

Network Working Group  
Internet Draft  
Expiration Date: May 2001

Luca Martini  
Nasser El-Aawar  
Giles Heron  
Level 3 Communications, LLC.

Daniel Tappan  
Eric C. Rosen  
Cisco Systems, Inc.

Steve Vogelsang  
John Shirron  
Laurel Networks, Inc.

Andrew G. Malis  
Vivace Networks, Inc.

Dimitri Stratton Vlachos  
Mazu Networks, Inc.

November 2000

## **Transport of Layer 2 Frames Over MPLS**

[draft-martini-l2circuit-trans-mpls-04.txt](#)

### Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

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/1id-abstracts.txt>.

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

## Abstract

This document describes methods for transporting the Protocol Data Units (PDUs) of layer 2 protocols such as Frame Relay, ATM AAL5, Ethernet, and providing a SONET circuit emulation service across an MPLS network.

## Table of Contents

<a href="#">1</a>	Specification of Requirements .....	<a href="#">2</a>
<a href="#">2</a>	Introduction .....	<a href="#">2</a>
<a href="#">3</a>	Tunnel Labels and VC Labels .....	<a href="#">3</a>
<a href="#">4</a>	Protocol-Specific Issues .....	<a href="#">4</a>
<a href="#">4.1</a>	Frame Relay .....	<a href="#">4</a>
<a href="#">4.2</a>	ATM .....	<a href="#">4</a>
<a href="#">4.2.1</a>	OAM Cell Support .....	<a href="#">4</a>
<a href="#">4.2.2</a>	ILMI Support .....	<a href="#">5</a>
<a href="#">4.3</a>	HDLC ( Cisco ) .....	<a href="#">5</a>
<a href="#">4.4</a>	PPP .....	<a href="#">5</a>
<a href="#">5</a>	LDP .....	<a href="#">6</a>
<a href="#">6</a>	Security Considerations .....	<a href="#">9</a>
<a href="#">7</a>	References .....	<a href="#">9</a>
<a href="#">8</a>	Author Information .....	<a href="#">9</a>

## [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

In an MPLS network, it is possible to carry the Protocol Data Units (PDUs) of layer 2 protocols by prepending an MPLS label stack to these PDUs. This document specifies the necessary label distribution procedures for accomplishing this using the encapsulation methods in [\[7\]](#). We restrict discussion to the case of point-to-point transport. QoS related issues are not discussed in this draft.

An accompanying document [\[8\]](#) also describes a method for transporting time division multiplexed (TDM) digital signals (TDM circuit emulation) over a packet-oriented MPLS network. The transmission system for circuit-oriented TDM signals is the Synchronous Optical Network (SONET)[\[5\]](#)/Synchronous Digital Hierarchy (SDH) [\[6\]](#). To



support TDM traffic, which includes voice, data, and private leased line service, the MPLS network must emulate the circuit characteristics of SONET/SDH payloads. MPLS labels and a new circuit emulation header are used to encapsulate TDM signals and provide the Circuit Emulation Service over MPLS (CEM). This encapsulation method is described in [8].

### **3. Tunnel Labels and VC Labels**

Suppose it is desired to transport layer 2 PDUs from ingress LSR R1 to egress LSR R2, across an intervening MPLS network. We assume that there is an LSP from R1 to R2. That is, we assume that R1 can cause a packet to be delivered to R2 by pushing some label onto the packet and sending the result to one of its adjacencies. Call this label the "tunnel label", and the corresponding LSP the "tunnel LSP".

The tunnel LSP merely gets packets from R1 to R2, the corresponding label doesn't tell R2 what to do with the payload, and in fact if penultimate hop popping is used, R2 may never even see the corresponding label. (If R1 itself is the penultimate hop, a tunnel label may not even get pushed on.) Thus if the payload is not an IP packet, there must be a label, which becomes visible to R2, that tells R2 how to treat the received packet. Call this label the "VC label".

So when R1 sends a layer 2 PDU to R2, it first pushes a VC label on its label stack, and then (if R1 is not adjacent to R2) pushes on a tunnel label. The tunnel label gets the MPLS packet from R1 to R2; the VC label is not visible until the MPLS packet reaches R2. R2's disposition of the packet is based on the VC label.

If the payload of the MPLS packet is, for example, an ATM AAL5 PDU, the VC label will generally correspond to a particular ATM VC at R2. That is, R2 needs to be able to infer from the VC label the outgoing interface and the VPI/VCI value for the AAL5 PDU. If the payload is a Frame Relay PDU, then R2 needs to be able to infer from the VC label the outgoing interface and the DLCI value. If the payload is an ethernet frame, then R2 needs to be able to infer from the VC label the outgoing interface, and perhaps the VLAN identifier. This process is unidirectional, and will be repeated independently for bidirectional operation. It is REQUIRED to assign the same VC, and Group ID for a given circuit in both directions.

Note that the VC label must always be at the bottom of the label stack, and the tunnel label, if present, must be immediately above the VC label. Of course, as the packet is transported across the MPLS network, additional labels may be pushed on (and then popped off) as needed. Even R1 itself may push on additional labels above the tunnel



label. If R1 and R2 are directly adjacent LSRs, then it may not be necessary to use a tunnel label at all.

This document does not specify a method for distributing the tunnel label or any other labels that may appear above it on the stack. Any acceptable method of MPLS label distribution will do.

This document does specify a method for assigning and distributing the VC label. Static label assignment MAY be used, and implementations SHOULD provide support for this. If signaling is used, the VC label MUST be distributed from R2 to R1 using LDP in the downstream unsolicited mode; this requires that an LDP connection be created between R1 and R2.

Note that this technique allows an unbounded number of layer 2 "VCs" to be carried together in a single "tunnel". Thus it scales quite well in the network backbone.

## **4. Protocol-Specific Issues**

### **4.1. Frame Relay**

The MPLS edge LSR MAY provide a Frame Relay LMI to the CE device.

If the MPLS edge LSR detects a service affecting condition as defined in [2] Q.933 Annex A.5 cited in IA FRF1.1, it MUST withdraw the label that corresponds to the frame relay DLCI. The Egress LSR SHOULD generate the corresponding errors and alarms as defined in [2] on the Frame relay VC.

### **4.2. ATM**

#### **4.2.1. OAM Cell Support**

OAM cells MAY be transported on the VC LSP. A router that does not support transport of ATM cells MUST discard incoming MPLS frames on an ATM VC LSP that contain a control word with the T bit set. [7] A router that supports transport of OAM cells MUST follow the procedures outlined in [9] [section 8](#) for mode 0 only in addition to the applicable procedures specified in [6].

A router that does not support transport of OAM cells across an LSP MAY provide OAM support on ATM PVCs using the following procedures:

If an F5 end-to-end OAM cell is received from a VC by a LSR with a loopback indication value of 1 and the LSR has a label mapping for the VC, the LSR MUST decrement the loopback indication value and loop back the cell on the VC. Otherwise the loopback cell MUST be



discarded by the LSR.

The LSR MAY optionally be configured to periodically generate F5 end-to-end loopback OAM cells on a VC. In this case, the LSR must only generate F5 end-to-end loopback cells while a label mapping exists for the VC. If the VC label mapping is withdrawn the LSR MUST cease generation of F5 end-to-end loopback OAM cells. If the LSR fails to receive a response to an F5 end-to-end loopback OAM cell for a pre-defined period of time it MUST withdraw the label mapping for the VC.

If an ingress LSR receives an AIS F5 OAM cell, fails to receive a pre-defined number of the End-to-End loop OAM cells, or a physical interface goes down, it MUST withdraw the label mappings for all VCs associated with the failure. When a VC label mapping is withdrawn, the egress LSR SHOULD generate AIS F5 OAM cells on the VC associated with the withdrawn label mapping.

#### **4.2.2. ILMI Support**

An MPLS edge LSR MAY provide an ATM ILMI to the CE device.

If an ingress LSR receives an ILMI message indicating that the CE has deleted a VC, or if the physical interface goes down, it MUST withdraw the label mappings for all VCs associated with the failure. When a VC label mapping is withdrawn, the egress LSR SHOULD notify its client of this failure by deleting the VC using ILMI.

#### **4.3. HDLC ( Cisco )**

If the MPLS edge LSR detects that the physical link has failed it MUST withdraw the label that corresponds to the HDLC link. The Egress LSR SHOULD notify the CE device of this failure by using a physical layer mechanism to take the link out of service.

#### **4.4. PPP**

If the MPLS edge LSR detects that the physical link has failed it MUST withdraw the label that corresponds to the PPP link. The Egress LSR SHOULD notify the CE device of this failure by using a physical layer mechanism to take the link out of service.

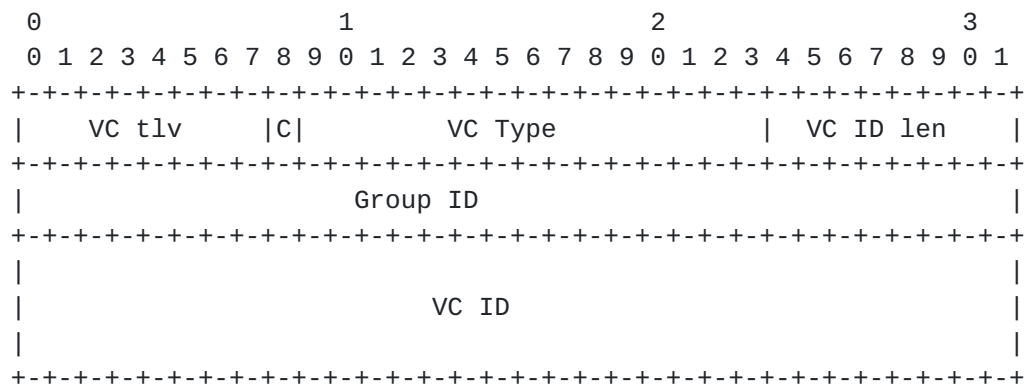




## 5. LDP

The VC label bindings are distributed using the LDP downstream unsolicited mode described in [1]. The LSRs will establish an LDP session using the Extended Discovery mechanism described in [1, [section 2.4-2.5](#)], for this purpose a new type of FEC TLV element is defined. The FEC element type is 128. [[note1](#)]

The Virtual Circuit FEC TLV element, is defined as follows:



### - VC Type

A 15 bit quantity containing a value which represents the type of VC. Assigned Values are:

VC Type	Description
0x0001	Frame Relay DLCI
0x0002	ATM VCC transport
0x0003	ATM VPC transport
0x0004	Ethernet VLAN
0x0005	Ethernet
0x0006	HDLC ( Cisco )
0x0007	PPP
0x8008	CEM [ <a href="#">8</a> ]

The highest order bit is used to flag the presence of a control word as follows:

- bit 15 = 1 control word present on this VC.
- bit 15 = 0 no control word present on this VC.

### - VC ID length

Length of the VC ID field in octets. If this value is 0, then it references all VCs using the specified group ID



- Group ID

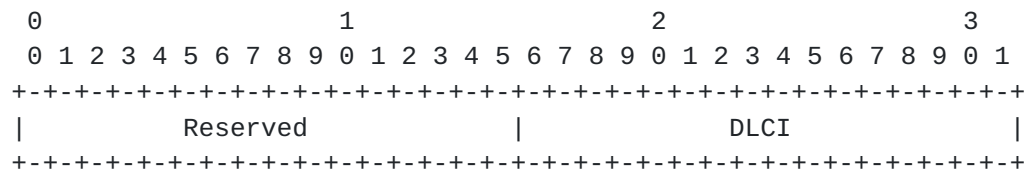
An arbitrary 32 bit value which represents a group of VCs that is used to augment the VC space. This value MUST be user configurable. The group ID is intended to be used as either a port index , or a virtual tunnel index. In the latter case a switching function at ingress will map a particular circuit from a port to a circuit in the virtual tunnel for transport to the egress router.

- VC ID

Identifies a particular VC. The interpretation of the identifier depends on the VC type:

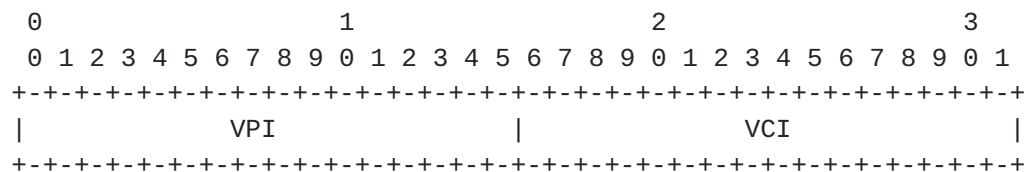
- \* Frame Relay

A 32-bit value representing a 16-bit DLCI value as follows:



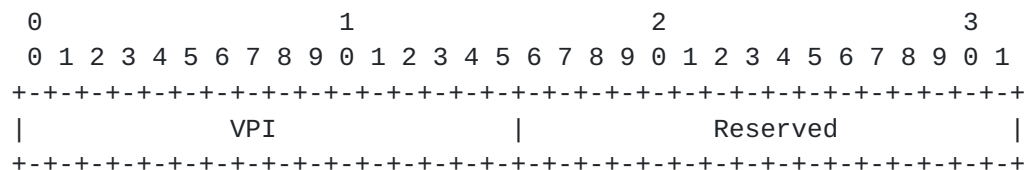
\* ATM VCC Transport

A 32-bit value representing a 16-bit VPI, and a 16-bit VCI as follows:



\* ATM VPC Transport

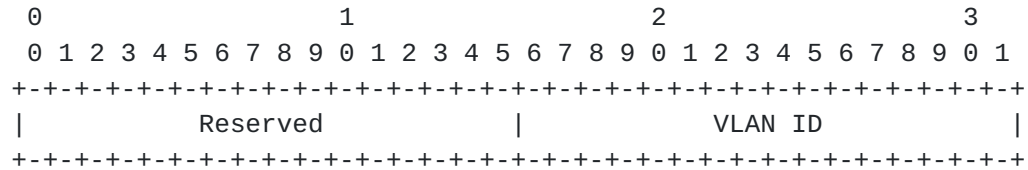
A 32-bit value containing a 16-bit VPI as follows:





- \* Ethernet VLAN

A 32 bit value representing 16bit vlan identifier as follows:



- \* Ethernet

A 32 bit port identifier.

```
* HDLC ( Cisco )
```

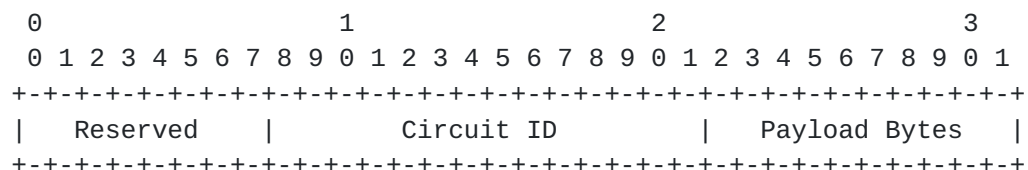
A 32-bit port identifier

\* PPP

A 32-bit port identifier

\* CEM[8]

A 32-bit value used follows:



Circuit ID: An assigned number for the SONET circuit being transported.

Payload Bytes(N): the number of TDM payload bytes contained in all packets on the CEM stream, from 48 to 1,023 bytes. All of the packets in a given CEM stream have the same number of payload bytes. Note that there is a possibility that the packet size may exceed the SPE size in the case of an STS-1 SPE, which could cause two pointers to be needed in the CEM header, since the payload may contain two J1 bytes for consecutive SPEs. For this reason, the number of payload bytes must be less than or equal to 783 for STS-1 SPEs. The reserved fields in the above specifications MUST be set to 0 in the FEC TLV, and ignored when received.



## **6. Security Considerations**

This document does not affect the underlying security issues of MPLS.

## **7. References**

- [1] "LDP Specification", [draft-ietf-mpls-ldp-11.txt](#) ( work in progress )
  - [2] ITU-T Recommendation Q.933, and Q.922 Specification for Frame Mode Basic call control, ITU Geneva 1995
  - [3] "MPLS Label Stack Encoding", [draft-ietf-mpls-label-encaps-08.txt](#) ( work in progress )
  - [4] "IEEE 802.3ac-1998" IEEE standard specification.
  - [5] American National Standards Institute, "Synchronous Optical Network Formats," ANSI T1.105-1995.
  - [6] ITU Recommendation G.707, "Network Node Interface For The Synchronous Digital Hierarchy", 1996.
  - [7] "Encapsulation Methods for Transport of Layer 2 Frames Over MPLS", [draft-martini-l2circuit-encap-mpls-00.txt](#) ( Work in progress )
  - [8] "SONET/SDH Circuit Emulation Service Over MPLS (CEM) Encapsulation", [draft-malis-sonet-ces-mpls-01.txt](#) ( Work in progress )
  - [9] "Frame Based ATM over SONET/SDH Transport (FAST)," 2000.
- [note1] FEC element type 128 is pending IANA approval.

## **8. Author Information**

Luca Martini  
Level 3 Communications, LLC.  
1025 Eldorado Blvd.  
Broomfield, CO, 80021  
e-mail: [luca@level3.net](mailto:luca@level3.net)





Nasser El-Aawar  
Level 3 Communications, LLC.  
1025 Eldorado Blvd.  
Broomfield, CO, 80021  
e-mail: nna@level3.net

Giles Heron  
Level 3 Communications  
66 Prescott Street  
London  
E1 8HG  
United Kingdom  
e-mail: giles@level3.net

Dimitri Stratton Vlachos  
Mazu Networks, Inc.  
125 Cambridgepark Drive  
Cambridge, MA 02140  
e-mail: d@mazunetworks.com

Dan Tappan  
Cisco Systems, Inc.  
250 Apollo Drive  
Chelmsford, MA, 01824  
e-mail: tappan@cisco.com

Eric Rosen  
Cisco Systems, Inc.  
250 Apollo Drive  
Chelmsford, MA, 01824  
e-mail: erosen@cisco.com

Steve Vogelsang  
Laurel Networks, Inc.  
2607 Nicholson Rd.  
Sewickley, PA 15143  
e-mail: sjv@laurelnetworks.com



John Shirron  
Laurel Networks, Inc.  
2607 Nicholson Rd.  
Sewickley, PA 15143  
e-mail: [jshirron@laurelnetworks.com](mailto:jshirron@laurelnetworks.com)

Andrew G. Malis  
Vivace Networks, Inc.  
2730 Orchard Parkway  
San Jose, CA 95134  
Phone: +1 408 383 7223  
Email: [Andy.Malis@vivacenetworks.com](mailto:Andy.Malis@vivacenetworks.com)

