Internet Draft Document   Vach Kompella
Category: Standards Track   Florin Balus
Expires: August 2008   Alcatel-Lucent

Andrew Malis
Verizon

July 2008

Generic Ethernet Pseudowire
draft-vkompella-pwe3-ethernet-pw-bis-00.txt

Status of this Memo

Copyright Notice

Abstract

This draft proposes a simpler approach to handling various encapsulations of Ethernet packets over a pseudowire, over the existing Ethernet Pseudowire definition in [RFC4448].

Conventions used in this document

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].

1. Introduction

In this draft, we describe a methodology for encapsulating Ethernet packets in pseudowires that eliminates the need for multiple encapsulation formats and pseudowire code-points when Ethernet formats change. From a service provider perspective it simplifies the introduction of new technology by minimizing the required network upgrades. This draft extends the concepts already published in [RFC4448].

[RFC4448] defines two different pseudowire types for Ethernet packets, one where the packet transported over the pseudowire must have a service delimiting VLAN tag (Ethernet tagged type), and the other where it does not need to have one (Ethernet type) as per [IANA PWE3]. The original justification for the tagged PW type came from the need to accommodate routers that could not handle standard Ethernet functions like imposing, stripping or rewriting VLAN tags.

This draft has been written in response to the following statements from [RFC4448] and consequent concerns that as new Ethernet formats (such as Provider Backbone Bridging, PBB I-tag [802.1ah]) are defined, that corresponding pseudowires have to be defined:

- For an Ethernet VLAN PW, VLAN tag rewrite can be achieved by NSP at the egress PE, which is outside the scope of this document ([RFC4448], Section 3).

- The Ethernet or Ethernet VLAN PW only supports homogeneous Ethernet frame type across the PW; both ends of the PW must be either tagged or untagged. Heterogeneous frame type support achieved with NSP functionality is outside the scope of this document ([RFC4448], Section 3).

This document proposes a Generic Ethernet PW (GE-PW) that extends the definition of Ethernet PWs and their usage in the existing L2VPNs (VPWS, VPLS) and simplifies the development of future solutions that require transport of Ethernet frames over a

pseudowire.

2. The Generalized Ethernet Pseudowire

To define the Generalized Ethernet Pseudowire (GE-PW) we need to describe the following functions:

- the packets eligible to be placed in a GE-PW
- the processing functions outside the scope of the GE-PW
- the processing functions within the scope of the GE-PW

2.1. Eligible Packets in the GE-PW

Any ethernet packet defined by the IEEE is an eligible packet for the GE-PW, regardless of the type of tags it contains - e.g. IEEE 802.1Q, 802.1ad, 802.1ah tags. The Ethernet header is defined with a set of well defined code points that can be used by the NSP to identify the type of tag and the following header fields and to process the frame accordingly.

A new Interface Parameter sub-TLV is defined to describe the capabilities of the NSP function to adapt different ethernet encapsulations as packets arrive from the pseudowire and are delivered to the attachment circuit. Section 4 describes the applicability and the format for the new sub-TLV.

2.2. Processing outside the scope of the GE-PW

When an ethernet packet is received on an attachment circuit (AC), it may be processed before being passed to the PW termination function. This processing is done by the NSP function which has the responsibility of removing service delimiting encapsulations on the packet, and identifying the PW that the packet is bound for.

When a packet arrives on a PW, the PW service label is used to identify the particular service instance and subsequently the NSP function that is applied to the ethernet packet at the egress PE. The NSP function puts on the appropriate ethernet encapsulation before passing the packet on to the attachment circuit(s) associated with the NSP.

The processing of the Ethernet frames at the ingress and egress NSPs are outside the scope of the GE-PW.

2.3. Processing within the scope of the GE-PW

A packet passed to the PW termination function by the ingress NSP at the ingress PE is encapsulated with the appropriate PW label and is passed to the PSN function for encapsulation and transport to the

egress PE. A packet arriving on a pseudowire egress has its PW label popped, and the resulting packet is passed to the appropriate NSP instance.

3. Detailed processing steps

There are four main processing points in using a pseudowire: the ingress NSP function, the ingress PW termination function, the egress PW termination function, and the egress NSP function.

Although just the PW termination functions are relevant for GE-PW, this section discusses how the GE-PW can be used with different types of existing NSPs to emulate existing and future services.

In order to more clearly illustrate how a GE-PW may be used to provide pseudowire services, we introduce the following terms. These terms are to be used as a guideline to understanding the behavior of a GE-PW, and in no way define an implementation:

- In-NSP(packet, options): the ingress NSP function, includes the ingress attachment circuit (AC) function
- In-PW(packet, options): the ingress PW termination function
- Out-PW(packet, options): the egress PW termination function
- Out-NSP(packet, options): the egress NSP function, includes the egress AC function

In addition, we define the pseudowire context (PWC) as the construct which has, at ingress, the following information:

- the incoming encapsulation of the packet
- the In-NSP function to be applied on the packet
- the In-PW termination function to be applied

- the ingress PSN encapsulation information to be applied

and at egress:

- the Out-PW termination function to be applied
- the Out-NSP function to be applied on the packet
- the outgoing encapsulation of the packet on the attachment circuit

Using the above functions, we will define actions at the various stages of the packet through the network that affect its behavior. In particular, we will show the capabilities in describing the two existing Ethernet pseudowires, and introduction of a number of

standard VLL service types that are enabled as a consequence of implementing the GE-PW methods.

As the packet arrives from the customer, at the In-NSP function, the NSP may remove the FCS and any extraneous packet header information that is locally significant [RFC4448]. The optional method described in [RFC4720] can be used to achieve payload integrity transparency. In particular, the function of the In-NSP is to render the packet ready for consumption by the remote Out-NSP function. In other words, the In-NSP must replace, remove or impose VLAN tags, PBB I-tag or any required Ethernet headers so that the packet is recognizable by the egress NSP. It performs also functions specific to the type of native service that is rendered at the ingress PE, for example MAC switching, MAC learning, packet replication for VPLS. Finally, the In-NSP function also identifies the pseudowire that is associated with the attachment circuit over which the packet arrived. If, for example, the packet arrived on a VLAN tagged port, and the VLAN tag identifies the attachment circuit, then the In-NSP is responsible for recovering the VLAN tag and identifying the attachment circuit. It is out of the scope of this document to define how attachment circuits are represented and associated with pseudowires.

The resulted Ethernet frame, as delivered by the In-NSP is not modified in any way by the In-PW function. The In-PW termination simply imposes the PW encapsulation for the packet. This may be introducing the optional control word and its fields, depending on how the pseudowire was configured and which options were negotiated. Following this step, the PW label for the packet is imposed. Subsequently, the appropriate forwarding engine component imposes the PSN encapsulation. How the PSN encapsulation is determined and

imposed for a particular pseudowire is out of the scope of this document.

Once the packet is delivered to the network, the PSN encapsulation directs the packet towards the egress endpoint of the pseudowire. At the egress, the PSN encapsulation is removed. Following this, the packet is delivered to the Out-PW termination function. The Out-PW function removes the PW encapsulation. This involves removing both the PW label and the optional control word (if negotiated and present). In addition, the PW label is used to identify an Out-NSP function associated with one or more attachment circuit(s) that is/are the outgoing interface(s) to the customer. The packet is handed to the Out-NSP function which will complete processing and hand the packet to the egress attachment circuit.

The Out-NSP function has the responsibility for processing the packet in order to prepare it for the egress attachment circuit.

This would involve, e.g., the insertion, deletion or replacement of different Ethernet tags or other Ethernet encapsulation so that it can accommodate different types of ACs, for example port, tagged with one VLAN, tagged with multiple VLANs (QinQ [802.1ad]) or PBB I-tags.

There are a number of functions that are out of the scope of this specification. The primary reason they are out of scope is that they are implementation dependent, and do not relate to the standard definition of the Generalized Ethernet Pseudowire (GE-PW). For example, whether the incoming attachment circuit has a pointer to the In-NSP function or an index into a table, or how the pseudowire is identified within the forwarding engine are not material to the definition of the GE-PW. The definition of what the NSP function has to perform is also a matter of local configuration.

4. Control Plane

All the PW Setup and Maintenance procedures described in [RFC4447] apply to the GE-PW. There is no need for a new PW type. The Ethernet type (0x0005) may be used [IANA PWE3] as the new GE-PW procedures are backwards compatible with an existing implementation using Ethernet type PW - see section 6. The function of an Ethernet tagged PW can be also emulated in the NSPs with a GE-PW.

There might be some network scenarios where the required NSP capabilities need to be signaled between PEs. This might be the case for certain implementations that need to know what kind of NSP they need to instantiate for certain PW. Also in some scenarios the Service Providers might need to make sure the capabilities of the related NSPs match.

An optional NSP capabilities sub-TLV for the Interface Parameters TLV is defined to allow the signaling of NSP capabilities between Layer 2 PEs.

The NSP Capabilities sub-TLV MAY be included in the related LDP messages for PW setup and maintenance if the transmitting PE needs to verify that the remote NSP is capable of performing certain functionality. Note that there are cases when the NSP functionalities do not need to match. For example in a PBB deployment using HVPLS, a PBB BEB may be connected via an Ethernet PW to a PBB BCB [PBB-VPLS]. The NSP for the PBB BEB performs full PBB functions (I and B-components) while the NSP for the related PBB BCB performs just regular Ethernet switching using the Backbone Header (B-component) [PBB-VPLS].

On reception of the NSP Capabilities sub-TLV a Layer 2 PE MUST reject the PW Setup if it does not support or if the related NSP is not properly configured to support any of the required NSP capabilities. If the Layer 2 PE does not understand the NSP Capabilities sub-TLV, it should continue the processing of the interface parameters while silently discarding the unknown interface parameter as per [RFC4447] section 5.5.

The format of the NSP sub-TLV follows the standard format defined for all Interface Parameter sub-TLVs [RFC4447]:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Sub-TLV Type  |    Length     |    Variable Length Value      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Variable Length Value                 |
|                             "                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The type for NSP Capabilities sub-TLV is to be assigned by IANA. The first 2 Bytes of the value field define the capability whereas the rest of the value field is used to indicate parameters specific to

the required capability ID. A list of code points for the capabilities is to be added as the applications are being defined. Examples of possible capabilities ID are VPLS, PBB-VPLS BEB, PBB-VPLS BCB with or without I-tag swapping [PBB-VPLS].

5. The Control Word

The OPTIONAL control word specified in [RFC4385] MAY be used for the GE-PW.

6. Backwards Compatibility with existing Ethernet PW implementations

The NSP of a GE-PW capable PE that needs to interoperate with older implementations may be manually configured to fully emulate the behavior of an existing Ethernet PW NSP as described in section 4.1 and 4.4 of [RFC4448].

The NSP Capabilities sub-TLV may be also used to automatically identify the old Ethernet PW implementation. The GE-PW capable PE may use the NSP Capabilies indication to identify the regular Ethernet PW implementations.

The GE-PW capable PE MUST always include the NSP capability sub-TLV in the PW setup message. If the PW signaling received from the remote PE does not contain the NSP Capability sub-TLV, the local PE switches to regular Ethernet PW mode. If the other Layer 2 PE does not understand the NSP Capabilities sub-TLV, it should continue the processing of the interface parameters while silently discarding the unknown interface parameter as per [RFC4447] section 5.5.

7. Emulating existing Ethernet Services

The rationale for redefining the two variants of the Ethernet pseudowires is that the GE-PW is a more powerful construct that subsumes both of them, prevents the future proliferation of other Ethernet PW types and allows the definition of many more services than are allowed by strictly following the existing pseudowire definitions.

The following examples demonstrate that the right set of NSP and PW functions yield not just the known ethernet VPWS, VPLS behaviors induced by the two existing Ethernet pseudowires, but additional models used by service providers, that are not strictly compliant to the existing pseudowires.

7.1. GE-PW applicability to Ethernet VPWS

The GE-PW allows for emulation of an Ethernet point-to-point service between different types of Ethernet Attachment Circuits by simply emulating in the ingress NSP the required behavior. Specifically the attachment circuits (ACs) may be defined at the two PEs to represent the whole port, one or more VLAN(s) (tagged, one service delimiting VLAN) or a VLAN combination (QinQ/[802.1ad], tagged with two service delimiting VLANs) and are mapped to the point-to-point NSP function in each PE through local association.

As a result the packet arriving on the local AC is classified as belonging to an ingress NSP who may be configured to remove/rewrite/add one or more VLAN tags before forwarding the packet to the GE-PW termination function that will encapsulate it for transport over PSN core.
Similarly the egress NSP can manipulate encapsulation received over GE-PW, adding, replacing or removing one or more tags of different types as desired to accommodate different types of egress Ethernet interfaces and AC definitions, for example port, tagged with one VLAN, tagged with multiple VLANs (QinQ [802.1ad]).

7.2. GE-PW Applicability to VPLS

The GE-PW can be used to serve the interconnect needs for the VPLS Forwarders. The previous section discussed how the GE-PW can replace the existing Ethernet PW types from a point to point forwarding perspective, handling also the VLAN tags on a per local AC basis. For VPLS the NSPs will have to replicate the same functions but for multiple ACs, handling also the Ethernet switching functions (e.g. MAC switching, MAC Learning, Packet Replication) as described in [RFC4664] and [RFC4762].

7.3. GE-PW for point-to-point transport of PBB over Pseudowires

When a PW interconnects two PBB access networks (see Figure 1) there may be a need to use the PBB ISID Tag (I-tag) [802.1ah] to perform the mapping to/from ACs. The PBB I-tag can be processed as an extended tag that may need to be changed at ingress / egress PE. The related NSPs can be configured to identify the I-tag (using the Ethernet type assigned by [802.1ah]) and to change different I-tag fields, for example the 24 bits ISID or the QoS information. The GE-PW termination function will not need to change to accommodate this new type of service. The required PW encapsulation described in the previous sections will be added before the PBB Backbone MAC header.

For example assuming the PE depicted in Figure 1 [RFC4448] is connected to a PBB network (PBBN) and the I-tag field is used to defined the AC.

```
+-------------------------------------------+
|                PE                          |
+---+   +-+  +------+  +------+  +------+  +-+
|   |   |P|  |      |  |PW ter|  | PSN  |  |P|
|   |<==|h|<=| NSP  |<=|minati|<=|Tunnel|<=|h|<== From PSN
|   |   |y|  |      |  |on    |  |      |  |y|
| P |   +-+  +------+  +------+  +------+  +-+
| B |   |                                    |
| B |   +-+  +------+  +------+  +------+  +-+
| N |   |P|  |      |  |PW ter|  | PSN  |  |P|
|   |==>|h|=>| NSP  |=>|minati|=>|Tunnel|=>|h|==> To PSN
|   |   |y|  |      |  |on    |  |      |  |y|
+---+   +-+  +------+  +------+  +------+  +-+
|                                            |
+-------------------------------------------+
```

Figure 1: PBB PW

The ingress PE may be configured to check on the incoming PBB packets the ISID value of the I-tag to determine the AC and subsequently the ingress NSP. The ingress NSP may be configured to change the values of different I-tag fields (for example the ISID, PCP, DEI bits).

Next the GE-PW termination functions will handle the PW header that will transport it throughout the PSN core and towards the egress NSP.

The egress NSP may perform different functions depending on the type

of egress AC: e.g.

- if the access network is part of a different PBB service domain it may re-write the I-tag field.
- if the access domain is a QinQ domain it may remove the PBB header altogether assuming it supports this capability.

7.4. PBB-VPLS Applicability

The GE-PWs can be also applied to provide the interconnect between PBB-VPLS entities as described in [PBB-VPLS]. On top of the I-tag identification, processing functions described in the previous section, the related NSPs can emulate the PBB components described in [802.1ah], for example the Backbone (B) and Customer (I) components. For example, in a PBB-VPLS PE, as the In-NSP function receive the packet on the local AC, it performs the Customer MAC switching and possibly the mapping to the Backbone MAC address specific to the I-component [802.1ah]. Subsequently the same In-NSP function will perform the Backbone MAC switching function characteristic to the B-component [802.1ah]. If the packet is to be transported over the PW infrastructure it will be handled to the in-PW function as per [PBB-VPLS]. From now on the processing steps described in section 3 apply. Same for the Out-PW termination function in the other PE. The PW service label is used by the Out-PW function to identify the Out-NSP to which the packet is handled. There are two possibilities for the Out-NSP.

If the PE is the final termination point for PBB, the Out-NSP runs both I and B components (PBB BEB node in [802.1ah]). It will handle first the Ethernet switching functions for Backbone MAC header followed by the switching functions for Customer MAC header. As the packet is handled to the egress AC the regular AC processing functions described in section 3 apply.

If an HVPLS architecture is used the next PE in the chain may be just an intermediate switching point for Backbone MAC header. The related Out-NSP function runs just the PBB B-component (PBB BCB node in [802.1ah]). It will handle just the Ethernet switching functions

for Backbone MAC header. The possible results may be switching the packet back to another In-PW function for further forwarding towards the PBB-VPLS PE (PBB BEB).

[PBB-VPLS] and [PBB-Interop] describe the requirements, interoperability and solution in detail.

Other applications are possible and require just the definition of the required NSP functions.

8. Management Model

There will be a new object to describe the NSP compatibility capabilities [PW MIB].

9. Acknowledgements

The authors gratefully acknowledge the contributions of Nabil Bitar, and Dimitri Papadimitriou.

10. References

Normative References

[RFC4447] Martini, L., Ed., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.

[RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.

[IANA PWE3] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", BCP 116, RFC 4446, April 2006.

[RFC4385] Bryant, S., Swallow, G., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, February 2006.

[RFC4720] Malis, A., Allan, D.,and N. Del Regno, "Pseudowire Emulation Edge-to-Edge (PWE3) Frame Check Sequence Retention", RFC 4720, November 2006

[802.1ad] IEEE 802.1ad-2005 "Virtual Bridged Local Area Networks, Amendment 4: Provider Bridges", December 7, 2005

[802.1ah] IEEE Draft P802.1ah/D4.2 "Virtual Bridged Local Area Networks, Amendment 6: Provider Backbone Bridges", Work in Progress,

[RFC4762] Lasserre, M. and Kompella, V. (Editors), "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.

[PW MIB] T. Nadeau, and D. Zelig "Pseudowire (PW) Management Information Base (MIB)", draft-ietf-pwe3-pw-mib-14.txt, January 2008 (work in progress).

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

Informative References

[RFC4664] Andersson, L. and Rosen, E. (Editors),"Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, September 2006.

[PBB-VPLS] F. Balus, et Al. "VPLS Extensions for Provider Backbone Bridging", draft-balus-l2vpn-vpls-802.1ah-02.txt, February 2008 (work in progress).

[PBB-Interop] A. Sajassi, et Al. "VPLS Interoperability with Provider Backbone Bridges", draft-sajassi-l2vpn-vpls-pbb-interop-02.txt, November 2007 (work in progress).

11. Security Considerations

No new security issues arise out of the extensions proposed here.

12. IANA Considerations

A new IANA allocation is required to describe NSP compatibility capabilities.

13. Authors' Addresses

Andrew Malis
Verizon
andrew.g.malis@verizon.com

Vach Kompella
Alcatel-Lucent

vach.kompella@alcatel-lucent.com

Florin Balus
Alcatel-Lucent
florin.balus@alcatel-lucent.com

14. Full Copyright Statement

15. Intellectual Property

ipr@ietf.org.

16. Acknowledgments

Funding for the RFC Editor function is provided by the IETF
Administrative Support Activity (IASA).