Network Working Group	S. Bryant, Ed.
Internet-Draft	M. Morrow
Intended status: Informational	G. Swallow
Expires: December 31, 2009	Cisco Systems
	R. Cherukuri
	Juniper Networks,
	T. Nadeau
	N. Harrison
	B. Niven-Jenkins
	BT
	June 29, 2009

TOC

Application of Ethernet Pseudowires to MPLS Transport Networks
draft-ietf-pwe3-mpls-transport-04

Status of this Memo

Abstract

A requirement has been identified by the operator community for the transparent carriage of the MPLS(-TP) network of one party over the MPLS(-TP) network of another party. This document describes a method of satisfying this need using the existing PWE3 Ethernet pseudowire standard RFC4448.

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 RFC2119 (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.) [RFC2119].

Table of Contents

1. Introduction
2. PWE3 Configuration
3. OAM
    3.1. VCCV profile 1: BFD without IP/UDP Headers
    3.2. VCCV profile 2: BFD with IP/UDP Headers
4. MPLS Layer
    4.1. External Configuration
    4.2. Control Plane Configuration
5. Congestion Considerations
6. Security Considerations
7. IANA Considerations
8. Acknowledgements
9. References
    9.1. Normative References
    9.2. Informative References
§ Authors' Addresses

---

1. Introduction

TOC

The operator community has identified the need for the transparent carriage of the MPLS(-TP) network of one party over the MPLS(-TP) network of another party [I-D.ietf-mpls-tp-requirements] (Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, “MPLS-TP Requirements,” August 2009.). This document describes one mechanism to satisfy this requirement using existing IETF standards such as PWE3 Ethernet pseudowire standard [RFC4448] (Martini, L., Rosen, E., El-Aawar, N., and G. Heron, “Encapsulation Methods for Transport of Ethernet over MPLS Networks,” April 2006.) . The mechanism described here fulfills the MPLS-TP requirements for transparent carriage (MPLS-TP requirements 30 & 31) of the Ethernet data plane.

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 recognised 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.

Much of the notation used in this document is defined in [RFC3985] (Bryant, S. and P. Pate, “Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture,” March 2005.) to which the reader is referred for definitions.

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

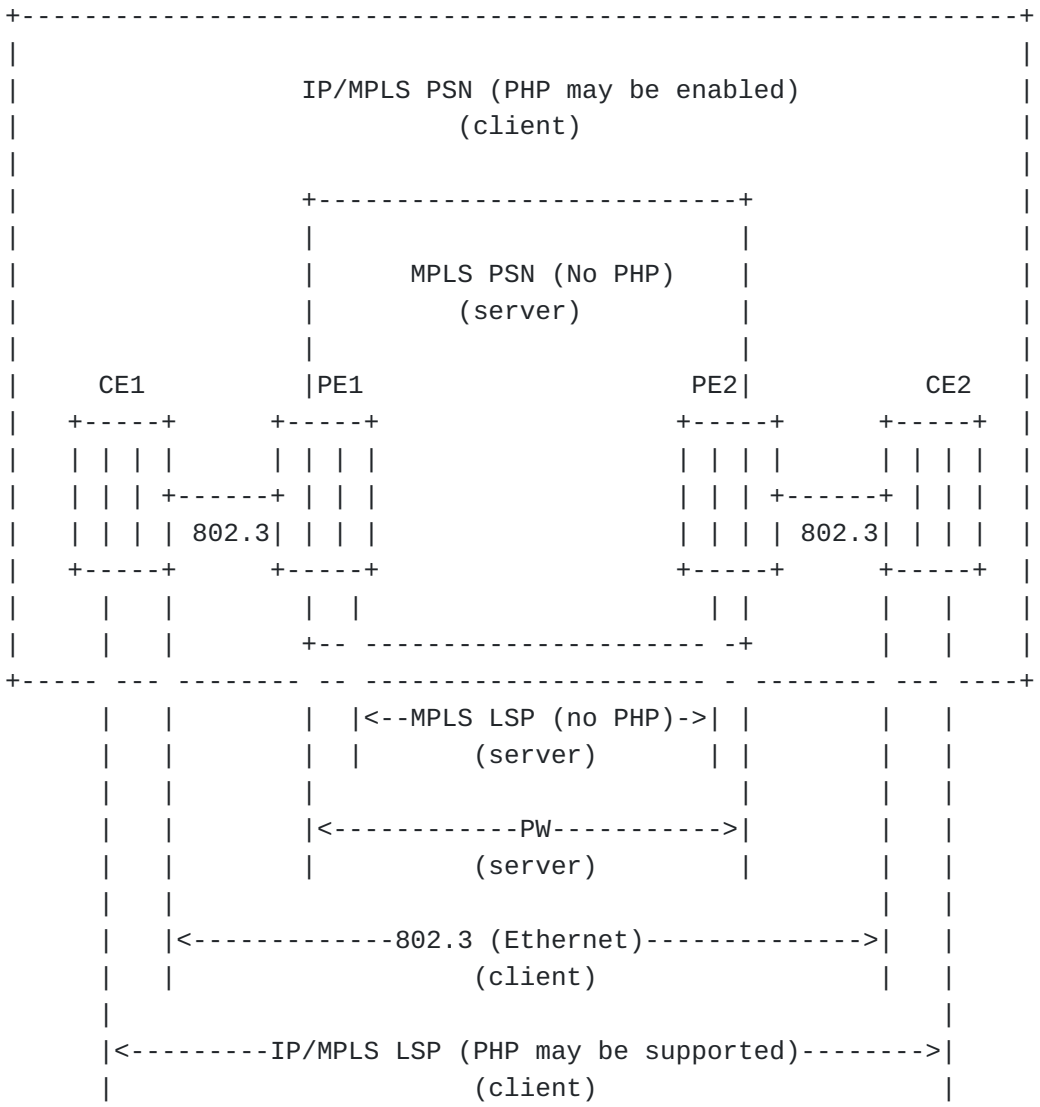

Figure 1: Application Ethernet over MPLS PW to MPLS Transport Networks
An 802.3 (Ethernet) circuit is established between CE1 and CE2. This circuit may be used for the concurrent transport of MPLS packets as well as IPv4 and IPv6 packets. The MPLS packets may carry IPv4, IPV6, or Pseudowire payloads, 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 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 draft call the server network a transport network.

2. PWE3 Configuration

TOC

The PWE3 encapsulation used by this specification to satisfy the transport requirement is Ethernet [RFC4448] (Martini, L., Rosen, E., El-Aawar, N., and G. Heron, “Encapsulation Methods for Transport of Ethernet over MPLS Networks,” April 2006.). 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] (Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, “Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP),” April 2006.) is not required by the profile of the PWE3 Ethernet pseudowire functionality defined in this document.
The Pseudowire Label is statically provisioned.

---

3. OAM

TOC

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 Multi-path (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 note 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 [RFC5085] (Nadeau, T. and C. Pignataro, “Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires,” December 2007.), using Bidirectional Forwarding Detection (BFD) [I-D.ietf-bfd-base] (Katz, D. and D. Ward, “Bidirectional Forwarding Detection,” January 2010.) for connection verification (CV). The method of using Bidirectional Forwarding Detection (BFD) as a CV method in VCCV is described in [I-D.ietf-pwe3-vccv-bfd] (Nadeau, T. and C. Pignataro, “Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV),” July 2009.). 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

TOC

When PE1 and PE1 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 [I-D.ietf-pwe3-vccv-bfd] (Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)," July 2009.).
[RFC5085] (Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires," December 2007.) specifies that the first nibble is set to 0x1 to indicate a channel associated with a pseudowire [RFC4385] (Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN," February 2006.).
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 [I-D.ietf-pwe3-vccv-bfd] (Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)," July 2009.).

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

TOC

When PE1 and PE1 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 [I-D.ietf-pwe3-vccv-bfd] (Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)," July 2009.).
[RFC5085] (Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires," December 2007.) specifies that the first nibble is set to 0x1 to indicate a channel associated with a pseudowire [RFC4385] (Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN," February 2006.).
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] (Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)," April 2006.).

---

4. MPLS Layer

TOC

The architecture of MPLS enabled networks is described in [RFC3031] (Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture," January 2001.). 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] (Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," October 2004.)

4.1. External Configuration

TOC

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 [I-D.ietf-mpls-tp-requirements] (Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "MPLS-TP Requirements," August 2009.) to address the needs in a transport environment.
The MPLS encapsulation is specified in [RFC3032] (Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding," January 2001.). 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 bi-directional point-to-point connections.
The transport LSPs SHOULD support unidirectional point-to-multipoint connections.
The forward and backward directions of a bi-directional connection SHOULD follow a symmetrically routed (reciprocal) LSP in the server network.
Equal cost multi-path (ECMP) load balancing MUST NOT be configured on the transport LSPs utilized by the PWs described in sections 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] (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,” May 2002.).
For the MPLS EXP field [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,” May 2002.) [RFC5462] (Andersson, L. and R. Asati, “Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field,” February 2009.) only the pipe and short-pipe models are supported.

4.2. Control Plane Configuration

TOC

In this section we describe the control plane configuration when[RFC3209] (Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, “RSVP-TE: Extensions to RSVP for LSP Tunnels,” December 2001.) “RSVP-TE: Extensions to RSVP for LSP Tunnels” or the bi-directional support in GMPLS [RFC3471] (Berger, L., “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description,” January 2003.) “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description" and[RFC3473] (Berger, L., “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions,” January 2003.) “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions” 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] (Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, “RSVP-TE: Extensions to RSVP for LSP Tunnels,” December 2001.).

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:

*Penultimate hop popping by the LSRs MUST be disabled on LSPs providing PWE3 transport network functionality [I-D.ietf-mpls-rsvp-te-no-php-oob-mapping] (Ali, Z. and G.

Swallow, “Non PHP Behavior and out-of-band mapping for RSVP-TE LSPs,” March 2010.).

*Both E-LSP and L-LSP MUST be supported as defined in [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,” May 2002.).

*The MPLS EXP [RFC5462] (Andersson, L. and R. Asati, “Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field,” February 2009.) field is supported according to [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,” May 2002.) for only when the pipe and short-pipe models are utilized.

---

5. Congestion Considerations

TOC

This draft describes a method of using the existing PWE3 Ethernet pseudowire [RFC4448] (Martini, L., Rosen, E., El-Aawar, N., and G. Heron, “Encapsulation Methods for Transport of Ethernet over MPLS Networks,” April 2006.) 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 draft.

---

6. Security Considerations

TOC

This draft is 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] (Bryant, S. and P. Pate, “Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture,” March 2005.) and the Ethernet pseudowire security considerations of[RFC4448] (Martini, L., Rosen, E., El-Aawar, N., and G. Heron, “Encapsulation Methods for Transport of Ethernet over MPLS Networks,” April 2006.).
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] (Rosen, E., Viswanathan, A., and R. Callon, “Multiprotocol Label Switching Architecture,” January 2001.).

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 worst, 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. IANA Considerations

TOC

There are no IANA actions required by this draft.

8. Acknowledgements

TOC

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.

9. References

TOC

9.1. Normative References

TOC

[I-D.ietf-bfd-base]	Katz, D. and D. Ward, “Bidirectional Forwarding Detection,” draft-ietf-bfd-base-11 (work in progress), January 2010 (TXT).
[I-D.ietf-mpls-rsvp-te-no-php-oob-mapping]	Ali, Z. and G. Swallow, “Non PHP Behavior and out-of-band mapping for RSVP-TE LSPs,” draft-ietf-mpls-rsvp-te-no-php-oob-mapping-04 (work in progress), March 2010 (TXT).
[I-D.ietf-pwe3-vccv-bfd]	Nadeau, T. and C. Pignataro, “Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV),” draft-ietf-pwe3-vccv-bfd-07 (work in progress), July 2009 (TXT).

[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC3031]	Rosen, E., Viswanathan, A., and R. Callon, “Multiprotocol Label Switching Architecture,” RFC 3031, January 2001 (TXT).
[RFC3032]	Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, “MPLS Label Stack Encoding,” RFC 3032, January 2001 (TXT).
[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 (TXT).
[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 (TXT).
[RFC3471]	Berger, L., “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description,” RFC 3471, January 2003 (TXT).
[RFC3473]	Berger, L., “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions,” RFC 3473, January 2003 (TXT).
[RFC3945]	Mannie, E., “Generalized Multi-Protocol Label Switching (GMPLS) Architecture,” RFC 3945, October 2004 (TXT).
[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 (TXT).
[RFC4446]	Martini, L., “IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3),” BCP 116, RFC 4446, April 2006 (TXT).
[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 (TXT).
[RFC4448]	Martini, L., Rosen, E., El-Aawar, N., and G. Heron, “Encapsulation Methods for Transport of Ethernet over MPLS Networks,” RFC 4448, April 2006 (TXT).
[RFC5085]	Nadeau, T. and C. Pignataro, “Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires,” RFC 5085, December 2007 (TXT).
[RFC5462]	Andersson, L. and R. Asati, “Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field

	Renamed to "Traffic Class" Field," RFC 5462, February 2009 (TXT).

9.2. Informative References

TOC

[I-D.ietf-mpls-tp-requirements]	Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "MPLS-TP Requirements," draft-ietf-mpls-tp-requirements-10 (work in progress), August 2009 (TXT).
[RFC3985]	Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture," RFC 3985, March 2005 (TXT).

Authors' Addresses

TOC

	Stewart Bryant (editor)
	Cisco Systems
	250, Longwater, Green Park,
	Reading RG2 6GB, UK
	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
	Thomas D. Nadeau
	BT
Email:	tom.nadeau@bt.com

	Neil Harrison
	BT
Email:	neil.2.harrison@bt.com
	Ben Niven-Jenkins
	BT
	208 Callisto House, Adastral Park
	Ipswich, Suffolk IP5 3RE
	UK
Phone:	
Fax:	
Email:	benjamin.niven-jenkins@bt.com
URI:	