PAPER *Special Issue on Internet Technology and its Applications*

VCID Notification for Label Switching

Kenichi NAGAMI[†], *Nonmember*, **Hiroshi ESAKI**[††], *Member*, *and* **Noritoshi DEMIZU**[†††], *Nonmember*

SUMMARY Label switching technology enables high performance and flexible layer-3 packet forwarding using the fixed length label information mapped to the corresponding layer-3 packet stream. A Label Switching Router (LSR), applying label switching technology, forwards layer-3 packets either by their layer-3 address information or by their label information mapped to the layer-3 address information.

In order to apply label switching technology to ATM links, we have introduced a new identifier called a Virtual Connection ID (VCID). The VCID value is shared between the neighboring LSR nodes.

This paper proposes a VCID notification procedure and evaluates the performance of LSP (Label Switching Path) establishment with the VCID notification procedure. The prototype system can establish 142.9 LSPs in one second with PVC operation, and 10.53 LSPs in one second with SVC operation. Of the entire LSP establishment procedure, the VCID notification procedure contributes only 3 ms, corresponding to only 3% of the entire LSP establishment procedure with SVC operation.

Evaluation using a real traffic trace between Japan and the U.S. shows that the required maximum number for LSP establishment is 17 per second. With SVC operation, the prototype system will not be able to achieve sufficient performance with regard to LSP establishment. Since the contribution by the ATM signaling is large, i.e., 88 ms (=92.6%), we need to improve the performance of ATM signaling with SVC operation.

key words: MPLS, LSR, Router, IETF, ATM

1. Introduction

Computer networks applying TCP/IP technology, i.e., the Internet and intranet, are growing both in geographical and numerical scale, as well as in traffic volume. Also, recent and upcoming applications often require a specific quality of service (QoS) or class of service (CoS) rather than the conventional best-effort services. Label switching technology will satisfy these requirements, i.e., high speed and large processing capability while providing QoS and CoS, with reasonable implementation cost. A Label Switching Router (LSR) forwards layer-3 packets using either their layer-3 address information or their label information mapped to the layer-3 address information.

Several architectures for LSRs, e.g., ARIS (Aggregate route-based IP switching) [5], CSR (Cell Switch Router) [2], IP Switch [1], and TSR (Tag Switch Router) [4], have been proposed with their own protocol designs. The protocol to establish the mappings between a layer-3 packet stream and the corresponding label is generally acalled an LDP (Label Distribution Protocol).

The existing LDP, discussed in the MPLS Working Group of the IETF, assumes that the ingress and egress label values of the link between neighboring LSRs are the same. When we use an ATM link as a link between neighboring LSRs, the label values (i.e. VPI/VCI) at both ends of the link (ATM-VC) are different in general. This is because VPI and VCI are rewritten with new values at every ATM switch node between neighboring LSRs. In order to apply label switching to the ATM link, we introduce a VCID (Virtual Connection IDendifier) [6] and VCID notification procedure. VCID is an identifier representing a particular ATM VC used for label switching, and is commonly shared between neighboring LSR nodes.

This paper specifies and evaluates the VCID notification procedure over an ATM link. However, the proposed procedure is generally applicable to any type of datalink that has different ingress and egress datalink label values.

Section 2 describes an overview of the LSR architecture and an LDP. Section 3 describes VCID procedures with various datalink configuration and operational policies. Section 4 describes the performance evaluation of LSP establishment with the proposed VCID procedure, using the prototype system. Section 5 gives a brief conclusion.

2. Overview of LSR Architecture

Label switching technology enables layer-3 packet forwarding either using the layer-2 label (e.g., VPI/VCI) or using the shim label [8]. The shim label is inserted between the layer-2 header and layer-3 header. The label is used as information to forward the layer-3 packets without the analysis of the layer-3 address (e.g., IP address). This means that the label represents the forwarding direction of the received layer-3 packets in the network. By using the label instead of the actual layer-3 address for packet forwarding, the LSR does not need to look up anything in the best match policy-based routing table, in which the search key is the layer-3 address with a variable length network prefix.

An example of packet forwarding at the LSR is

Manuscript received
Manuscript revised

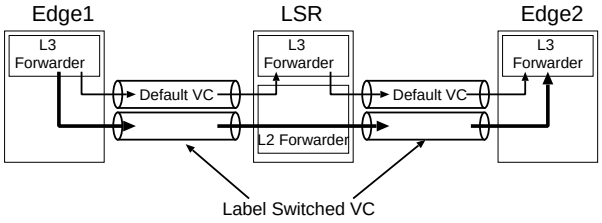

Fig. 1 LSR

shown in figure 1. The LSR has a layer-3 forwarding engine and a layer-2 forwarding engine. The LSR is connected with two neighboring LSR-edge routers (Edge1 and Edge2), as shown in the figure.

LSRs are connected with two types of VC. One is the Default VC, and the other is the Label Switched VC.

The Default VC is used for conventional hop-by-hop packet forwarding. For example, when Edge1 sends a packet to Edge2 using conventional packet forwarding, Edge1 sends the packet through the default VC. LSR receives it and transmits it to the layer-3 forwarding engine through the layer-2 forwarding engine. The layer-3 forwarding engine of the LSR looks up the destination address of the packet in the routing table, and sends it to Edge2 through the default VC.

The Label Switched VC is used for cut-through packet forwarding (i.e., label switching). For example, when Edge1 sends a packet to Edge2 using cut-through packet forwarding, Edge1 sends the packet through the label switched VC. LSR receives it and forwards it using only the layer-2 forwarding engine to send it through the Label Switched VC. The LSR forwards these packets faster than in conventional hop-by-hop forwarding because the LSR does not need to look up the layer-3 packet. The conjunction of the Label Switched VC is called as the Label Switched Path (LSP).

The LSR must establish the mapping table between the layer-3 packet stream and the corresponding label in order to achieve cut-through packet forwarding (label switching). The Label Distributed Protocol (LDP) establishes the mapping between the layer-3 packet stream and the corresponding label.

One of the major implementations of the LSR is the ATM Label Switching Router (ATM-LSR). The ATM-LSR system uses the ATM cell as a transmission layer-2 unit, and uses the VPI/VCI as a label for label switching. Therefore, the ATM-LSR can contain the cell switch module within itself in order to perform cut-through packet forwarding using the VPI/VCI label.

The existing LDPs, as discussed in the MPLS Working Group of the IETF, assumes that the ingress and egress label values of the link between neighboring LSRs are the same. ATM-LSR systems should be able to use an ATM link to interconnect the ATM-LSRs. When we use an ATM link as a link between neighboring LSRs, the label values (i.e. VPI/VCI) at both ends of the link (ATM-VC) are different in general. This is because the VPI and VCI are rewritten with new values at every ATM switch node between neighboring LSRs.

In order to apply label switching to the ATM link, we introduce a VCID (Virtual Connection IDendifier) and VCID notification procedure. VCID is an identifier representing a particular ATM VC used for label switching, and is commonly shared between neighboring LSR nodes. The follwing section describes the detailed VCID procedures.

3. VCID Procedures

3.1 Overview of VCID Notification Procedures

The ATM-LSR system uses several types of layer-2 (transparent point-to-point link/VP/PVC/SVC) and several VCID notification procedures. Here, a transparent point-to-point link is defined as a link that has the same VPI/VCI label at both ends of a VC. For example, two nodes are directly connected (i.e., without intervening ATM switches) or are connected through a VP with the same VPI value at both ends of the VP.

There are two types of VCID notification procedures: inband and out of band. In the inband procedure, the VCID notification procedure control messages are transferred over the label switched VC. In the out of band procedures, the VCID notification procedure control messages are transferred over some other VC (other than the label switched VC), such as over an ATM signaling VC or over the default VC.

We summarize below the various types of links and the applicable VCID notification procedures. The detailed VCID notification procedure is described in the next subsection.

- Transparent point-to-point link: no VCID notification
  VCID notification procedure is not necessary because the label (i.e., VPI/VCI) is the same at each end of the VC.
- VP:
  - Inband notification
    VCID notification is needed because the VPI values at each end of the VC may not be the same. Inband VCID notification is applied in this configuration
  - No notification
    If a node has only one VP to a neighboring node, the VCID notification procedure is not mandatory. The VCI can be used as the VCID. This is because the VCI value is the same at each end of the VP.

Type of VC	no VCID	inband	outband (small)	outband
Point-to-point	Yes	–	–	–
VP	Yes	Yes	–	–
PVC	–	Yes	–	–
SVC	–	Yes	Yes	Yes

Fig. 2 VCID notification procedures and link configuration types

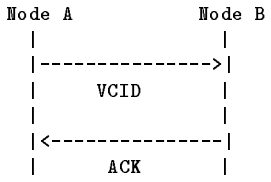

Fig. 3 Sequence of message exchange in inband notification with point-to-point link

- PVC: inband notification
  Inband VCID notification is applied in this configuration because the labels at each end of the VC may not be the same.
- SVC: there are three possibilities
  - Out of band notification
    When an ATM signaling message has a field which is large enough to carry a VCID value (e.g., GIT [7]), then the VCID is carried directly in that field.
  - Outband notification using a small-sized field
    When an ATM signaling message has a field which is not large enough to carry a VCID value, this procedure is applied.
  - Inband notification
    When an ATM signaling message can not carry user information, this procedure is applied.

Figure 2 shows a summary of applicable VCID notification procedures for each type of link configuration.

3.2 Inband Notification Procedures

3.2.1 Inband Notification with Point-to-Point VC

VCID notification is performed by the exchange of control messages through the point-to-point label switched VC that is newly used as an LSP for the corresponding layer-3 packet stream. The label switched VC can be both an SVC and a PVC. Figure 3 shows the sequence of message exchange in the inband notification procedure. The detailed procedure for VCID notification between nodes A and B is described below.

1. Node A establishes a label switched VC for Node B. (by signaling or management)
2. Node A selects a VCID value.
3. Node A sends a VCID control message containing the selected VCID value by node A through the newly established label switched VC to Node B.
4. Node A establishes the mapping information between the outgoing label (VPI/VCI) of the new label switched VC and the selected VCID value.
5. Node B receives the VCID control message from node A through the label switched VC, and establishes the mapping information between the VCID in the message and the incoming label (VPI/VCI) of the corresponding label switched VC.
6. Node B sends a VCID ACK message to node A.
7. When Node A receives the VCID ACK message, node A and node B use the commonly shared VCID value representing the label switched VC.

3.2.2 Inband Notification with Point-to-Multipoint VC

VCID notification is performed by the exchange of control messages through the point-to-multipoint label switched VC that is newly used as an LSP for the corresponding layer-3 packet stream. This procedure is applied to multicast packet transmission service. The upstream node must assign the VCID value in order to ensure a common VCID value among leaf LSR nodes. The procedure involving the upstream node and two leaf nodes (i.e., downstream nodes) is described below. Two types of procedures are described: (1) the first downstream node joins the multicast service, (2) a new downstream node joins the existing multicast service tree. Figure 4 shows the sequence of message exchange in inband notification with point-to-multipoint VC.

First, the procedure for establishing the first VC is described. This is the case where the first leaf node joins the multicast service.

1. The upstream node assigns a VCID value to the new label switched VC.
2. The upstream node sends a VCID control message containing the VCID value and the address of the upstream node through the label switched VC used as an LSP.
3. The upstream node establishes the mapping information between the outgoing label for the label switched VC and the corresponding VCID value.
4. The downstream node receiving the VCID control

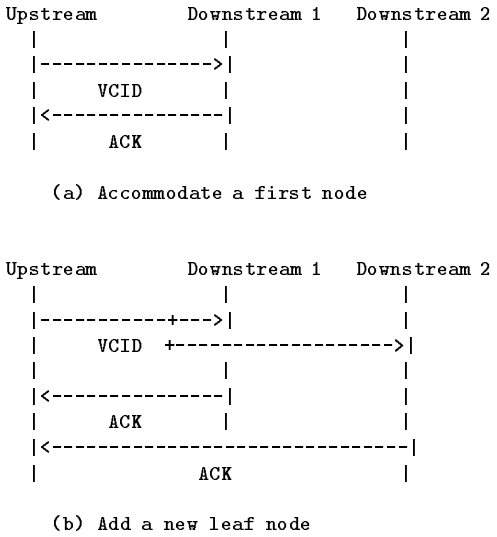

Fig. 4 Message sequence of inband notification with point-to-multipoint VC

message checks the address of the upstream node. If the address of the upstream node does not have the same network prefix as that of the downstream node, the VCID control message is discarded. Otherwise, the downstream node establishes the mapping information between the VCID in the control message and the corresponding label switched VC.

5. The downstream node sends a VCID ACK message to the upstream node.

6. When the upstream node receives the VCID ACK message, the upstream node and the downstream node share the same VCID value.

Second, the procedure for adding a leaf to the existing point-to-multipoint VC is described.

1. The upstream node adds the downstream node, using ATM signaling.

2. The VCID value which has already been assigned to the point-to-multipoint label switched VC is used.

3. The upstream node sends a VCID control message containing the VCID value and the address of the upstream node through the point-to-multipoint label switched VC used as an LSP. This VCID control message is transferred to all leaf nodes.

4. The downstream nodes receiving the VCID control message check the address of the upstream node. If the address of the upstream node does not have the same network prefix as that of the downstream node, the VCID control message is discarded. Otherwise, the downstream nodes establish the mapping information between the VCID in the control message and the corresponding point-to-multipoint label switched VC.

5. The downstream node sends a VCID ACK message to the upstream node.

6. When the upstream node receives the VCID ACK message from a newly joining downstream node, the upstream node and all downstream nodes share the same VCID value.

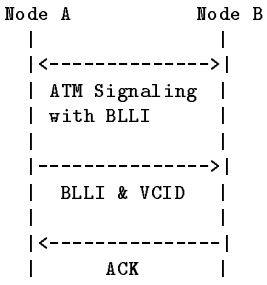

Fig. 5 Sequence of message exchange in outband using small-sized field

3.3 Outband Notification Using a Small-Sized Field

This method can be applied when a label switched VC is established using ATM signaling and the ATM signaling message has a user-to-user field which is not large enough to carry a VCID value.

The SETUP message of the ATM Forum UNI 3.1/4.0 has a 7-bit mandatory field for the use of the user. This is a user specific field in the layer-3 protocol field in the BLLI IE (Broadband Low Layer Information Information Element).

The BLLI value is used as a temporary identifier for a label switched VC during a VCID notification procedure. This mechanism is defined as "outband notification using a small-sized field". The BLLI value of a new label switched VC must not be assigned to other VCs during the procedure, to avoid identifier conflict. When the mapping among the BLLI value, a VCID value, and the corresponding label switched VC is established, the BLLI value can be released to be reused for a new label switched VC. VCID values can be assigned independently of BLLI values. Figure 5 shows a sequence of message exchange in outband notification using a small-sized field.

A point-to-multipoint label switched VC can also be established using ADD_PARTY of ATM signaling. ADD_PARTY adds a new leaf to an existing label switched point-to-multipoint VC. In this procedure, the BLLI value of ADD_PARTY must be the same as the value used to establish the first point-to-multipoint label switched VC between an upstream node and a downstream node for multicast service. The same BLLI value can be used in different label switched VCs only

when the multiple label switched point-to-multicast VC does not add a new leaf at the same time. As a result, the BLLI value used in signaling must be determined by the root node of the multicast tree.

3.3.1 Outband Notification Using a Small-Sized Field with Point-to-Point VC

This subsection describes the procedures for establishing a point-to-point label switched VC, with regard to unicast service.

1. An upstream LSR establishes a label switched VC for the downstream LSR using ATM signaling and provides a temporary value in the BLLI field that it is not currently used by any other (incomplete) VCID notification transaction regarding this particular peer.
2. The upstream LSR sends the mapping information between the BLLI and VCID values to notify the downstream LSR.
3. The downstream LSR establishes the mapping information between the label switched VC with the BLLI value and the VCID.
4. The downstream LSR sends a VCID ACK message to the upstream LSR.
5. When the upstream LSR receives the VCID ACK message, the mapping between the label switched VC and the VCID is established. The label switched VC is ready to be used for label switching. At this time, the BLLI value employed in this particular transaction is released for reuse.

3.3.2 Outband Notification Using a Small-Sized Field with Point-to-Multipoint VC

This subsection describes the procedure for establishing the first label switched VC for a multicast service tree and the procedure for adding a new label switched VC leaf to an existing label switched VC tree, including the notification of its VCID for a multicast stream using point-to-multipoint VCs.

In this procedure, an upstream LSR determines both the VCID and BLLI values for the particular multicast service. The reason that the BLLI value must be determined by an upstream LSR is described above.

First, the procedure for establishing the first label switched VC is described.

1. An upstream LSR establishes a label switched VC by ATM signaling between the downstream LSR with a unique BLLI value.
2. The upstream LSR sends the BLLI value and VCID using a dedicated message for the VCID notification procedure to notify the downstream LSR.
3. The downstream LSR establishes the mapping information between the label switched VC with the BLLI value and the VCID.
4. The downstream LSR sends a VCID ACK message to the upstream LSR.
5. When the upstream LSR receives the VCID ACK message, the label switched VC is ready to be used and the BLLI value can be released for another label switched VC.

Second, the procedure for adding a leaf to the existing point-to-multipoint VC is described.

1. The upstream LSR establishes a label switched VC by ATM signaling toward the newly added downstream LSR with the BLLI value that was used during the first signaling procedure for the multicast service that the downstream node is joining. If another label switched VC is using the BLLI value at the same time, the upstream waits for the completion of the signaling procedure that is using this particular BLLI value.
2. Go to step 2 of the procedure for the first label switched VC.

3.4 Outband Notification

This method can be applied when a VC is established using an ATM signaling message and the message has a field (e.g., GIT [7]) which is large enough to carry a VCID value. Figure 6 shows a message sequence of an outband notification. The message format is described in [7].

A point-to-multipoint label switched VC can also be established using ADD_PARTY of ATM signaling. ADD_PARTY adds a new label switched VC leaf to an existing label switched VC. The VCID is carried in a GIT field in ADD_PARTY.

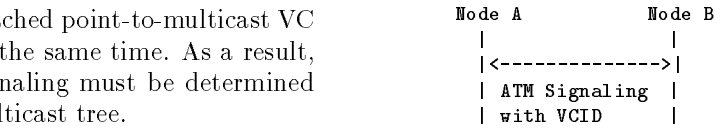

Fig. 6 Message sequence of outband notification

4. Performance Evaluation of VCID Procedure

In this section, we evaluate the performance of LSP establishment, including a label distribution and a VCID notification procedure.

4.1 CSR Prototype System

We implement the VCID notification procedure on the CSR system [2], which is one implementation of LSR. The CSR node consists of a Pentium 166-MHz PC and a Fore Systems ASX-200WG ATM switch. The VCID notification procedure and FANP (Flow Attribute Notification Protocol) [3], which is one label distribution protocol, are developed on the PC.

VCID notification and FANP software modules are written in the C language. The VCID notification module consists of 457 lines. The FANP module consists of 6742 lines. With regard to the number of lines of source code in the software module, the VCID notification module is only about 7% the size of the FANP module. From the viewpoint of source code volume, the contribution of the proposed VCID notification procedures is sufficiently small.

4.2 LSP Establishment Performance

The network topology for an evaluation is shown in Fig. 7. R1 and R2 are edge nodes of the CSR. They are Pentium 166-MHz PCs. They also include the VCID notification procedure and the FANP. The CSR and edge nodes are connected with an ATM switch, which is a Toshiba AX-1500 ATM switch.

We evaluate the performance of inband VCID notification using an ATM SVC, in which R1 establishes an LSP to R2. LSP establishment consists of three procedures: ATM signaling, VCID notification, and label distribution.

1. ATM signaling procedure
   When R1 wants to establish an LSP to R2, R1 establishes a label switched VC to the CSR using ATM signaling.

2. VCID notification procedure
   Next, R1 sends a VCID inband message to the CSR. The CSR sends a VCID ACK message to R1.

3. Label distribution procedure
   R1 sends a label distribution message to the CSR. The CSR sends a label distribution ACK message to R1.

We next evaluate the required processing time for each procedure. Figure 8 shows the required processing time for each procedure. ATM signaling requires 88 ms, the VCID notification procedure requires 3 ms, and the label distribution procedure requires 4 ms. In total, LSP establishment requires 95 ms. If R1 does not use VCID notification, LSP establishment requires 92 ms. The required LSP establishment processing time with the proposed VCID notification is only 3% longer than that without the VCID notification.

Fig. 7 Network topology

Procedure	Time
ATM signaling	88 ms
VCID Notification	3 ms
Label Distribution	4 ms

Fig. 8 Evaluation for LSP setup time

As a result, in an SVC environment, 10.52 LSPs can be established per one second. In a PVC environment, 142.9 LSPs can be established per one second, since we do not need the ATM signaling that consumes 88 ms (92.6%) of the processing time.

4.3 Performance Evaluation for Actual Network Environment

We estimate the required LSP establishment performance using a real traffic trace. The traffic trace used for this evaluation is the traffic at the Ethernet segment which is located between the WIDE (Widely Integrated and Distributed Environment) project internet backbone network in Japan and the U.S. The monitoring host traces and records the traffic using the tcpdump application on the monitoring host. Measurements were obtained for about 2 hours (7437 seconds), from 14:08 JST November 4, 1997, to 16:11 JST November 4, 1997. The traffic traces are recorded, separately, for each direction.

The parameters applied in the estimation are assumed as follows.

- The granularity of the layer-3 packet stream is a host-pair. Therefore, layer-3 packets which have the same IP source address and destination IP address are sent through the same LSP.

- The label-establishment trigger is invoked when the HTTP/FTP/TELNET/NNTP packets are sent out.

- The label-released trigger is invoked when no packet is sent out during the Disconnect Timer Interval. The value of the Disconnect Timer Interval is 60 seconds.

Using the traffic trace from Japan to the U.S., the required LSP establishment performance is 8.86 [LSP/sec] on average and 17 [LSP/sec] maximum. Using the traffic trace from the U.S. to Japan, the required LSP establishment performance is 2.75 [LSP/sec] on average and 12 [LSP/sec] maximum.

In the PVC environment, the current implementation will be applicable to actual networks. This

is because the prototype implementation can achieve 142.9 [LSP/sec] LSP establishment performance. In the SVC environment, the current prototype implementation can be applied with average performance, but can not be applied with maximum performance. The major factor responsible for insufficient LSP establishment performance is obviously the ATM signaling. When we apply label switching to actual networks with an SVC environment, the performance of ATM signaling must be improved.

5. Conclusion

This paper proposes the VCID notification procedures that are required to apply label switching technology using ATM links to various ATM based datalink configurations. Point-to-point links and single VP links do not require the VCID notification procedure, but multiple VP, PVC, and SVC links do. Multiple VPs and PVCs perform inband VCID notification. The SVC allows three VCID notification procedures.

In this paper, we alse evaluate the performance of LSP establishment with the VCID notification procedure. In prototype system implementation, the VCID notification procedure consumes 3 ms. The performance of LSP establishment is evaluated with the required LSP establishment performance estimated by a real traffic trace in an operating network.

- PVC environment
  In the PVC environment, though the VCID notification procedure accounts for about 50% of the processing time, the LSP establishment processing time is only 7 ms. The prototype system can establish 142.9 LSPs per second and the actual network requires 17 LSPs per second. Therefore, the prototype system can provide a sufficient LSP establishment performance.

- SVC environment
  In the SVC environment, the VCID notification procedure accounts for only 3% of the processing time, since the major contributor is ATM signaling (88 ms, 92.6%). The prototype system can establish 10.52 LSPs per second, which is less than the 17 LSPs per second required by the actual network. Therefore, the prototype system is not able to achieve sufficient LSP establishment performance, and the performance of ATM signaling must be improved with SVC operation.

References

[1] P. Newman, T. Lyan, G. Minshall, "Flow labelled: Connectionless ATM Under IP", Engineer Conference, Networld+Interop'96 Las Vegas, April, 1996.

[2] Y. Katsube, K. Nagami, H. Esaki, "Toshiba's Router Architecture Extensions for ATM : Overview", IETF RFC2098, Feb. , 1997.

[3] K. Nagami, Y. Katsube, Y. Shobatake, A. Mogi, S. Matsuzawa, T. Jinmei, H. Esaki, "Toshiba's Flow Attribute Notification Protocol (FANP)", April, 1997.

[4] Y. Rekhter, B. Davie, D. Katz, E. Rosen, G. Swallow, "Cisco Systems's Tag Switching Architecture Overview", IETF RFC2105, Feb. , 1997.

[5] A. Viswanthan, N. Feldman, R. Boivie, R. Woundy, "ARIS: Aggregate route-based IP switching", IETF Internet-Draft, March, 1997.

[6] K. Nagami, N. Demizu, H. Esaki, P. Doolan, "VCID Notification over ATM link", IETF Internet-Draft, Aug., 1998.

[7] M. Suzuki, "The Assignment of the Information Field and Protocol Identifier in the Q.2941 Generic Identifier and Q.2957 User-to-user Signaling for the Internet Protocol", IETF Internet-Draft, June, 1998

[8] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", IETF Internet-Draft, March, 1998.