Internet Engineering Task Force (IETF)
Request for Comments: 5994
Category: Informational
ISSN: 2070-1721

S. Bryant, Ed.
M. Morrow
G. Swallow
Cisco Systems
R. Cherukuri
Juniper Networks
T. Nadeau
Huawei Technologies
N. Harrison
BT
B. Niven-Jenkins
Velocix
October 2010

Application of Ethernet Pseudowires to MPLS Transport Networks

Abstract

Ethernet pseudowires are widely deployed to support packet transport of Ethernet services. These services in-turn provide transport for a variety of client networks, e.g., IP and MPLS. This document uses procedures defined in the existing IETF specifications of Ethernet pseudowires carried over MPLS networks.

Many of the requirements for the services provided by the mechanisms explained in this document are also recognized by the MPLS transport profile (MPLS-TP) design effort formed jointly by the IETF and ITU-T. The solution described here does not address all of the MPLS-TP requirements, but it provides a viable form of packet transport service using tools that are already available.

This document also serves as an indication that existing MPLS techniques form an appropriate basis for the design of a fully-featured packet transport solution addressing all of the requirements of MPLS-TP.

Status of This Memo

Table of Contents

1. Introduction

Ethernet pseudowires are widely deployed to support packet transport of Ethernet services. These services in-turn provide transport for a variety of client networks, e.g., IP and MPLS. This document uses procedures defined in the existing IETF specifications of Ethernet pseudowires carried over MPLS networks.

Many of the requirements for the services provided by the mechanisms explained in this document are also recognized by the MPLS transport profile (MPLS-TP) design effort formed jointly by the IETF and ITU-T [RFC5654]. For example, the ability to operate solely with network management control, the ability to use Operations, Administration, and Maintenance (OAM) that does not rely on IP forwarding, and the ability to provide light-weight proactive connection verification (CV) functionality.

The solution described in this document does not address all of the MPLS-TP requirements, but it provides a viable form of packet transport service using tools that are already available.

The key purpose of this document is to demonstrate that there is an existing IETF mechanism with known implementations that satisfies the requirements posed by the operator community. It is recognized that it is possible to design a more efficient method of satisfying the requirements, and the IETF anticipates that improved solutions will be proposed in the future as part of the MPLS-TP effort. Indeed, the solution described in this document is not intended to detract from the MPLS-TP effort. Instead, it provides legitimacy for that work by showing that there is a real demand from networks that are already deployed, and by indicating that the MPLS-TP solutions work is based on sound foundations.

Much of the notation used in this document is defined in [RFC3985] to which the reader is referred for definitions.

The architecture required for this mechanism is illustrated in Figure 1.

```
     +----------------------------------------------------------------+
     |                                                                |
     |                  IP/MPLS PSN (PHP may be enabled)              |
     |                            (client)                            |
     |                                                                |
     |                  +---------------------------+                 |
     |                  |                           |                 |
     |                  |      MPLS PSN (No PHP)    |                 |
     |                  |         (server)          |                 |
     |                  |                           |                 |
     |     CE1          |PE1                     PE2|           CE2   |
     |   +-----+      +-----+                   +-----+      +-----+  |
     |   | | | |      | | | |                   | | | |      | | | |  |
     |   | | | +------+ | | |                   | | | +------+ | | |  |
     |   | | | | 802.3| | | |                   | | | | 802.3| | | |  |
     |   +-----+      +-----+                   +-----+      +-----+  |
     |     |   |        |  |                      | |        |   |    |
     |     |   |        +-- ---------------------- -+        |   |    |
     +----- --- -------- -- ---------------------- - -------- --- ----+
           |   |        |  |<--MPLS LSP (no PHP)->| |        |   |
           |   |        |  |       (server)       | |        |   |
           |   |        |                           |        |   |
           |   |        |<------------PW----------->|        |   |
           |   |        |          (server)         |        |   |
           |   |                                             |   |
           |   |<-------------802.3 (Ethernet)-------------->|   |
           |   |                   (client)                  |   |
           |                                                     |
           |<---------IP/MPLS LSP (PHP may be supported)-------->|
           |                       (client)                      |
```

Figure 1: Application Ethernet over MPLS PW to MPLS Transport
Networks

An 802.3 (Ethernet) circuit is established between CE1 and CE2. This circuit may be used for the concurrent transport of MPLS packets as well as IPv4 and IPv6 packets. The MPLS packets may carry IPv4, IPV6, or pseudowire payloads, and Penultimate Hop Popping (PHP) may be used. For clarity, these paths are labeled as the client in Figure 1.

An Ethernet pseudowire (PW) is provisioned between PE1 and PE2 and is used to carry the Ethernet from PE1 to PE2. The Ethernet PW is carried over an MPLS Packet Switched Network (PSN), but this PSN MUST NOT be configured with PHP. For clarity, this Ethernet PW and the MPLS PSN are labeled as the server in Figure 1. In the remainder of this document, call the server network a transport network.

1.1. Requirements Language

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].

2. PWE3 Configuration

The PWE3 encapsulation used by this specification to satisfy the transport requirement is Ethernet [RFC4448]. This is used in "raw" mode.

The Control Word MUST be used. The sequence number MUST be zero.

The use of the Pseudowire Setup and Maintenance Label Distribution Protocol [RFC4447] is not required by the profile of the PWE3 Ethernet pseudowire functionality defined in this document.

The pseudowire label is statically provisioned.

3. Operations, Administration, and Maintenance (OAM)

Within a connection, traffic units sent from the single source are constrained to stay within the connection under defect-free conditions. During misconnected defects, a connection can no longer be assumed to be constrained, and traffic units (and by implication also OAM packets) can 'leak' unidirectionally outside a connection. Therefore, during a misconnected state, it is not possible to rely on OAM, which relies on a request/response mechanism, and, for this reason, such OAM should be treated with caution if used for diagnostic purposes.

Further, when implementing an Equal Cost Multipath (ECMP) function with MPLS, use of the label stack as the path selector such that the OAM and data are not in a co-path SHOULD be avoided, as any failure in the data path will not be reflected in the OAM path. Therefore, an OAM that is carried within the data-path below the PW label (such as Virtual Circuit Connectivity Verification (VCCV)) is NOT vulnerable to the above failure mode. For these reasons, the OAM mechanism is as described in [RFC5085], which uses Bidirectional Forwarding Detection (BFD) [RFC5880] for connection verification (CV). The method of using BFD as a CV method in VCCV is described in [RFC5885]. One of the VCCV profiles described in Section 3.1 or Section 3.2 MUST be used. Once a VCCV control channel is provisioned and the operational status of the PW is UP, no other profile should be used until such time as the PW's operational status is set to DOWN.

3.1. VCCV Profile 1: BFD without IP/UDP Headers

When PE1 and PE2 are not IP capable or have not been configured with IP addresses, the following VCCV mechanism SHOULD be used.

The connection verification method used by VCCV is BFD with diagnostics as defined in [RFC5885].

[RFC5085] specifies that the first nibble is set to 0x1 to indicate a channel associated with a pseudowire [RFC4385].

The Version and the Reserved fields are set to zero, and the Channel Type is set to 0x7 to indicate that the payload carried is BFD without IP/UDP headers, as is defined in [RFC5885].

3.2. VCCV Profile 2: BFD with IP/UDP Headers

When PE1 and PE2 are IP capable and have been configured with IP addresses, the following VCCV mechanism may be used.

The connection verification method used by VCCV is BFD with diagnostics as defined in [RFC5885].

[RFC5085] specifies that the first nibble is set to 0x1 to indicate a channel associated with a pseudowire [RFC4385].

The Version and the Reserved fields are set to 0, and the Channel Type is set to 0x21 for IPv4 and 0x56 for IPv6 payloads [RFC4446].

4. MPLS Layer

The architecture of MPLS-enabled networks is described in [RFC3031]. This section describes a subset of the functionality of the MPLS-enabled PSN. There are two cases that need to be considered:

1. The case where external configuration is used.

2. The case where a control plane is available.

Where the use of a control plane is desired, this may be based on Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945].

4.1. External Configuration

The use of external provisioning is not precluded from being supported by the current MPLS specifications. It is however explicitly described in this specification to address the requirements specified by the ITU [RFC5654] to address the needs in a transport environment.

The MPLS encapsulation is specified in [RFC3032]. All MPLS labels used in the server layer (Figure 1) MUST be statically provisioned. Labels may be selected from either the per-platform or the per-interface label space.

All transport Label Switched Paths (LSPs) utilized by the PWs described in Section 2 MUST support both unidirectional and bidirectional point-to-point connections.

The transport LSPs SHOULD support unidirectional point-to-multipoint connections.

The forward and backward directions of a bidirectional connection SHOULD follow a symmetrically routed (reciprocal) LSP in the server network.

Equal Cost Multipath (ECMP) load balancing MUST NOT be configured on the transport LSPs utilized by the PWs described in Section 2.

The merging of Label Switched Paths is prohibited and MUST NOT be configured for the transport LSPs utilized by the PWs described in Section 2.

Penultimate hop popping by the transport Label Switched Routers (LSRs) MUST be disabled on transport LSPs.

Both EXP-Inferred-PSC LSPs (E-LSP) and Label-Only-Inferred-PSC LSPs (L-LSP) MUST be supported as defined in [RFC3270].

For the MPLS EXP field [RFC3270] [RFC5462], only the pipe and short-pipe models are supported.

4.2. Control Plane Configuration

In this section, we describe the control plane configuration when [RFC3209] or the bidirectional support in GMPLS ([RFC3471] and [RFC3473]) are used to configure the transport MPLS PSN. When these protocols are used to provide the control plane, the following are automatically provided:

1. There is no label merging unless it is deliberately enabled to support Fast Re-route (FRR) [RFC3209].

2. A single path is provided end-to-end (there is no ECMP).

3. Label Switched Paths may be unidirectional or bidirectional as required.

Additionally, the following configuration restrictions required to support external configuration MUST be applied:

- o Penultimate hop popping [RFC3031] by the LSRs MUST be disabled on LSPs providing PWE3 transport network functionality.

- o Both E-LSP and L-LSP MUST be supported as defined in [RFC3270].

- o The MPLS EXP [RFC5462] field is supported according to [RFC3270] only when the pipe and short-pipe models are utilized.

5. Congestion Considerations

This document describes a method of using the existing PWE3 Ethernet pseudowire [RFC4448] to solve a particular network application. The congestion considerations associated with that pseudowire and all subsequent work on congestion considerations regarding Ethernet pseudowires are applicable to this RFC.

6. Security Considerations

This RFC provides a description of the use of existing IETF Proposed Standards to solve a network problem, and raises no new security issues.

The PWE3 security considerations are described in [RFC3985] and the Ethernet pseudowire security considerations of [RFC4448].

The Ethernet pseudowire is transported on an MPLS PSN; therefore, the security of the pseudowire itself will only be as good as the security of the MPLS PSN. The server MPLS PSN can be secured by various methods, as described in [RFC3031].

The use of static configuration exposes an MPLS PSN to a different set of security risks to those found in a PSN using dynamic routing. If a path is misconfigured in a statically configured network, the result can be a persistent black hole, or much worse, a persistent forwarding loop. On the other hand, most of the distributed components are less complex. This is however offset by the need to provide fail-over and redundancy in the management and configuration system and the communications paths between those central systems and the LSRs.

Security achieved by access control of media access control (MAC) addresses, and the security of the client layers, is out of the scope of this document.

7. Acknowledgements

The authors wish to thank Matthew Bocci, John Drake, Adrian Farrel, Andy Malis, and Yaakov Stein for their review and proposed enhancements to the text.

8. References

8.1. Normative References

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

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.

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

[RFC4446] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", BCP 116, RFC 4446, April 2006.

[RFC4447] Martini, L., 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., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.

[RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, February 2009.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC5885] Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, June 2010.

8.2. Informative References

[RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

[RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.

Authors' Addresses

Stewart Bryant (editor)
Cisco Systems
250, Longwater, Green Park
Reading RG2 6GB
UK

EMail: stbryant@cisco.com

Monique Morrow
Cisco Systems
Glatt-com
CH-8301 Glattzentrum
Switzerland

EMail: mmorrow@cisco.com

George Swallow
Cisco Systems
1414 Massachusetts Ave.
Boxborough, MA 01719

EMail: swallow@cisco.com

Rao Cherukuri
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089

EMail: cherukuri@juniper.net

Thomas D. Nadeau
Huawei Technologies
Central Expressway
Santa Clara, CA 95050

EMail: thomas.nadeau@huawei.com

Neil Harrison
BT

EMail: neil.2.harrison@bt.com

Ben Niven-Jenkins
Velocix
326 Science Park
Milton Road, Cambridge CB4 0WG
UK

EMail: ben@niven-jenkins.co.uk