

Network Working Group
Internet Draft
Expiration Date: January 2001

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

Dimitri Stratton Vlachos
Daniel Tappan
Eric C. Rosen
Cisco Systems, Inc.

Steve Vogelsang
John Shirron
Laurel Networks, Inc.

July 2000

Transport of Layer 2 Frames Over MPLS

[draft-martini-l2circuit-trans-mpls-02.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/lid-abstracts.txt>.

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

Abstract

This document described a method for transporting the Protocol Data Units (PDUs) of layer 2 protocols such as Frame Relay, ATM AAL5, and ethernet across an MPLS network.

Table of Contents

1	Specification of Requirements	2
2	Introduction	2
3	Tunnel Labels and VC Labels	3
4	Optional Sequencing and/or Padding	4
5	Protocol-Specific Issues	5
5.1	Frame Relay	5
5.2	ATM	5
5.2.1	F5 OAM Cell Support	6
5.2.2	CLP Bit	6
5.2.3	PTI Field in ATM Cell Mode	6
5.3	Ethernet VLAN	7
5.4	Ethernet	7
6	LDP	7
7	Security Considerations	10
8	References	10
9	Author Information	10

[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 and encapsulation procedures for accomplishing this. We restrict discussion to the case of point-to-point transport. QoS related issues are not discussed in this draft.

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 desirable, but not 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.

The MPLS network should be configured with a MTU that is at least 12 bytes larger than the largest packet size that will be transported in the LSPs. If a packet, once it has been encapsulated, exceeds the LSP MTU, it MUST be dropped.

[4. Optional Sequencing and/or Padding](#)

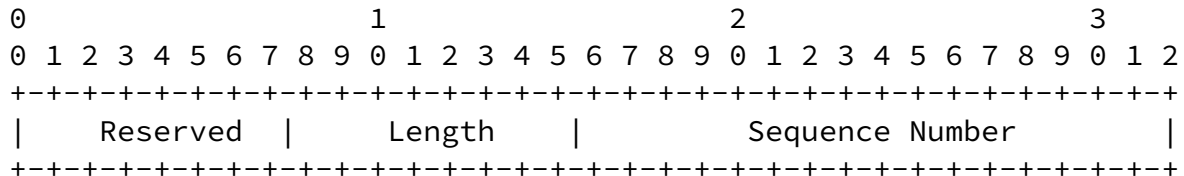
Sometimes it is important to guarantee that sequentiality is preserved on a layer 2 virtual circuit. To accommodate this requirement, we provide an optional control word which may appear immediately after the label stack and immediately before the layer 2 PDU. This control word contains a sequence number. R1 and R2 both need to be configured with the knowledge of whether a control word will be used for a specific virtual circuit.

Sometimes it is necessary to transmit a small packet on a medium where there is a minimum transport unit larger than the actual packet size. In this case, padding is appended to the packet. When the VC

label is popped, it may be desirable to remove the padding before forwarding the packet.

To facilitate this, the control word has a length field. If the packet's length (without any padding) is less than 256 bytes, the length field **MUST** be set to the packet's length (without padding). Otherwise the length field **MUST** be set to zero. The value of the length field, if non-zero, can be used to remove any padding.

The control word is defined as follows:



The first 8 bits are reserved for future use. They **MUST** be set to 0

when transmitting, and **MUST** be ignored upon receipt. The length byte is set as specified above.

The next 16 bits are the sequence number that is used to guarantee ordered packet delivery. For a given VC label, and a given pair of LSRs, R1 and R2, where R2 has distributed that VC label to R1, the sequence number is initialized to 0, and is incremented by one for each successive packet carrying that VC label which R1 transmits to R2.

The sequence number space is a 16 bit unsigned circular space. PDUs carrying the control word **MUST NOT** be delivered out of order. They may be discarded or reordered.

[5. Protocol-Specific Issues](#)

[5.1. Frame Relay](#)

A Frame Relay PDU is transported in its entirety, including the Frame Relay Header. The sequencing control word is **OPTIONAL**.

The BCN and FCN signals are carried unchanged across the network in the frame relay header. These signals do not appear in the MPLS header, and are unseen by the MPLS network.

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

The ingress LSR MAY consider the DE bit of the Frame Relay header when determining the value to be placed in the EXP fields of the MPLS label stack. In a similar way, the egress LSR MAY consider the EXP field of the VC label when queuing the packet for egress.

[5.2.](#) ATM

Two modes are supported for ATM transport, ATM Adaptation Layer 5 (AAL5) and ATM cell.

In ATM AAL5 mode the ingress LSR is required to reassemble AAL5 CPCS-PDUs from the incoming VC and transport each CPCS-PDU as a single packet. No AAL5 trailer is transported. The sequencing control word is OPTIONAL.

Martini, et al.

[Page 5]

Internet Draft [draft-martini-l2circuit-trans-mpls-02.txt](#)

July 2000

In ATM cell mode the ingress LSR transports each ATM cell payload as a single packet. No ATM cell header is transported. The sequencing control word is OPTIONAL.

[5.2.1.](#) F5 OAM Cell Support

F5 OAM cells are not transported on the VC LSP.

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 silently discarded by the LSR.

A 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 must generate AIS F5 OAM cells on the VC associated with the withdrawn label mapping.

5.2.2. CLP Bit

The ingress LSR MAY consider the CLP bit when determining the value to be placed in the EXP fields of the MPLS label stack.

The egress LSR MAY consider the value of the EXP field of the VC label when determining the value of the ATM CLP bit.

5.2.3. PTI Field in ATM Cell Mode

ATM cell mode is intended for transporting non-AAL5 traffic only. The ingress LSR must transport cells with a PTI of 0. Cells with a PTI other than 0 are not transported on the LSP. The egress LSR must set the PTI to 0 for cells switched from a VC LSP to an outgoing VC.

5.3. Ethernet VLAN

For and ethernet 802.1q VLAN the entire ethernet frame without the preamble or FCS is transported as a single packet. The sequencing control word is OPTIONAL. If a packet is received out of sequence it MUST be dropped. The VLAN 4 byte tag is transported as is, and MAY be overwritten by the egress LSR. The ingress LSR MAY consider the user priority field [4] of the VLAN tag header when determining the value

to be placed in the EXP fields of the MPLS label stack. In a similar way, the egress LSR MAY consider the EXP field of the VC label when queuing the packet for egress. Ethernet packets containing hardware level CRC, Framing errors, or runt packets MUST be discarded on input.

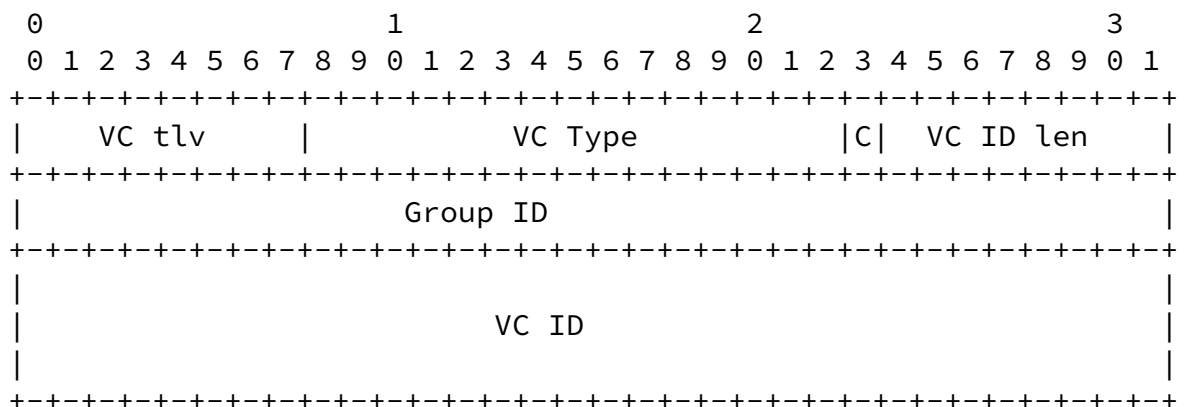
5.4. Ethernet

For simple ethernet port to port transport, the entire ethernet frame without the preamble or FCS is transported as a single packet. The sequencing control word is OPTIONAL. If a packet is received out of sequence it MUST be dropped. As in the Ethernet VLAN case, ethernet packets with hardware level CRC, framing, and runt errors are discarded.

6. 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:



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

VC TypeDescription

0x0001 Frame Relay DLCI
0x0002 ATM AAL5 PVC
0x0003 ATM Cell
0x0004 Ethernet VLAN
0x0005 Ethernet
0x0006 HDLC (Cisco)
0x0007 PPP

The highest order bit is used to flag the presence of a sequencing 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

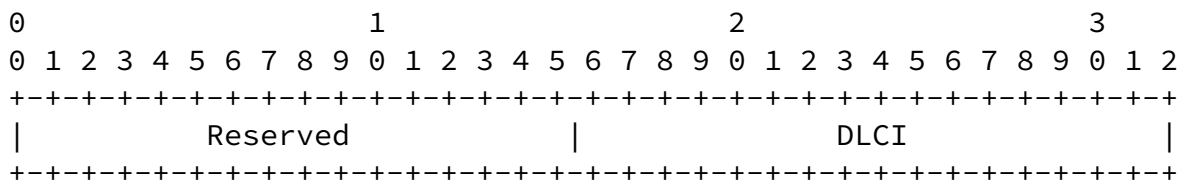
An arbitrary 32 bit value which represents a group of VCs that is used to augment the VC space. This value MUST be user configurable. The group ID is intended to be used as either a port index , or a virtual tunnel index. In the latter case a switching function at ingress will map a particular circuit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

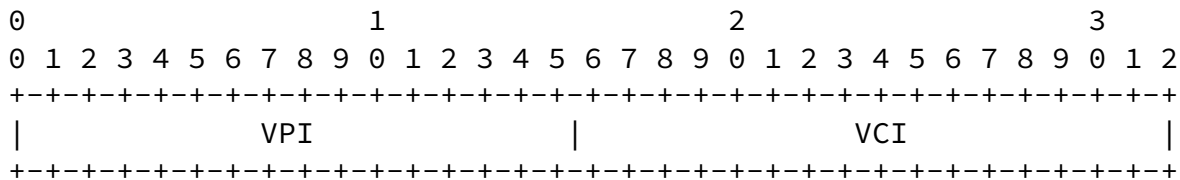
* Frame Relay

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



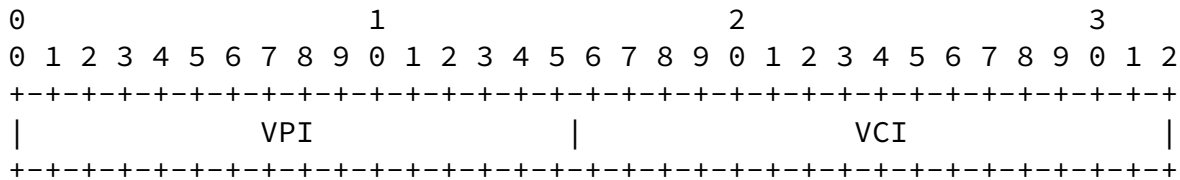
* ATM AAL5 PVC

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



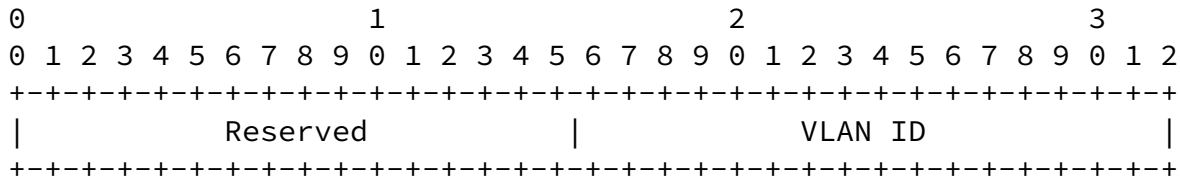
* ATM Cell

A 32-bit value representing a 16-bit VPI, and a 16-bit VCI 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 (details TBD).

* PPP

A 32-bit port identifier (details TBD).

The reserved fields in the above specifications MUST be set to 0 in the FEC TLV, and ignored when received. In some cases the the

[7.](#) Security Considerations

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

[8.](#) References

[1] "LDP Specification", [draft-ietf-mpls-ldp-07.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-07.txt](#) (work in progress)

[4] "IEEE 802.3ac-1998" IEEE standard specification.

[note1] FEC element type 128 is pending IANA approval.

[9.](#) Author Information

Luca Martini

Level 3 Communications, LLC.
1025 Eldorado Blvd.
Broomfield, CO, 80021
e-mail: luca@level3.net

Martini, et al.

[Page 10]

Internet Draft [draft-martini-l2circuit-trans-mpls-02.txt](#)

July 2000

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

Dimitri Stratton Vlachos
Cisco Systems, Inc.
250 Apollo Drive
Chelmsford, MA, 01824
e-mail: dvlachos@cisco.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: sjv@laurelnetworks.com

Martini, et al.

[Page 11]