

Internet-Draft
Expiration Date: August 1999

Nancy Feldman
IBM

Bilel Jamoussi
Nortel Networks

Sridhar Komandur
Ascend Communications

Arun Viswanathan
Lucent Technologies

Tom Worster
GDC

February 1999

MPLS using ATM VP Switching

[<draft-feldman-mpls-atmvp-00.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

The MPLS Architecture [[ARCH](#)] discusses the way in which ATM switches may be used as Label Switching Routers. The MPLS ATM VC switching document [[ATMVC](#)] details the way in which ATM-LSRs are used, specifically in the presence of non-merge and VC-merge capable switches. This document is a companion document to [[ATMVC](#)]; it extends the ATM procedures for the support of VP-switching on ATM-LSRs. It is assumed the reader is familiar with the contents of [[ATMVC](#)].

1. Introduction

The MPLS Architecture [[ARCH](#)] discusses the way in which ATM switches may be used as Label Switching Routers. The MPLS ATM VC Switching [[ATMVC](#)] document details the way in which ATM-LSRs are used, specifically in the presence of non-merge and VC-merge capable switches. This document extends the ATM procedures in [[ATMVC](#)] for supporting the distribution of labels as VPI values. This document assumes the reader is familiar with the ATM terms and procedures as defined in [[ATMVC](#)].

To connect all ingresses to all the egresses in a switched network requires $O(n^2)$ Label Switched Paths (LSPs). In order to scale to $O(n)$ (rather than $O(n^2)$), MPLS makes use of the concept of stream merge. This allows the creation of multipoint-to-point LSP trees, in which multiple incoming streams with the same FEC may be combined into a single outgoing stream. To merge multiple incoming VCs to a single outgoing VC requires that ATM switching hardware have the capability of preventing cell interleaving [[ARCH](#)]. However, much of the legacy ATM switching hardware cannot support VC merging.

Several legacy ATM switches support what is known as "VP-switching". In ATM switches that support VP-switching, it is possible to create connections (cross-connects) that use only the VPI field in the ATM header to switch cells, and the VCI field is not used in the lookup that determines the outgoing label and/or the outgoing interface. This capability is termed as VP-merge. VP-merge is the process by which a switch receives cells on several incoming VPIs and transmits them on a single outgoing VPI. In this merging style, each ingress LSR in a merged multipoint-to-point LSP use a unique VCI value when injecting packets (cells) into the multipoint-to-point connection. At the merging nodes, where multiple upstream VPs are merged into a single outgoing VP, cells from different sources get interleaved, but the unique VCI

values used by the sources help to maintain the identity of all PDUs so that they may be properly reassembled at the egress node.

Another use of VP-switching is in label stacking [[LBLSTK](#)]. The VPI/VCI label pair in ATM can be likened to be two separate labels, such that, the lower label is VCI and the top label is VPI. The cells are switched based on the top label (VPI; that is VP-switching), till the top label is popped (that is, the VP switching ends), and the cells are switched based on the bottom label (that is VCI; or VPI/VCI together).

Yet another use of VP-switching is in tunneling via ATM networks that aren't MPLS capable. This procedure is described in [[ATMVC](#)], and is repeated in this document for completeness.

This document provides procedures for exchanging VPIs as labels to create LSPs for use in the situations described above.

2. 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](#) [[RFC2119](#)].

3. Definitions

The following definitions are from [[ATMVC](#)], with appropriate modifications for the inclusion of VP-switching:

A Label Switching Router (LSR) is a device which implements the label switching control and forwarding components described in [[ARCH](#)].

A label switching controlled ATM (LC-ATM) interface is an ATM interface controlled by the label switching control component. When a packet traversing such an interface is received, it is treated as a labeled packet. The packet

```
--0__=HMHuhIMHG1DD2XKsmZDmsmAiX7Ua0k9QxAH1mQKJSw69tdNnfPmPkdrq
Content-type: text/plain; charset=iso-8859-1
Content-Disposition: inline
Content-Transfer-Encoding: quoted-printable
```

=C6s top label is inferred

from the contents of the VCI field only, the VPI field only, or the combined contents of the VPI and VCI fields. Any two LDP

peers which are connected via an LC-ATM interface will use LDP negotiations to determine which of these cases is applicable to that interface. Moreover, only one of the above cases can be active at a time on an LC-ATM interface.

An ATM-LSR is a LSR with a number of LC-ATM interfaces which forwards cells between these interfaces using labels carried in the VCI, VPI, or VPI/VCI field.

4. Use of VPI/VCIs

Label switching is accomplished by associating labels with Forwarding Equivalence Classes, and using the label value to forward packets. An ATM-LSR that is performing VP switching carries the label in the VPI field.

We say that two LSRs are "directly connected" over an LC-ATM interface if all cells transmitted out that interface by one LSR will reach the other, and there are no ATM switches between the two LSRs.

When two LSRs are directly connected via an LC-ATM interface, they jointly control the allocation of VPIs/VCIs on the interface connecting them. They may agree to use the VCI, VPI, or VPI/VCI field to encode a single label.

The default VPI/VCI value for the non-MPLS connection is VPI 0, VCI 32. Other values can be configured, as long as both parties are aware of the configured value.

The VPI value 0 MUST NOT be used with VP switching.

Any VCI value in the range 0-65535 inclusive MAY be used as a label with VP switching.

With the exception of these reserved values, the VPI/VCI values used in the two directions of the link MAY be treated as independent spaces.

The allowable ranges of VPIs and VCIs are communicated through LDP.

4.1. VP-merge Connection

Directly connected ATM-LSRs may use the VPI field to carry the MPLS label. The choice of this mode of operation and the valid VPI ranges must be negotiated via LDP. However, the VPI value of 0 MUST NOT be used as an MPLS label indicator. In addition, if two LDP peers using VP-merge are connected via a LC-ATM interface, the VPI value 0 is reserved for non-MPLS connections. This non-MPLS connection is used to carry LDP packets between the two peers, and MAY also be used (but is not required to be used) for other unlabeled packets (such as OSPF packets, etc.). The LLC/SNAP encapsulation of [RFC 1483](#) [[RFC1483](#)] MUST be used on the non-MPLS connection.

[4.2.](#) Connections via an ATM Virtual Path

Feldman et al

Expires August 1999

[Page 4]

Sometimes it can be useful to treat two LSRs as adjacent (in some LSP) across an LC-ATM interface, even though the connection between them is made through an ATM "cloud" via an ATM Virtual Path. In this case, the VPI field is not available to MPLS, and the label MUST be encoded entirely within the VCI field.

In this case, the default VCI value of the non-MPLS connection between the LSRs is 32. The VPI is set to whatever is required to make use of the Virtual Path.

A VPI/VCI value whose VCI part is in the range 0-32 inclusive MUST NOT be used as the encoding of a label.

With the exception of these reserved values, the VPI/VCI values used in the two directions of the link MAY be treated as independent spaces.

The allowable ranges of VPI/VCIs are communicated through LDP. If more than one VPI is used for label switching, the allowable range of VCIs may be different for each VPI, and each range is communicated through LDP.

5. Label Distribution and Maintenance Procedures

In general, an ATM-LSR configured for VP-merge MAY use either "ordered control" OR "independent control" label distribution, as well as "unsolicited downstream" OR "downstream-on-demand" label allocation. (see [[ARCH](#), [LDP](#)]).

Due to the limited number of VPs that may be created in a VP-merge environment, there may be a shortage of VPs if the number of IP-prefixes in the domain is greater than the maximum number of VPs available (as determined by the allocated VPI bits). This can be mitigated in the following ways:

- All ATM-LSRs in a domain may be configured for BGP or link state egress aggregation. In this case, a single LSP is sustained for all prefixes which share a common egress point, thereby significantly reducing the number of necessary VPs.
- All ATM-LSRs in a domain may be configured for "ordered control" label distribution and "unsolicited downstream" label allocation.

This allows the egress nodes to control which LSPs are created in the domain, thereby permitting a limited VP space to be allocated wisely.

- All ATM-LSRs in a domain SHOULD be configured for Conservative Label Retention (see [[ARCH](#)]). This eliminates the use and maintenance of unnecessary labels.

5.1. VP-merge-capable ATM Switches

VP-merge is the process by which a switch receives cells on several incoming VPis and transmits them on a single outgoing VPi. The ATM-LSRs neither look into the VCI field for switching the cells nor modify its content on the outgoing interface. VP-merge MAY be used only when the ENTIRE ATM-LSR domain is capable of VP-merge. If the entire ATM-LSR domain is not capable of VP-merge, then one of the applicable procedures described for either non-merge or VC-merge must be used (see [ATMVC]).

To use VP-merge in an ATM-LSR domain, each of the LSRs in the Edge Set MUST be assigned a unique VCI value. VP-Merge capable Ingress Edge LSRs MUST provide a means of configuring a unique VCI value. While other unique VCI allocation mechanisms may be possible, a description of such mechanisms is outside the scope of this document. The egress must be cognizant of the VCI ranges allocated at the ingresses, so that it aware of the VCI values it must reassemble on.

5.2. Limitations of VP-merge

VP-merge, besides its usefulness, comes with several limitations. These limitations must be carefully considered before deploying a network that uses VP-merge.

All ATM-LSRs in an ATM-LSR domain must be capable of VP-merge. That is, VP-merge does not interoperate with VC-merge or non-merge ATM-LSRs.

- There should be sufficient number of VPI bits supported by the ATM-LSRs in use. If hybrid mode (ships in the night mode) is used, some LC-ATM can have only a maximum of 8-bits for VPI.
- LSRs in an ATM-LSR domain using VP-merge must be directly connected. That is, the LC-ATM interface of a VP-merge ATM-LSR must directly connect to another VP-merge ATM-LSR without any intervening ATM switches. Therefore, two VP-merge domains cannot be connected via a public ATM network except through the IP layer.

- To use VP-merge, each Ingress Edge ATM-LSR MUST be assigned a unique VCI value. Currently, the only specified method to assign these VCI values is via configuration.

6. Egress aggregation with VP-merge

Aggregation is the concept of merging multiple FECs with a common destination into a single FEC, sharing the same LSP. Aggregation may be achieved through information available in BGP or link-state routing protocols, where a set of IP destinations are known to traverse the same IP path to a common egress point. However, some route protocols, such as RIP, do not provide knowledge of the egress point. In these situations, if egress aggregation is desired, the egress LSR must be configured to advertise the same label for all (or a set of) IP prefixes reachable through that LSR.

In the case of aggregation via BGP or a link-state routing protocol, the egress node is advertised in the FEC TLV "IP Prefix" with a mask of /32 [[LDP](#)]. In the other form of egress aggregation,

hence referred to as "generic" egress aggregation, each prefix may be advertised individually as a separate FEC via the TLV "IP Prefix", yet given the same label. However, when VP-merging is used, this latter form of generic egress aggregation could cause cell interleaving problems if each destination does not follow the exact same IP-path through the network. Thus, an LSR desiring generic egress aggregation must use the FEC Aggregation-List TLV (see [section 6.1](#)).

In the Aggregation-List TLV, each set of IP prefixes which are to share a common LSP are distinguished by a unique identifier. The aggregating egress LSR transmits a Mapping Message with an Aggregation-List TLV containing a unique identifier, commonly the router-id of the egress LSR, as well as a list of IP prefixes that are to share the same label (as found in the Label TLV; see [[LDP](#)]). Each LSR which receives the Aggregation-List from a peer verifies in the Forwarding Information Base (FIB) that each IP prefix in the given list uses that peer as its nexthop; any IP prefix which fails the test MUST be pruned off of the aggregation list. The LSR then forwards the remaining list to its upstream neighbors.

The Aggregation-List is tightly coupled with its peer and assigned label. Once it has been received and accepted, any further Mapping Message containing that aggregation identifier MUST be received from the same peer, and MUST contain the same label. An LSR MUST NOT accept an Aggregation-List with the same unique identifier from any other peer, or with a different label, until the original Aggregation-List has been cancelled via the Withdraw Message [[LDP](#)]. By constraining the Aggregation-List to associate with only one peer and label, it guarantees that all IP prefixes

to a common egress LSR will share the same LSP, and thus avoids any cell interleaving problems.

Note that a Withdraw Message containing an Aggregation-List TLV with zero prefixes is treated as a "wildcard", and removes ALL

Prefix Len

Feldman et al

Expires August 1999

[Page 8]

One octet unsigned integer containing the length in bits of the address prefix that follows.

Prefix

A variable length field containing an address prefix whose length, in bits, was specified in the previous (Prefix Len) field. The last prefix in this TLV must be padded to a byte boundary.

7. Stacking

ATM switches typically support the use of the VPI and VCI fields as a two level hierarchy. This capability may be used to support label stacking in MPLS. In a VCI/VPI label stacking scenario, aggregation happens as LSPs in VCs in the outer cloud are multiplexed onto VPs to be sent across the inner cloud. When a VP reaches its egress from the inner cloud it is demultiplexed and the VC LSPs are segregated for distribution to their destinations. In this stacking model the VP LSPs are point-to-point and VC merge may be deployed in the outer cloud.

Procedures for using VP switching in this mode will be discussed in a future revision.

8. Encapsulation

The MPLS encapsulation to be used when sending labeled packets to or from ATM-LSRs is as specified in [[ATMVC](#)].

9. Optional Loop Detection: Distributing Path Vectors

The optional loop detection procedures are as specified in [[ATMVC](#)]. Rules specific to VC-merge are also applicable to VP-merge.

10. TTL Manipulation

The TTL handling procedures are as specified in [[ATMVC](#)].

11. Security Considerations

The use of the procedures and encapsulations specified in this document does not have any security impact other than that which

may generally be present in the use of any MPLS procedures or

encapsulations (see [[ARCH](#)]).

12. Intellectual Property Considerations

The IETF has been notified of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information consult the online list of claimed rights.

13. Acknowledgements

This document shamelessly borrows from the ATM companion document [[ATMVC](#)].

14. References

- [ARCH] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", Work in Progress, February 1999.
- [ATMVC] B. Davie, J. Lawrence, K. McCloghrie, Y. Rekhter, E. Rosen, G. Swallow, P. Doolan, "MPLS using ATM VC Switching", Work in Progress, November, 1998.
- [LBLSTK] E. Rosen, Y. Rekhter, D. Tappan, D. Farinacci, G. Fedorkow, T. Li, A. Conta, "MPLS Label Stack Encoding", Work in Progress, September 1998.
- [LDP] L. Anderson, P. Doolan, N. Feldman, A. Fredette, B. Thomas, "LDP Specification", Work in Progress, January 1999.
- [RFC1483] J. Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5", [RFC 1483](#), Telecom Finland, July 1993.
- [RFC1700] J. Reynolds, J. Postel, "Assigned Numbers", October 1994.
- [RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", [RFC 2119](#), [BCP 14](#), March 1997.

15. Authors' Addresses

Nancy Feldman

IBM Corp.

Feldman et al

Expires August 1999

[Page 10]

30 Saw Mill River Rd.
Hawthorne, NY 10532
Phone: +1 914-784-3254
Email: nkf@us.ibm.com

Bilel Jamoussi
Nortel Networks
P O Box 3511 Station C
Ottawa, ON K1Y 4H7
Canada
Phone: +1 613 765-4814
Email: jamoussi@NortelNetworks.com

Sridhar Komandur
Ascend Communications
1 Robbins Rd
Westford, MA 01886
Phone: +1 978-952-7164
Email: sridhar.komandur@ascend.com

Arun Viswanathan
Lucent Technologies
101 Crawford Cornet Rd.
Holmdel, NJ 07733
Phone: +1 732-332-5163
Email: arunv@lucent.com

Tom Worster
GDC
199 W Newton St.
Boston, MA 02116 USA
Phone: +1 617 247 2624
Email: tom.worster@gdc.com

