

Internet Draft

Document: [draft-fedyk-gmpls-ethernet-ivl-00.txt](#)

Category: Standards Track

Don Fedyk, David Allan

Nortel

October 2005

GMPLS control of Ethernet IVL Switches

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire in April, 2006.

Copyright Notice Copyright (C) The Internet Society (2005).

Abstract

This memo describes how GMPLS signaling may be applied to appropriately configured Ethernet Independent VLAN Learning capable switches in order to configure Ethernet switched paths.

[1](#). Introduction

Ethernet switches are growing in capability. As a consequence the role of Ethernet is rapidly expanding in networks that were the domain of other technologies such as SONET/SDH TDM and ATM. The question of how Ethernet will evolve and what capabilities it can offer in these areas is still under development.

GMPLS Control of Ethernet IVL Switches

This draft explores some unique capabilities that Ethernet has and blends them with some of the values of control planes developed in the IETF.

The techniques for repurposing Ethernet switching outlined in this draft have been introduced elsewhere as a complement to 802.1ah Provider Backbone Bridging under the title "Provider Backbone Transport".

2. Terminology

In addition to well understood GMPLS terms, this memo uses terminology from IEEE 802.1 and introduces a few new terms:

B-MAC	Backbone MAC
B-VID	Backbone VLAN ID
B-VLAN	Backbone MAC
C-MAC	Customer MAC
C-VID	Customer VLAN ID
DA	Destination Address
ESP	Ethernet Switched Path
IVL	Independent VLAN Learning
PBB	Provider Backbone Bridge
PBT	Provider Backbone Transport
SA	Source Address
S-VID	Service VLAN ID

3. IVL Overview and ability to configure Ethernet Switched Paths

Ethernet as specified today is a system. How spanning tree, data plane flooding and MAC learning combine to populate forwarding tables and produce resilient any-to-any behavior in a bridged network is well understood.

What is less obvious is that the resulting behavior is purely a consequence of this particular combination of functions combined

with what the underlying hardware can do, and that by simply disabling some Ethernet functionality, it is possible to employ alternative control planes and obtain different forwarding behaviors.

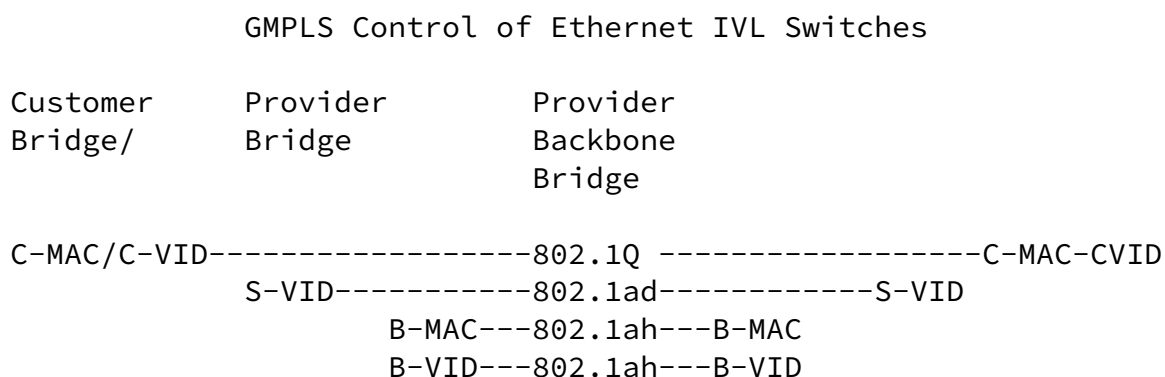


Figure 1 802.1 MAC/VLAN Hierarchy

Recent work in [PWE] and IEEE 802 are leading to a separation of Ethernet functions permitting an increasing degree of carrier control. So while the Ethernet service to the customer appears the same, the carrier components and behaviors have become decoupled from the customer presentation.

One example of this is the 802.1ah work in hierarchical bridging. Once the carrier is in exclusive control of their Ethernet sub-network and all customer specific state pushed to the edges of that sub-network, the ability of the carrier to exploit latent Ethernet behavior is facilitated. One key capability is to overcome limitations imposed by bridging.

Bridging offers a simple solution for any-to-any connectivity within a VLAN partition but attempting to use degenerate forms of bridged connectivity for point to point services puts strain on the control plane for forwarding and learning and may not offer

the engineerability that providers require in offering P2P services. It is easier to modify Ethernet to scale engineered P2P as a special case than to specify forms of any-to-any that are so scalable that consumption of any-to-any transport instances for point to point connectivity is inconsequential.

Ethernet Switches may perform Individual B-VLAN Learning (IVL) based unicast forwarding on the basis of a B-VID/B-MAC tuple. This means the forwarding hardware performs a full 60 bit lookup (B-VID (12) + B-MAC DA (48)) only requiring uniqueness of the full 60 bits for forwarding to resolve correctly.

We can redefine the semantics of the constituent elements and get complete route freedom for each 60 bit entry so long as the overall requirement for global uniqueness is maintained. For compatibility and flexibility reasons, it makes sense to preserve the global uniqueness and semantics of MAC addresses as interface names, but we can redefine the semantics associated with administration and use of VLAN values.

We partition the B-VID space and allocate a range of B-VIDs (say 'n' B-VIDs) as only significant when combined with a destination B-MAC address. We can then consider a B-VID in that range as an individual instance identifier for one of a maximum of 'n' P2P connections or MP2P multiplexed connections (subsequently termed

GMPLS Control of Ethernet IVL Switches

"shared forwarding" to distinguish from simple merges) terminating at the associated destination MAC address. While a B-VID in the allocated range is not unique on an Ethernet subnetwork basis, the B-VID/DA tuple is, and procedures can be designed to delegate administration of the allocated VID range to the node/interface identified by the DA MAC.

This technique results in a single unique and invariant identifier or label associated with the path termination and not a sequence of local identifiers associated with the individual link terminations. One consequence of this particular allocation is there can be up to 4094 labels to any one destination MAC address, a space that is consider large for P2P applications and overkill

when shared forwarding is leveraged. In practice, most network scaling requirements may be met via delegation of only a small portion of the VID space, resulting in minimal impact of the number of bridging VLANs that can be supported in addition to this behavior.

In order to use this unique 60 bit label, we then disable the normal mechanisms by which Ethernet populates the forwarding table for the allocated range of B-VIDs, and use a separate control or management plane to populate the forwarding table. When we do that for a specific label across a contiguous sequence of IVL capable Ethernet switches, this will create a unidirectional connection or Ethernet Switched Path (ESP). A bi-directional path is composed of two unidirectional paths. The technique does not require the VID to be common in both directions. Just as is the case for regular Ethernet these paths are preferred to be symmetrical such that a single bidirectional connection is composed of unidirectional paths that have common routing in the network.

There are a few modifications required to standard Ethernet to make this approach robust:

1. In Standard Ethernet, discontinuities in forwarding table configuration in the path of a connection will normally result in packets being flooded as "unknown". In the proposed point to point case this is unnecessary and potentially catastrophic in meshed topologies, therefore "unknown" flooding must be disabled for the allocated B-VID range. Flooding is similarly disabled for broadcast/multicast traffic.
2. B-MAC learning is not required for the ESP, and may interfere with management/control population of the forwarding tables. For this reason Provider B-MAC learning is disabled for the allocated B-VID range.
3. Blocking must be disabled to achieve complete configured route freedom. Similarly a configured path is assumed to be loop free. Spanning tree is disabled for the allocated B-VID range.

GMPLS Control of Ethernet IVL Switches

4. Robustness is enhanced with the addition of data plane OAM to provide both fault and performance management. How the hardware performs unicast packet forwarding of known MAC addresses is unmodified from how Ethernet currently works, so the OAM currently defined [802.1ag, Y.17ethoam] can also be reused without modification of the protocols, but with slightly altered semantics (primarily w.r.t. implementation scaling variations). An additional benefit of global path identifiers for dataplane forwarding is the tight coupling of the 60 bit unique connection ID (B-VID + B-MAC:DA) and the associated OAM packets. It is a simple matter to determine lost or misdirected packet because the unique connection ID cannot be altered.

Ethernet VLAN tags include priority tagging in the form of the 802.1p priority bits. When combined with configuration of the paths via management or control plane, this produces the Ethernet equivalent of MPLS TE E-LSPs and L-LSPs. Priority tagged Ethernet PDUs self-identify the required queuing discipline independent of the configured connectivity. This significantly simplifies the task of signaling scalable connectivity.

4. Deployment Scenarios

This technique of GMPLS controlled Ethernet switching is applicable to all deployment scenarios considered by the design team [[CCAMP-ETHERNET](#)].

5. Path creation and maintenance

One simple mode of path creation described earlier is configuration. Node by node a path can be created by simple configuration or by a set of commands originating from a management station.

One improvement to node by node configuration is to look at single ended provisioning and signaling. Since this is the domain of CCAMP and GMPLS we will discuss the applicability of GMPLS to this problem.

5.1.1 Using a GMPLS Control Plane for IVL

GMPLS has been adapted to the control of optical switches for the purpose of management optical paths. GMPLS signaling is well suited to setup paths with labels but it does require a minimal IP control plane and IP connectivity so it is suited to certain scenarios where a large number of paths or dynamic path management of IVL is required.

In many situations for IVL the addition of a complete GMPLS

control plane may be overkill for the switches or the application. So we will decompose the problem into Signaling, Routing, Link discovery and Path management. While we discuss the options it will become apparent that using all functions of GMPLS is less of

GMPLS Control of Ethernet IVL Switches

an operational overhead than any other combination. However, using only some components of GMPLS can lead to more provisioned parameters than a purely static system.

Link discovery is one of the foundations of GMPLS. It is also a capability that has been specified for Ethernet in IEEE 802.1AB. All link discovery is optional but the benefits of running link discovery in large systems are significant. A recommendation is that 802.1AB could be run with an extension to feed information into an LMP based discovery. The LMP based discovery would allow for a complete functional LMP for the purpose of GMPLS topology discovery. LMP requires an IP control plane (as does GMPLS). 802.1AB does not have the ability to carry all of the LMP messages. So the LMP implementation would be compatible with 802.1AB but add the extensions for LMP discovery. See Figure 3.

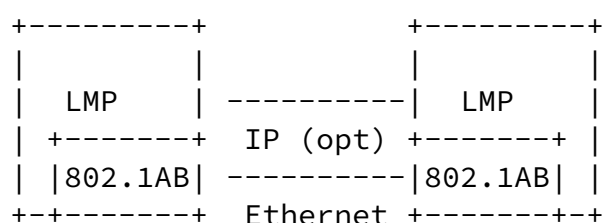


Figure 3 Link Discovery Hierarchy

[5.1.2](#) Control Plane Network

In order to have a GMPLS control plane, an IP control plane consisting of an IGP with TE extension needs to be established. This foundation of routing and traffic engineering parameters is used to establish control channels with only minimal capability to forward IP packets.

This IP control plane can be provided as a separate independent

network or integrated with the Ethernet switches.

If it is a separate network, it may be provided as a Layer 2 connected VLAN where the Ethernet switches are connected via a bridged network or HUB. It may also be provided by a network that provides IP connectivity where an IPVPN provides the IP connectivity.

If it is integrated with the switches it may be provided by a bridged VLAN that uses the Data bearing channels of the network for adjacent nodes. This ties the fate of IVL service and the IP control plane links, but then the same Ethernet hardware is already being shared.

[5.1.3](#) Signaling

GMPLS signaling is well suited to the set up of Ethernet Switched Paths on IVL capable Ethernet switches. GMPLS signaling uses link identifiers in the form of IP addresses either numbered

GMPLS Control of Ethernet IVL Switches

or unnumbered. If LMP is used the creation of these addresses can be automated. If LMP is not used there is an additional provisioning requirement to add GMPLS link identifiers. For large implementations LMP would be beneficial. As mentioned earlier the primary benefit of signaling is the control of a path from an endpoint. GMPLS can be used to create bidirectional or unidirectional paths, bi-directional paths being the preferred mode of operation for numerous reasons (OAM consistency etc.).

Signaling enables the ability to dynamically establish a path and to adjust the path in a coordinated fashion after the path has been established. Signaling also allows multi-vendor interoperability since the signaling is based on GMPLS signaling protocols. This allows the network to adapt to changing conditions or failures with a single mechanism. Signaling can be used in pure static paths as well, but the benefit of maintaining a GMPLS control plane for a pure static path application could be questioned.


```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     MAC Address                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Note that there is no syntax in signaling to force the label in the UPSTREAM_LABEL and GENERALIZED_LABEL objects to be passed unchanged, and so the semantics of the new label type are that the label is passed unchanged. This has similarity to how a wavelength label is handled at an intermediate node that cannot perform wavelength conversion, and is described in [\[GMPLS-RSVP\]](#).

[5.1.5](#) Ethernet Service

Ethernet Switched Paths that are setup either by configuration or signalling can be used to provide Ethernet service to customers of the IVL network. The Metro Ethernet Forum has defined some services in MEF.6 and MEF.11 (e.g., Ethernet Private Line), and these are also aligned with ITU-T G.8011.x Recommendations. Of particular interest are the bandwidth profile parameters in MEF.10 and whose associated bandwidth profile algorithm are based on [DS-COLOR]. Consideration should be given to supporting these in any signalling extensions for ESPs. This will be addressed in a future version of this specification.

[5.1.6](#) GMPLS Routing

GMPLS routing is IP routing with the TLV extensions for the purpose of carrying around GMPLS TE information. The TE information is populated with TE resources from LMP or from configuration if LMP is not available. GMPLS Routing is an optional piece but it is highly valuable in maintaining topology for path management.

[5.1.7](#) Path Computation

There has been a lot of recent activity in the area of path computation. Originally MPLS path computation was performed locally in a TE database on a router. While this is effective when a few paths are required in a primarily connectionless environment, if a large number of coordinated paths are required, it is advantageous to have a more sophisticated path computation engine capable of more elaborate paths. The longer lived the paths and the higher bandwidth the greater the computational budget required in order to get well engineered and cost effective paths. This is highly dependent on topology, since in simple topologies path computation is trivial.

GMPLS Control of Ethernet IVL Switches

Path computation in GMPLS generates explicit route objects (EROs) that can be used directly by GMPLS signaling.

[5.1.8](#) Combinations of GMPLS Features

The combinations of LMP, Routing, Signaling and Path computation can be all supported on a switch or a subset of GMPLS pieces can be supported.

Signaling is the minimal function that might be supported on a small switch. The ability to process Signaling messages with an ERO may be all that is desired on an end point.

Routing is the next important function since it provides the forwarding of signaling functions. However it is possible to design switches without routing that could proxy their routing to other larger switches. From the routing perspective there would be little difference in the routing database but the small switches would not have to perform routing operations. The information for the proxied routing might be configured or it might be data filled by an automated procedure. Rather than creating a new protocol it is probably better to put in a full OSPF or IS-IS control plane.

LMP is optional as mentioned earlier. The primary benefit of LMP over 802.1AB is LMP's capability for optimizing routing information for the purpose of link bundling on large switches. LMP has the capability to compress the information required to represent a large number of parallel resources automatically.

Path computation is one function that is probably best done on a centralized database for this application. Depending on the physical topology the explicit route object (ERO) may be trivial to calculate or more elaborate. Some form of path protection either based on Fast Reroute techniques or local computation may also be desirable for network outages but the targeted service for this is long lived relatively static paths.

[5.2](#) Addresses, Interfaces, and Labels

PBT is currently envisioned to have no explicit UNI or E-NNI which

simplifies the addressing of the control plane. This specification uses an addressing scheme and a label space for the ingress/egress connection: the hierarchical GMPLS Node Address/Port ID and the Ethernet MAC/VID tuple for the delegated VID range as a label space.

GMPLS Control of Ethernet IVL Switches

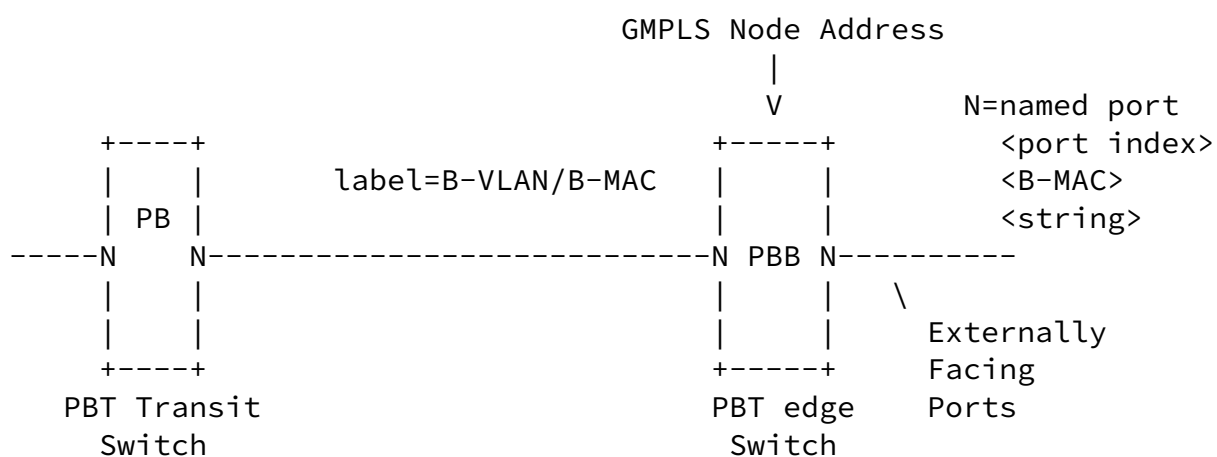


Figure 4 Ethernet/GMPLS Addressing & Label Space

Depending on how the service is defined and set up one or more of these labels may be used for actual setup. It is also to be noted that although it is possible for a terminating node to offer any 60 bit label value that can be guaranteed to be unique, the convention of using MAC addresses to name specific ports is retained, an Ethernet port name being common to both PBT and bridging modes of

operation. One implication of this is that a port index and a MAC address of a port on the node may be effectively interchangeable for signaling purposes, although incorrect information can result in an incorrect association between a GMPLS Node Address and the set of MAC named interfaces local to that node.

For a GMPLS based system the GMPLS Node Address/logical port is the logical signaling identifier for the control plane via which Ethernet layer label bindings are solicited. In order to create a point to point path for IVL an association must be made between the ingress and egress node. But the actual IVL label distributed via signaling and instantiated in the switch forwarding tables contains the egress interface name (B-MAC) of the port on the egress PBB. See Figure 4.

GMPLS uses identifiers in the form of 32 bit number which are in the IP address notation but these are not IP addresses. An IP routing control plane for the propagation of TE information may be supported. The provider B-MAC addresses are exchanged by the LLDP and by LMP if supported. Actual label assignment is performed by the signaling initiator and terminator.

A particular port on a Provider network node would have:

- A B-MAC
- A 32 bit IPv4 Node address or 128 bit IPv6 address plus 32 bit port Identifier
- One (or more) Mnemonic String Identifiers

GMPLS Control of Ethernet IVL Switches

This multiple naming convention leaves the issue of resolving the set given one of the port identifiers. On a particular node, mapping is relatively straight forward. The preferred solution to this is to use the GMPLS IP node address for signaling resolution. In so doing, the problem of setting up a path is reduced to figuring out what node supports a B-MAC address and then finding the corresponding GMPLS IP node address and performing all signaling and routing with respect to the GMPLS node address.

There are several options to achieve this:

- Provisioning
- Auto discovery protocols that carry MAC address (e.g. 802.1ab)
- Augmenting Routing TE with B-MAC Addresses
- Name Servers with identifier/address registration

This will be clarified in a subsequent version of this draft.

[6. Specific Procedures](#)

[6.1 PT to PT connections](#)

The Data Plane for IVL has three key fields. B-VID, B-MAC DA and B-MAC SA. A connection instance is uniquely identified by the B-MAC DA, B-VID and B-MAC SA for the purpose of the provider network terminations. The B-VID and B-MAC DA tuple identifies the forwarding multiplex at transit switches and a simple degenerate form of the multiplex is P2P (only one SA/VID/DA connection uses the VID/DA tuple).

This results in unique labels end to end and no merging or multiplexing of tunnels. The data streams may merge but the forwarding entries are unique allowing the connection to be de-multiplexed downstream.

[6.2 PT to PT connections with shared forwarding](#)

The B-VID/B-MAC DA can be considered to be a shared forwarding identifier for a multiplex consisting of some number of P2P connections distinctly identified by the B-MAC SA/B-VID/B-MAC DA tuple.

VLAN tagged Ethernet packets include priority marking. This means that the queuing discipline applied is selectable on a per flow basis and is decoupled from the actual steering of the packet at any given node. As noted earlier, GMPLS signaled paths can have similar properties to MPLS traffic engineered E-LSPs[MPLS-DS]. What this means is that multiple Ethernet switched paths to a destination may be considered functionally equivalent. This is a useful property when considering setup of shared forwarding Ethernet switched paths. A terminating node will frequently have more than one suitable candidate path with which new path requests may be multiplexed on the data plane (common VID/DA, unique SA).

[6.3 Dynamically established P2P symmetry with shared forwarding](#)

Similar to how a destination node may select a B-VID/B-MAC DA from the set of existing shared forwarding multiplexes rooted at the destination node, the originating node may also do so for the reverse path. Once the initial ERO has been computed, the originating node may select the optimal (by whatever criteria) existing shared forwarding multiplex for the new destination to merge with and offer its own B-VID/B-MAC DA tuple for itself as a

destination. This is identified via use of the UPSTREAM LABEL object.

[6.4](#) Planned P2P symmetry

Normally the originating node will not have knowledge of the set of shared forwarding path rooted on the destination node.

Use of a Path Computation Server or other planning style of tool with more complete knowledge of the network configuration may wish to impose pre-selection of the a more optimal shared forwarding multiplexes to use for both directions. In this scenario the originating node uses the SUGGESTED LABEL and UPSTREAM LABEL objects to indicate complete selection of the shared forwarding multiplexes at both ends. This may also result in the establishment of a new B-VID/B-MAC DA path as the SUGGESTED LABEL object may legitimately refer to a path that does not yet exist.

[7.](#) Error conditions

The following errors have been identified as being unique to these procedures in addition to those already defined:

[7.1](#) Invalid B-VID value

The originator of the error is not configured to use the B-VID value in conjunction with GMPLS signaling. This may be either the initiating or terminating node.

[7.2](#) Invalid ERO

The ERO offered has discontinuities with the identified VID/MAC path in either the UPSTREAM LABEL or SUGGESTED LABEL objects.

[8.](#) Security Considerations

TBD.

[9.](#) IANA Considerations

TBD.

10. References

10.1 Normative References

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

10.2 Informative References

[CCAMP-ETHERNET] Papadimitriou, D. et.al, "A Framework for Generalized MPLS (GMPLS) Ethernet", internet draft, [draft-papadimitriou-ccamp-gmpls-ethernet-framework-00.txt](#) , June 2005

[DS-COLOR] Aboul-Magd, O. et.al. "A Differentiated Service Two-Rate, Three-Color Marker with Efficient Handling of in-Profile Traffic", IETF [RFC 4115](#), July 2005

[GMPLS-ARCH] E. Mannie, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", [RFC 3495](#).

[GMPLS-SIGNALING] Berger, L. (editor), "Generalized MPLS - Signaling Functional Description", January 2003, [RFC3471](#).

[GMPLS-ROUTING] Kompella, K., Rekhter, Y., "Routing Extensions in Support of Generalized MPLS", work in progress.

[LMP] Lang. J. Editor, "Link Management Protocol (LMP)" work in progress.

[IEEE 802.1ab] "IEEE Draft Standard for Local and Metropolitan Area Networks, Station and Media Access Control Connectivity Discovery"

[IEEE 802.1ag] "IEEE standard for Connectivity Fault Management", Work in Progress, July 2005.

[IEEE 802.1ah] "IEEE standard for Provider Backbone Bridges", Work in Progress, July 2005.

[MEF.6] The Metro Ethernet Forum MEF 6 (2004), "Ethernet Services Definitions - Phase I"

[MEF.10] The Metro Ethernet Forum MEF 10 (2004), "Ethernet Services Attributes Phase 1"

[MEF.11] The Metro Ethernet Forum MEF 11 (2004), "User Network Interface (UNI) Requirements and Framework"

GMPLS Control of Ethernet IVL Switches

(MPLS) Support of Differentiated Services" IETF [RFC 3270](#), May 2002

[PATH-COMP] Farrel, A. et.al., "Path Computation Element (PCE) Architecture", Internet draft, [draft-ietf-pce-architecture-02.txt](#), September 2005

[Y.17ethoam] ITU-T Recommendation, "Draft Recommendation Y.17ethoam - OAM Functions and Mechanisms for Ethernet based Networks " work in progress, May 2005.

11. Author's Address

Don Fedyk
Nortel Networks
600 Technology Park Drive
Billerica, MA, 01821
Email: dwfedyk@nortel.com

David Allan
Nortel Networks
3500 Carling Ave.
Ottawa, Ontario, CANADA

Phone: 1-613-763-6362
Email: dallan@nortel.com

12. 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.

13.Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND

Fedyk et al.

Expires April 2006

Page 14

GMPLS Control of Ethernet IVL Switches

THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

14.Copyright Statement

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

15.Acknowledgments

The authors would like to thank Nigel Bragg, Stephen Shew and Sandra Ballarte for their extensive contributions to this document.

