

MPLS AND GMPLS NETWORKING CONTROL AND
MANAGEMENT TECHNOLOGIES

By

PRAVIN RASIAH

Bachelor of Technology

Cochin University of Science and Technology

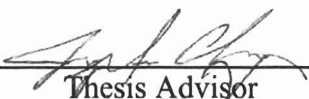
Cochin, India

1999

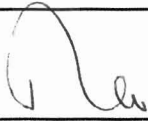

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2002


MPLS AND GMPLS NETWORKING CONTROL AND
MANAGEMENT TECHNOLOGIES

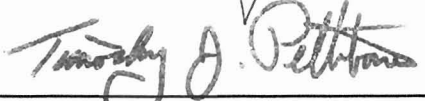
Thesis Approved:



Thesis Advisor





Dean of the Graduate College

PREFACE

Numerous networking architectures have evolved in order to provide large bandwidth and high-speed communication. MultiProtocol Label Switching (MPLS) is an Internet Engineering Task Force (IETF) specified Wide Area Network (WAN) framework that uses routing and forwarding of traffic flows applying Traffic Engineering (TE) features. Generalized MultiProtocol Label Switching (GMPLS) is an extension of MPLS that supports label switching for a variety of networks in addition to packet switched networks. This thesis provides major contributions to Operation, Administration and Maintenance (OAM) implementations in GMPLS and multicasting services through MPLS networking.

The first aspect of research in this thesis is based on both the control and user plane implementation of OAM processes. Extensions to the Resource ReSerVation Protocol with Traffic Engineering extensions (RSVP-TE) [8] and the Label Distribution Protocol (LDP) [21] as well as the user plane implementations of the OAM packets using the OAM Label Alert [13] have been proposed. These extensions will enable GMPLS [11] networking to conduct all OAM functionalities that Asynchronous Transfer Mode (ATM), Frame Relay (FR), Synchronous Optical NETwork (SONET), Synchronous Digital Hierarchy (SDH), Optical Transport Network (OTN), MPLS, etc. are capable of and thereby provide common OAM control plane operations throughout the GMPLS network.

Secondly, extensions to the MPLS signaling protocols such as RSVP-TE [8] and LDP [21] to support MPLS network multicasting functionalities have been proposed. The concepts presented in this paper support the delivery of multicasting traffic in accordance to the TE features provided in the MPLS specifications [12]. These extensions to the signaling protocols will enable MPLS networking to conduct all required multicasting features that Internet Protocol (IP) multicasting protocols (e.g., Distance Vector Multicasting Routing Protocol (DVMRP), Multicasting Open Shortest Path First (MOSPF), Protocol Independent Multicasting – Dense Mode (PIM-DM), Protocol Independent Multicasting – Sparse Mode (PIM-SM), and Core Based Tree (CBT)) are capable of, while adding on the feature benefits of MPLS TE.

ACKNOWLEDGMENTS

I would like to thank God Almighty for guiding me throughout my life and helping me become a better individual.

I wish to express my sincere appreciation and gratitude to my advisor Dr. Jong-Moon Chung for his supervision, constructive guidance, inspiration, and friendship. I owe my valuable experience and knowledge to his inspiring insight and patience. I would like to take this opportunity also to thank Dr. R. K. Yarlagadda and Dr. Ramakumar for their invaluable support, encouragement and guidance throughout my Master's program here at the Oklahoma State University.

I would specially like to thank Mauricio A. Subieta Benito, Grace Y. Cho, Harleen Chhabra, and Hooi Miin Soo of the Advanced Communication Systems Engineering Laboratories (ACSEL) and Oklahoma Communication Laboratory for Networking and Bioengineering (OCLNB) for their contribution and help for the research developed in Chapter 6.

I would like to thank my family and friends for their patience and support. Last but not the least; special thanks to my parents and my sister for their love and prayers.

TABLE OF CONTENTS

Chapter	Page
1. INTRODUCTION AND THESIS ORGANIZATION.....	1
1.1. Problem Statement and Justification for Research.....	2
1.2. Thesis Organization.....	5
2. INTRODUCTION TO MPLS AND GMPLS.....	6
2.1. Introduction to MPLS.....	6
2.2. IP Forwarding vs. MPLS Forwarding.....	7
2.3. Advantages of Label Forwarding Techniques Over IP Forwarding.....	9
2.4. LERs and LSRs.....	10
2.5. Forwarding Equivalence Class.....	10
2.6. MPLS Label Switched Paths.....	11
2.7. MPLS Operation.....	12
2.8. MPLS Evolution to GMPLS.....	13
2.9. LSP Establishment in GMPLS Network.....	14
3. INTRODUCTION TO OAM.....	15
3.1. ATM based OAM Techniques.....	15
3.1.1. ATM Layer Operational Flows.....	16
3.1.2. VPC/VCC Performance Monitoring.....	17
3.1.3. VPC/VCC Failure Reporting.....	18
3.1.4. VPC/VCC Continuity Check and Loopback Testing.....	19
3.1.5. Traffic Management Functions.....	19
3.1.6. ATM Layer QoS Measurement.....	20
3.1.7. ATM OAM Cell Format.....	21
3.2. Frame Relay OAM Techniques.....	22
3.2.1. Frame Relay OAM Message Format.....	22
3.3. SONET/SDH OAM Techniques.....	25
3.3.1. OAM Support in Section Overhead.....	25
3.3.2. OAM Support in Line Overhead.....	26
3.3.3. OAM Support in STS Path Overhead.....	26
3.3.4. OAM Support in VT Path Overhead.....	27
3.3.5. SONET Alarms for Fault Notification.....	27
3.4. OTN OAM Techniques.....	28

3.4.1. Representation of Optical Network.....	28
3.4.2. Operation and Maintenance Concept.....	29
3.4.2.1. Performance Monitoring.....	29
3.4.2.2. Failure Detection.....	29
3.4.2.3. Failure Information and Fault Localization.....	30
4. INTRODUCTION TO IP MULTICASTING ROUTING PROTOCOLS.....	31
4.1. Multicasting Routing Protocol.....	31
4.1.1. Flooding.....	32
4.1.2. Spanning Tree.....	33
4.1.3. Reverse Path Broadcasting.....	33
4.1.3.1. Example of RPB.....	34
4.1.4. Truncated Reverse Path Broadcasting.....	35
4.1.5. Reverse Path Multicasting.....	36
4.1.5.1. Procedure for Operation of RPM.....	36
4.1.6. Steiner Tree.....	37
4.2. Review of Existing IP Multicast Routing Protocols.....	39
4.2.1. Multicast Extensions for Open Shortest Path First.....	39
4.2.2. Distance Vector Multicasting Routing Protocol.....	40
4.2.3. Protocol Independent Multicast- Dense Mode.....	41
4.2.4. Protocol Independent Multicast- Sparse Mode.....	42
4.2.5. Core Based Tree.....	43
5. OAM IMPLEMENTATION IN GMPLS.....	45
5.1. Extensions to GMPLS Control Plane to Include OAM Functionalities.....	47
5.1.1. Extensions to RSVP-TE for Enabling GMPLS OAM Functionalities.....	48
5.1.2. Extension to LDP for Enabling GMPLS OAM Functionalities.....	50
5.1.3. GMPLS Control Plane OAM Functionalities and Procedures.....	51
5.1.3.1. Fault Management Procedures in GMPLS.....	52
5.1.3.2. Loopback Procedures in GMPLS.....	52
5.1.3.3. Activation/Deactivation of PM and CC in GMPLS.....	53
5.2. Extension to GMPLS User Plane to Include OAM Functionalities.....	54
6. MULTICASTING SERVICES THROUGH MPLS NETWORKING.....	57
6.1. Extension to RSVP-TE and LDP for Multicasting in MPLS Networks.....	57
6.1.1. RSVP-TE Extensions for Multicasting in MPLS Networks.....	58
6.1.1.1. The Multicasting Session and Tree Objects.....	58
6.1.1.2. RSVP-TE Multicasting Message Extensions.....	59

6.1.1.3. Multicasting Extensions to the RSVP-TE Path Resv Messages...	61
6.1.1.4. Multicasting Extensions to the Hello Message.....	62
6.1.2. LDP Extensions for Multicasting in MPLS Networks.....	62
6.1.2.1. LDP Multicasting Message.....	62
6.1.2.2. Hello Message Extensions.....	63
6.1.2.3. Notification Message Extensions.....	65
6.1.2.4. Multicast Extensions to Label Request Message.....	65
6.1.2.5. Multicast Extensions to the Label Mapping Message.....	66
6.1.3. Multicast Distribution Tree Construction Using RSVP-TE and LDP.....	66
6.1.3.1. Root-Initiated Tree Calculation.....	67
6.1.3.2. Leaf-Initiated Tree Calculation.....	69
6.1.3.3. Dynamic Updates to the Tree.....	70
7. CONCLUSION.....	73
8. REFERENCES.....	76

LIST OF FIGURES

Figure	Page
2.1 MPLS support of multiprotocols.....	7
2.2 IP forwarding.....	8
2.3 MPLS forwarding.....	9
2.4 Example for binding between label and FEC.....	11
2.5 GMPLS network-common control and management plane.....	13
2.6 Establishment of LSP in GMPLS.....	14
3.1 Performance monitoring scheme in ATM.....	17
3.2 Failure notifications in ATM.....	18
3.3 Alarm indications.....	18
3.4 Loopback for diagnosis of ATM networks.....	19
3.5 ATM OAM cell format.....	21
3.6 Frame relay header and OAM message format.....	22
3.7 OAM information field format.....	23
3.8 Example of a SONET/SDH network.....	25
3.9 Example of an OTN network.....	29
3.10 Failure information and notification.....	30
4.1 Example of RPB.....	35
4.2 Procedures for RPM.....	37
4.3 Steiner tree.....	38
4.4 Example of PIM-DM.....	41
4.5 Example of CBT.....	44
5.1 RSVP-TE OAM message format.....	48
5.2 RSVP-TE OAM object format.....	48
5.3 LDP extensions for OAM.....	51
5.4 RSVP-TE/LDP execution of AIS/RDI.....	52
5.5 LDP/RSVP-TE execution for loopback.....	53
5.6 RSVP-TE/LDP execution of activation/deactivation of PM or CC.....	54
5.7 GMPLS OAM packet.....	55
5.8 GMPLS OAM payload for performance monitoring.....	55
6.1 RSVP-TE extension objects format.....	59
6.2 RSVP-TE message extensions.....	61
6.3 LDP message extensions.....	63
6.4 Root-initiated tree calculations.....	68
6.5 Leaf-initiated tree calculations.....	70

LIST OF TABLES

Table	Page
3.1 QoS parameters for performance monitoring.....	20
3.2 Different IF types in the OAM message.....	24
5.1 OAM functionalities for GMPLS.....	49

CHAPTER I

INTRODUCTION AND THESIS ORGANIZATION

Telecommunications is a multi-billion dollar industry which has been growing in leaps and bounds over the past decade. The need for higher bandwidth and faster communication techniques has fueled the innovation of numerous networking architectures to solve this ever increasing demand. One such networking architecture which has attracted a lot of interest is Multiprotocol Label Switching commonly known as MPLS [12]. MPLS evolved out of networking architectures such as the Aggregate Route based IP Routing (ARIS) by IBM, Tag Switching by Cisco, IP Switching by Ipsilon, and Cell Switching Router by Toshiba. These network architectures have a lot in common (they employed packet forwarding by using label switching), but were implemented differently. In order to resolve this, the Internet Engineering Task Force (IETF) set up the MPLS working group to standardize the architecture.

Generalized Multiprotocol Label Switching (GMPLS) [11] was introduced as an extension to MPLS in order to support label switching not only for packet switched networks but also for time, wavelength, and space switched networks. Since GMPLS and MPLS are in their initial stages of standardization, they provide motivation for research.

The main aspect of research that this thesis deals with is the control and management of MPLS and GMPLS.

The two main control and management aspects of MPLS and GMPLS investigated in this thesis are:

- GMPLS support for Operation, Administration and Maintenance
- MPLS networking for enhanced multiplatform multicasting services

Operation, Administration and Maintenance (OAM) procedures have been used in telecommunication networks in order to provide in-service fault detection, fault localization and also performance management. ATM [14][28], FR [32], Gigabit Ethernet and MPLS [13][25] are provisioned with a set of OAM procedures in order to enable easy management and administration of their networks.

Multicasting is a process of transmitting data from one source to multiple destinations or from multiple sources to multiple destinations. The importance and applicability of multicasting applications have been recognized and numerous start up companies and industrial giants are cashing into this booming market. In order to make this possible, multicasting networks have to enable multicasting application transmission to be provisioned at the underlying networking layers.

1.1. Problem Statement and Justification for Research

The first topic of research introduced in this thesis is the GMPLS support for OAM operations. It has been studied that the current standards for GMPLS are not capable of provisioning OAM procedures for management and administration. A new protocol layer called the Link Management Protocol (LMP) has been introduced for establishing, managing, and releasing of connections between two GMPLS capable nodes

[20], but this does not cater to all diagnostic and troubleshooting techniques for GMPLS. In [23] the Notify message is proposed which is used for indicating alarms in case of degraded links. This proposal also does not address all the OAM functionalities. Therefore, it is very important that OAM standards have to be addressed for alarm initiation, fault localization and performance monitoring. In ATM [31], FR [32], and MPLS [12] [25] the user plane implementation of OAM techniques have been employed. These strategies involve transmission of OAM packets (frames or cells) along with user data. SONET, SDH and OTN perform their OAM functionalities using transfer of alarm and failure signals in their headers (path, line, or section) [33] [24]. As GMPLS provides a common control plane (signaling and routing) that is made available to these underlying switching topologies that switch packets, frames, time, and wavelengths, it gives us the advantage of using both the control and user plane for OAM support.

This research extends two aspects of the implementation of GMPLS OAM protocols and procedures. The first aspect is based on the control plane using RSVP-TE [27] [8] and LDP [22] and the other aspect is based on the user plane implementation using the OAM Label Alert [13] to indicate the OAM packet. The extensions to the control plane signaling protocols and user plane OAM packets proposed in this thesis will enable GMPLS networking to conduct all the required OAM functionalities that other networking architectures such as ATM, FR, MPLS, Gigabit Ethernet, SONET, and OTN are capable of and will thereby enable GMPLS to provide common maintenance and administration throughout the multiplatform network.

The second aspect of research discussed is the multicasting services deployed in MPLS. It has been studied that the current standards of MPLS are not capable of

providing MPLS based multicasting traffic services. The objective of this research is to enable MPLS networking with all the functionalities and services required for multicasting, where the addressing and management topology of current IP multicasting users and groups will be done by the Internet Group Management Protocol (IGMP), thereby, keeping the management and addressing the same as currently done in IP, but enabling the TE advantages of MPLS to exist over the multicasting network connections.

This research extends two aspects of implementation of MPLS multicast protocols and procedures based on the RSVP-TE [27] [8], and implementations based on LDP [22]. The extensions provided to RSVP-TE for the Label Switching Path (LSP) tunnel applications in RFC 3209 [8] are restricted to unicast label switched paths, and multicast LSP establishment is left for further study. Similarly, the standards for LDP [22] also lack support for point-to-multipoint connection as well as multicast services, and are left for future study. Based on the limited strategies for multicasting services for MPLS, several extensions to the RSVP-TE and LDP signaling protocols are proposed to support multicasting services. The extensions to the signaling protocols proposed in this thesis will enable MPLS networking to conduct all required multicasting features that IP multicasting protocols are capable of, while adding on the feature benefits of MPLS TE to the service operations. Through these extensions, MPLS multicasting services can be provided based on the traditional IP-based multicast routing protocols, such as, DVMRP, MOSPF, PIM-DM, PIM-SM, and CBT, or independently as a set of stand-alone MPLS multicasting operations.

1.2. Thesis Organization

MPLS and GMPLS are networking technologies that are under extensive development and their importance in future broadband networks have been identified. Chapter 2 provides an extensive review of the two technologies.

In order to understand the protocol engineering aspects of the development, this thesis also reviews all existing standards for multicasting and OAM. Chapter 3 discusses and reviews OAM and its deployment in different networks including ATM, FR, Gigabit Ethernet, SONET, SDH, and OTN. The introduction to IP multicasting topologies and a detailed review of the common multicasting routing algorithms are given in Chapter 4, Section 4.1 and Section 4.2.

Chapter 5 discusses OAM implementations in GMPLS. The extensions to the GMPLS control plane to employ OAM operations are shown in Section 5.1. Extensions for the GMPLS user plane are indicated in Section 5.2. Also the OAM functionalities and procedures in GMPLS are shown in Section 5.1.3.

A solution to Multicasting service deployment in MPLS networks is discussed in Chapter 6. This chapter also includes extensions to RSVP-TE (in Section 6.1.1), LDP (Section 6.1.2) for multicasting services. The strategies for deploying a multicasting distribution tree using RSVP-TE and LDP are provided in Section 6.1.3. The final conclusion to the research is provided in Chapter 7.

CHAPTER II

INTRODUCTION TO MPLS AND GMPLS

MPLS is an emerging network architecture that is becoming a popular choice for high speed backbone networks. It provides an overlay model to the existing IP backbone networks. MPLS uses data link techniques such as ATM, FR, etc. for label switching of packets along specific paths discovered using IP routing techniques. MPLS also provides better management of the network for connection setup and QoS control.

GMPLS is a rapidly evolving network architecture used for label switching over different underlying topologies such as ATM, FR, SDH, SONET, and WDM. This networking architecture extends the features of MPLS to the time, wavelength, and fiber switched networks. Standards proposed for GMPLS are extensions to standards proposed for MPLS.

This chapter gives an overview of the working procedures for MPLS [12] and GMPLS [11] and provides a basic understanding of the two technologies.

2.1. Introduction to MPLS

MPLS is a switching topology that uses small labels to forward packets in the MPLS domain. This forwarding mechanism makes it possible for very high-speed

communication. Packets are forwarded in the MPLS domain through LSPs, by mapping the packets to Forward Equivalence Classes (FECs) [12]. The FEC gives the next hop for the packet to the Label Switch Router (LSR). This easy and robust forwarding scheme gives MPLS its benefits.

MPLS is so called because it can be used to work with multiple protocols above and below it. The forwarding scheme employed by MPLS can be used by any one of the network layer protocols such as IP, IPX, etc., thereby making MPLS support multiple network layer protocols. Another main feature of MPLS is the enabling of label switching across any of the data link layer protocols. This makes MPLS a true multiprotocol solution. This aspect of MPLS is illustrated in Figure 2.1.

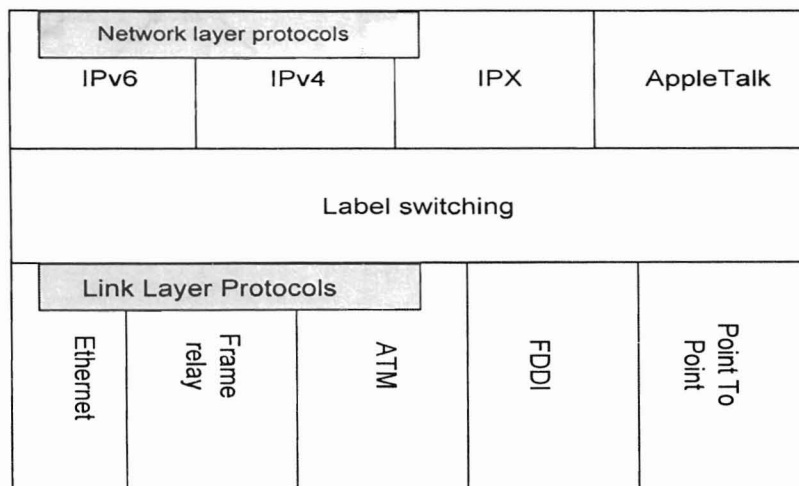


Figure 2.1 MPLS support of multiprotocols.

2.2. IP Forwarding vs. MPLS Forwarding

In connectionless networks, packets are routed by means of the routers that forward them after analyzing the IP header and performing a routing algorithm on each of the packets. This causes increased overload on the routers themselves. In IP routers, the FEC of a particular router is assigned at each hop, in other words, the FEC is determined

by checking the destination address at each hop and then running the routing algorithm to determine the shortest path to the destination and forwarding it along that path. This procedure is illustrated in Figure 2.2. In Figure 2.2, we can see that when an IP packet reaches a router (say R2), the router checks its destination IP address and finds out the shortest path to destination and forwards the packet on an outgoing interface.

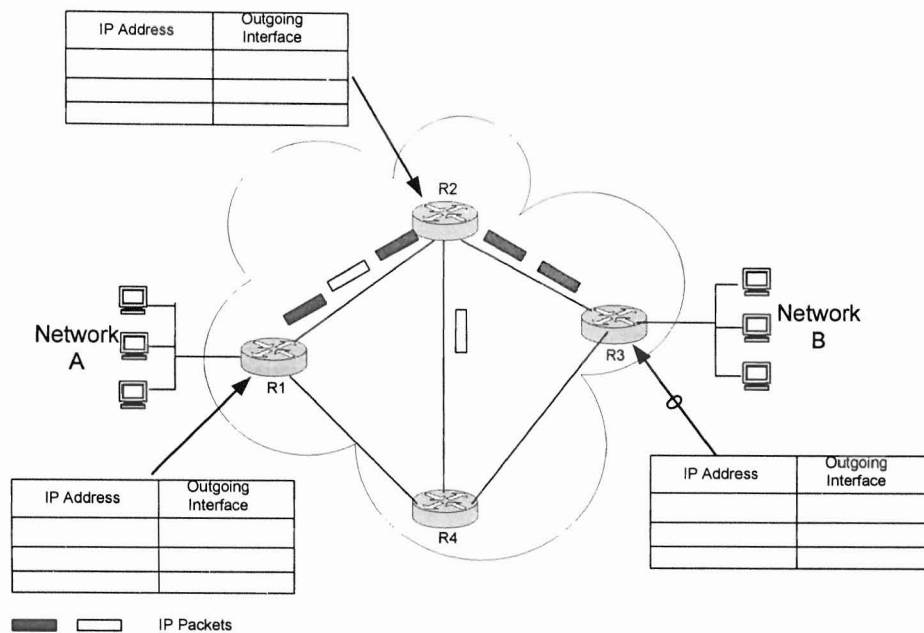


Figure 2.2 IP forwarding.

In MPLS the next hop is chosen by two functions. The first function classifies an incoming packet as an FEC and the second function is the mapping of an FEC to a particular hop. In this way packets from the same source have the same FEC and traverse the same path. In MPLS, the FEC is assigned only once at the Label Edge Router (LER) ingress node and no further header analysis is done at the subsequent nodes. The packet belonging to a particular FEC is forwarded by means of labels. The FEC to which a packet belongs is encoded into a label. This label is used at each node to determine the

next hop as well as the outgoing label. An illustration of the MPLS forwarding technique is given in Figure 2.3.

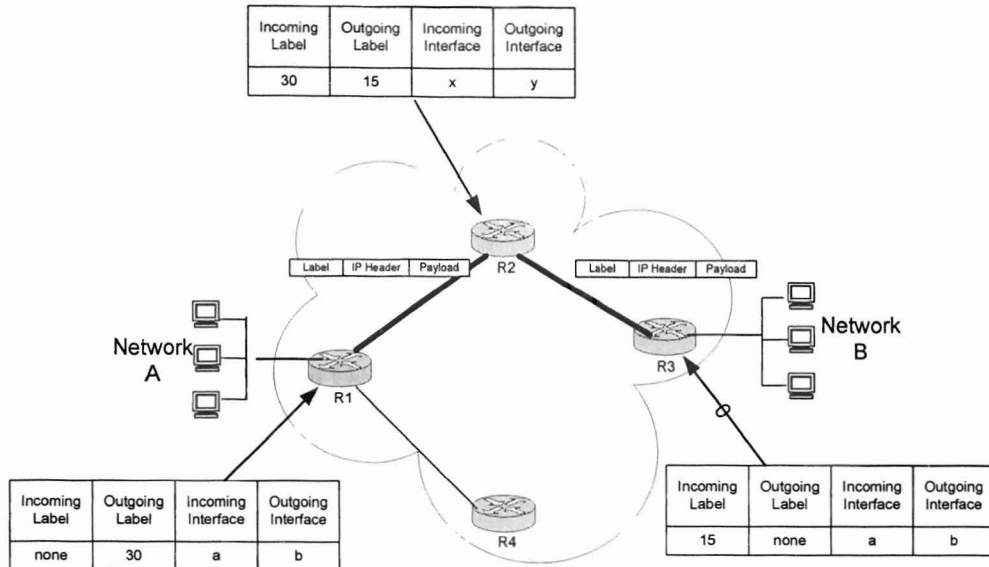


Figure 2.3 MPLS forwarding.

2.3. Advantages of Label Forwarding Techniques Over IP Forwarding

The advantages of using label forwarding over normal IP based forwarding are shown below:

- Forwarding in MPLS can be done by means of switches that examine the label and then decide on the next hop for the packet. In IP based routing the router analyzes the network layer header and then makes the decision to forward the packet.
- LER ingress node assigns a particular FEC to a packet with whatever information it has about the packet even though the information is not assembled from the network layer header.

- Assignment of a packet to a particular FEC could become complicated but this would not affect the LSRs because their main function is only to forward the packets.
- The use of explicitly defined routes can be done in MPLS. A particular route could be selected in MPLS at the LER ingress node when a packet arrives or even before the packet arrives. This could be done as a policy or because of TE requirements.
- The label could represent not only an FEC but also a class of service. This makes MPLS compatible with QoS, DiffServ, etc.

2.4. LERs and LSRs

The Label Edge Router (LER) and the Label Switching Router (LSR) are MPLS capable devices that participate in the MPLS mechanisms.

- LERs are high speed switches that are present at the edge of the MPLS domain. They are used to interface dissimilar networks out of which one must be MPLS and the other could be ATM, Frame Relay, etc.
- LSRs are high speed switching devices that are present in the core of the MPLS domain. They are used in the creation of LSPs.

An example of LSRs and LERs are illustrated in Figure 2.4. R1 and R3 are LERs while R2 is an LSR.

2.5. Forwarding Equivalence Class

The FEC represents a set of packets that have to be handled in the same way for their transportation. In MPLS, FEC for a particular data stream is assigned only once and

this is not like the FEC used in IP networks, which is used every time a forwarding decision has to be performed. A label is used to identify the FEC. It's a fixed length identifier that is used to associate a particular packet to a FEC. In Figure 2.4, there are three LSRs R1, R2 and R3, where R1 wants to send packets to R3. Here R1 and R2 have to agree for a binding between FEC "F" and label value "L" and then R2 and R3 have to agree for a binding between FEC "F" and label value "K". Once the binding has taken effect, R1 sends packets to R2 by attaching a label to it. Label "L" now becomes R1's outgoing label and R2's incoming label. The value of L is local between two LSRs and should be uniquely used to represent a particular FEC.

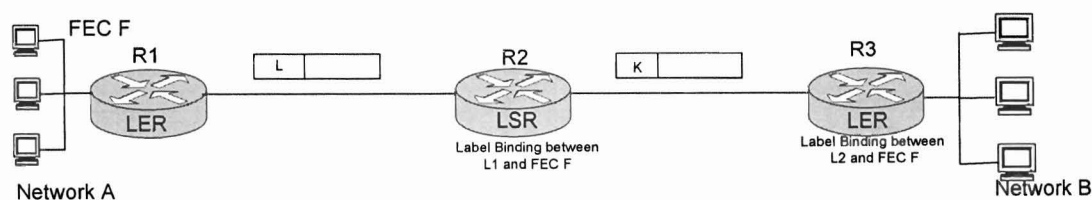


Figure 2.4 Example for binding between label and FEC

When R1 (upstream LSR) wants to send packets to R2 (downstream LSR) then a label binding has to be done to bind the FEC to the label. The label binding is done at the downstream router. The distribution of this label binding is from downstream to upstream. In some scenarios, the LSRs will be able to choose the labels only from a certain numeric range.

2.6. MPLS Label Switched Paths

A MPLS domain consists of MPLS enabled nodes. The LSP has to be setup before data transmission. An LSP is a predetermined path through which the data traverses based on its FEC.

The two kinds of LSPs are Hop-by-Hop routed LSP and Explicit routed LSP. In Hop-by-Hop routed LSP, the decision of the next hop is done by every LSR in that LSP. This is similar to routing in IP networks with the only difference being that this hop selection is done prior to data transmission and is maintained as an entry in the Next Hop Label Forwarding Entry (NHLFE). In Explicit routed LSPs, the source specifies the LSRs that have to be included in the LSP. This is in the form of a source driven scheme.

LSP ingress is the point at which the data stream enters an MPLS domain. Its here that the label stack is attached to the packet and forwarded to the subsequent LSRs. The point at which the LSP ends and data leaves the MPLS domain is called an LSP egress node.

2.7. MPLS Operation [12]

The procedures that have to be followed for a packet to be forwarded in a MPLS domain are given below:

1. Label creation and distribution: Before data is transmitted, the FEC for that data flow has to be defined. The downstream LSR does the label/FEC binding which then is distributed to the upstream LSR.
2. Table creation: Once the label binding information is transmitted to the upstream LSR, both LSRs have to maintain a table called the Label Information Table (LIB). The LIB has information about the mapping between the label and the FEC. It also has entries for the input label and input port and output label and output port.
3. The table is updated whenever there is a renegotiation of the FEC/label binding between the two LSRs.

4. Creation of LSP.
5. Label insertion and table lookup: The LER ingress node uses the LIB to find out the label for a particular FEC and also determines the next hop. The subsequent LSRs only use the LIB for determining the next hop.
6. Forwarding of packets according to the labels.

2.8. MPLS Evolution to GMPLS

It was noted that during the initial stages of MPLS development that MPLS was developed for the edge networks and not for the core of the network. The core of the network used Dense Wavelength Division Multiplexing (DWDM), Add Drop Multiplexers (ADMs), Optical Cross Connects (OXC), etc. These devices that formed the core of the network handled very high bandwidths. The Generalized MPLS (GMPLS) protocol suite was proposed in order to include these devices also. What evolved was a common control plane that traversed different networks (packet, time, space, and fiber).

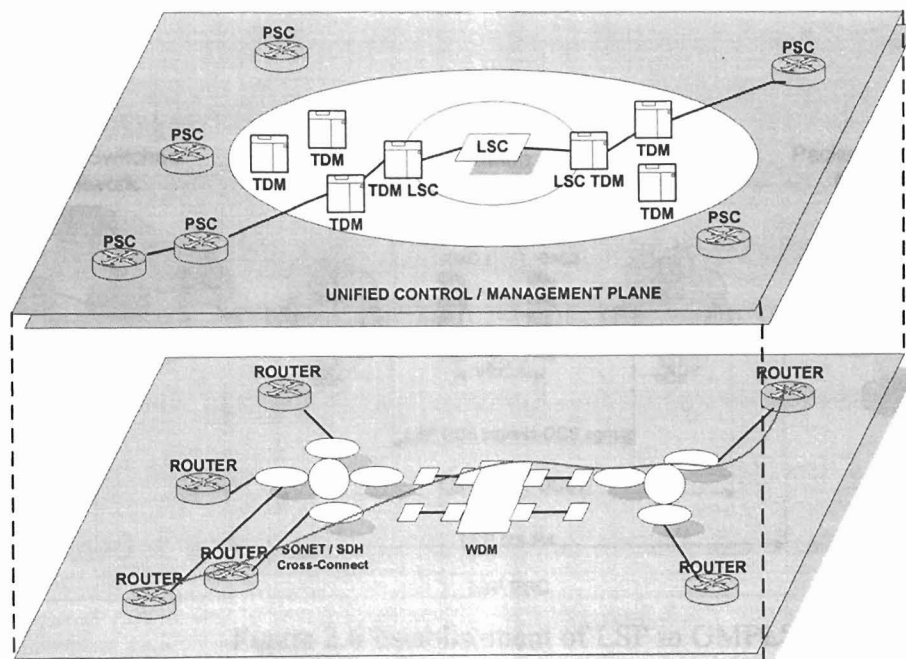


Figure 2.5 GMPLS network-common control and management plane.

An example of a GMPLS network with its layered structure is shown in Figure 2.5. It can be seen that the LSP starts from a Packet Switched network (PSC) and ends in a PSC. The example of Figure 2.5 was developed to show the possible various protocol and physical interfaces that a GMPLS structure will have to consider.

2.9. LSP Establishment in GMPLS Network

The LSP setup in GMPLS is similar to the LSP setup in MPLS networks with the only difference being that here it takes place over dissimilar networks. In Figure 2.6, we see that the two PSC networks are connected to each other by means of two TDM networks and a WDM network. The WDM network connects the two TDM networks to each other.

Figure 2.6 illustrates the formation of an LSP between R1 and R6. When an LSP has to be setup between R1 and R6, the common GMPLS control plane provides functionalities to set this connection up. RSVP-TE and LDP extensions for GMPLS are used for LSP setup.

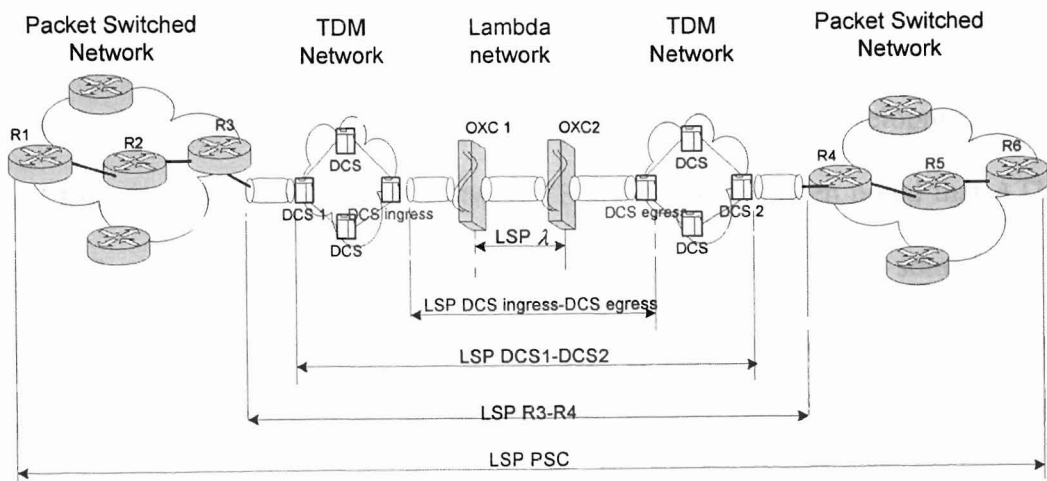


Figure 2.6 Establishment of LSP in GMPLS.

CHAPTER III

INTRODUCTION TO OAM

OAM has been used in Telecommunication networks in order to provide in-service fault detection, fault localization, and also performance management. The various standards for OAM in different network architectures are discussed and studied in this chapter. For example, OAM which has been widely deployed in Asynchronous Transfer Mode (ATM) networks has been discussed in detail in Section 3.1. OAM procedures for FR, SONET/SDH, and OTN are also discussed in the sections of this chapter.

3.1. ATM Based OAM Techniques

OAM techniques for ATM have been discussed in detail in [31] [14]. These papers address management tools for ATM during the early stages of development of the Broadband Integrated Services Digital Network (B-ISDN). It has been noted that ATM, which is a high speed, high performance network architecture also needs some kind of performance and management tool in order to detect, avoid and correct certain performance glitches that could occur. In [31], the authors specifically deal with the Virtual Channel (VC) and Virtual Path (VP) performance, fault and traffic management, and also addresses them to the Consultative Committee on International Telegraphy and

Telephones (CCITT) and T1 standards committee. The authors of [14] suggest that in ATM networks, due to congestion, there could be a considerable decrease in the QoS. In order to keep track of the QoS, the OAM functionalities check for the performance of certain monitored parameters. Therefore, in-service OAM monitoring has been proposed to help the measurement of performance data of every individual connection. It also helps in monitoring and evaluating these QoS parameters without discontinuing the service.

The network provider should be concerned about the management of Private Virtual Connections (PVCs) and also virtual connections that require some kind of end-to-end QoS guaranteed to their connection during the Service Level Agreement (SLA). These management tools come in handy for supervision and management during these kinds of scenarios.

3.1.1. ATM Layer Operation Flows [28] [31]

In broadband networks, operation information has to be propagated between various nodes. The ATM management layer should be able to support VPC/VCC management. In order to make this possible, failure detection, test requests and performance data will have to be sent between the different nodes that have the VPC/VCC. Such a mechanism that transmits operation information at the VP and the VC levels is referred to as the F4 and the F5 flows. The F4 and F5 flows have OAM cells that transmit this operation information. There are two types of VPC/VCC operation flows:

- End-to-end Operation flow: The OAM cell flows (F4 and F5) are transmitted along the entire length of the VPC or the VCC.
- Segment Operation flow: The OAM cell flows are transmitted only along a particular segment of the VPC or the VCC.

3.1.2. VPC/VCC Performance Monitoring

Performance monitoring by the use of OAM cells is used so that the performance of the VPC/VCC can be monitored at the initial point when degradation of the performance starts. The performance of the VPC/VCC can be obtained by monitoring the OAM cells at intermediate nodes or the end node.

This has been implemented by using a special code that is inserted into an OAM cell along with the block size of user data over which this code was calculated. The user data is transmitted first and then the OAM cell is sent. By monitoring the contents of the OAM cell, the far end node determines the performance and then generates a performance report, which is then sent back as an OAM cell. By analyzing the OAM performance report, intermediate nodes can get an insight to the performance. Figure 3.1 below gives a brief idea of how performance monitoring is conducted using OAM cells.

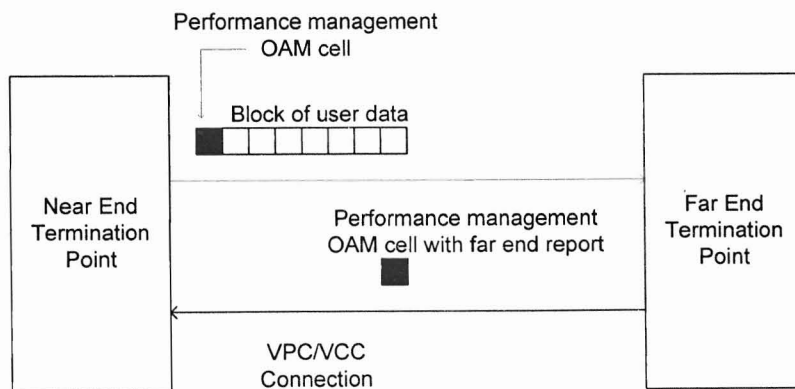


Figure 3.1 Performance monitoring scheme in ATM [28].

There are a number of performance parameters that are measured by using the above scheme. Some of them are cell loss ratio, cell misinsertion ratio, cell error ratio, cell transfer delay, etc.

3.1.3. VPC/VCC Failure Reporting

A failure of a VPC/VCC should be indicated to the nodes included in the circuit. An example of this failure notification procedure in ATM networks is shown in figure 3.2. A VP/VC failure could occur due to failure in the lower layer physical link or due to problems experienced at the ATM layer such as corruption of the VPI/VCI translation tables (Step 1).

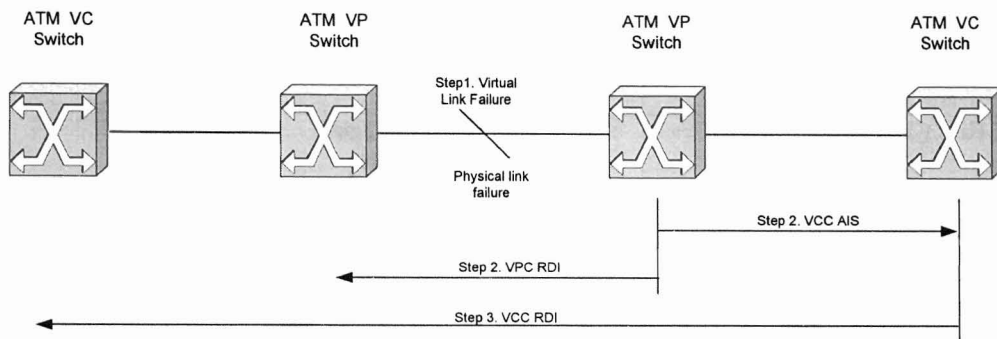


Figure 3.2 Failure notifications in ATM [28].

Failure reporting is done by the use of the VP/VC-Alarm Indication Signal (AIS), which propagates from the node that detects failure to the downstream nodes (Step 2). In response to this the far end node sends a VP/VC – Far End Received Failure (FERF) or Remote Defect Indication (RDI) notification back to the upstream nodes (Step 3).

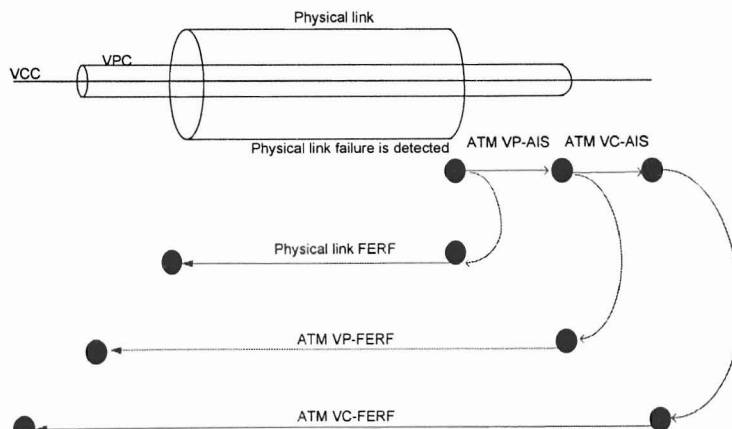


Figure 3.3 Alarm indications [28].

3.1.4. VPC/VCC Continuity Checking and Loopback Testing

Failures in the VPC/VCC are not readily detected like the failures that occur at the physical layer. There should be some mechanism to detect malfunctions at the ATM layer (malfunctioning of the VPCs and VCCs). Nodes cannot distinguish an idle period as a connection or a connection failure. In order to check for normal working of the VP and VC connections OAM cells are generated periodically from the upstream node to the downstream node. The presence of the OAM cell, which arrives periodically, notifies for a proper connection. This procedure is called continuity checking.

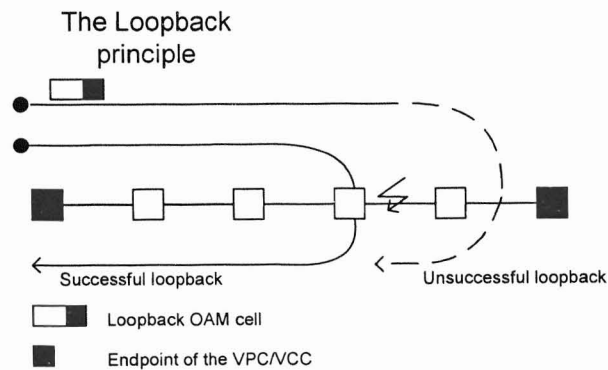


Figure 3.4 Loopback for diagnosis of ATM networks [28].

The loop back capabilities have been included for pre-service connection testing, cell delay measurement, and identification of faults. OAM cells that are added into the flow at a particular node are looped back from pre-defined nodes. The information of loop back is present in the OAM cell.

3.1.5. Traffic Management Functions

OAM cells have been proposed for backward congestion control in ATM networks in order to alleviate the congestion problems. The purpose of sending an OAM

cell for congestion notification instead of an explicit ATM cell is that we can add in additional information about the cause and level of congestion back to the source.

3.1.6. ATM Layer QoS Measurement

There are two kinds of measurement techniques for monitoring the performance of the ATM network. They are:

- In-service measurement methods
- Out of service measurement methods

Performance monitoring OAM cells are used for providing in-service performance monitoring. The user data along with the OAM forward performance cells are used to determine the transfer performance data of the user connection in the ATM network. The transfer performance data can be used to determine the QoS parameters of the connection.

The in-service measurement methodology is shown in Table 3.1.

QoS parameters	Methodology
Maximum Cell Transfer Delay (MaxCTD)	This can be calculated by monitoring the TSTP field of the forward performance OAM cells. From this the CTD distribution can be found out.
Peak to Peak Cell Delay Variation (Peak to Peak CDV)	The Peak to Peak CDV can be estimated by subtracting the minimum CDV with the maximum CDV over a given period of time.
Cell Loss Ratio (CLR), Cell Misinsertion Ratio(CMR), Severely Errored Cell Block Ratio (SECBR)	Estimates of the CLR,CMR, and SECBR are found out by noting the number of cells between two OAM cells i.e. within TUC0+1 and TUC0 timestamps.
Cell Error Rate (CER)	The CER is estimated using the Block Error Detection Code information i.e. the BEDC field in the OAM cell.

Table 3.1 QoS parameters for performance monitoring.

3.1.7. ATM OAM cell Format

The OAM cell format for ATM is given in detail in Figure 3.5.

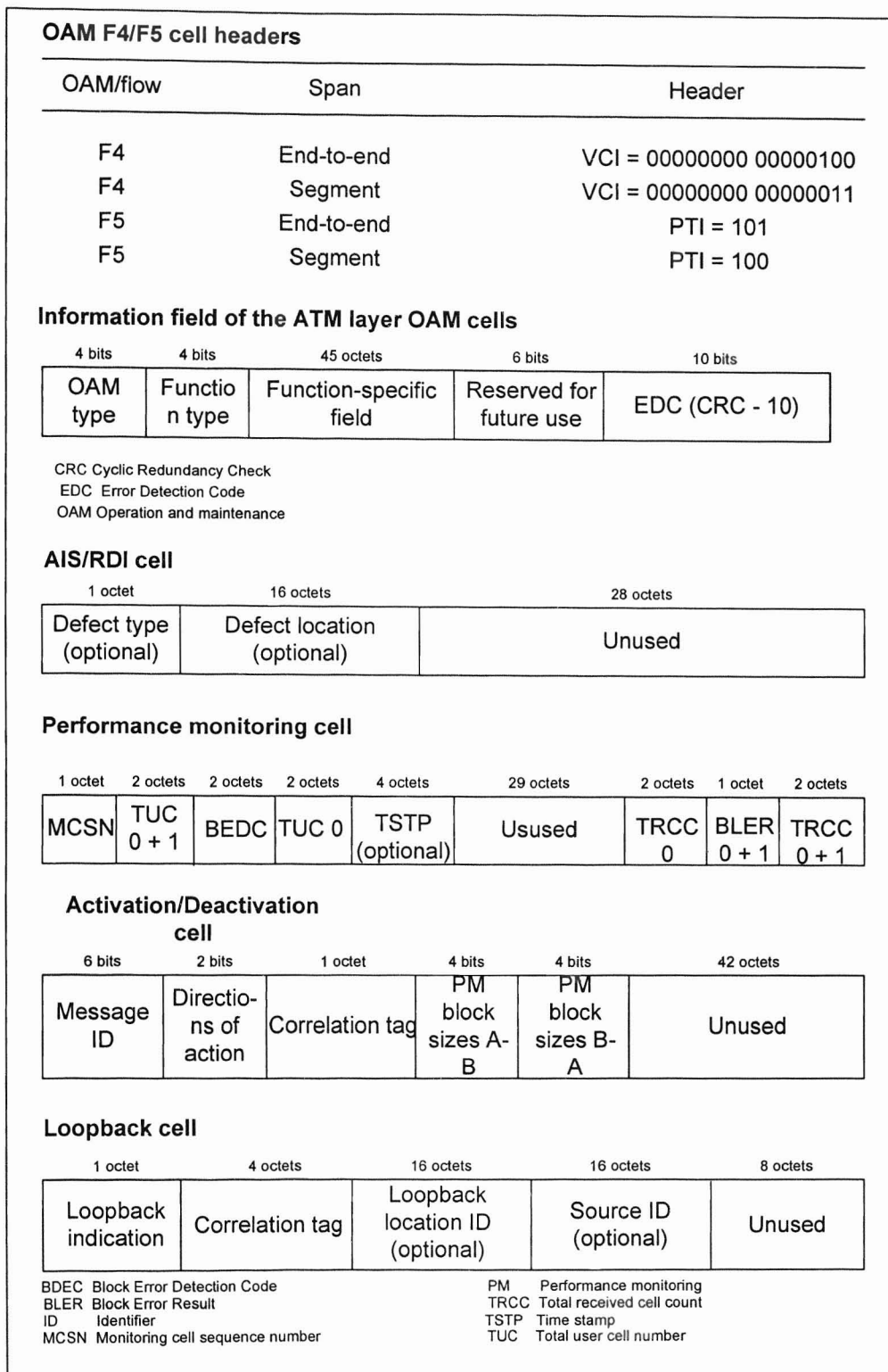


Figure 3.5 ATM OAM cell format [28].

3.2. Frame Relay OAM Techniques

The standards for OAM in frame relay networks have been setup by the Frame Relay Forum in the “Frame Relay Operation, Administration, and Maintenance Implementation Agreement” document [32]. A short description of the OAM standards for Frame relay services is provided in this section.

3.2.1. Frame Relay OAM message format

The OAM message format for Frame Relay is shown in Figure 3.6.

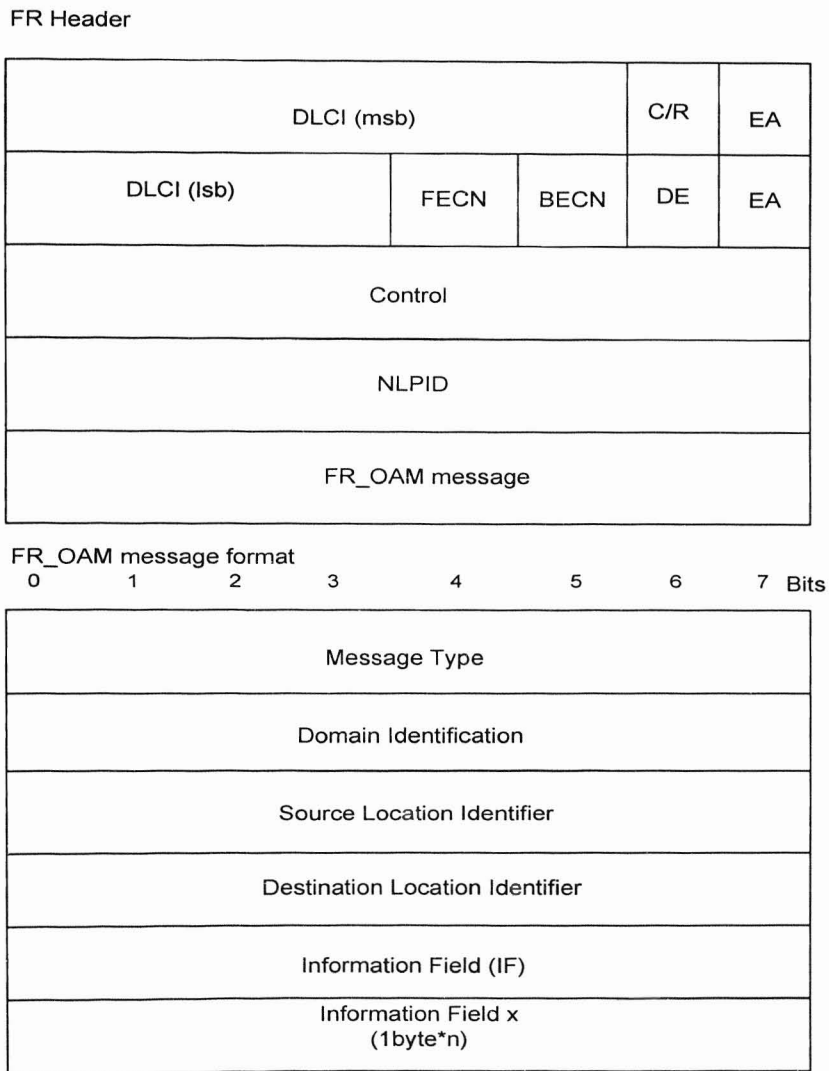


Figure 3.6 Frame relay header and OAM message format [32].

- **Message Type Field:** Identifies the message being sent. The different message types are the Hello message, Service verification, Non-latching Loopback, Latching Loopback, and Diagnostic indication.
- **Domain Identification Field:** It identifies the administrative domain to which the message belongs. The different domains are User defined identifier, OUI identifier, IPv4 network identifier, X.121 identifier, E.164 identifier and Private domain identifier.
- **Source Location Identifier:** Identifies the source of the OAM message along with the administrative domain.
- **Destination Location Identifier:** Identifies the destination of the OAM message along with the administrative domain.

OAM Information Field: Identifies the Type-Length-Data entities. An OAM message may contain one or more OAM information fields.

OAM Information field format

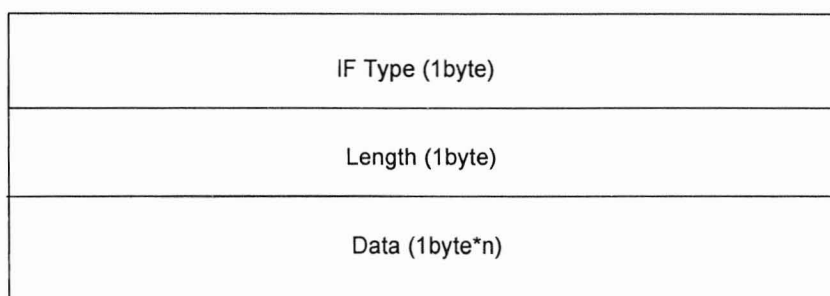


Figure 3.7 OAM information field format

IF Type Field Value: Defines what kind of information is being sent by the OAM information field. Some of the information types are Frame transfer delay, Frame transfer delay results, Frame delivery ratio sync, etc.

Information Field Length: Provides the length of the total OAM information field including the Type-Length-Data.

IF Type Value	Usage	Hello Message	Service Verification Message	Latching loopback Message	Non-latching loopback Message	Diagnostic Indication
0x01	Capabilities	Mandatory	N/A	N/A	N/A	N/A
0x02	Frame transfer delay	Optional	N/A	N/A	N/A	N/A
0x03	Frame transfer delay results	N/A	Optional	N/A	N/A	N/A
0x04	Frame delivery ratio sync	N/A	Optional	N/A	N/A	N/A
0x05	Frame delivery ratio result	N/A	Optional	N/A	N/A	N/A
0x06	Data delivery ratio sync	N/A	Optional	N/A	N/A	N/A
0x07	Data delivery ratio result	N/A	Optional	N/A	N/A	N/A
0x08	Non-Latching loopback	N/A	N/A	N/A	Mandatory	N/A
0x09	Latching loopback	N/A	N/A	Mandatory	N/A	N/A
0x0A	Diagnostic indication	N/A	N/A	N/A	N/A	Mandatory
0x0B	Full source address	Optional	N/A	N/A	N/A	N/A
0x0C	Opaque	Optional	Optional	Optional	Optional	Optional
0x0D	Pad	Optional	Optional	Optional	Optional	Optional

Table 3.2. Different IF types in the OAM message [32].

3.3. SONET/SDH OAM Techniques

SONET and SDH have substantial overhead information fields that can be used to provide OAM functionalities. The Path level overhead is added to DS-1 signals that travel end to end. There can be two kinds of path overheads; these are the STS and VT path overheads. Line overhead is used for the STS-I signals that travel between multiplexers, and finally the Section overhead is used for communication between sections, eg. adjacent network elements in the optical network. This section introduces the OAM features present in SONET (OAM implementations in SDH are similar) [33].

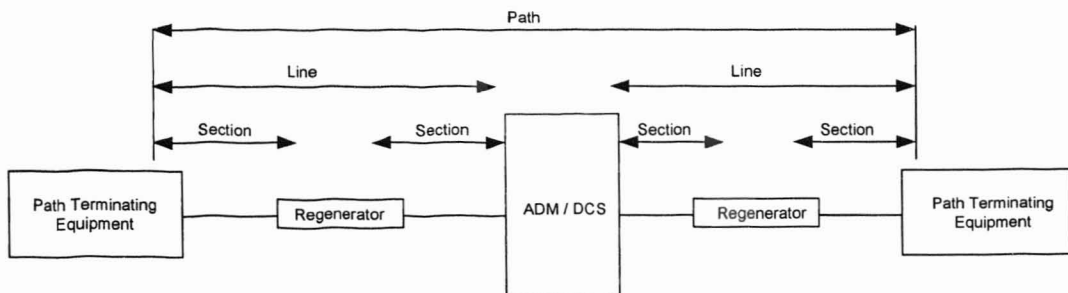


Figure 3.8 Example of a SONET/SDH network.

Figure 3.8 illustrates the network architecture of a SONET network. It also clearly defines a path, section and line.

3.3.1. OAM Support in Section Overhead

The section overhead can be used for the following purposes:

- Performance monitoring of the STS-I signal.
- Local order wire.
- Data communication channel to carry OAM information.
- Framing.

We shall identify the bytes in the overhead that are used directly for OAM purposes:

- B1 – 1 byte for BIP-8 used for checking transmission errors.
- D1, D2 and D3 – These are used for control, monitoring, administration, and other communication needs and is used mainly for OAM purposes.

3.3.2. OAM Support in Line Overhead

The OAM functionalities supported by the line overhead are:

- Performance monitoring.
- Automatic protection switching.
- Line maintenance.

The bytes used for OAM purposes are:

- H1, H2 – Used to indicate path Alarm Indication Signal (AIS-P).
- B2 – Used for BIP-8 for parity check.
- K1, K2 – Used for bidirectional Automatic Protection Switching (APS) and also for detecting AIS-L and RDI signals.
- D4-D12 – Used for alarm, control, monitoring, administration and other OAM functions.
- M0 – Line error indication function for indicating the error count detected by the LTE using the BIP-8.

3.3.3. OAM Support in STS Path Overhead

The STS POH is used for:

- Performance monitoring of the STS SPE.
- Path status.
- Path trace.

The overhead bytes used for OAM:

- J1 – 64 byte or 16 byte, which is user programmable and used by the intermediate terminals to verify continuity from the intended transmitter.
- B3 - BIP-8 used for parity check.
- G1 – This is used to send the status and performance back to the originating path terminal equipment. Therefore the path performance can be monitored in both directions. Bits 1 to 4 are used for REI-P, bits 5 6 and 7 are used for RDI-P signal and bit 8 is undefined.

3.3.4. OAM Support in VT Path Overhead

The OAM functionalities for the VT are included in the V5 frame. The structure of the V5 frame is given as:

- 2 bits for BIP-2.
- 1 bit for REI-V.
- 1 bit for RFI-V.
- 3 bits for the signal label to give content of the VT SPE.
- 1 bit for RDI-V.

3.3.5. SONET Alarms for Fault Notification

Some of the alarms that are used whenever faults occur in the SONET network are given below:

- Loss Of Signal (LOS) – occurs when a BER of 1 in 10³ is noticed. It is then that the LOS state is set. It is cleared after receiving two consecutive non-errored frames.

- Out Of frame (OOF) – When three or four frames with A1 and A2 invalid framing patterns arrive.
- Loss of Frame.
- Loss Of Pointer – SP – LOP and VR – LOP are set.
- AIS – Sent to the downstream nodes in order to show that alarm has been raised and to avoid unnecessary repetition of alarms. AIS-L, SP – AIS, and VP – AIS.
- Remote Error Indication (REI) – Notification to the transmitting node that errors block has been received. REI-L, REI-P, and REI-V.
- Remote Defect Indication (RDI) – Signal returned to the transmitter in case of LOS, LOF or AIS. RDI-L, RDI-P, and RDI-V.
- Remote Failure Indication (RFI) – RFI sent to far end transmission system to initiate an APS.
- Loss of Sequence Synchronization – Pseudo-random number from the synchronization receiving side is not in sync with the object under test

3.4. OTN OAM Techniques

3.4.1. Representation of Optical Network

The optical network can be represented as shown in Figure 3.9. It has three sections called the Optical Path layer, Optical Frequency section layer and Optical Regeneration section layer.

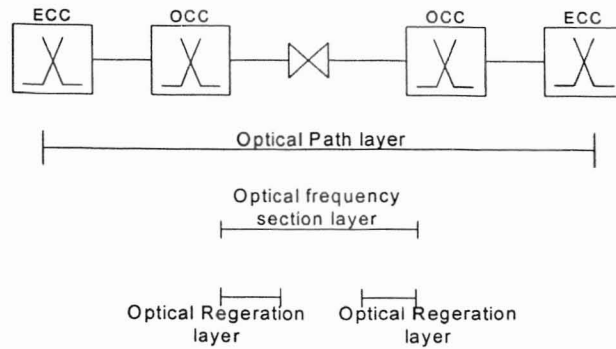


Figure 3.9 Example of an OTN network.

3.4.2. Operation and Maintenance Concept:

The five main areas of network management are Fault management, Configuration management, Accounting management, Performance management, and Security management. Fault and Performance management are covered by the OAM concept as described in [24].

3.4.2.1. Performance Monitoring

Performance monitoring is done similar to that done in SDH networks. The BIP is calculated over a block of data. This is then transported to the receiver which cross checks the value of the block with the BIP it receives. This gives the Bit Error Rate (BER) of the system.

3.4.2.2. Failure Detection

There can be different kinds of failures in the optical transport layer. They could be fiber cuts, failure of the Optical Regenerators (OR), failure of the OXCs, drifting of the Optical Frequency Division Multiplexed (OFDM) channels or even failure of the Over Head channels.

3.4.2.3. Failure Information and Fault Localization

The use of AIS and RDI for downstream and upstream transfer of failure information is used. An illustration of the two alarm signals is given in Figure 3.10.

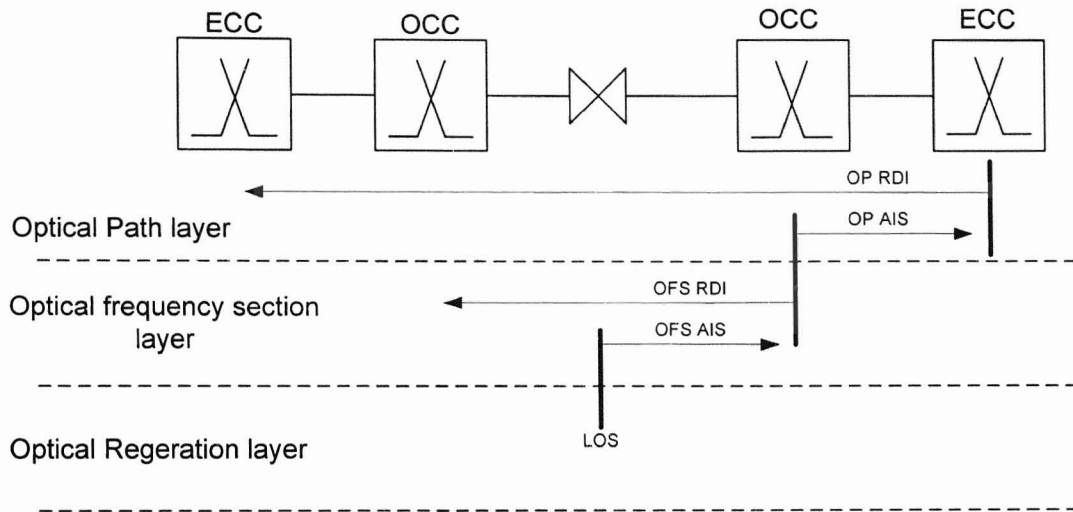


Figure 3.10 Failure information and notification.

multicast packet forwarding. This section explores a number of different algorithms that may potentially be employed by multicast routing protocols:

- Flooding.
- Spanning Trees.
- Reverse Path Broadcasting (RPB).
- Truncated Reverse Path Broadcasting (TRPB).
- Reverse Path Multicasting (RPM).
- Steiner Tree (ST).

4.1.1. Flooding

The flooding procedure begins when a router receives a packet that is addressed to a multicast group.

- The router employs a protocol mechanism to determine whether this is the first time it has seen this particular packet or whether it has seen the packet before.
- If it is the first reception of the packet, the packet is forwarded on all interfaces except the one on which it arrived, guaranteeing that the multicast packet reaches all routers in the internetwork. If the router has seen the packet before, the packet is simply discarded.

A flooding algorithm is very simple to implement since a router does not have to maintain a routing table and only needs to keep track of most recently seen packets. However, flooding technology does not scale for Internet-wide applications since it generates an excessively large number of duplicate packets and uses all available paths across the internetwork, instead of just a limited number. Also, the flooding algorithm

makes inefficient use of router memory resources since each router is required to maintain a distinct table entry for each recently seen packet.

4.1.2. Spanning Tree

A more efficient solution than flooding would be to select a subset of the Internet topology that forms a spanning tree. The spanning tree defines a tree structure where only one active path connects any two routers on the Internet.

- Once the spanning tree has been built, a multicast router simply forwards each multicast packet to all interfaces that are part of the spanning tree except the one on which the packet originally arrived.
- Forwarding along the branches of a spanning tree guarantees that the multicast packet will not loop and that it will eventually reach all routers in the internetwork.

4.1.3. Reverse Path Broadcasting

An even more efficient solution than building a single spanning tree for the entire Internet would be to build a group-specific spanning tree for each potential source subnetwork. These spanning trees would result in source-rooted delivery trees emanating from the subnetwork directly connected to the source station. Since there are many potential sources for a group, a different spanning tree is constructed for each active (source, group) pair.

For each (source, group) pair, if a packet arrives on a link that the local router considers to be the shortest path back to the source of the packet, then the router forwards the packet on all interfaces except the incoming interface. Otherwise, the packet is

discarded. The interface over which the router expects to receive multicast packets from a particular source is referred to as the “parent” link. The outbound links over which the router forwards the multicast packet are called the “child” links. The basic algorithm can be enhanced to reduce unnecessary packet duplication if the router making the forwarding decision can determine whether a neighboring router on a potential child link considers the local router to be on its shortest path back to the entire source. If this is the case, the packet is forwarded to the neighbor otherwise it is discarded. The information needed to make this “downstream” decision is relatively easy to derive from a link-state routing protocol since each router maintains a topological database for the entire routing domain. If a distance-vector routing protocol is employed, a neighbor can either advertise its previous hop for the (source, group) pair as part of its routing update messages or “poison reverse” the route. Either of these techniques allows an upstream router to determine if a downstream neighboring router considers it to be on the downstream router’s shortest path back to the source.

4.1.3.1. Example of RPB

Figure 4.1 is an example of RPB. In this example, we will look at the RPB algorithm from Router B’s perspective. Router B receives the multicast packet from Router A on link 1. Since Router B considers link 1 to be the parent link for the (source, group) pair, it forwards the packet on link 4, link 5 and the local leaf subnetworks if they have group members. Router B does not forward the packet on link 3 because it knows from routing protocol exchanges, that Router C considers link 2 as its parent link for the (source, group) pair. If Router B were to forward the packet on link 3 it would be

discarded by Router C since it would arrive on a non-parent link for the (source, group) pair.

One of the major limitations of the RPB algorithm is that it does not take into account multicast group membership when building the distribution tree for a (source, group) pair, as a result, datagrams may be unnecessarily forwarded to subnetworks that have no members in the destination group.

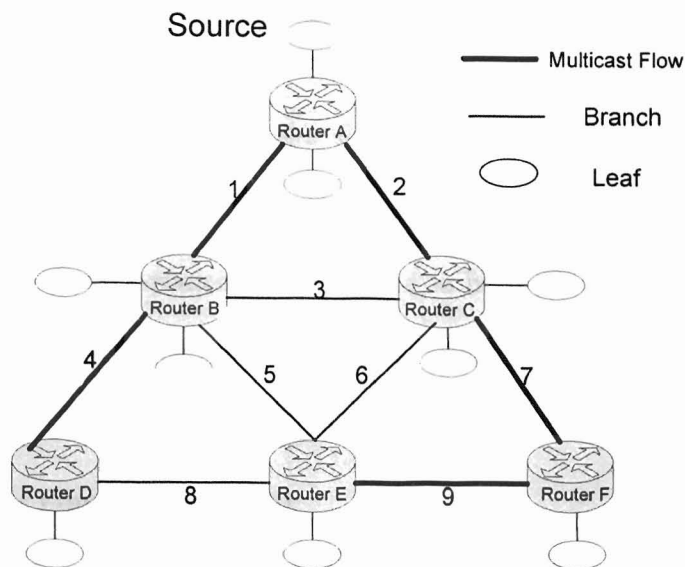


Figure 4.1 Example of RPB.

4.1.4. Truncated Reverse Path Broadcasting

TRPB was developed to overcome the limitations of RPB. With the help of IGMP, multicast routers determine the group memberships on each leaf subnetwork and avoid forwarding datagrams onto a leaf subnetwork if it does not have a member of a destination group present. The spanning delivery tree is “truncated” by the router if a leaf subnetwork does not have group members.

TRPB removes the limitations of RPB, but it solves only part of the problem. It eliminates unnecessary traffic on leaf subnetworks but it does not consider group memberships when building the branches of the distribution tree.

4.1.5. Reverse Path Multicasting

RPM is an enhancement of RPB and TRPB. RPM creates a delivery tree that spans only:

- Subnetworks with group members, and
- Routers and subnetworks along the shortest path to subnetworks with group members

RPM allows the source-routed spanning tree to be pruned so that datagrams are only forwarded along branches that lead to members of the destination group.

4.1.5.1. Procedure for Operation of RPM

When a multicast router receives a packet for a (source, group) pair, the first packet is forwarded following the TRPB algorithm to all the routers in the internetwork. Routers that are at the edge of the network and have no further downstream routers in the TRPB tree are called leaf routers. If there is a group member on one of its leaf subnetworks, a leaf router forwards the packet based on its IGMP information. If none of the subnetworks connected to the leaf router have group members, the leaf router may transmit a “prune” message on its parent link informing the upstream router that it should not forward packets for the particular (source, group) pair on the child interface receiving the prune message. Prune messages are sent only one hop back towards the source.

An example of the RPM algorithm is illustrated in Figure 4.2. In this figure we demonstrate the execution of prune messages to extend the RPB and TRPB algorithms to the RPM algorithm.

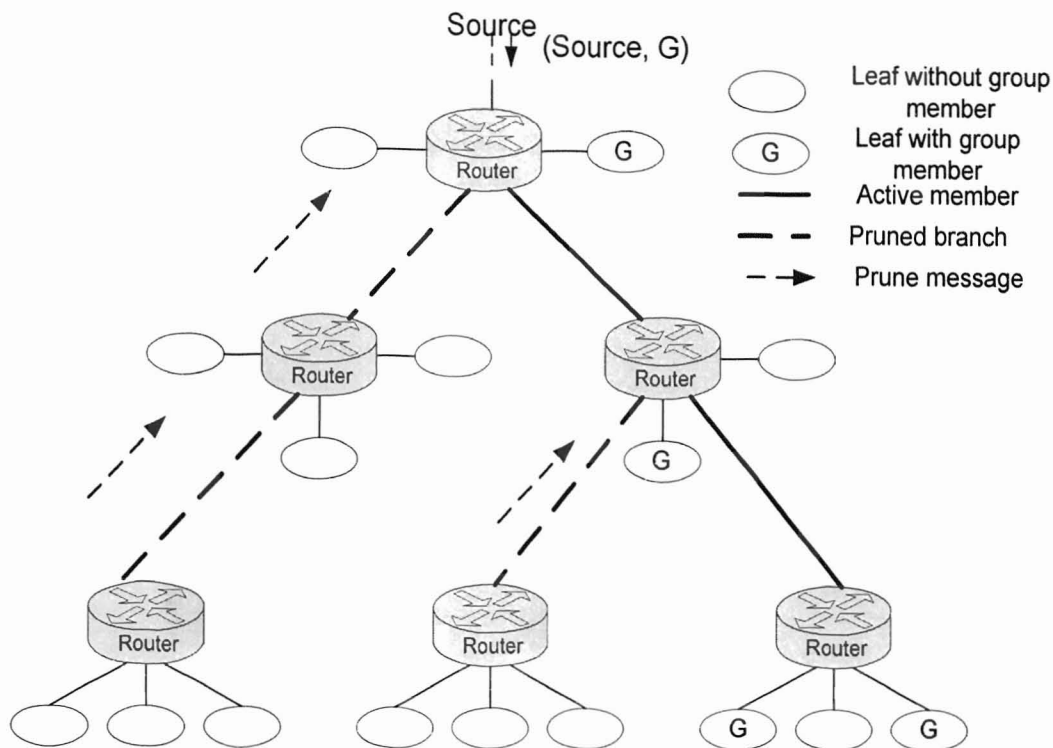


Figure 4.2 Procedures for RPM.

4.1.6. Steiner Tree

In the RPB family of algorithms (RPB, TRPB, and RPM) the shortest path between the source node and each destination node is used for delivering multicast packets, guaranteeing that multicast packets are delivered as fast as possible. However, none of these algorithms try to minimize the use of network resources. In Figure 4.3 the

RPB tree and the ST are shown, assuming that C is the source and A and D are the recipients.

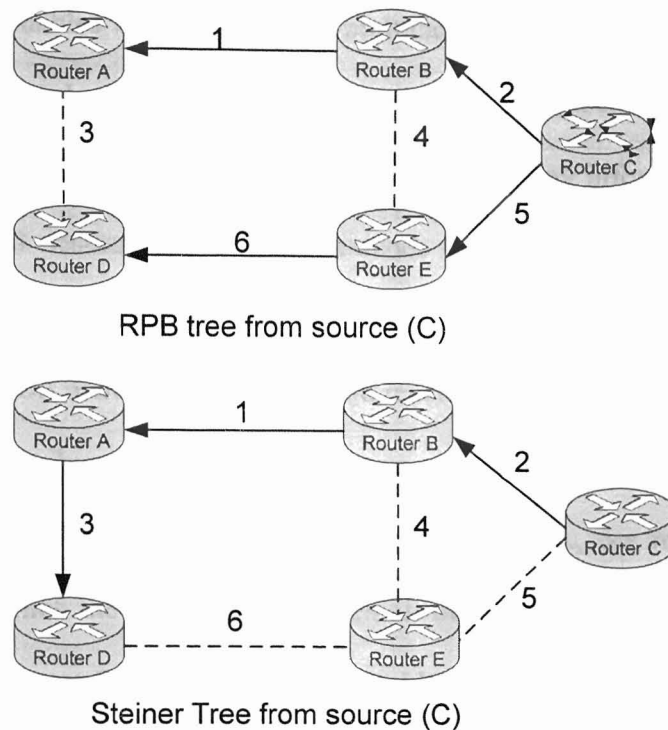


Figure 4.3 Steiner tree.

It can be easily observed that the second tree uses lesser number of links. Although this tree is slower than the RPB tree (because packets need to pass three hops for reaching D instead of two hops required in RPB tree), it uses fewer links. This type of tree is called ST. Although STs minimize the number of links used for constructing a delivery tree, difficulties in computing these trees has made these trees of little practical importance. STs are also unstable since the form of ST changes with a node joining or leaving a multicast group.

4.2. Review of Existing IP Multicast Routing Protocols

4.2.1. Multicast Extensions for Open Shortest Path First

MOSPF is a modification of the Open Shortest Path First (OSPF) protocol in its version 2, which enables the routing of IP multicast datagrams. Since OSPF is a link-state routing protocol, it provides a database that describes the topology of the Autonomous System (AS). This topology provides the routing information for datagrams to reach their final destinations. MOSPF extends OSPF with multicasting abilities by implementing IGMP mechanism that monitors multicast group memberships. By requesting IGMP host membership queries and receiving IGMP host membership reports, MOSPF routers distribute the group information and locations by flooding a new Link State Advertisement (LSA). Upon receiving this new LSA, the routers calculate the shortest path by which the datagram can reach all of the indicated group members. This in-turn forms a Shortest Path Tree (SPT) that holds the source as the root, and the group members as terminal branches. If members have to leave or be added on to the multicast group, MOSPF prunes or grafts its SPT to include or exclude members upon the updated LSA. Therefore, these trees are built on demand, and separate trees are built for each combination. The results of the different trees are then cached for later use for datagrams that have the same source and destination. Due to the amounts of link state information that MOSPF would have to carry if it were to create a tree for the entire Internet, its functions are limited within an AS. Group members outside of the AS implement DVMRP to allow MOSPF to function with members outside of the AS [19].

4.2.2. Distance Vector Multicasting Routing Protocol

The DVMRP is an extension of the Distance-vector Routing protocol, which is a routing protocol to support multicasting in IP networks. DVMRP combines distance vector algorithms with TRPB for the construction of source-based trees [10].

TRPB is an algorithm that computes the shortest path between the multicasting router and the source. According to the RPB algorithm, each packet received on the interface with the shortest distance to the source is forwarded to all other interfaces. TRPB uses the IGMP messages in the Local Area Network (LAN) domain to determine the presence of multicast receivers in the LAN. The multicast enabled router then forwards the packet onto a LAN that has group membership for that particular source.

The basic operations of DVMRP are flooding, pruning, and grafting. The first datagram to multicast to a particular multicast group is flooded throughout the whole network. The leaf routers that do not have any group members connected to them issue a Prune message to the upstream router to indicate that there are no more hosts needing the multicast information. As each router has no more downstream interfaces to multicast to, the Prune message keeps traveling upstream until all the necessary interfaces are disconnected. The Graft message is used to cancel this prune that has been performed on the branch. This supports dynamic group membership by allowing hosts to join the multicast tree at any time. Any router that finds new group members attached to it, issues a Graft message to its upstream router on the interface that has the shortest reverse path to the source and to which a prune was previously issued. This graft procedure allows new hosts to join onto the multicast tree at any given time.

4.2.3. Protocol Independent Multicast-Dense Mode

The PIM-DM uses the existing unicast routing table to perform the RPF instead of maintaining a separate multicast routing table. It is protocol independent, meaning that regardless of the protocol in which the unicast routing table was created, PIM can use its routing table to perform multicasting. This method of using the existing unicast routing table protocol avoids sending multicast updates to the other routers, significantly reducing the amount of overhead that other protocols use for this purpose.

An example of the PIM-DM multicasting protocol is illustrated in Figure 4.4. The three steps shown below describe the process in which multicast trees are created.

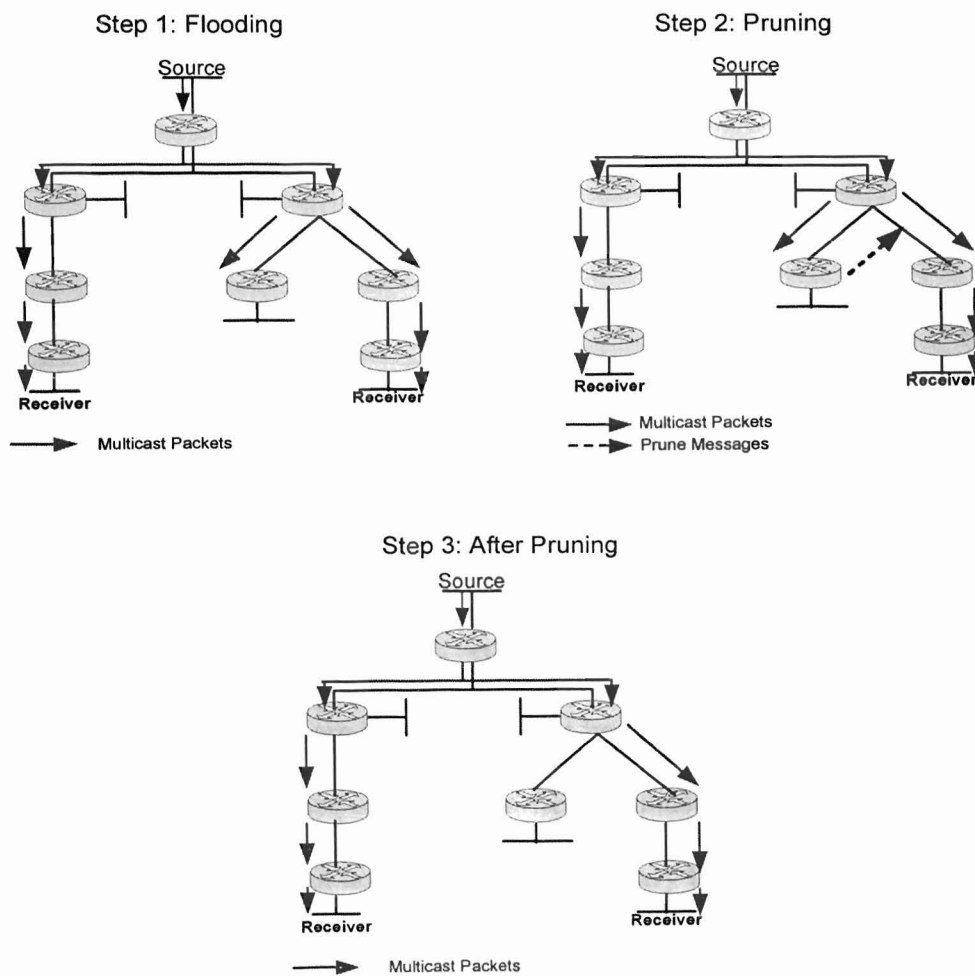


Figure 4.4 Example of PIM-DM.

Dense mode indicates that the protocol only employs SPTs to deliver source to multicasting group [(source, group)] multicast traffic using a push-principle. The push-principle floods the multicast traffic to all points in the network that is associated with a cost (bandwidth, router CPU, etc.). In order to avoid consumption of network resources, routers send Prune messages back up the source distribution tree to stop unwanted multicast traffic. The SPTs are built dynamically by means of the flood and prune mechanisms as soon as a multicast source begins transmitting based on its Neighbor Table. The initial state is determined by assuming that each neighbor is on the SPT, and all other neighbors, creating a Broadcasting Tree, because a router sends the multicast traffic to all neighbors in a broadcast-like fashion.

PIM uses a neighbor discovery mechanism to establish the neighbor adjacencies using a PIM Hello message. In order to forward multicast traffic, RPF is performed on the incoming interface of a router using the information in the unicast routing table [6].

4.2.4. Protocol Independent Multicast-Sparse Mode

Similar to the PIM-DM, the PIM-SM uses the unicast routing table to perform its multicasting functionalities, but unlike it PIM-SM only delivers multicast traffic to the nodes that explicitly request it. This is achieved through PIM Joins, which are sent hop-by-hop towards the root node of the tree. The root node of a tree is the Rendezvous Point (RP) router in the case of a shared tree, or the first-hop router that is directly connected to the multicast source in the case of a STP. To free up unused routes, a Prune message is sent up the tree toward the root node to avoid unnecessary traffic. The main point for this protocol is that in the Explicit Join Model forwarding state in the routers is set as a result of the Join messages, which differs from the flood-and-prune mechanism used in PIM-

DM. The operation of PIM-SM is based on a single, unidirectional shared-tree whose root node is called the RP. The routers that are directly connected to the receiver, or last-hop routers, that need to receive the traffic from a specific multicast group make up the tree. When a router has no more receivers attached, the router prunes itself.

One of the primary advantages of PIM-SM is that it does not limit the retrieval of multicast traffic via the shared tree. Just as it is possible to use the Explicit Join mechanism to join the shared tree, whose root is the RP, this mechanism can be used to join the SPT whose root is a particular source. This reduces the network latency and possible congestion at the RP. The drawback, however, is that all the routers must create and maintain the addresses along the SPT, which is resource consuming for the routers.

Because PIM-SM uses a unidirectional-shared tree, multicast traffic can only flow downstream, so multicast sources must somehow get their traffic to the RP so that the traffic can flow down the shared tree. PIM-SM accomplishes this by having the RP join the SPT back to the source so it can receive the source's traffic. Because PIM-SM uses the Explicit Join model, multicast traffic is better constrained to only those portions of the network where it is actually desired, avoiding the inefficiencies found in the flood-and-prune protocols such as PIM-DM. Therefore, it is better suited for multicast networks spanning WAN links [7].

4.2.5. Core Based Tree

The CBT protocol is designed based on the concept of bi-directional shared trees. It is used to build and maintain shared multicast trees, which involve routers that are leading to interested hosts [2] [3]. In a shared tree, all sources of a particular group share the same tree. There is a Primary core and one or more Secondary cores. The Primary

cores may be different for different groups. The biggest advantage of the shared tree approach is that it offers more favorable scaling characteristics than all other multicast algorithms [2].

IGMP messages are used to indicate group membership to the Designated Router (DR) (a router to which the host is connected). The DR on receiving the membership report sends a JOIN_REQUEST message towards the group's core router. This request is sent on a hop-by-hop basis. The core router on receiving the JOIN_REQUEST will send the JOIN_ACK message to confirm the join to the multicast tree. A CBT tree is "pruned" downstream-to-upstream whenever there are no more hosts receiving multicasting data from that group. The CBT router prunes itself from its parent router when its child interface list for a group becomes NULL [2] [3]. The above mentioned procedures are illustrated in Figure 4.5.

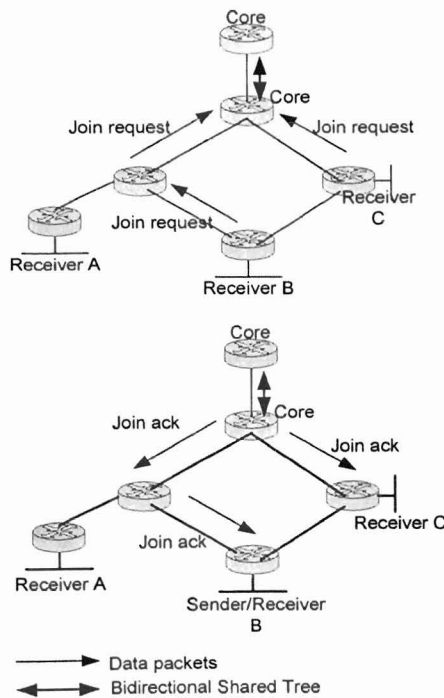


Figure 4.5 Example of CBT.

CHAPTER V

OAM IMPLEMENTATION IN GMPLS

In chapter 3 we have discussed the OAM functionalities in different networking architectures namely ATM, FR, SONET and OTN. The OAM functionalities of the different architectures have been studied and their modes of operation have been identified. Extensions to GMPLS common control plane as well as extensions to the GMPLS user plane have to be proposed in order to make GMPLS capable of handling a common OAM control that spans all network topologies. The section below describes the extensions to GMPLS user and control plane. The identification and classification of OAM functionalities for their deployment in either the user or control plane is also justified in the section below.

Table 5.1 gives an elaborate description of the different OAM functionalities for the different network architectures such as ATM, FR, SONET, SDH, and OTN. It specifies the signals that each of the networking architectures employ to carry out the OAM functions. It also clearly classifies these OAM functionalities for GMPLS into control plane and user plane implemented functionalities.

OAM Functions	ATM	Frame Relay	SONET/SDH	OTN	GMPLS User Plane	GMPLS Control Plane
Fault Management	<ul style="list-style-type: none"> AS RDI Loopback 	<ul style="list-style-type: none"> Non-latching loopback Latching loopback 	<ul style="list-style-type: none"> AIS RDI REI LOS OOF LOF LOP LSS 	<ul style="list-style-type: none"> AIS RDI 		RSVP-TE/LDP OAM message with FM object/TLV and Loopback Object/TLV
Performance Management	<ul style="list-style-type: none"> Forward Monitoring Backward Monitoring 	<ul style="list-style-type: none"> Service level verification 	<ul style="list-style-type: none"> Forward Monitoring Backward Monitoring (G1 in STS Path overhead) 	<ul style="list-style-type: none"> Performance Monitoring 	OAM packets	
Activation / Deactivation	<ul style="list-style-type: none"> PM Forward Monitoring PM Backward Monitoring 	N/A	N/A	N/A		RSVP-TE/LDP OAM message with Activation-deactivation Object/TLV
Automatic Protection Switching	N/A	N/A	<ul style="list-style-type: none"> APS using RFI in the VT Path overhead APS using K1, K2 in Line overhead 	Available		RSVP-TE/LDP OAM message with FM Object/TLV
QoS Parameters	<ul style="list-style-type: none"> MaxCTD Peak-to-Peak CDV CLR CMR SECBR CER 	<ul style="list-style-type: none"> FTD FDR DDR 	N/A	N/A		QoS signaling for a connection at setup using the Sender_TSPEC Object in RSVP-TE or the Traffic Parameter TLV in LDP
Continuity Check (CC)	<ul style="list-style-type: none"> Available 	N/A	<ul style="list-style-type: none"> Continuity check for Path 	N/A	OAM packets	

AIS - Alarm Indication Signal
 REI - Remote Error Indication
 OOF - Out Of Frame
 LOP - Loss Of Pointer
 APS - Automatic Protection Switching
 RFI - Remote Failure Indication
 CLR - Cell Loss Ratio
 CMR - Cell Misinsertion Ratio
 CER - Cell Error Ratio

RDI - Remote Defect Indication
 LOS - Loss Of Signal
 LOF - Loss Of Frame
 LSS - Loss of Sequence Synchronization
 PM - Performance Management
 MaxCTD - Maximum Cell Transfer Delay
 Peak-to-Peak CDV - Peak-to-Peak Cell Delay Variation
 SECBR - Severely Errored Cell Block Ratio

Table 5.1 OAM functionalities for GMPLS.

5.1. Extensions to GMPLS Control Plane to Include OAM Functionalities

The OAM functionalities that have to be deployed by the control plane have been identified in Table 5.1. The Fault Management, Activation / Deactivation, Automatic Protection Switching (APS), and Signaling QoS have been identified as potential OAM functionalities that have to be deployed by using the GMPLS control plane i.e. the label distribution protocols such as LDP, CR-LDP and RSVP-TE. Extensions to these protocols have to be proposed in order to equip them with OAM functionalities. Also the process of OAM deployment in the control plane is discussed.

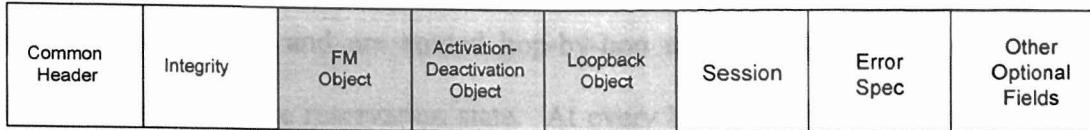
Fault Management OAM functionalities include alarms such as AIS, RDI, REI, LOF, OOF, LOS, loopback, Automatic Protection Switching (APS), etc. These functions are signals used to trigger off or indicate a particular state. It has been identified that these signals could be triggered off as control signals using RSVP-TE, LDP or CR-LDP.

Performance Management includes functionalities such as forward and backward monitoring of user traffic as well as QoS signaling. In order to ensure that monitoring of user traffic is done in real time, the forward and backward monitoring of user traffic has to be included as GMPLS user plane extensions. Signaling of QoS parameters should be included as a GMPLS control plane function as the required traffic parameters for a user connection can be populated to the different nodes for online QoS estimation using forward and backward monitoring.

The other OAM functionalities such as Activation / Deactivation and Hello messages for peer discovery have to be deployed using the control plane. Like the fault management function, these functions are also signals, which indicate or trigger off a particular state such as an automatic protection state triggered off by the APS function when transmitted to a particular node in the network.

5.1.1. Extensions to RSVP-TE for Enabling GMPLS OAM Functionalities

In order to make OAM functionalities available in GMPLS, we have proposed the inclusion of a RSVP-TE OAM message. The format of the OAM message is shown in Figure 5.1.



FM Object - Fault Management
 PM Object - Performance management
 Activation-Deactivation Object - Activation or deactivation
 Session Object - Identifies a particular connection

Figure 5.1 RSVP-TE OAM message format.

The three objects proposed here are the FM Object, Loopback object and the Activation-Deactivation object. The formats for the three objects are shown in Figure 5.2.

FM Object

Alarm Type (2 bytes)	Must be Zero
-----------------------	--------------

FM object:
 This object indicates the kind of alarm generated by some fault in the network.
 Class= FAULT_MANAGEMENT

Loopback Object

Loopback direction (2bytes)	Must be Zero
Correlation Tag (4bytes)	
Loopback location (IPv4/IPv6 address)	

Loopback object:
 This object indicates loopback functionalities for diagnostic purposes
 Class= LOOPBACK

Activation-Deactivation Object

Type	Direction (1byte)	Correlation tag (2bytes)
PM block size (A-B) (4bytes)		
PM block size (B-A) (4bytes)		

Activation-Deactivation object:
 This object activates or deactivates two nodes for PM
 Class= ACTIVATION_DEACTIVATION

Figure 5.2 OAM object format.

The FM object is included in the OAM message in order to indicate an alarm in the network. The different alarms that the FM object indicates are the AIS, RDI, REI, LOS, LOF, LOP etc. OAM messages with FM object report faults in the LSP. According to the kind of alarm generated the OAM message with FM object could travel upstream towards the ingress and are routed hop-by-hop using the path state or could travel downstream using the reservation state. At every hop, the IP destination address is the unicast address of the previous hop as suggested in RFC2205.

OAM FM messages do not modify the state of any of the nodes through which they traverse. The ERROR_SPEC object is defined as in RFC 2205 and it gives information about the error that has occurred and includes the IP address of the node that detected the error. The OAM message with FM object could be sent to the ingress or the egress nodes as indicated above or could be sent to some other intermediate node (for fault management) which could be specified during the path message. This process could be similar to the Notify object in the Path message to indicate a node for sending notification messages as described in [23].

Loopback object is used during the diagnostic phase. The loopback direction indication gives the direction of the signal and also includes the kind of loopback. The loopback could have a forward and a reverse direction. If the received OAM message has a “reverse loopback” indication, then the message isn’t looped. The Loopback location identifies the node at which loopback should take place. The correlation tag along with the sender descriptor allows a particular OAM message to be uniquely identified by the source node. If loopback is initiated in the forward direction from ingress to a particular node then the OAM message is sent downstream in a similar fashion to the Path message.

The reverse loopback is sent from the loopback location to the source in the same way as a Resv message.

The Activation-Deactivation object is used for activating or deactivating the Performance Monitoring (PM) and Continuity Check (CC) done at the GMPLS user plane. It also is used for negotiating the block size of data over which performance monitoring is done periodically (in forward and reverse directions). The activation/deactivation OAM messages are sent upstream or downstream in a similar fashion to the Path message or Resv message.

5.1.2. Extensions to LDP for Enabling GMPLS OAM Functionalities

In the previous section we have introduced the extensions for RSVP-TE in order to enable OAM functionalities for the common control plane. In this section we will identify the extensions for LDP in order to make it OAM capable.

We introduce the OAM message in order to carry out all the operation, maintenance and administration functionalities in the network. The format of the OAM message and the Type Length Value (TLV) extensions are shown in Figure 5.3.

The FM TLV, Loopback TLV and the Activation-Deactivation TLV are the new TLVs proposed for performing performance monitoring, loopback diagnostic test and activation and deactivation of the performance monitoring. The Session field is added to each TLV in order to indicate the OAM message for a particular session.

OAM Messenger

0	Message Type = OAM message (15 bits)	Message Length (2 bytes)
Message ID (4 bytes)		
FM TLV (variable length)		
Activation-Deactivation TLV (variable length)		
Loopback TLV		
Optional TLVs		

FM TLV

0	0	Alarm Type	Length (2 bytes)
v	Session		Reserved
Defect Location			

Loopback TLV

0	0	A/D	Length (2 bytes)
v	Loopback direction		Session
Source ID			
Loopback ID			

Activation-Deactivation TLV

0	0	Type	Length (2 bytes)
v	Direction (7 bits)		Session
Source ID			
PM Block size (A-B) (4 bytes)			
PM Block size (B-A) (4 bytes)			

Figure 5.3 LDP extensions for OAM.

5.1.3. GMPLS Control Plane OAM Functionalities and Procedures

In the previous section, the extensions to LDP and RSVP-TE have been defined in order to make OAM functionalities possible. The three main OAM functionalities that are defined are Fault Management (fault monitoring and fault localization), Loopback and Activation-Deactivation for Performance monitoring. The procedures for the three functionalities will be defined in this section.

5.1.3.1. Fault Management Procedures in GMPLS

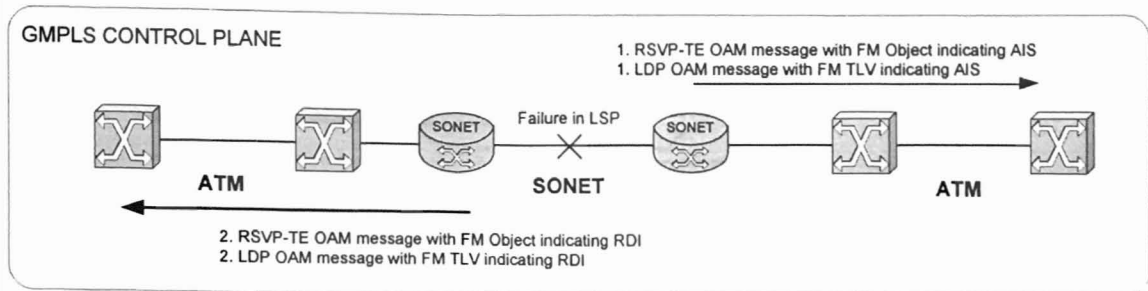


Figure 5.4 RSVP-TE/LDP execution of AIS/RDI.

Figure 5.4 explains how AIS/RDI functionalities are employed using LDP and RSVP-TE. When there is a failure in the link or other physical layer failures, the AIS has to be triggered. In Figure 5.4, there is a failure between LSR C and LSR D. This failure triggers two RSVP-TE OAM message with the FM Object or an LDP OAM message with a FM TLV in the case of LDP. The FM object/TLV indicates that the signal is AIS and the other an RDI. AIS is sent downstream to LSR F while RDI is sent upstream to LSR A. The AIS and the RDI have the defect location ID which indicates the location at which the fault has occurred.

5.1.3.2. Loopback Procedures in GMPLS

The next OAM functionality that is explained below is the loopback test that is executed for diagnostic purposes. This loopback test can be executed for on demand connectivity verification, pre-service connectivity verification and for fault localization.

In Figure 5.5, the diagnostic functionalities are carried out. RSVP-TE OAM message with Loopback object or an LDP OAM message with Loopback TLV can be used to activate the loopback diagnostic tests in the GMPLS network.

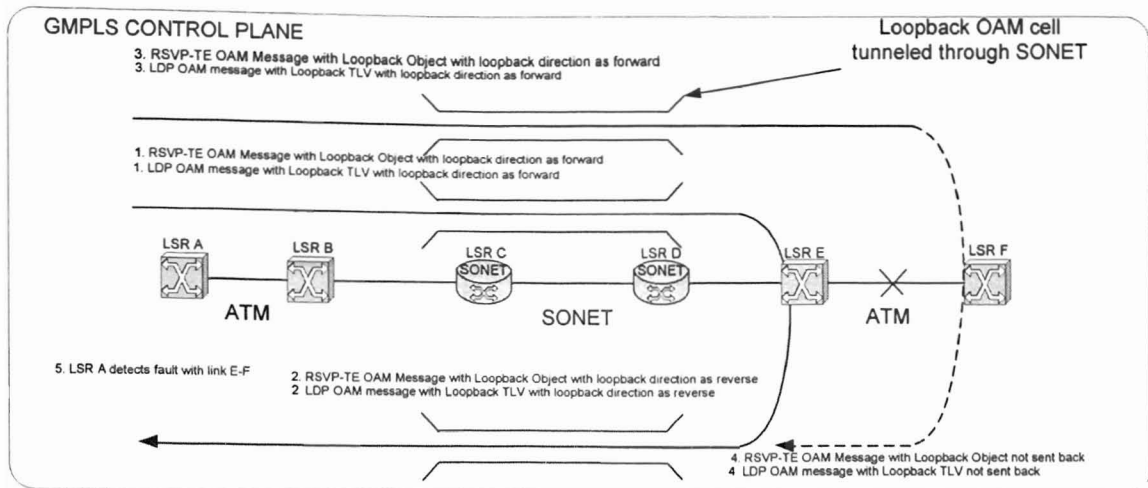


Figure 5.5 LDP/RSVP-TE execution of loopback

LSR A sends the LDP/RSVP-TE OAM message with the necessary extensions. The Loopback direction in the FM Object and FM TLV is set to “forward direction”. The loopback of the OAM message occurs at the Loopback destination ID specified in the FM Object/TLV. Once the RSVP-TE/LDP OAM message is looped back, the Loopback direction is set to “reverse direction”. If the source node receives the loopback message, then the loopback is successful. If a particular loopback doesn’t return, then the fault can be located at that loopback node to which the loopback message was sent.

5.1.3.3. Activation/Deactivation of PM and CC in GMPLS

The RSVP-TE/LDP OAM message with Activation-Deactivation Object/TLV is used for activating or deactivating the Performance monitoring and the Continuity check functionalities. The type field suggests the kind of functionality (Continuity check or Performance monitoring) that is being activated/deactivated.

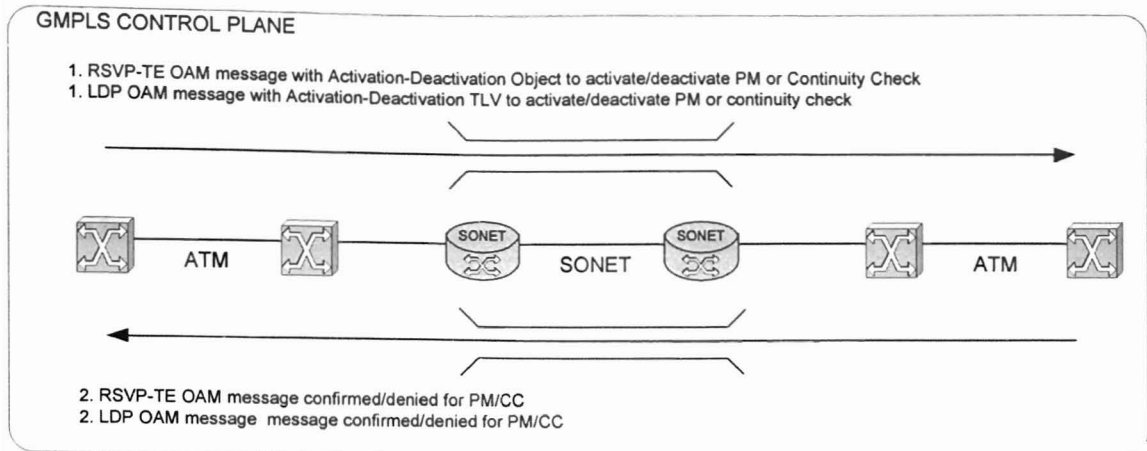


Figure 5.6 RSVP-TE/LDP execution of activation/deactivation of PM or CC.

The direction field indicates the direction in which the PM/CC is activated/deactivated. The correlation tag correlates between a request and a response. A request for activation/deactivation can be initiated by ingress LSR to the egress LSR. This OAM message traverses the same path as that of the connection from the ingress to the egress. At the egress, the request could be confirmed/denied and this is reported to the ingress LSR using the RSVP-TE/LDP OAM message with confirmed/denied FM Object/TLV. Figure 5.6 illustrates the procedure for Activation/Deactivation of PM and CC OAM packets.

5.2. Extensions to GMPLS User Plane to Include OAM Functionalities

The two OAM functionalities deployed in the user plane are:

- Performance Monitoring of the connection.
- Continuity Check of the connection.

Performance monitoring in GMPLS is executed by using both the control and user plane for GMPLS. In the previous section, we have seen that by using the Activation-Deactivation Object/TLV we were able to activate or deactivate PM between two nodes

in the network. The actual performance monitoring for a connection should be done in real time along with the flow of data for that connection. OAM packets are inserted into the data flow. These OAM flows inserted along with the data flow should have the format shown in Figure 5.7.

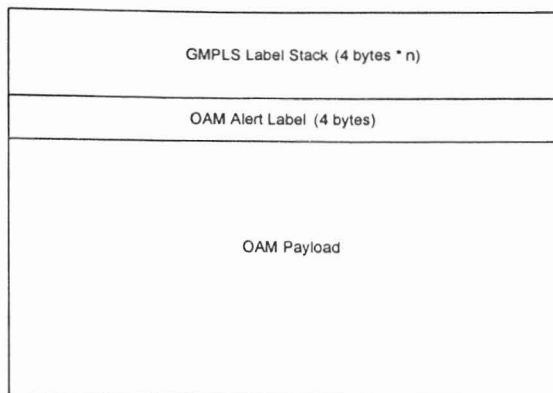
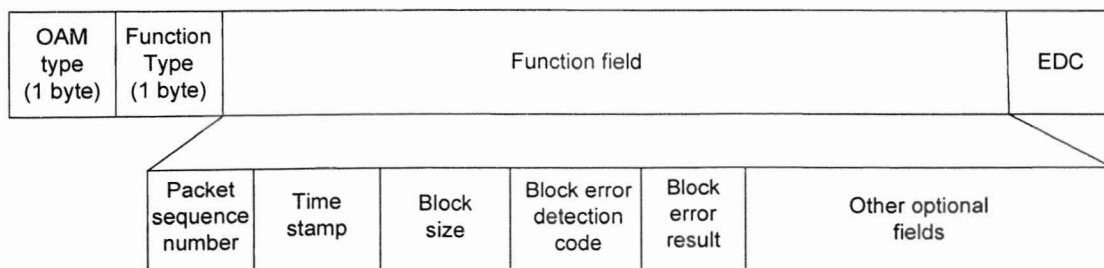


Figure 5.7 GMPLS OAM packet.

The OAM Alert label used in GMPLS is similar to its implementation in RFC 3429. One of the reserved label values in RFC 3032 is assigned to the OAM Alert Label. This label has the value as 14 as proposed in RFC 3429.

The OAM type distinguishes whether the packet is a PM OAM packet or a CC OAM packet. The Performance monitoring packet has the format illustrated in Figure 5.8.



OAM Type - Specifies Performance Monitoring packet / Continuity check packet
 Function type - Specifies Forward monitoring or backward reporting for Performance monitoring packets

Figure 5.8 GMPLS OAM payload for performance monitoring

The performance monitoring packet works in conformance to the ATM standards for performance monitoring. Once Performance monitoring in GMPLS is enabled using the RSVP-TE/LDP OAM messages, OAM Performance monitoring packets are transmitted from the ingress node to the egress node in the GMPLS network after a block of user data packets. In GMPLS, the performance of the network is estimated at the egress LSR and the report is sent back to the ingress as a backward reporting OAM packet containing the result in the Block error result field.

The Continuity Check OAM packets are injected into the network at regular time intervals. The presence of continuity check OAM packets ensure that the LSRs do not mistake the connection to be down even when there is no data transmitted.

CHAPTER VI

MULTICASTING SERVICES THROUGH MPLS NETWORKING

This section provides the extensions to RSVP-TE and LDP for the implementation of MPLS multicast protocols and procedures. The multicasting extensions for MPLS networking proposed in this chapter will enable the MPLS networking layer to support full multicasting services corresponding to the IP based multicasting algorithms, such as, DVMRP, CBT, PIM-DM, and PIM-SM which were explained in Chapter 4. These procedures and protocol extensions also allow the advantages of TE to be available for MPLS multicasting operations.

6.1. Extensions to RSVP-TE and LDP for Multicasting in MPLS Networks

The proposed extensions to RSVP-TE and LDP are focused towards enabling the RSVP-TE and LDP signaling mechanisms to establish, maintain, and terminate multicasting connections. RSVP applies IP datagram forwarding as a transport mechanism [27]. This can be accomplished by adding message types and new C_type identifiers to the existing objects of the messages to establish multicasting based operations using the objects of the RSVP-TE messages. The LDP signaling mechanism can be achieved by introducing a new message called the Multicasting message (Figure

6.3). All the multicasting functions of Join, Leave and Destroy can be implemented by this multicasting message.

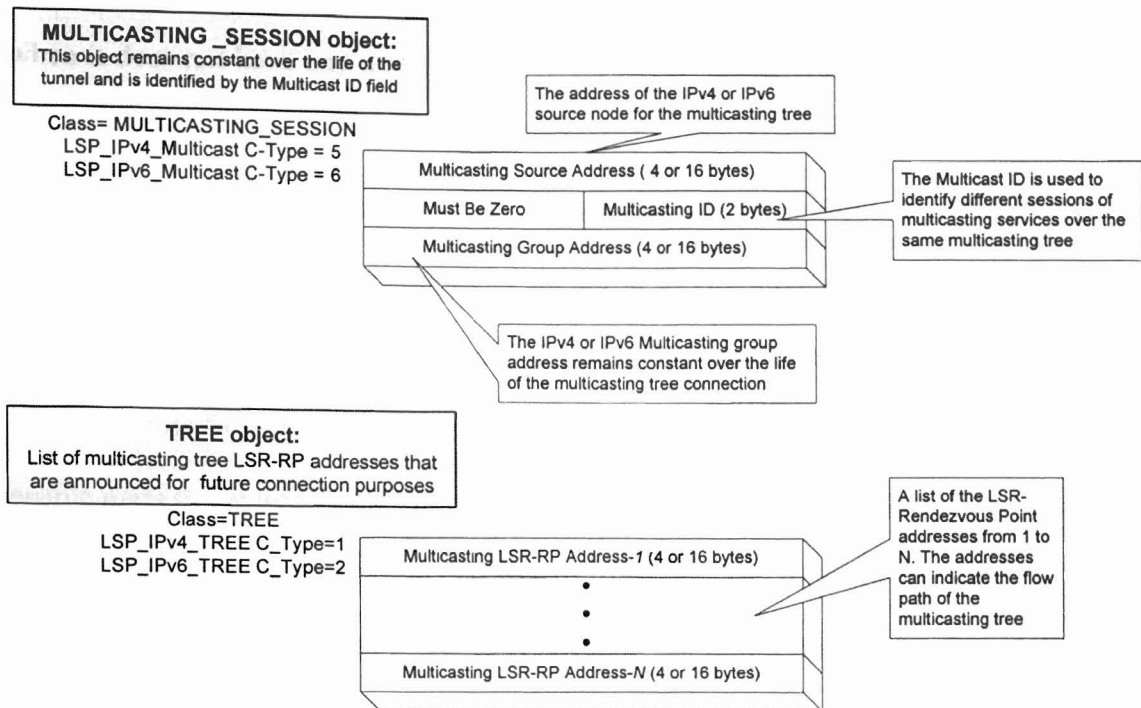
6.1.1. RSVP-TE Extensions for Multicasting in MPLS Networks

6.1.1.1. The Multicasting Session and Tree Objects

RSVP defines a "session" to be a data flow with a particular destination and transport-layer protocol [27]. Hence, a Multicasting Session object and the Tree object need to be defined for multicasting tree control operations as optional objects additional to the Session object (Figure 6.1).

In the Multicasting Session object, the Multicasting Source Address is the IPv4 address or the IPv6 address of the source node for the multicasting tree. The Multicasting Session object remains constant over the life of the tunnel and is identified by the Multicast ID field. The IPv4 or the IPv6 Multicasting Group Address is a Multicasting group address used in the session that remains constant over the life of the multicasting tree connection.

The Tree object is a list of multicasting tree LSR-RP addresses that are announced, such that LSRs that are not part of the multicasting tree can identify these LSR-RPs for future connection to the multicasting tree. An LSR-RP is any intermediate LSR that has two or more branches that forward traffic or signaling downstream or aggregate upstream. In addition, for the Join message sent from the root LSR, the Tree object can be used to establish an explicit traffic path when connecting to a new host. In this case, the Tree object list of LSR-RPs will indicate this flow path, where the last LSR-RP of the Tree object list will be the one to connect to the new host.



Common object headers are not shown in the figures

Figure 6.1 RSVP-TE extension objects format.

6.1.1.2. RSVP-TE Multicasting Message Extensions

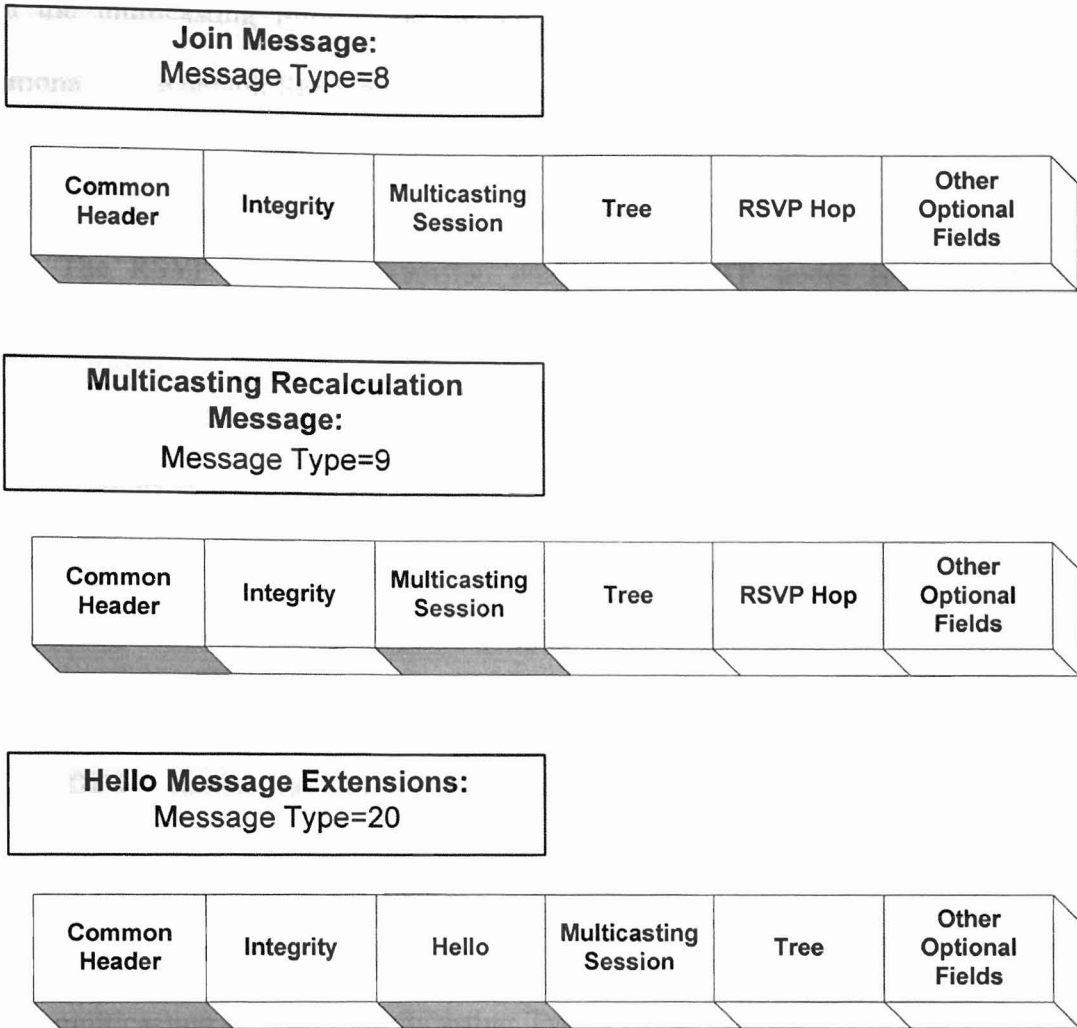
Using these object fields, four message types are defined (Join, Leave, McRecal, and Destroy) (Figure 6.2).

Join Message: When a new host wants to join an existing multicast tree, the new host will issue a Join message in making a request to join the multicasting tree. In the case where a non-connected LSR is making this request, the LSR will send a Join message to a multicasting LSR of the multicasting tree requesting for a connection. Following this, a Join message will be sent from that multicasting LSR to the root LSR of the multicasting tree requesting a connection and to update the Multicasting Information DataBase (MIDB) with this request.

Leave Message (ResvTear message): A leaf LSR may initiate a Leave message when it does not have any more hosts actively participating in the multicast session. The Leave message used in multicasting is conducted by the ResvTear message that is sent to the upstream LSR-RP, containing the multicasting session information. If the LSR-RP sees that all attached multicasting users are not in service anymore, then the LSR-RP will send a ResvTear message (Leave message) to the upstream LSR requesting a disconnection from the multicasting tree. In this fashion, the Leave functionality will enable parts of the tree that are not used to disconnect itself in modular units.

Multicast Recalculation (McRecal) Message: The recalculation message is used to inform the root LSR that it has to recalculate the multicast tree for that multicasting group, and it carries the updated information downstream to keep the MIDB of the intermediate nodes updated. The McRecal message should be routed like the corresponding Resv message or a ResvTear message, and its IP destination address will be the multicast address of the previous hop.

Destroy Message (PathTear message): The root initiates the Destroy function for a Multicast tree when it wants to terminate a multicast session. Since multiple multicasting trees may overlap in a network, the multicasting session information is included to identify the multicasting source and multicasting group to destroy. The destroy function is conducted through the PathTear message and is sent from the root to all of its downstream LSRs. The PathTear message that is inherent in RSVP-TE removes all the entries in the LSP as well as all reservations.



The highlighted object fields are optional to the message

Figure 6.2 RSVP-TE message extensions.

6.1.1.3. Multicasting Extensions to the RSVP-TE Path and Resv Messages

The Path message originates from the LSR-RP to the LSR that desires to join the multicasting tree. The Path message may include the multicasting information in addition to the Session object, which will enable LSRs that are not part of the multicasting tree to know the multicasting source and group information. The Resv message is also extended

with the multicasting information, which can be treated as optional data that are additional to the default Session object.

6.1.1.4. Multicasting Extensions to the Hello Message

The RSVP-TE Hello extension enables RSVP-TE nodes to detect when a neighboring node is not reachable. The Hello mechanism is intended for use between immediate neighbors, and therefore, a Hello message and a Hello Acknowledgement message are exchanged between two RSVP-TE neighbors. The extensions to the Hello message format for multicasting applications include the multicasting session and tree information, which are also considered optional fields (Figure 6.2).

6.1.2. LDP Extensions for Multicasting in MPLS Networks

6.1.2.1. LDP Multicasting Message

The Multicasting message is used to conduct a Join, Leave, or Destroy command of the multicasting tree. The Multicasting Type-Length-Value (TLV) contains the three main command types: Join, Leave, and Destroy. The Multicasting TLV is used as a mandatory TLV in the Multicasting message and is used as an optional TLV in messages supporting multicasting operations (Figure 6.3). The Tree TLV is used to provide a list of selected multicasting tree LSR-RP addresses that are listed for identification purposes (as in the Hello messages) or for selected path traversing (as in the Notification messages) (Figure 6.3).

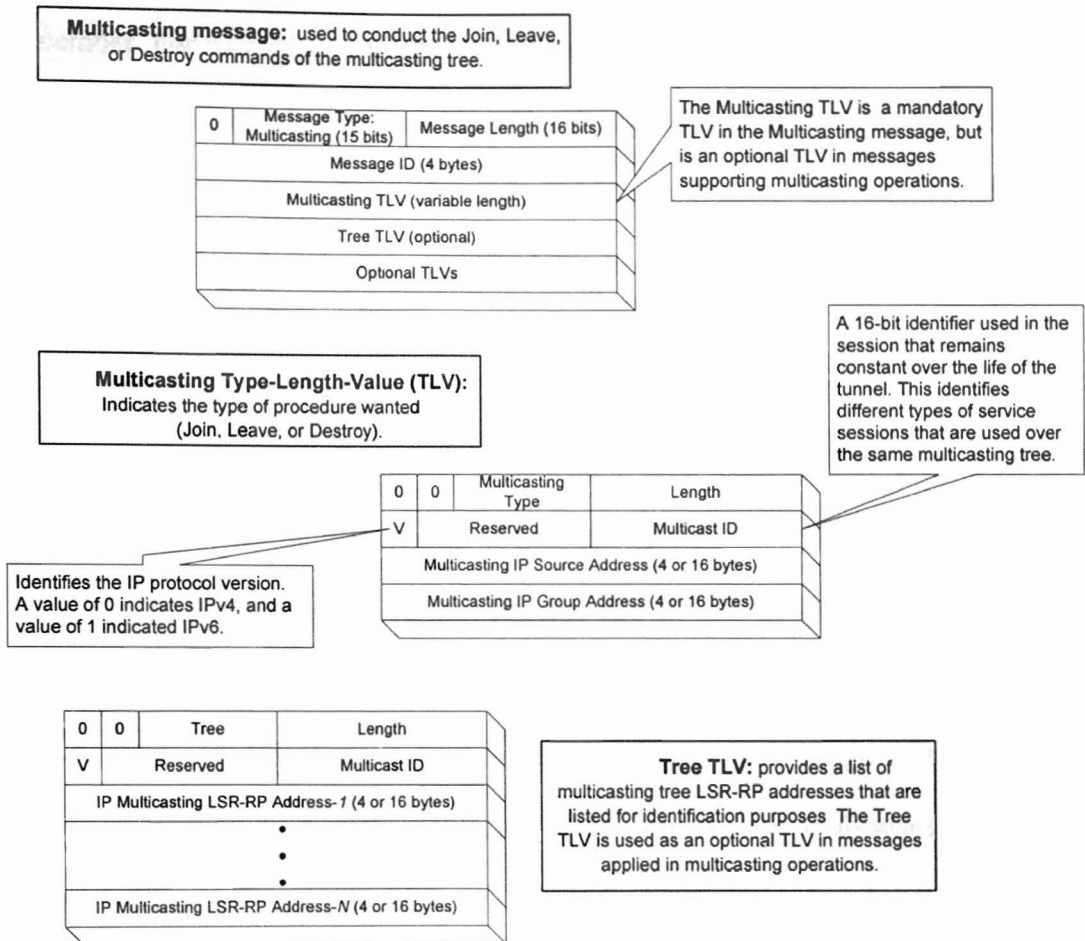


Figure 6.3 LDP message extensions.

Join Command: When a leaf-initiated Join operation must be performed, a Multicast message is sent (by means of a unicast transmission) to an LSR of the existing multicast tree to create a new LSP connection to the leaf. The Join command is initiated when an LSR has identified a multicasting group that it wants to join in order to receive group specific multicasting data.

Leave Command: The Leave command allows members of the multicast group to detach themselves from the group, stopping all information from the group to reach the pruned member. For example, a member may decide to stop participating in the group. In this case, a Leave command must be initiated by the receiving LSR to the upstream LSR of the multicasting tree to indicate the end of its participation. Correspondingly, the upstream LSR will send a Leave command to the root LSR such that the MIB information can be updated.

Destroy Command: The Destroy command is used when the LSP created for the multicasting group is no longer needed. When the destroy procedure takes place, all branches within the multicast tree are torn down to end all data flow for the entire group. The Destroy command is issued through the Multicasting message. As the multicasting group is no longer needed, the root LSR sends a Multicasting message with a Destroy command indication to its directly connected LSRs, which will be forwarded to other downstream LSRs. The receiving LSRs will identify this command and will disconnect from its upstream LSR. This procedure continues until the Destroy command reaches the last LSR of the tree, which then disconnects from their upstream LSRs.

6.1.2.2. Hello Message Extensions

The multicasting extensions to the Hello message include the Multicasting TLV and the Tree TLV. The Multicasting TLV is used to inform the LSRs of the multicasting source and group IP address of the multicasting tree. The Tree TLV provides a list of multicasting tree LSR-RP addresses, such that LSRs that are not part of the multicasting

tree can identify these LSR-RP addresses for future connection purposes. Both TLVs are optional.

6.1.2.3. Notification Message Extensions

Notification messages are used to provide advisory information of a significant event to an LDP peer node [22], for example, the outcome of processing an LDP message or the state of the LDP session is informed using the Notification message. The multicasting extensions to the Notification message include the Multicasting TLV and the Tree TLV.

Based on applications of multicasting, Join and Leave messages require the root LSR to respond to their request with either permission for connecting to the multicasting tree or requesting procedures to disconnect using label release procedures. The multicasting extensions to the Notification message serve this purpose of communication between the root LSR and the LSR-RP that needs to conduct the connection or disconnection establishments to or from the multicasting tree.

6.1.2.4. Multicast Extensions to the Label Request Message

The Label Request Message allows the construction of the distribution trees [22]. The extensions to the Label Request message include an optional Multicasting TLV for multicasting applications of either responding to a Join command of a multicasting message or when the multicasting tree wants a LSR to join a multicasting tree.

6.1.2.5. Multicast Extensions to the Label Mapping Message

The label mapping procedure allocates single labels for each LSP that is requested by a Label Request Message. The multicasting extensions allow more than one outgoing label mapping for a specific incoming interface. If necessary, a multicasting TLV may be used as an option. More details of the Label Mapping message are provided in the following sections.

6.1.3. Multicasting Distribution Tree Construction Using RSVP-TE and LDP

In order to provide a mechanism for a point-to-multipoint LSP tree to be constructed, the calculation of the tree must be done by a mechanism that does not rely on external protocols or mechanisms. The calculation of the tree must be done before any other mechanisms for delivering data have been established. The tree can be constructed based on two different contexts: root-initiated tree calculation and leaf-initiated tree calculation. The root-initiated tree calculation should be used when a new source of multicasting traffic is going to start delivering traffic to all the members of a group. The leaf-initiated tree calculation is implemented when a new member of a group wishes to receive the multicasting traffic, implementing a request-driven scheme.

In addition, the multicasting extensions made to RSVP-TE and LDP are enabled to be independent of traditional IP-based multicast routing protocols, such as, DVMRP, PIM, CBT, etc. However, IGMP mechanisms will be used to provide the functionalities for establishing and maintaining multicasting group memberships.

Some key issues in constructing the tree are the TE parameters that provide the DiffServ [5] [15], and QoS features. With these considerations in mind, the tree can be calculated and constructed:

1. By means of source controlled routing.
2. By means of using traditional algorithms such as the Distance Vector (DV) or Link State (LS) algorithms that are based on distance oriented metric values.
3. By means of enhanced versions of the DV or LS algorithms that employ both TE parameters and distance oriented cost values as metric values in the tree calculations.

6.1.3.1. Root-Initiated Tree Calculation

The calculation and construction of the tree is performed based on the LSRs that have active members (hosts) that wish to receive multicasting traffic. We assume that they have been either identified in advance or they will be joining dynamically.

In order to properly calculate the tree, the first step is to find out if there is an IGMP group definition table constructed. This information is assumed to be found in every node in a Network Management System (NMS) database with listings of the group memberships, or in an LSR. With this information, the MIDB is created with the purpose of keeping track of all multicasting sources, distribution groups, and destinations. In addition, the MIDB allows the multicast information to be decentralized, making it more effective for dynamic group membership control. If this information is not available, the Hello mechanisms with multicasting extensions can be used to provide information to the MIDB.

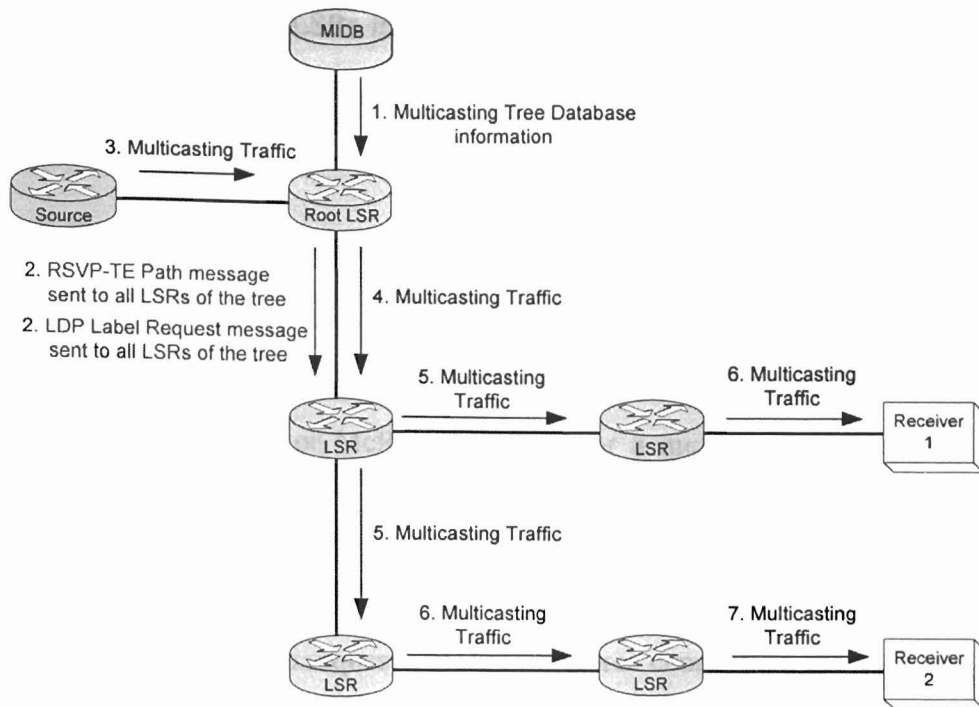


Figure 6.4 Root-initiated tree calculations.

With the information contained in the MIDB, the calculation of the tree can be done quickly based on one of the methods described above. Once the tree is calculated, a Path message is generated if RSVP-TE is used (or a Label Request message is generated if LDP is used). Once all the messages have been delivered and all the reservations are defined, the multicast traffic can be delivered to all the predefined nodes.

In Step 1 of Figure 6.4, the information gathered in the MIDB is used by the root LSR to construct the distribution tree. In Step 2, the root LSR will send out an RSVP-TE Path message (or an LDP Label Request message) to the LSRs of the multicasting tree to establish an LSP connection. This will be confirmed by an RSVP-TE Resv message (or an LDP Label Mapping message) at each link of the multicasting tree (not shown in the figure). After the tree connection is completed the source starts multicasting traffic

through the root LSR to the LSRs in the multicasting tree composed by Steps 3 through 7.

5.1.3.2. Leaf-Initiated Tree Calculation

A multicasting distribution tree is not static; hence, mechanisms that allow dynamic calculation of the tree have to be defined. In order to recalculate the tree, a node has to issue a Join, Leave, or McRecal message, or a multicasting message with the corresponding commands. The receiver-initiated tree is based on a predefined source already generating multicast content, and the information contained within the MIDB.

1. When a node wants to request a multicasting tree connection, the node will send a RSVP-TE Join message (or an LDP Multicasting message with a join command) to an LSR-RP MPLS router of the multicasting tree.
2. This LSR-RP router will send an RSVP-TE Join message (or an LDP Multicasting message with a join command) to the root LSR. This join request will verify the connection permission, and a recalculation of the tree may occur if necessary.
3. The multicasting root LSR will issue an RSVP-TE Join message (or an LDP Notification message) to the node that is requesting a connection. This LSR-RP will then issue an RSVP Path message (or an LDP Label Request message) to the new host for connection establishment to the tree.
4. If the procedure is not successful, the receiver making the join request will wait for an arbitrary timeout and retry.

In Figure 6.5, step 1 shows the RSVP-TE Join message and the LDP Multicasting message with join command indicating the new receiver's intention to receive

multicasting traffic. Step 2 shows the RSVP-TE Join message (or the LDP Multicasting message with Join command) generated from the LSR-RP that received the Join message requesting the root LSR to establish a multicasting connection. This in turn triggers the recalculation of the multicasting tree. Steps 3 and 4 show an RSVP-TE Join message (or an LDP Notification message) issued by the root LSR. At the LSR-RP, this Join message will be converted to a RSVP-TE Path message (or an LDP Label Request message) to enable an LSP connection to the multicasting tree (Step 5). Step 6 shows the RSVP-TE Resv message (LDP Label Mapping message) including the label to be used.

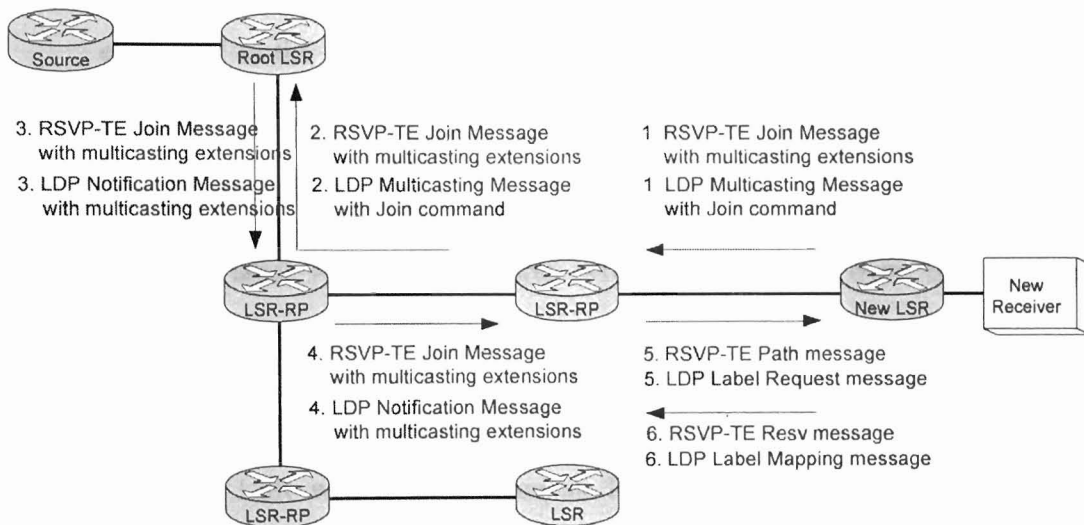


Figure 6.5 Leaf-initiated tree calculations.

6.1.3.3. Dynamic Updates to the Tree

Given the dynamic nature of the trees, constant adjustments to the tree have to be performed. When an LSR does not have any more multicasting receivers, it will issue an RSVP-TE Leave message (or an LDP Multicasting message with a Leave command),

which will be sent to the upstream LSR in order to stop the multicast traffic. The upstream LSRs will perform the following procedures:

1. If RSVP-TE is used, if the LSR that is leaving is not the last LSR within the tree, then the upstream LSR will erase the label mapping entry for the downstream LSR, and it will issue a McRecal message to the root LSR such that the root LSR can recalculate the tree should it be necessary. The information on the MIDB will also be updated. (If LDP is used and if the LSR that is leaving is not the last LSR within the tree, then the upstream LSR will send a Leave command to its upstream LSR-RP, where this LSR-RP will then issue a Leave command to the root LSR. The root LSR can recalculate the tree should it be necessary. The information in the MIDB will also be updated. The root LSR will send a Notification message to the LSR-RP. Following this, the LDP standard label release procedures will be conducted over the connection between the LSR-RP and the leaving LSR [22].)
2. If the LSR that is leaving is the last leaf LSR on the multicasting tree, then the upstream LSR will in turn issue an RSVP-TE Leave message (or an LDP Multicasting message with leave command) to the root LSR. The MIDB table will be updated. The root LSR in turn (after receiving the message and recalculating the tree) will trigger a recalculation procedure via an RSVP-TE McRecal message (or an LDP Notification message) so that all the nodes can use the updated information in the MIDB.
3. If the leaving LSR is the last one connected to the root LSR (that is, the root LSR is the last node on the tree), then the RSVP-TE Leave message procedure

(or the LDP Label Release procedures) will be conducted to disconnect the LSR. A destroy procedure will be used only when necessary, otherwise, the multicasting session will be kept alive, such that new users can establish a new multicasting tree if necessary.

4. In the event that the source has no more multicast content to multicast, it will issue an RSVP-TE Destroy message (or an LDP Multicasting message with a Destroy command) that will notify all the LSRs in the tree that the label mappings have to be released, and all the multicasting traffic forwarding should be stopped. This will also trigger a purge procedure within the MIDB, clearing all the entries for the specific multicasting group.

CHAPTER VII

CONCLUSION

The two control and management aspects investigated in this thesis are the OAM implementation in GMPLS and MPLS networking for enhanced multiplatform multicasting services.

OAM procedures can be implemented in GMPLS by extending the RSVP-TE and LDP protocol suites to perform OAM functionalities. Also extensions to the user plane OAM packets have been proposed. The addition of these OAM functionalities will enable carrier and backbone companies to manage and administer the GMPLS network easily and effectively.

The OAM implementation in GMPLS has extensions proposed for both the control and user planes. This thesis proposes the use of the OAM message in the control plane in order to indicate faults, for diagnostic tests and for simple performance monitoring. The extensions for RSVP-TE propose the inclusion of an OAM message that is capable of fault handling, performance monitoring and diagnostic testing using Objects such as the FM Object, Loopback Object, and the Activation-Deactivation Object. A set of operational procedures describing fault detection, fault localization, and diagnostic testing have also been illustrated. Similarly all of the above procedures indicated in

RSVP-TE have also been included in LDP. An OAM message has been proposed for LDP TLVs, such as the, FM TLV, Loopback TLV, and Activation-Deactivation TLV have been extended for these purposes. In the user plane we use the OAM alert label [13] for continuity checking and performance monitoring. We also tailor the OAM message in order to do performance monitoring over different networks such as ATM, FR, etc.

GMPLS with its common control plane concept requires the establishment of a common OAM control for all the different network architectures that could interface in the GMPLS domain. This proposal thus provides a simple, robust, and scalable OAM architecture for GMPLS which has not been proposed earlier.

The second topic deals with enhancing MPLS to provide multicasting services. Multicasting based on traffic engineering constraints is feasible by means of extending the capabilities of the RSVP-TE and LDP. RSVP-TE messages have been added and extended to enable MPLS with the required functionalities for multicasting services. Also, by using the LDP Traffic Engineering constraints, the necessary guarantees for end-to-end traffic delivery can be provided, allowing service providers or carrier companies to ensure customer data transmission to be effective and allow service agreements to be maintained for multicasting services.

The extensions proposed for RSVP-TE and LDP allows MPLS networking to conduct multicasting services independent of traditional IP-based multicast routing protocols. However, IGMP was used to serve the mechanism for establishing and maintaining multicasting group memberships.

For RSVP-TE, this thesis established Join, McRecal, and Leave messages for multicasting, and a multicasting distribution tree that is conformant with the constraints

defined by the MPLS TE parameters was created through RSVP-TE signaling. Furthermore, a set of operational procedures have been developed for the management of multicasting trees by combining the multicasting extensions for RSVP-TE with the IGMP features for multicasting group database control, and the Hello message with the MIDB extensions. For LDP, specifically, by including the Multicasting TLV, and through the development of Join, Leave, and Destroy commands as a part of the Multicasting message in conjunction with IGMP capabilities, the complete construction of a multicasting distribution tree using only IP multicast addressing information is possible.

By implementing functions needed for multicasting within the signaling protocols rather than using the multicasting routing protocols as middleware between the IP and data link layers, the overheads are dramatically reduced. Additionally, the proposed conceptualization allows all traffic engineering features of MPLS networking to be flexibly provided within the multicasting services in a very efficient, scalable, and straightforward fashion that has not been fully attempted in previous works.

CHAPTER VIII

REFERENCES

- [1] A. Adams, J. Nicholas, and W. Siadak, "Protocol Independent Multicast – Dense Mode (PIM-DM): Protocol Specification (Revised)," work in progress, Internet Draft, The Internet Society, Feb 2002.
- [2] A. Ballardie, "Core Based Trees (CBT)," RFC 2201, The Internet Society, Sept. 1997.
- [3] A. Ballardie, "Core Based Trees (CBT version 2) Multicast Routing," RFC 2189, The Internet Society, Sept. 1997.
- [4] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y. Rekhter, "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements," *IEEE Communications Magazine*, vol. 39, no. 1, January 2001.
- [5] B. Jamoussi, O. Aboul-Magd, P. Ashwood-Smith, F. Hellstrand, K. Sundell, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Halpern, J. Heinanen, T. Kilty, A. Malis, and P. Vaananen, "Constraint-Based LSP Setup using LDP," work in progress, Internet Draft, The Internet Society.
- [6] B. Williamson, *Developing Multicasting Networks*. Vol. I, Cisco Press, 2000.
- [7] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L. Wei, "Protocol Independent Multicast – Sparse Mode (PIM-SM) Protocol Specification," RFC 2362, The Internet Society, June 1998.

- [8] D. O. Awduche, L. Berger, D.-H. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209, The Internet Society, Dec. 2001.
- [9] D. Ooms, B. Sales, W. Livens, A. Acharya, F. Griffoul, and Fanfare, "Framework for IP Multicast in MPLS," work in progress, Internet Draft, The Internet Society.
- [10] D. Waitzman, C. Partridge, and S. Deering, "Distance Vector Multicast Routing Protocol," RFC 1075, The Internet Society, Nov. 1988.
- [11] E. Mannie, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," work in progress, Internet Draft, The Internet Society.
- [12] E. Rosen, and A. Viswanathan, "Multiprotocol Label Switching Architecture," RFC 3031, Internet Society, Jan. 2001.
- [13] H. Ohta, "Assignment of the 'OAM Alert Label' for Multiprotocol Label Switching Architecture (MPLS) Operation and Maintenance (OAM) Functions," RFC 3429, The Internet Society, Nov. 2002.
- [14] J. Garcia-Hernandez, and M. Ghanbari, "In-service monitoring of the quality of service in ATM networks using OAM cells," *Proc. IEE - Communications*, 146:2 (1999) pp. 102-106 1350-425.
- [15] J.-M. Chung, E. Marroun, H. Sandhu, and S.-C. Kim, "VoIP over MPLS Networking Requirements," *Proc. IEEE International Conference on Networking 2001 (ICN'01)*, Colmar, France, July 9-13, 2001.
- [16] J.-M. Chung, (Invited Paper) "Analysis of MPLS Traffic Engineering," *Proc. IEEE Midwest Symposium on Circuits and Systems 2000 (MWSCAS'00)*, East Lansing, Michigan, USA, Aug. 8-11, 2000.

- [17] J.-M. Chung, Subieta M. A Benito, H Chhabra, G. Y. Cho, and P. Rasiah, "RSVP Extensions for MPLS Multicasting Services," work in progress, Internet Draft, The Internet Society.
- [18] J.-M. Chung, M. A. Subieta Benito, H. Chhabra, G. Y. Cho, and P. Rasiah, "LDP Extensions for MPLS Multicasting Services," work in progress, Internet Draft, The Internet Society.
- [19] J. Moy, "Multicast Extensions to OSPF," RFC 1584, The Internet Society, Mar. 1994.
- [20] J. P. Lang, K.Mitra, J.Drake, K.Kompella, Y.Rakhter, L.Berger, D.Saha, D.Basak, H.Sandik, A.Zinin, and B.Rajagopalan, "Link Management Protocol (LMP)," work in progress, Internet Draft, The Internet Society.
- [21] K. C. Miller, *Multicast Networking and Applications*. Reading, MA: Addison-Wesley, 1999.
- [22] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas, "LDP Specifications," RFC 3036, The Internet Society, Jan. 2001.
- [23] L. Berger, "Generalized MPLS Signaling-RSVP-TE Extensions," work in progress, Internet Draft, The Internet Society.
- [24] M. Bischoff, M.N. Huber, O. Jahreis, and F. Derr, "Operation and Maintenance for an All-Optical Transport Network," *IEEE Communications Magazine*, November 1996, pp. 136 - 142.
- [25] N. Harrison, P. Willis, S. Davari, E.G. Cuevas, B. Mack-Crane, E. Franze, H. Ohta, T. So, S. Goldfless, and F. Chen, "Requirements for OAM in MPLS Networks," work in progress, Internet Draft, The Internet Society.
- [26] P.A. Smith, and L. Berger, "Generalized MPLS Signaling-CR-LDP Extensions," work in progress, Internet Draft, The Internet Society.

- [27] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification," RFC 2205, The Internet Society, Sept. 1997.
- [28] R. Handel, M.N. Huber, and S. Schroder, *ATM Networks Concepts, Protocols, Applications*. Singapore: Addison Wesley Longman, 1999.
- [29] R.H. Glitho, and S. Hayes, "Telecommunications Management Network: Vision vs. Reality," *IEEE Communications Magazine*, vol. 33, no. 3, Mar. 1995.
- [30] R. Whitmann, and M. Zitterbart, *Multicasting Communication Protocols and Applications*. San Diego, CA: Morgan Kaufman, 1999.
- [31] S. C. Farkouh, "Managing ATM-based Broadband Networks," *IEEE Communications Magazine*, vol. 31, no. 5, pp. 82 – 87, May 1993.
- [32] The Frame Relay Forum, "Frame Relay Operations, Administration, and Maintenance Implementation Agreement," Frame Relay Forum Technical Committee, Mar. 2001.
- [33] T.H.Wu and N.Yoshikai, *ATM Transport and Network Integrity*. San Diego, CA: Academic Press, 1997.
- [34] W. Stallings, *High-Speed Networks TCP/IP and ATM Design Principles*. Upper Saddle River, New Jersey: Prentice Hall, 1998.

✓

VITA

Pravin Rasiah

Candidate for the Degree of

Master of Science

Thesis: MPLS AND GMPLS NETWORKING CONTROL AND MANAGEMENT
TECHNOLOGIES

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Madras, India, on December 5th, 1977, son of J. R. Rasiah and Meena Rasiah.

Education: Received a Bachelor of Technology degree in Electronics Engineering from the Cochin University of Science and Technology University, India in May 1999. Completed the requirements for the Master of Science degree with a major in Electrical Engineering at the Oklahoma State University in December 2002.

Experience: Employed as a Research Assistant by the ACSEL & OCLNB Laboratories of the School of Electrical and Computer Engineering, Oklahoma State University (Aug. 2000 to Present).

Employed as a Teaching Assistant for Telecommunication Systems 1 at the School of Electrical and Computer Engineering, Oklahoma State University (Aug. 2001 - Aug. 2002).

Employed as a Customer Service Engineer in Wipro Infotech Pvt. Ltd., India (Aug. 1999 – July 2000)