 +-------+   +-------+
     |2        |4
     |         |
 +-------+   +-------+
 |  RB5 |   |  RB1 |
 +-------+   +-------+
```

On this tree, R2 through R4 directly attach to pseudonode E1, so they can safely use PseNickE1 as ingress nickname when they ingress native frames from E1 and propagate the TRILL-encapsulated frames along this tree.

To calculate such a tree in the campus, if it has confirmed the pseudonode nickname function can be enabled on its link, the DRB MUST announce in its pseudonode LSP the pseudonode nickname and a "tree root" priority for this nickname as well as the willing to use this tree when the pseudonode (i.e., all the RBridges on this link) ingressing a multi-destination packet.

If such a tree are not actually calculated by all the RBridges in TRILL campus for some reasons, e.g., lower "tree root" priority, the RBridges on this link, which don't support the optional method given in Section 4.2.2, SHOULD reset pseudonode nickname capability in their TRILL Hellos. Then the DRB on this link disables the pseudonode nickname function on this link.
This method minimizes the change in ingress RBridges, caused by pseudonode nickname, but cannot guarantee that the special tree are actually calculated, which limits the application of pseudonode nickname function in multi-destination frames in TRILL campus. So an optional method is given in Section 4.2.2.

4.2.2. Changes of Processing Behavior

From the viewpoint of a distribution tree, for a pseudonode, all the adjacencies on the link (represented by the pseudonode) can be divided into two categories:

*Adjacencies that directly attach the link's pseudonode, for examble, RB3 and RB4 in Figure 2;

*The rest of adjacencies, which do not directly attach the pseudonode, for examble, RB2 in Figure 2.

For an adjacency in the first category, it is safe to use the pseudonode nickname to do TRILL-encapsulation to native frames from the link, and to propagate the TRILL data frame along the chosen tree. But in the second category, an RBridge MUST change its processing of such a native frame, to some extent, if it desires to use the pseudonode nickname to do TRILL-encapsulation and propagate the TRILL frame safely along the chosen tree. The changes of frame processing on an RBridge in the second category are as following:

*When doing TRILL-encapsulation to a native frame, the RBridge MUST NOT send this frame in native form or TRILL form out of any ports except the port from which this native frame is received, if it uses the link's psuedonode nickname as ingress nickname to encapsulate the native frame into TRILL form. Only the TRILL-encapsulated frame is sent back to the link from the only port;

*When recieving a multi-destination TRILL data frame, if the ingress nickname of the frame is the pseudonode nickname of one of its links, the RBridge processes the frame as specified in Section 4.6.2.5 of [RFC6325], but MUST NOT forward this frame in native form to the link represented by the ingress nickname (even if it is the appointed forwarder for that link for the frame's VLAN).

With the help of the first change, the TRILL-encapsulated frame will be received by the RBridges in the first categories. But the frame still fails Tree Adjacency Check on these RBridges, which causes the drop of this frame (see Section 4.5.2 of [RFC6325]). So some bogus adjacencies on this tree must be added into these RBridges' forwarding tables, that is:

*For the RBridges in the first category, add all the RBridges in the second category as bogus adjacencies into their forwarding tables. These bogus adjacencies are only used for the frame to pass the Tree Adjacency Check on the receiving RBridges, and MUST NOT be used for TRILL-encapsulated frame proporgating.

With the help of the above three changes of frame processing, the multi-detination TRILL-encapsulated frames can be ingressed into the chosen tree from the psudonode and propagated along the tree safely. For example, in Figure 2, when RB2 encapsulates a native frame (received from H2) to multi-destination TRILL frame where the ingress nickname is PseNickE1, it does not sent the native frame to H1 and the TRILL frame to RB5 but does send this TRILL frame back to pseudonode E1. Then this TRILL frame will be ingressed into this tree by RB3 and RB4. RB1 will recieve this frame from the correct port, i.e., the port on the right side, and propagate it out of the left port. At last, when receiving this frame, RB2 forwards it to RB5 and may decapsulate it into native form and forward to H1, but does not forward the native frame onto link E1.

This method is optional. It does not need special distribution trees to be calculated, but changes the frame processing on RBridges on this link, which imposes some changes of frame processing on silicon.

5. Link Partition

When RBridges on a link cannot receive the DRB's hellos during holding time, a new DRB will be elected. Some issues can cause RBridges to receive no hellos from the DRB, for example, the DRB down or link partitioned and DRB on the other part of the original link. In order to improve the stability of remote RBridges' forwarding table, the new DRB should reuse the link's pseudonode nickname if it finds such a nickname has been used on this link (i.g., from its neighbors' hellos).

In the issue of link partition, both of the DRBs on the two parts try to reuse the original link's psudonode nickname, which causes nickname collision. A method to resolve such collision is given in Section 5.1.

5.1. Solution to Nickname Collision

To avoid a new DRB to usurp a pseudonode nickname from another DRB that is still using this nickname, extra rules are given for the priority of the pseudonode nickname reused by a new DRB, that is:

*The priority of the nickname reused by a new DRB SHOULD be lower than the priority of this nickname found on this link (i.g., from its neighbors' hellos), before the DRB ensuring no other DRBs using the same nickname in their pseudonode LSPs.

- After ensuring no other DRBs using this pseudonode nickname, the DRB increases the priority of this nickname to its original found value.

When an RBridge is elected as new DRB and advertises its first pseudonode LSP where a pseudonode nickname is contained, it should start a non-cyclic waiting timer to detect such collision. If the timer expires and no such collision is found, the DRB can ensure that no other DRBs using the same pseudonode nickname.
Whenever receiving a pseudonode LSP originated by other DRB, a DRB looks up the LSP in its LSP database to see whether it also has an instance of this LSP. If it does not, or if the database copy is less recent, it installs the LSP into its database. For such a pseudonode LSP, its pseudonode nickname will be compared with the link's nickname used by the DRB. If the two nicknames are same, pseudonode nickname collision is detected. Then the prirority of this nickname, along with system ID of the RBridge (numerically higher = higher priority) as tiebreaker, in the received LSP is compared with the priority (of this nickname) used locally.
For the conflicting pseudonode nickname, the DRB performs the following extra steps to clear up this confliction:

- if the nickname priority in the received LSP is higher, the DRB SHOULD give up this pseudonode nickname and acquire a new one;

- else, the DRB continues to use this nickname, but it SHOULD update its pseudonode LSP (with a larger sequence number) to the TRILL campus.

If a DRB oughts to give up its pseudonode nickname and acquire a new one, it SHOULD inform other RBridges in TRILL campus to clear up some the MAC entries whose egress "RBrige" nicknames equal to this pseudonode nickname, or to modify the remaining time of these entries. Since DRB knows all the VLANs for which end-station service is enabled on its link, the mechanism of Appointed Forwarder Status Lost Counter (AFSLC) (see Section 4.8.3 of [RFC6325]) can be employed for this purpose. The DRB SHOULD update its pseudonode LSP, where the AFSLC for all these VLANs is increased at least one, before it uses the new acquired pseudonode nickname on this link and its pseudonode LSPs. Furthermore, if a DRB finds one or more such VLANs lost on its link, it SHOULD update its pseudonode LSP, in which the AFSLC for these VLANs is increased at least one, to inform other RBridges to handle the associated MAC entries in their forwarding tables.

6. TLV Extensions for Pseudonode Nickname

6.1. Pseudonode Nickname Capability in Hellos

The pseudonode nickname capability of an RBridge MUST be included in one subTLV of Port Capability TLV in the RBridge's TRILL Hello PDUs. This capability is included in Special VLANs and Flags (subTLV Type #1) [RFC6326]. This subTLV MUST appear exactly once in a Port Information TLV in every TRILL Hello PDU. The length of the value is four octets.
Pseudonode Nickname capability TLV:

```
+-+-+-+-+-+-+-+-+
|Type=VLAN Flags|                  (1 byte)
+-+-+-+-+-+-+-+-+
|   Length      |                  (1 byte)
+---------------+---------------+
|    Port ID                    |  (2 bytes)
+-------------------------------+
|     Sender Nickname           |  (2 bytes)
+--+--+--+--+-------------------+
|AF|AC|VM|BY|    Outer.VLAN     |  (2 bytes)
+--+--+--+--+-------------------+
|TR|PN|R |R |    Desig.VLAN     |  (2 bytes)
+--+--+--+--+-------------------+
```

The PN bit, if one, indicates that the sending RBridge supports and enables the pseudonode nickname capability. If an RBridge does not support or not enable this capability, the PN bit MUST be set zero. Other bits and fields refer to [RFC6326].
When receiving this subTLV from other RBridges on the link, the DRB can confirm whether all the adjacencies, in Report state [RFC6327], support and enable this capability. If not, DRB MUST NOT announce pseudonode nickname in its pseudonode LSPs to the TRILL campus, which can avoid the issue that remote traffic is forwarded to a RBridges without pseudonode nickname capability.

6.2. Pseudonode Nickname TLV

If the DRB has confirmed that pseudonode nickname capability can be enabled on this link, it will announce the pseudonode nickname to be used on this link in its hello PDUs and in its pseudonode nickname. The pseudonode nickname is carried in pseudonode Nickname TLV, which is formatted as following:
Pseudonode Nickname TLV:

```
+-+-+-+-+-+-+-+-+
|Type= PSEU-NICK|                         (1 byte)
+-+-+-+-+-+-+-+-+
|     Length    |                         (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   pseudonode NICKNAME RECORDS (1)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        ...................                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   pseudonode NICKNAME RECORDS (n)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where each pseudonode nickname record is of the form:

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Nickname.Pri |SType|R|R|R|R|R|       (2 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Tree Root Priority       |       (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Nickname          |       (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Type: Pseudonode Nickname Type, TBD (NICKNAME).

*Length: 6*N, where N is the number of pseudonode nickname records present.

*SType: An 3-bit unsigned integer sub-type for nickname. If this nickname is pseudonode nickname, value of this field is 1.

*Nickname.Pri: An 8-bit unsigned integer priority to hold a nickname as specified in Section 3.7.3 of [RFC6325].

*Tree Root Priority: This is an unsigned 16-bit integer priority to be a tree root as specified in Section 4.5 of [RFC6325].

*Nickname: This is an unsigned 16-bit integer as specified in Section 3.7 of [RFC6325].

6.2.1. Pseudonode Nickname TLV in Hellos

For an RBridge enabled pseudonode nickname capability on this link, it announces one pseudonode nickname TLV in Hellos if it knows nickname for the pseudonode, otherwise, it MUST NOT announce pseudonode nickname in its Hellos. From the adjacencies' hellos or the nickname stored locally, the new DRB can knows the pseudonode nickname already used on its link.

For an RBridge that is not DRB, it only processes the pseudonode nickname announced by DRB, and MUST overwrite its own pseudonode nickname with the DRB's pseudonode nickname if the two nicknames are different. DRB should process the pseudonode nickname TLV from all the adjacencies in the Report state on the link in order to obtain the pseudonode nickname that was being used on this link.
This TLV MUST appear no more than once in a Port Information TLV in every Hello PDU. Only one nickname record can be contained in this TLV, if this subTLV appears in Hello PDUs.

6.2.2. Pseudonode Nickname TLV in DRB's LSPs

For a DRB on a link, it MUST originate and flood a pseudonode LSP for this link if the bypass pseudonode bit is reset. All the adjacencies in the Report state on this link are contained in its pseudonode LSP. Furthermore, if a pseudonode nickname capability is enabled on this link, a pseudonode Nickname TLV MUST be contained in its pseudonode LSP.
For a pseudonode LSP, the only one record in this TLV contains the nickname for the pseudonode standing for the link. In this case, the value of Nickname.Pri varies from 1 to 255, which describes the DRB's priority to hold this nickname as specified in [RFC6325] Section 3.7.3.

7. IANA Considerations

TBD.

8. Security Considerations

TBD.

9. Acknowledgements

We would like to thank Mingjiang Chen for his contributions to this document. Additionally, we would like to thank Erik Nordmark, Les Ginsberg, Ayan Banerjee, Dinesh Dutt, Anoop Ghanwani, and Tissa Senevirathne for their good questions and comments.

10. References

10.1. Normative References

[RFC1195]	Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
[RFC2119]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
[RFC6325]	Perlman, R., Eastlake, D., Dutt, D., Gai, S. and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.
[RFC6326]	

	Eastlake, D., Banerjee, A., Dutt, D., Perlman, R. and A. Ghanwani, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 6326, July 2011.
[RFC6327]	Eastlake, D., Perlman, R., Ghanwani, A., Dutt, D. and V. Manral, "Routing Bridges (RBridges): Adjacency", RFC 6327, July 2011.

10.2. Informative References

[Mltrill]	Perlman, R, Eastlake 3rd, D and A Ghanwani, "RBridges: Multilevel TRILL", draft-perlman-trill-rbridge-multilevel-02.txt Work in Progress, April 2011.

Authors' Addresses

Hongjun Zhai Zhai ZTE 68 Zijinghua Road, Yuhuatai District Nanjing, Jiangsu 210012 China Phone: +86 25 52877345 EMail: zhai.hongjun@zte.com.cn

Fangwei Hu Hu ZTE 889 Bibo Road, Pudong District Shanghai, Shanghai 201203 China Phone: +86 21 68896273 EMail: hu.fangwei@zte.com.cn

Radia Perlman Perlman Intel Labs 2200 Mission College Blvd Santa Clara, CA 95054-1549 USA Phone: +1-408-765-8080 EMail: Radia@alum.mit.edu

Donald Eastlake 3rd Eastlake 3rd Huawei 155 Beaver Street Milford, MA 01757 USA Phone: +1-508-333-2270 EMail: d3e3e3@gmail.com