Network Working Group
Internet Draft
Intended status: Standards Track
Expiration Date: December 2007

Ping Pan
(Hammerhead Systems)
Shane Amante
Nasser El-Aawar
(Level-3)

June 2007

Setup and Manage PBB-based Tunnels with PWE3 Mechanism

draft-pan-pwe3-pbb-tunnel-00.txt

Status of this Memo

Abstract

This document describes a mechanism that enables providers to setup point-to-point tunnels over PBB networks for the purpose of aggregating customer flows. The interior network nodes will not be disturbed with this mechanism.

Table of Contents

1. Specification of Requirements

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.

2. Introduction

With the proliferation of Ethernet deployment in metro and backbone networks, managing Ethernet connections becomes increasingly important. Despite the popularity of MPLS, many carriers continue to build-out Ethernet-only networks, where no MPLS switching takes place at data plane inside the network.

To make the deployment scalable, the carriers have been deploying Q-in-Q [Q-in-Q] whereby each Ethernet frame can be encapsulated with multiple VLAN tags.

As the metro network continues to expand, the carriers need to improve scalability beyond Q-in-Q VLAN tagging. PBB (Provider Backbone Bridges) [PBB], also known as MAC-in-MAC, is a technique that encapsulates Ethernet frames with another Ethernet MAC and VLAN header. It therefore increases Ethernet framing hierarchy at data-plane.

Generally, the scaling improvement at data-plane needs to be coupled with the scalability at control-plane.

One simple yet popular application is to aggregate VLAN-tagged flows from customer locations into point-to-point tunnels from PBB backbone edge. It requires the network operators to be able to dynamically provision, monitor and manage the PBB-based tunnels, as well as mapping of customer VLAN flows into PBB-based tunnels.

PBT (Provide Bridge Transport) [PBT] has been proposed as the solution for the above application. However, PBT inherently has two shortcomings:

First, the currently defined PBT proposal has no dynamic provisioning mechanism, and won t scale for large scale backbone with static configuration.

Second, the currently defined PBT proposal requires the internal nodes to be programmed for tunnel setup. It is debatable if it will bring the justification for control-plane simplicity.

In this document, we first define PBB-based tunnels, as the virtual tunnels between PBB edge nodes. We then propose of using the well-defined PWE3 protocols to setup and manage the PBB-based tunnels.

The proposal does not require the changes on the internal nodes.

2.1. Assumptions

Our proposal is based on the following assumptions:

1. The Ethernet metro/core network is relatively stable. The network will not experience frequent link and node failures, or subsequent long converging and looping problems introduced by STP.

2. Network providers will over-provision the network to overcome per-flow QoS issues.

2.2. Perspectives

There should be little doubt on the effectiveness of MPLS/IP

to-multi-point alike. In particular, PWE3 is a well-defined and widely deployed mechanism in setting up and managing edge-to-edge data connections. In particular, the PWE3 control-plane can operate over just about any type of networks.

There should also be obvious that the carriers have incentive in deploying PBB (Provider Backbone Bridges, or, MAC-in-MAC) to expand their Ethernet-only networks. Inside the PBB networks, the carriers have the option of using spanning tree protocols to interconnect nodes. It should be noted that the MAC-in-MAC mechanism itself is a data-plane function.

Hence, instead of operating the networks with either MPLS or PBB/PBT, we believe that an alternative is to deploy MPLS/PWE3 protocols at places where they fit (such as, network edge), while using PBB for simple and cheap Ethernet packet transport within the backbone.

In our proposal, we define the protocol extension in using target LDP (PWE3) at network edge to exchange MAC-in-MAC and VLAN information between network edges. The backbone network itself will operate with PBB only.

3. Background

In this section, we will outline some of the relevant technologies that we work with.

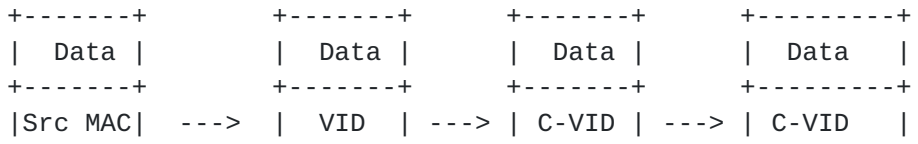
3.1. PWE3

PWE3 [RFC3985] is to create edge-to-edge connections between two edge nodes, and has been widely deployed to interconnect user traffic over the MPLS/IP backbone. It comes with a number of important features: VCCV [VCCV] provides an efficient OAM capability for PW s, and MHOP [MHOP] and PW switching [PW Switching] enable the PWE3 to be deployed over multiple networks.

Further, Pseudo-wire itself is a very flexible concept, and can operate over any data network, including MPLS, TDM and Ethernet (a.k.a. dry-martini).

3.2. PBB

PBB extends the Ethernet stacking as shown in Figure-1:



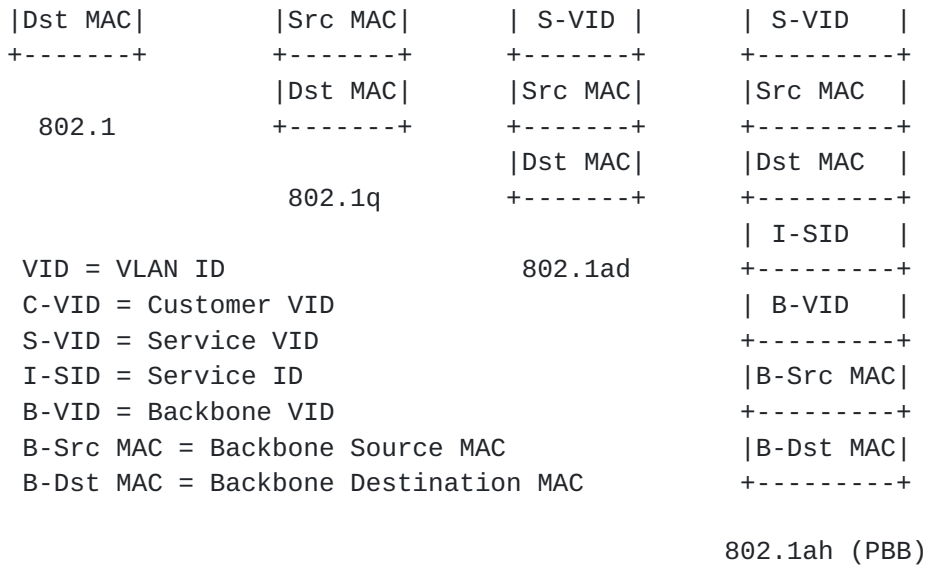


Figure 1: A brief history of Ethernet Stacking

As mentioned previously, Q-in-Q (or 802.1ad) has been widely deployed. PBB leads to a reasonable evolution for Ethernet backbone expansion.

Since PBB is to encapsulate a new MAC header to each packet, the network intermediate nodes can be constructed with any out-of-shelf Ethernet switches and employee STP or other means for connectivity.

In summary, if the carriers have already adapted native Ethernet services, the deployment of PBB should not introduce much overhead and cost.

3.3. PBT

By definition, PBB offers a bridged network, where each node can communicate with other nodes simultaneously (a.k.a. any-to-any connectivity). This may not be a desirable behavior for some of the backbone applications.

One important backbone application is to create point-to-point tunnels between edge nodes to aggregate user flows.

PBT (Provider Bridge Transport) is designed to support point-to-point tunnels over PBB networks. Instead of using STP to interconnect edge nodes, PBT defines one or multiple B-VID bits in the PBB stack for the purpose of tunneling functionality. The intermediate nodes will not perform STP flooding on the frames that have the special B-VID bits. To setup the PBT tunnels, PBT requires all the intermediate nodes to manage the B-VID entries. It has been proposed of using

GMPLS to setup and maintain the PBT tunnels.

The advantage here is that it enables the operators to run traffic engineering on every PBT tunnel. Also the PBT path setup can be independent from STP, therefore, is immune from the undesirable long convergence problems.

The disadvantages in PBT are that each PBB nodes need to be provisioned, and the lack of dynamic provisioning mechanism.

3.4. PBB-based Tunnels

If the PBB backbone is over-provisioned and relatively stable, and the carrying data is best-effect, then it does not seem necessary to provision all the nodes, as recommended in PBT.

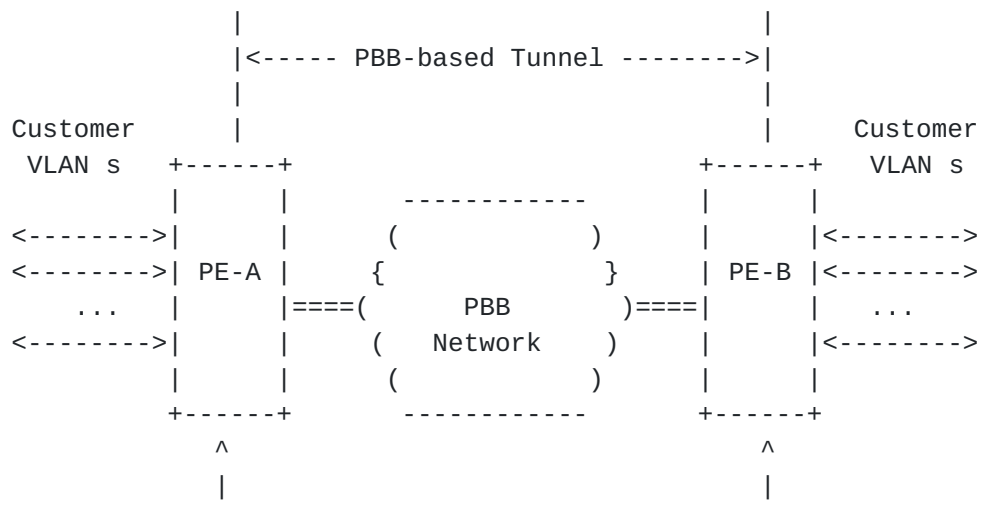
In the context of our proposal, we define PBB-based tunnels as the following:

If two edge nodes have connectivity over the PBB backbone, then the operators can establish bi-directional virtual tunnels between two nodes for the purpose of aggregating user traffic flows. Such tunnels are called PBB-based tunnels .

Note that the interior nodes do not need to do anything special. The connectivity between edge nodes can be achieved by dynamic protocols, such as STP, or static configuration, or other means.

4. Operation Outline

The objective here is to establish PBB-based tunnels over the PBB core, and aggregate customer flows through the point-to-point tunnels.



IP-A	IP-B
MAC-A	MAC-B
I-SID-A	I-SID-B
B-VID-A	B-VID-B

Figure 2: Operation example

The operation is illustrated in Figure 2. The service provider needs to transport multiple customer VLAN flows between PE-A and PE-B. The PBB backbone may apply STP or other means to discover the connectivity between PE-A and PE-B.

The operators will configure IP address to both PE nodes. ARP will resolve the MAC addresses, that is, MAC-A and MAC-B. The operators will need the I-SID and B-VID information to aggregate customer flows with MAC-in-MAC encapsulation.

We propose of using RFC4447 [PWE3-LDP] to exchange I-SDI and B-VID information. Essentially, we will use the mechanism for setting up PW s to create PBB-based tunnels between PE-A and PE-B.

Specifically, the tunnel setup is based on the Generalized ID (GID) FEC (type 129). This enables the operators setting up and managing the tunnels from local PE s. The PBT-based tunnels will use a new AII to identify PBB-based tunnels, which contains MAC, B-VID and I-SID information.

In the example, the operators at PE-A will send a LDP mapping message with SAII = {MAC-A, I-SID-A, B-VID-A}, and TAII = {MAC-B, I-SID-B, B-VID-B} to PE-B. Upon successful reception, PE-A and PE-B will install the PBB information on data path. When a VLAN packet from a customer site arrives on PE-A, PE-A will encapsulate a backbone Ethernet header with the appropriated MAC, I-SID and B-VID. Upon the reception, PE-B will remove the header, and continue to the native Ethernet switching.

Each PBB-based tunnel can carry traffic from multiple customer sources.

When a PBB-tunnel is no longer in use, the operator will apply the standard PWE3 procedure to remove it.

For OAM support, the operators have the option to apply either native Ethernet OAM schemes (802.3ah, 802.1ag) or VCCV (LSP-ping, BFD), or both.

For protection support, the operators can either rely on STP to converge, or statically establish backup connections between PE-1 and

PE-2. The method in establishing backup Ethernet connections is beyond the scope of this draft.

In terms of QoS, we adapt the most common practice in Ethernet networks, over-provisioning.

Finally, other than the PE nodes need to be IP-enabled to support ARP and target LDP, the rest of the network can remain Ethernet-only.

5. Protocol Extensions

The document specifies the protocol requirements and definitions for setting up PBB-based tunnels from network edge.

5.1. Generalized ID FEC

RFC4447 defines the mechanism for setting up point-to-point Pseudowires between two PE nodes. During Pseudowire setup, the PE nodes will exchange messages containing information about the PW type and an endpoint identifier used in the selection of the packet forwarder.

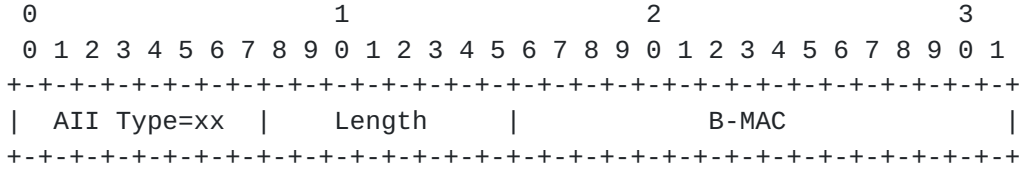
Endpoint identifier can be represented in two formats: PWid FEC (type 128) and Generalized ID (GID) FEC (type 129). The PWid FEC requires to be configured on the local and remote PE prior to PW setup.

The GID FEC defines each PW as a combination of Source and Destination Attachment Individual Identifiers (AII). Each AII is globally unique. This arrangement enables GID to be used for individual configuration, as well as provisioning models where the local PE s can learn (or discover) about the remote PE s prior to launching PW s. For sizeable PBB-based tunnel deployment, the GID FEC presents a more flexible and scalable solution.

We require GID to be used in setting up PBB-based tunnels.

5.2. AII Type: PBB-based Tunnels

We define a new AII Type for PPB-based Tunnels. The encoding is shown in Figure 3.



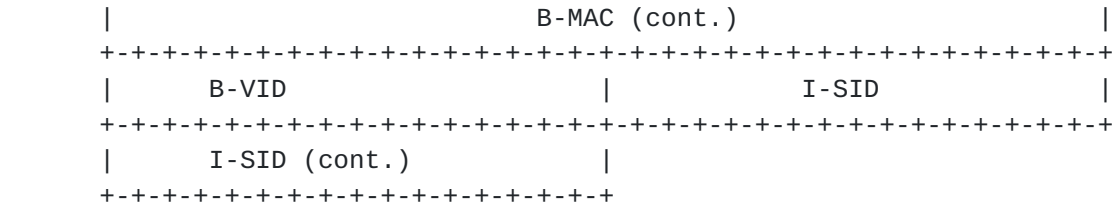


Figure 3: AII Type: PBB-based Tunnel

B-MAC: Backbone MAC address

B-VID: Backbone VLAN ID

I-SID: Service Instance ID

All above parameters have been defined in [PBB].

5.3. Pseudowire Status

We will use the defined Pseudowire Status [PWE3-IANA] for PBB-based tunnels.

5.4. OAM

We believe that VCCV needs to specific the OAM types. Other than LSP-ping and BFD, VCCV may need to expand to advertise 802.1ag and 802.3.ah as capabilities [VCCV].

6. Applications

6.1. Interconnection PBB Domains with PW MHOP

(TBD)

6.2. Carrying multi-service traffic over PBB-based tunnels

(TBD)

6.3. Interworking with PBT

(TBD)

7. Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

8. Full Copyright Statement

9. IANA Considerations

This document has defined an AII type for PBB-based tunnels.

10. Normative References

[PBB] IEEE 802.1ah, " Virtual Bridged Local Area Networks Amendment 6: Provider Backbone Bridges"

[PBT] IEEE 802.1Qay, Provider Backbone Bridge Traffic Engineering

[PWE3-LDP] L. Martini, et al., Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC4447, April 2006

[PWE3-IANA] L. Martini, IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3) , RFC4446, April 2006

11. Informative References

[Q-in-Q] IEEE 802.1ad, "Virtual Bridged Local Area Networks: Provider Bridges"

[RFC3985] Bryant, et al., "PWE3 Architecture", RFC3985, March 2005.

[VCCV] T. Nadeau, et al., Pseudo Wire Virtual Circuit Connectivity Verification (VCCV) , March 2007.

[PW Switching] L. Martini, et al., Pseudo Wire Switching , February 2007.

[MHOP] L. Martini, et al., Dynamic Placement of Multi Segment Pseudo Wires , October 2006.

12. Author Information

Ping Pan
Hammerhead Systems
ppan@hammerheadsystems.com

Shane Amante
Level 3 Communications
Shane.Amante@Level3.com

Nasser El-Aawar
Level 3 Communications
Nasser.El-Aawar@Level3.com