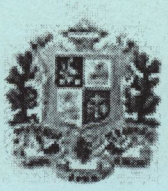


MULTI-PROTOCOL LABEL SWITCHING

BUILDING MPLS-BASED VIRTUAL PRIVATE NETWORKS AND SERVICES FOR SERVICE PROVIDERS CORE NETWORK

By

Mr. Charnvithya Sresthadatta

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Telecommunication Management

Assumption University

August, 2002

Project Title: Building Scalable Service Provider IP Network By MPLS Technology

Name: Mr.Charnvithaya Sresthadatta

Project Advisor: Dr.Sudhiporn Patumtaewapibal

Academic Year: 2002

The Faculty of Engineering, Assumption University had approved this final report of the six credit course, TM 6900 Master Project, submitted in partial fulfillment of the requirements for the degree of Master of Science in Telecommunications Management.

Approval Committee :

(Dr.Sudhiporn Patumtaewapibal)
Chairman / Advisor

(Asst.Prof.Dr.Kittiphan Techalcittiroj)
Member

(Assoc.Prof.Dr.Kobchai Dejhan)
MUA Representative

August 2002

Abstract

Businesses and organizations worldwide are embracing IP-based intranets and extranets to deliver their mission-critical applications. IP technologies give businesses the flexibility and ubiquity they need to distribute information to customers, suppliers, remote sites, and industry partners. While businesses recognize that their networks are a strategic asset, it is often less expensive and easier for them to outsource all or part of their networking to service providers, who are dedicated to provisioning and managing reliable, scalable network services that business customers need. Renting a service provider infrastructure and expertise allows companies to focus on their primary business.

Multiprotocol Label Switching (MPLS) gives providers the ability to offer highly scalable, advanced IP services end to end, with simpler configuration and management for both providers and end customers. Using MPLS, carriers can deliver the IP services that businesses demand, across either switched or routed networks. MPLS is the enabling technology that protects today's valuable revenue sources (Frame Relay and multiservice ATM), while paving the way for tomorrow's expanded services portfolio, of which private IP networking is the star.

Acknowledgements

I wish to thank my project advisor, Dr. Sudhiporn Patumtaewapibal, Dean of school of engineering, ABAC who has helped and guided me through the last steps of my master's degree.

And I am grateful to my instructor Dr. Kittiphan Techakittiroj, Chairperson of computer department, ABAC for his support and valuable comments. I would also like to special thank the members of school of engineering office for their support and arrangement.

I would like to gratefully acknowledge Mr. Ditaphon Chantraurai and Mr. Montri Mosa, my good friends. They help me as the consultant in this project. I would not have completed it without them.

Most importantly, I wish to express my deepest gratitude to my parents, father and mother for their love, encouragement, understanding and supports throughout my academic carrier.

Table of Contents

List of Figures

List of Tables

CHAPTER 1: Introduction

The demand for IP based applications along with rapid growth, the need for quality of service, reliability, and security have made traffic engineering an essential consideration in the design and operation of large public Internet backbone networks.

Moreover, the rapid proliferation of the Internet has established the IP protocol suite as the predominant networking technology. With IP's incumbent nature it is also predicted that the convergence of voice and video and data will happen over IP-based networks. To meet this end, the IETF (Internet Engineering Task Force) has chartered the MPLS (Multiprotocol Label Switching) working group to establish a standards based approach in developing MPLS as a means to achieve the benefits of this new technology that allows for increased performance in linking, routing and switching functions.

MPLS attempts to enhance traffic engineering over IP-based networks by combining elements of the Open System Interconnection Model (OSI), specifically between the Link Layer (Layer 2) and the Network Layer (layer 3). This framework for an integrated layer 2 and 3 routing paradigm is referred to as label switching. Packets are routed based on a size label, compared to the traditional IP network layer destination based routing. The Layer 3 protocols can be of the existing network layer protocols such as IP, IPX, Apple Talk, etc. Therefore, the label-switching scheme is referred to as 'multi-protocol'. MPLS's traffic engineering capabilities support existing forwarding platforms in both IP and ATM networks.

1.1 Problem Statement

The general objective of this study is to describe the business perspective for the IP netowrks service provider. In order to, it is important to first understand the basis on which, IP network technologies, service and applications. Only then can one explore the financial of the system and identify its inefficiencies, if any, and hence propose economically sound solutions to the problems. Specific objectives include:

1. To understand the background of IP network service provider, and how it works;
2. To indicate the current situation and trend of IP network technologies and their characteristics especially MPLS;
3. To describe how to migrate to MPLS and its applications such as VPN;
4. To present calculation of NPV, IRR, and Payback period method of new implemented technology, MPLS by service provider.

1.2 Structure of the project

The rest of the Chapters are organized as follow. Chapter 2 describes the background of IP network service of service providers and comparations between traditonal IP, ATM technology and MPLS. Chapter 3, this chapter describes the beginning and technology overview of MPLS. Chapter 4 describes the architectures of security of the MPLS architecture such as VPN. Chapter 5, describes how to migration to MPLS, Case study and Investment for service providers. Finally, which is the last chapter it suggests the recommendation and conclusion in Chapter 6.

1.3 Definitions

The authors find it important to introduce the reader to certain words and concepts. Below is therefore a brief explanation of the words and concepts most important to the reader of this paper.

Edge Label Switch Router (Edge LSR) – The edge device that performs initial packet processing and classification and applies the firs label. This device can be either a router or a switch with built-in routing.

Label Switch Router (LSR) – The core device that switches labeled packet according to precomputed switching Tables. It can also be a switch or a router.

Label Distribution Protocol (LDP) – Provides communications between edge and core devices. It assigns labels in the edge and core devices to establish Label Switched Paths (LSPs) in conjunction with routing protocols such as OSPF, IS-IS, Enhanced Interior Gateway Routing Protocol, or BGP.

Label Switched Path (LSP) -- Path defined by all labels assigned between end points. And LSP can be dynamic or static.

Label – A label is a header used by an LSR to forward packets. The header format depends upon network characteristics. In router networks, the label is a separate, 32-bit header. In ATM Networks the label is placed into the virtual channel identifier/virtual path identifier (VCI/VPI) cell header. In the core, LSRs read only the label, not the packet header. One key to the scalability of MPLS is that labels have only local significance between tow devices that are communicating.

Net Present Value (NPV) - The present of a project's future cash flow less the cost of the initial investment (Brigham, Gapenski, Ehrhardt Financial Management Theory and Practice 1999).

Internal Rate of Return (IRR) - The discount rate, which equates the present value of a project's expected cash inflows to the present value of the project costs (Brigham, Gapenski, Ehrhardt Financial Management Theory and Practice 1999).

Payback Period - The number of years required for a project's after tax cash inflows to accumulate to an amount that covers the initial investment (Brigham, Gapenski, Ehrhardt Financial Management Theory and Practice 1999).

CHAPTER 2: Multi-Protocol Label Switching, the introduction of new technology MPLS to service providers

2.1 Introduction

Service providers today are experiencing unprecedented demand for IP services. According to a report by Forrester Research, Inc., the entire U.S. Internet Services industry in 1996 amounted to $1.3 billion. By 1998, just the business segment of the market reached nearly $4 billion. For Thailand there are more than 2.3 million users of Internet in year 2001 and growth rate about 100% of Internet traffic from year 2001 to 2002. Service providers have responded to this demand by building out nation- and world-spanning networks. Since 1996, some service providers have had to double their backbone bandwidth nearly every four months just to keep up. Looking ahead, according to Forrester Research projections, the market for IP services in the U.S. will reach $57 billion by 2003, rivaling the amount businesses will spend on long-distance calls.

As of October 2001:		
Internet users: (3./ users/100)		2.3 million
Internet hosts: (4,584 hosts/10000)		million
Internet domains (.th):	7,000	d nit] ins
Internet domains (total estimate)	14,000	do nain
Schools online with SchoolNet:	3,259	schools
International bandwidth:	570	Mbps
Domestic backbone bandwidth:	921	Mbps
Domestic exchange traffic:	1,147	Gbytes/day

Table 2-1: Internet usage of Thailand

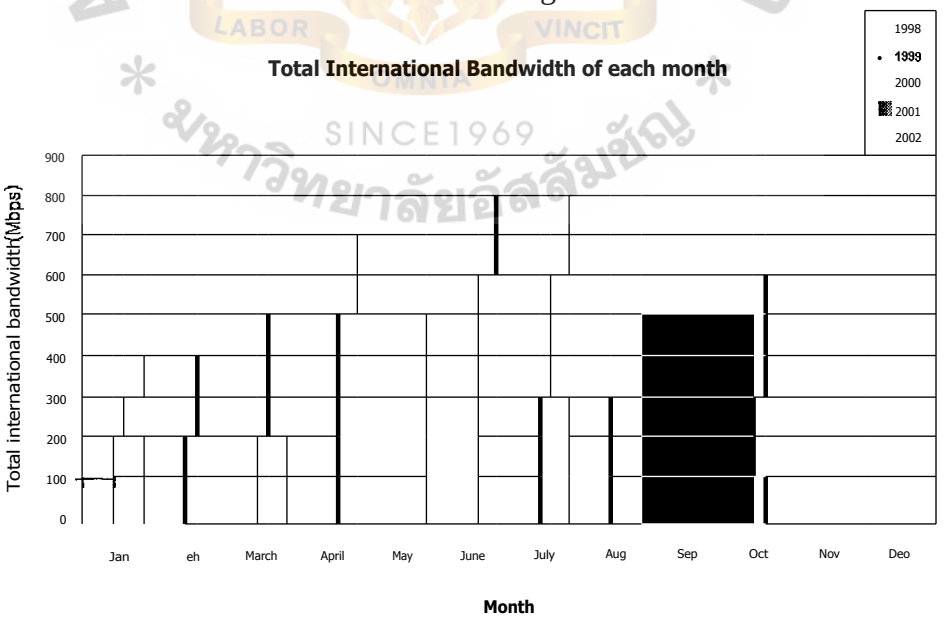


Figure 2-1 Chart of International Internet Bandwidth of Thailand (www.nectec.or.th)

Factors contributing to the demand for bandwidth include:

- Broadband access is reaching the American residential market. By 2003, 15% of American homes will have purchased broadband access to the Internet.
- More and more business is being done on the Internet. Firms in every industry, indeed all organizations, are looking at the Internet as a way to improve business processes and/or reduce the cost of doing business with partners. The big three automakers - GM, Ford and Daimler- Chrysler, for example - have announced they'll form an online market with each other and their suppliers to drive down costs. In many cases, completely new approaches to traditional business practices are being introduced (e.g., Amazon.com, eBay, Yahoo! just to name a few).
- The growth rate of traffic on the Internet is about 100% per year, a much higher growth rate than for traffic on other networks. If present growth trends continue, data traffic in the U. S. will overtake voice traffic around the year 2002, and will be dominated by the Internet2 (some analysts believe this may have already occurred).
- In addition to the U.S., the number of Internet users around the world is growing rapidly as well. The Computer Industry Almanac3 has reported that by the year 2002, 490 million people around the world will have Internet access, or 79.4 per 1,000 people worldwide. That ratio grows to 118 people per 1,000 by year-end 2005. And despite the fact that the U.S. has an overwhelming lead in Internet users - nearly 43 percent of the total 259 million worldwide - The U.S. will have only one-third of the total Internet users in 2002, and that number is expected to decline to 27 percent by the end of 2005.
- Much of the current explosion in demand is fuelled by applications that exploit the "best effort" nature of today's Internet. But there is growing demand for services that require a higher level of capability, specifically higher predictability from the Internet. Such services include: commercial Layer 3 VPNs, intranets, extranets, out-sourced firewalling and encryption and Voice Over IP (VoIP) services. As the Internet changes into a new public network that supports these new applications, growth in demand for these "higher level" services is expected to accelerate.
- Service Level Agreements (SLAs) written to meet Layer 2, Layer 3, and even Layer 4 parameters are being requested by customers to support new, emerging bandwidth-hungry applications. Once the Internet is able to support these requirements, demand for these applications is expected to increase dramatically.
- Web servers and, increasingly, application servers are not being located on an organization's own site, but rather at locations with high bandwidth and good connectivity to the core of the Internet.
- The demands of bandwidth from Applications Service Providers alone is expected to more than quadruple the current level of bandwidth demand in order for services to meet the performance requirements of their customers.

In this hyper-growth environment, service providers must find a way to accommodate the dramatic growth in network traffic and the number of users. To do so in a cost-effective way, they must add management capabilities and higher predictability to their IP networks. Predictability is critical to optimizing network capacity and providing premium revenue-generating IP services. At the same time, both the capital costs of growing the network and operational costs of offering an ever-widening selection of services must be minimized. Further, the solutions that service providers choose today must have a clearly articulated migration plan to incorporate future opportunities and technologies.

In many large IP networks, service providers have deployed connection-oriented ATM network cores to optimize bandwidth utilization and increase IP service predictability via ATM traffic engineering and infrastructure resilience. Though this has worked well, and indeed has enabled the current explosive growth of the Internet, there is a desire to simplify network operations by reducing the number of control planes operating in the network (currently ATM and IP).

One emerging technology, Multi-Protocol Label Switching (MPLS), has been widely identified as a new tool to help service providers meet the often-conflicting challenges of increased predictability, growth in revenue, and cost reduction. As will be discussed in this white paper, it is MPLS's connection-oriented nature that provides an ability to increase IP service predictability, create differentiated IP services, and potentially reduce operation costs in IP-centric and multi-service networks.

To achieve all of this, MPLS combines a variety of functions from both IP and ATM. Specifically, MPLS adds enhancements to IP routing protocols to make them connection-oriented. In short, MPLS aims to provide a Connection-Oriented Link Layer (COLL) for IP that results in reliable and predictable forwarding of IP traffic, and that enables traffic engineering, congestion management, optimized end-to-end transmission recovery, and differentiated IP services. Further, MPLS, when augmented with a QoS framework such as that specified in the IETF's Differentiated Services model, may enable deterministic QoS in IP-centric networks.

MPLS is the natural evolution required for networks to support predictable and optimized IP services, particularly in next-generation, IP-centric networks. The connection-oriented nature of MPLS aims to help service providers meet unprecedented customer demand and their own revenue and profit goals.

2.2 Internet Demand and Service Provider Network Evolution

In the last ten years, the Internet has grown from a small network of interconnected routers to a world-spanning network that global businesses are coming to rely on as a mission-critical tool. Table 2-2 shows a high-level overview of the control plane protocols involved, the data plane data-transfer format involved, IP's strengths that led in part to the success of the early Internet, and the limits of IP in today's rapidly expanding and changing Internet.

	IP
Network control plane	
Admission control	None
Routing	OSPF, IS-IS, BGP4
Path computation	None—per hop forwarding
Signaling	None—per hop forwarding
Connection Name	None
Connection ID	None
Explicit Routing	None
Network data plane	
Transmission unit	Packets (variable length)
Policing (for fairness)	None
Marking	None
Buffer allocation	Limited
Scheduling (for flow prioritization and fairness)	Limited-none set by protocol standards
Strengths in an IP-Centric Network	Flexibility; Rich suite of data-service protocols, UNIX OS integration; Multi-vendor standards based implementation
Limitations in an IP-Centric Network	Limited support for differentiated, predictable services; Connectionless hop-by-hop routing creates congestion (hyper-aggregation) and under-utilization of network resources.

Table 2-2: IP' s Characteristics

In just the last few years, the pace of the Internet's growth has seriously strained the capabilities of the traditional routed infrastructure. The concern for service providers has quickly become: *scaling the network to meet the growing demands, while improving service availability and minimizing end-to-end latency and operational costs.*

2.3 Traditional Routing

Every technology has its advantages and disadvantages. While being connectionless brings a number of well-known benefits to IP - for example, scalability and overall network resiliency - it has some drawbacks, most notably:

1) A tendency towards "hyper aggregation" of data on certain links, which leads to congestion,

2) a limited ability to alleviate hyper-aggregation by, for example, distributing traffic load over all available resources, and

3) an inability to provide "toll quality" service levels across a network end-to-end.

All three limitations are due to IP' s connectionless nature, whereby traffic is transported on a hop-by-hop basis, with routing decisions made at every node.

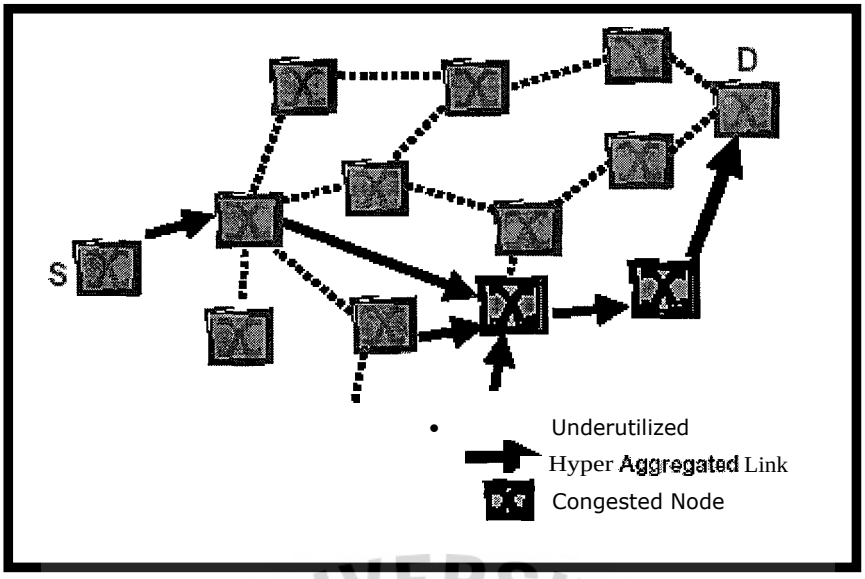


Figure 2-2: Hyper-Aggregation in an IP Network

Without a COLL, IP can create this kind of congestion in a network. Routers using the Open Shortest Path First (OSPF) routing protocol, for example, base routing decisions on the destination IP address of a packet's header, along with the least-cost path to that destination. All traffic then takes this least-cost path, congesting that path and leaving other paths through the network underutilized (see Figure 1-2). OSPF gives routers no end-to-end, overall view of the network and, therefore, the routers aren't aware of congestion in the network or of lightly loaded alternate routes and can't make the best use of all available network resources. Some of the larger ISPs claim that they lose up to forty percent of their network's capacity due to poor use of network resources by connectionless IP. By adding a COLL, they regain that capacity.

IP's hop-by-hop prioritization schema has another drawback: the inability to select paths through the network that guarantee the QoS requirements of latency-sensitive traffic flows. In an IP network, a real-time voice call or videoconference is routed by IP the same way as e-mail retrievals or busty file transfers, so all three may experience congestion under conditions of hyper-aggregation.

While that may be fine for non-real-time data traffic like e-mail, a voice call or videoconference call has requirements for low-latency that must be met end-to-end across the network, from source to destination. IP routing protocols can't *guarantee* that these requirements will be met. Therefore, service providers running Ipcentric networks can't make these guarantees on a network that scales over time in bandwidth, users, sites, and applications. They are limited to offering a "best effort" service. With no premium services to offer, IP service providers today are limited to charging only flat, commodity rates for the services they provide.

In addition, due to its connectionless nature, IP cannot guarantee fast per-flow reroute times for large-scale networks. Consequently, it's difficult for IP service providers to offer any guarantee of network availability to customers looking for the "dial tone" service they've come to take for granted from the other large-scale network supporting their business: the telephone network.

In summary, IP is one of the most powerful networking technologies ever created, due to its adherence to open standards and its flexibility as a networking technology applicable to the transfer of a wide range of data types. However, in service provider networks, an IP routing limit service providers' ability to engineer

and manage traffic in the network, and also limits the kind of service levels they can offer their customers. What's required is a Connection-Oriented Link Layer (COLL) that is aware of the end-to-end state of the network, routes traffic based on the requirements of the application/user sending the traffic, and allows service providers to load-balance traffic across all available links in order to optimize the use of network resources. All of these would allow service providers to use a given network infrastructure as efficiently as possible, while making and *meeting* commitments to their customers. Today, that's available with ATM and is emerging with MPLS.

The Benefits of a COLL include:
• Network devices that have knowledge of the network state, including the bandwidth available on each link, • Devices that have knowledge of areas of congestion. in the network. • Devices that use end-to-end load balancing for optimal network bandwidth • Virtual bandwidth partitioning that manages congestion and supports multi-tier service levels on common links, Optimized and predictable re-route times..

Table 2-3: Benefits of COLL technology

2.4 ATM in an IP-Centric Network

Many large-scale IP network operators have enhanced their IP service by incorporating ATM (Asynchronous Transfer Mode), a connection-oriented networking solution, in their networks. ATM optimizes network capacity through a Connection-Oriented Link Layer (COLL) that provides knowledge of the end-to-end state of the network in order to use the bandwidth on all available links optimally. Further, ATM's connection-oriented nature enables virtual bandwidth partitioning to avoid congestion, and optimizes reroute times in the case of network failures. In addition, with current backbone interface speeds of OC-12c/STM-4 and OC-48c/STM-16, ATM has helped service providers meet the growing demand for Internet capacity.

	IP	ATM
Network control plane		
Admission control	None	UNI
Routing	OSPF, IS-IS, BGP4	PNNI
Path computation	None—per hop forwarding	End-to-End, constraint-based and congestion-aware
Signaling	None—per hop forwarding	PNNI
Connection Name	None	Virtual -Connection
Connection ID	None	VPI, VCI
Explicit Routing	None	Designated. Transit Lists
Network data plane		
Transmission unit	Packets (variable length)	Cells or packets (ATM Forum FAST)
Policing (for fairness)	None	Yes, for multiple traffic contracts
Marking	None	Cells are marked conform or non-conform
Buffer **allocation**	Limited	Per flow reservations
Scheduling (for flow prioritization and fairness)	Limited-none set by protocol standards	Per port, per flow, per class
Strengths in an IP-Centric Network	Flexibility; Rich suite of data-service protocols, UNIX OS integration; Multi-vendor, standards based. implementation	Network predictability and reliability; Mature, field-hardened solutions; Connection Oriented, Layer2/Layer3 network partitioning; Optimized network bandwidth utilization; End-to-end load balancing
Limitations in an IP-Centric Network	Limited support for differentiated, Predictable services; Connectionless hop-by-hop muting creates congestion (hyper-aggregation) and under-utilization of network resources:	Additional control plane to manage; Lack of ATM integration in routers results in a large number of router adjacencies to manage.

Table 2-4: Shows a high-level comparison of IP and ATM characteristics

Figure 2-3 illustrates a frequent implementation of ATM in a service provider network. The core, or backbone, of many service provider networks is made up of an intelligent mesh of ATM switches. This Connection-Oriented Link Layer (COLL) core provides the primary transport for IP traffic. Surrounding the ATM core is a ring of IP routers. Behind those rings of core routers, there are likely several other types of networks. They may be the next tiers of the Internet or service provider network hierarchy. In some cases they may be PoP (Points of Presence) LANs, where subscribers access the network as well as Web-hosting content and any applications (e.g., ERP) hosted at that location.

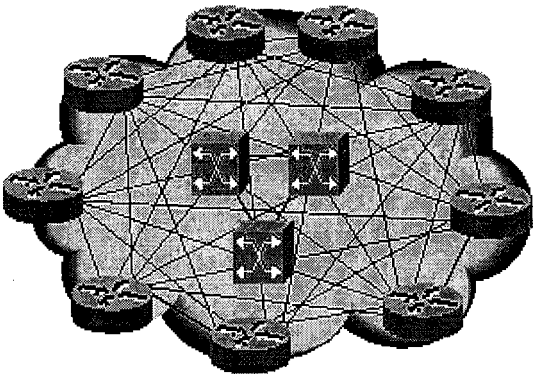


Figure 2-3: The Common Full-Mesh Configuration of Service Provider Networks

This network partitions Layer 2 and Layer 3 network functions, enabling the network to efficiently handle traffic at the networking layer best suited to the job. While the Layer 2 ATM switches provide the "big, fat pipes" to quickly and intelligently move data, the Layer 3 IP routers perform the IP routing and forwarding functions critical to the IP-centric network. In Figure 2-3, IP routers at the edge route data over the COLL provided by the ATM switches at the core. This offloads processor-intensive work from the routers, letting them simply map packets to the COLL of ATM virtual circuits - they now only have to deal with traffic as it enters or exits the core.

In the traditional routed architecture, routers within the core also handle "transit" traffic. The amount of transit traffic increases with the size of the network, thereby stressing the routers. Offloading this traffic significantly enhances the scalability of the network.

2.5 Constraint-Based and Congestion-Aware Routing

Further, the ATM COLL provides two key functions that enable more predictable IP flows: minimizing congestion, and optimizing network capacity, or constraint-based routing and congestion-awareness. A device using constraint-based routing routes traffic based on traditional network topology information along with a number of other constraints, including the capacity and utilization of links, the requirements of the flow itself (i.e. bandwidth, delay, and jitter) and other administrative constraints. It may be used to guarantee specific applications (like video-conferencing) a fixed amount of bandwidth end-to-end through the network. It may also be used to minimize latency and jitter for voice traffic and to provide very specific, guaranteed and quantifiable customer service levels. This ability to configure varying ranges of quality of service to different customers of the network is also an attractive method for service providers to offer their customers differentiated services.

A congestion-aware device uses traditional routing information, but also takes into account the current state of the network with respect to traffic loading on each and every link. Congestion awareness provides network nodes the base information required to dynamically load-balance traffic through the network, and optimizes traffic so as to avoid hyper-aggregation or under-utilization of links.

Constraint-based routing and congestion awareness are critical to ATM's connection-oriented performance. When data is sent across an ATM network, an end-to-end connection carries that data and takes into account:

1. The state of the network (available links, bandwidth available, guarantees available, etc.) and the most efficient routes through it.

2. The latency and bandwidth requirements of the application (or user) sending the data.

3. Preferred routes that have been configured previously by the network manager.

2.6 Service Availability and Network Resiliency

One of the ways service providers implement an ATM core is using Switched Permanent Virtual Circuits (SPVCs) between switches edge to edge. Using SPVCs, the network administrator simply designates the beginning and ending point of an ATM connection, while the rest of the connection in between is set up by the switches themselves, thanks to the high level of data each switch has about the network via PNNI.

Through PNNI, each switch has an extensive view of the network. Not only is it aware of the entire network topology, including the links, but also the speeds of those links, the current utilization of each, their delay, and other parameters that PNNI uses in routing updates. After performing sophisticated route computations with this information, PNNI selects an entire end-to-end path, a connection, which it encodes (by way of a Designated Transit List) into the signaling message. It then forwards application data out of the switch interface corresponding to the first switch entry on the path (this call set-up time in some vendor implementations is on the order of milliseconds. Once a call is set up, data transits a switch in microseconds). Any subsequent packets in that communication follow the same connection through the network.

SPVCs minimize operator involvement in configuring the network core and in provisioning services (in some vendor implementations, service provider customers themselves can determine and prioritize the switched routes used). In the event of a node or link failure, SPVC connections are automatically re-routed by the switches in milliseconds in the order of the connection priority set up by the administrator. Because the switches handle this rerouting themselves, it's very fast--on the order in some vendor's implementations of milliseconds. So fast that the ATM backbone of the service provider network has automatically detected a link failure and rerouted around it before the edge IP routers are even aware that there is a problem. This provides stability and reliability to the IP network that service provider customers can count on. Thus, ATM SPVCs, fast call set-up, and connection prioritization enable service providers to easily provision their network, offer high network availability to their customers and, potentially, offer premium-priced differentiated services.

In addition, because the ATM COLL provides congestion awareness, SPVCs also enable load balancing of traffic across multiple links in the network. Rather than try to send a number of large data transfers down a congested path, for example, SPVCs allow the network manager to balance multiple flows across various links in the network and transfer the information without creating congestion.

Since the flows are switched via various end-to-end connections, some vendors optimize fairness between flows via per virtual connection (per VC) queuing and scheduling. An ATM virtual connection is the end-to-end connection over which an application's data flows through the network. By queuing and scheduling traffic into and out of device buffers on a per-VC basis, each VC is treated by the network the way it needs to be. So, for example, data generated by a videoconference is given

high priority through device queues in order to provide smooth streaming images for the videoconference attendees. Per-VC queuing and scheduling can minimize delay for small packet flows at network congestion points, making sure delay-sensitive applications get treated in a way that optimizes their performance.

With per-VC queuing and buffering, it is also possible to partition network capacity and link bandwidth via connections that have specific, user-provided constraints. This partitioning allows operators to manage network traffic, protecting some high priority flows from the burstiness and congestion-creating volume of other flows.

This does not necessarily require policing of flows, but again is based on vendor implementation of per-VC queuing and prioritized scheduling.

2.7 A Limitation of IP Over ATM: Management of Multiple Control Planes

In an IP-centric network, network engineers and managers must configure, provision, and manage at least an IP topology. IP-over-ATM networks have an additional control plane to manage: the ATM control plane. This requires network managers to manage, provision, and control an ATM infrastructure in addition to their IP topology.

For some IP-centric operators, this control plane separation satisfies their desire to manage the backbone and infrastructure of the network separately from the IP service and access network, providing an additional level of reliability and stability to a critical part of the network (the backbone). For other IP-centric operators, the effort to manage this additional control plane either is not required. Perhaps highly predictable traffic flow is not required for the service profile being offered -- best-effort Internet access, for example, or is beyond the perceived benefits of the COLL. As we'll see in a moment, for these providers, MPLS has the potential to solve this problem.

Another limitation for some is the much publicized "cell tax" created when breaking down IP packets into 53-byte cells comprised of 48 byte payloads and 5 bytes of overhead. A new, open standard solution is "frame ATM", which uses variable-length ATM frames rather than fixed-length cells to transmit data. In the case of frame ATM, the cell tax is effectively "repealed." This is helpful for IP-centric providers who trunk traffic at 0C-3c or 0C-12c rates.

On the other hand, some operators are not concerned with this overhead for one of two reasons: either the traffic management advantages of ATM (including capturing underutilized network bandwidth) outweigh the additional overhead or, for the access portion of multi-service networks (those aggregating voice, video and data over the same links, for example, at speeds below 622 Mbps), they find the fixed 53-byte size of a traditional ATM cell is required to ensure all traffic receives the service it requires from the network, particularly in terms of delay.

2.8 MPLS: Connection-Oriented Networks

As a connection-oriented technology, MPLS:

- Enables IP-centric networks to be more predictable and efficient through load sharing of traffic across multiple links in a network, and by using a COLL that enables network resources to be used more efficiently,
- Enables network managers to avoid hyper-aggregation scenarios by providing some level of traffic engineering and traffic management,

- Enables service providers to offer higher levels of service to their customers by allowing high priority flows in the network to receive higher prioritization than others,
- May enable service providers to offer high network availability if the network elements can provide fast rerouting of Label Switched Paths, and
- Reduces the number of control planes to be managed in an IP-over-ATM network.

2.9 MPLS Makes IP-Centric Networks More Predictable and Reliable

MPLS introduces more predictability and reliability to IP-centric networks through traffic engineering and traffic management functionality enabled by MPLS's standards-based extensions to IP routing protocols. Those extensions effectively provide IP traffic the benefits of a COLL. The MPLS control plane sets up MPLS connections - known as Label Switched Paths (LSPs) - from ingress Label Switching Routers (LSRs) at the network edge through the core LSRs to form a connection across a service provider's network.

MPLS running on IP routers enables network managers to assign traffic to LSPs based on information about the end-to-end state of the network. That avoids the hyper-aggregation and under-utilization of IP routing by directing traffic away from congestion-prone links onto less-congested links. This allows alternate paths to be configured, either manually or automatically, across both IP and ATM portions of the network for fast rerouting in the case of node failure, providing high network availability.

In addition, as the MPLS standard matures and field-tested solutions emerge, MPLS may enable applications like voice and real-time video to get the network resources they require. An MPLS Forwarding Equivalence Class (FEC) enables customer traffic to be mapped to Label Switched Paths for high class of service traffic, thus enabling service provider networks to meet the latency and delay-tolerance requirements of these delay-sensitive applications. Thus, MPLS could enable service providers to make "high-margin promises" in a Layer 2 or Layer 3 Service Level Agreement (SLA), then confidently *meet* those promises in real-world networks.

2.10 MPLS Provides a Single Control Plane to Manage

When IP routers and ATM switches are both running MPLS, both the edge and core of the network may operate using the MPLS control plane. MPLS Label Switched Paths (LSPs) that pass through the entire network - from IP routers through ATM switches and back into routers - can be set up using only MPLS routing protocols. No longer will edge routers be running OSPF or IS-IS while the core ATM switches run PNNI. Now both sets of devices will operate using MPLS routing protocols so service providers have only one control plane to manage.

2.11 MPLS Signaling Protocols: RSVP, RSVP-TE, LDP and CR-LDP

As MPLS becomes a more robust standard and products emerge that run MPLS, different networking solution vendors will offer slightly different versions of MPLS. As customers begin to choose among these different MPLS offerings, one item of interest will be which signaling protocols a vendor supports.

The two primary options are RSVP (Resource Reservation Protocol) and LDP (Label Distribution Protocol). When choosing an MPLS vendor, it is especially important to select one that has implemented the open standards- based MPLS traffic engineering, traffic management, Quality of Service, and Constraint-Based Routing extensions not only to the IP routing protocols, but also to the signaling protocol.

An exciting possibility for MPLS signaling is its use to support the IETF IntServ (Integrated Services) QoS (Quality of Service) standard. Rather than simply using RSVP-TE or LDP to signal and maintain MPLS flows, these signaling protocols could be used for QoS management of traffic on aggregated trunks, enabling the MPLS COLL to provide QoS for IP traffic.

This work is fairly new, but looks to become quite important. Many vendors plan to integrate IntServ QoS and DiffServ with their RSVP-TE and LDP implementations.

	IP	MPLS	ATM
Network Control Plane			
Admission control	None	Not yet set forth in the MPLS standard	UNI
Routing	OSPF, IS-IS, BGP4	OSPF-TE, IS-IS-TE, BGP4-TE	PNNI
Path computation	None—per hop forwarding	End-to-End, constraint-based and congestion-aware	End-to-End, constraint-based and congestion-aware
Signaling	None—per hop forwarding	RSVP-TE, CR-LDP	PNNI
Connection Name	None	Label Switched Path	Virtual Connection
Connection ID	None	Label II)	VPI, VCI
Explicit Routing	None	Explicit Route Objects	Designated Transit Lists
Network data plane			
Transmission unit	Packets (variable length)	Packets and/or Cells	Cells or packets (ATM Forum FAST)
Policing (for fairness)	None	None	Yes, for multiple. traffic contracts
Marking	None	None	Cells are narked conform or non-conform
Buffer allocation	Limited	Not yet set forth in the MPLS standard	Per flow reservations
Scheduling (for flow prioritization and fairness}	Limited-none 'set by protocol standards	Not yet set forth in the MPLS standard	Per part, per flow, Or class
Strengths in an IF Centric Network	Flexibility; Rid] Suite of data-service protocols, UNIX OS integration; Multi-vendor, standards based implementati on	Efficient provisioning; Predictability and reliability; Support for differentiated services and SLAs; Optimized network bandwidth utilization; End-to-end load balancing	Network predictability and reliability; Mature, field-hardened solutions; Connection Oriented, Layer2/Layer3 network partitioning; Optimized network bandwidth utilization; End-to-end load balancing
Limitations in an IP-Centric Network	Limited support for differentiated services; Limited predictability, *Less* optimized networks; Connectionless hop-by-hop routing creates congestion (hyper-aggregation) and under-utilization of network resources.	An emerging standard; Little. eld experience; Lack of policing minimizes ability to guarantee fairness and throughput; No current plans for multi-service support or interoperability, therefore cannot currently be proposed as ubiquitous, common backbone for multi-service network operators..	Additional control plane-to manage; Lack of ATM integration in routers results in a large number of router adjacencies to :manage.

Table 2-5: Gives a high-level summary of the characteristics of IP, MPLS and ATM.

2.12 Operating an MPLS Network

Like ATM VCs before them, MPLS LSPs allow service providers to increase control over their networks and, potentially, also offer various levels of service (e.g., gold, silver, bronze) to customers. LSPs can be set up by the network administrator, for example, to reduce the chances of hyper-aggregation in the network. On the other hand, since LSRs are aware of bandwidth on all links in the network, LSRs can set up LSPs that route traffic in ways that make most efficient use of all the links in the

network, avoiding congestion entirely. In addition, LSRs can set up paths that are constrained by various user or application requirements. For example, if a user has purchased a "gold" level of service, high-priority LSPs could be set up for their data traffic to ensure that they always get the gold level of service from the network. In this way, MPLS may enable the incorporation of the IETF's DiffServ standard in service provider networks. LSPs can be set up to operate in several different ways:

• Point-to-Point switched paths, for example, can be used to connect all ingress nodes to all egress nodes to transport unicast traffic.

• A multipoint-to-Point LSP can connect all ingress nodes to a single egress node. This allows many microflows to take the same path through the network when they've been assigned to one Forwarding Equivalence Class (FEC).

• A multipoint-to-multipoint switched path can be used to combine multicast traffic from multiple sources into a single multicast distribution tree through the network.

• LSPs can also be tunnels, using label switching rather than network-layer encapsulation (L2TP, PPTP) as the means of moving packets through the tunnel.

In addition, MPLS allows streams of data to be forwarded as a unit, along a LSP. Thus a LSP through the network can be a single flow of user data or an aggregate flow of many users' data. MPLS uses the term "Forwarding Equivalence Class" or FEC to refer to a set of Layer-3 packets that are forwarded in the same manner by a particular MPLS node. The mapping of IP packet to a FEC occurs only once per LSP, at the ingress LSR of the path. The LSRs in the core of the network simply switch on the MPLS header already applied by the ingress LSR.

MPLS is also a possible mechanism for provisioning VPNs. Packets coming from a customer network, for example, could contain an encapsulated header with a VPN label. At the service provider VPN ingress node, the header could be removed and a VPN label applied for switching through the VPN. At the VPN egress, the VPN label would be removed and the original label would provide label switching through the customer site.

2.13 Misunderstandings About MPLS

One misunderstanding about MPLS is that it is "just IP". Another misunderstanding is that it is "just like ATM" and the two are mutually exclusive and/or that running both in a network is redundant. Neither statement could be further from the truth. Instead, MPLS is a Connection-Oriented Link Layer (COLL) over which IP can run. As such, it brings some of the benefits of connection-orientation to an IP-centric network. MPLS is a kind of "common ground" between IP and ATM (an OSI "Layer 2.5", if you will). It replaces neither IP nor ATM, but is a new networking tool that optimally solves a certain set of network problems. It combines functions of IP with functions of ATM to provide a new tool for 21st century networking solutions.

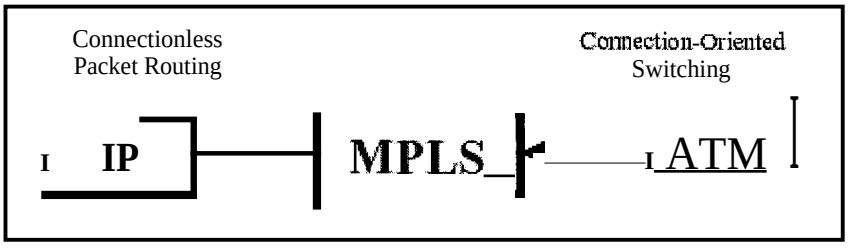


Figure 2-4: MPLS—A New Complimentary Networking Tool

MPLS, however, is a new standard and not yet field-tested. As such, it provides some level of COLL benefits today to IP services and networks, with more promised through future standardization. Some of those benefits include:

• In an all Packet Over SONET data network, MPLS can provide connections to enable a level of traffic engineering and traffic management that IP alone cannot provide. This is the type of network to which MPLS can add the most value.

• MPLS operates over frames or cells so it can operate in IP-over-ATM networks as well as IPover- MPLS networks. This "protocol agnosticism" enables service providers to continue to run the IP over ATM networks they have today and slowly add MPLS to parts of the network as their business plans require. Additionally, this can enable a step-by-step approach to migrating an IP over ATM network to all MPLS without entailing the risk of a hard cutover from one type of network to the other.

• MPLS has some congestion awareness and its routing protocols (IP protocols with traffic engineering extensions) are constraint-based. Hence, it should be able to provide some level of traffic engineering and management that reduces hyper-aggregation and should be able to provide some level of Class of Service prioritization to traffic. When combined with the IETF's Differentiated Services model, it may also enable service providers to guarantee a higher level of service to customers who require it.

• In an IP over ATM network, MPLS reduces the two control planes needed to run the network-IP and ATM- to one control plane, that of MPLS. This simplification of network operation promises lower operational costs for service providers, enabling them to achieve higher profit levels from the network infrastructures they've already put in place.

Additionally, there is room left for MPLS to grow to provide all the benefits of a COLL. For example:

• Unlike ATM, which has quantitative, "hard" QoS capabilities built into standards, as well as built into vendors' implementations, MPLS currently holds only the promise of providing IP traffic some level of "soft" QoS. Likely this will at first be some level of high or low prioritization given on a qualitative Class of Service (CoS) basis.

• Both ATM and MPLS are sensitive to calls per second setup performance for optimizing service availability during re-routes around failures. Best-of-breed vendor ATM implementations, for example, can perform a reroute in less than 50 milliseconds. It remains to be seen how quickly MPLS Label Switched Paths (LSPs) can be rerouted.

• The MPLS standards have not specified standardized COS/QOS parameters, as has already been done with ATM. As such, multi-vendor interoperability for MPLS-based circuits (LSPs) providing COS or QOS may not be available in the first phase of MPLS network rollouts.

MPLS is a new standard. Vendor implementations on routers will not provide all the traffic management benefits seen on some vendors' ATM or MPLS switches. These benefits include traffic policing, traffic shaping, and hierarchical scheduling, which enable network operators to further engineer, manage, and control traffic in their networks.

Because of this, it's important to remember, that all vendors' MPLS implementations will not be equal. Taking advantage of the benefits of the COLL provided by MPLS requires more than just the implementation of the MPLS routing protocol. It also requires hardware that can provide fast setup of connections (MPLS calls-persecond), per Label Switched Path and micro-flow queuing, and per LSP and micro-flow scheduling. Only vendor offerings that implement this level of MPLS COLL can offer all the advantages that MPLS promises.

2.14 Integrating MPLS Into an IP Over ATM Network

MPLS offers the promise of enormous competitive advantages to service providers. Integrating MPLS into existing networks, however, should be done with care and, if possible, in a smooth, incremental way. Using the industry-standard "Ships In the Night" mode of operation can make such a migration from today's IP over ATM networks to MPLS a smooth one. Ships In the Night also allows service providers to provide non-IP transit services such as private line, frame relay, voice and ATM over the same physical networking infrastructure.

Ship in the Night operation allows a step-by-step migration between IP over ATM and MPLS without requiring a build-out of a parallel physical network. It allows service providers to create two logical networks on top of one physical topology by allowing IP over ATM and MPLS to run concurrently on the same devices. This is a very powerful configuration option that permits the same physical port to be configured with both IP/ATM and MPLS control planes at the same time. The multiple control plane protocols are able to function over a single physical network and are completely aware of each other.

This means that service providers will be able to add MPLS incrementally to their existing networks and not have to build a completely separate network for it. So rather than having to go to MPLS all at once, a service provider can simply "turn on" MPLS in small portions of the network as test beds, see how operations are effected, and move on from there in a methodical, controlled way, while customer traffic runs through the network.

In addition, there are many network designs where the backbone will need to support both ATM and MPLS simultaneously for an indefinite amount of time. This may be due to multiple services sold or provisioned, or due to varying technology support among multiple network edge devices (especially where multiple vendors are concerned). Ship in the Night operation in this situation presents a compelling solution.

CHAPTER 3: The Beginning and Technology Overview of Multi-Protocol Label Switching

3.1 Introduction

This chapter describes the Multiprotocol Label Switching (MPLS) distribution protocol. MPLS is a high-performance packet forwarding technology that integrates the performance and traffic management capabilities of data link layer (Layer 2) switching with the scalability, flexibility, and performance of network-layer (Layer 3) routing. It enables service providers to meet challenges brought about by explosive growth and provides the opportunity for differentiated services without necessitating the sacrifice of existing infrastructure.

The MPLS architecture is remarkable for its flexibility:

- Data can be transferred over any combination of Layer 2 technologies
- Support is offered for all Layer 3 protocols
- Scaling is possible well beyond anything offered in today's networks.

Specifically, MPLS can efficiently enable the delivery of IP services over an ATM switched network. It supports the creation of different routes between a source and a destination on a purely router-based Internet backbone. Service providers who use MPLS can save money and increase revenue and productivity.

3.2 The Evolution of Multilayer Switching in the Internet

Multilayer switching describes the integration of Layer 2 switching and Layer 3 routing. Today, some ISP networks are built using an overlay model in which a logical IP routed topology runs over and is independent of an underlying Layer 2 switched topology (ATM or Frame Relay).

Layer 2 switches provide high-speed connectivity, while the IP routers at the edge—interconnected by a mesh of Layer 2 virtual circuits—provide the intelligence to forward IP datagrams. The difficulty with this approach lies in the complexity of mapping between two distinct architectures that require the definition and maintenance of separate topologies, address spaces, routing protocols, signaling protocols, and resource allocation schemes. The emergence of the multilayer switching solutions and MPLS is part of the evolution of the Internet to decrease complexity by combining Layer 2 switching and Layer 3 routing into a fully integrated solution.

3.3 Fundamental Building Blocks

Before beginning our discussion of multilayer switching in the Internet, it is important to understand the fundamental building blocks common to all multilayer switching solutions and MPLS:

- Separation of the control and forwarding components.
- Label-swapping forwarding algorithm.

3.4 Separation of Control and Forwarding Components

All multilayer switching solutions, including MPLS, are composed of two distinct functional components—a control component and a forwarding component

(see Figure 3-1). The control component uses standard routing protocols (OSPF, IS-IS, and BGP-4) to exchange information with other routers to build and maintain a forwarding table. When packets arrive, the forwarding component searches the forwarding table maintained by the control component to make a routing decision for each packet. Specifically, the forwarding component examines information contained in the packet's header, searches the forwarding table for a match, and directs the packet from the input interface to the output interface across the system's switching fabric.

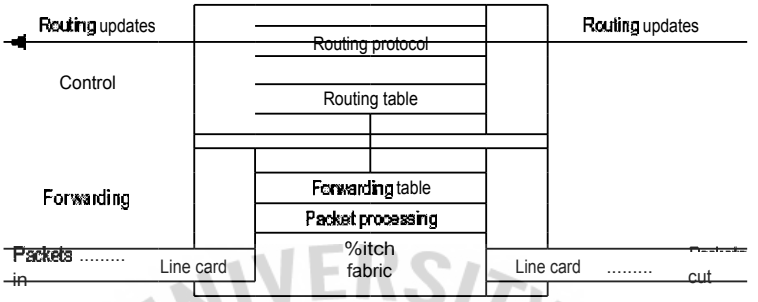


Figure 3-1: Routing Functional Components: Control and Forwarding

By completely separating the control component from the forwarding component, each component can be independently developed and modified. The only requirement is that the control component continues to communicate with the forwarding component by managing the packet-forwarding table. We will see that the deployment of an extremely simple forwarding algorithm, such as label swapping, can provide the extended forwarding capabilities needed to support new revenue-generating customer services.

3.5 Label-Swapping Forwarding Algorithm

The forwarding component of virtually all multilayer switching solutions and MPLS is based on a label-swapping forwarding algorithm. This is the same algorithm used to forward data in ATM and Frame Relay switches. Signaling and label distribution are fundamental to the operation of the label-swapping forwarding algorithm, but they are not discussed in this paper.

A label is a short, fixed-length value carried in the packet's header to identify a Forwarding Equivalence Class (FEC). A label is analogous to a connection identifier, such as an ATM VPI/VCI or a Frame Relay DLCI, because it has only link-local significance, does not encode information from the network layer header, and maps traffic to a specific FEC. An FEC is a set of packets that are forwarded over the same path through a network even if their ultimate destinations are different. For example, in conventional longest-match IP routing, the set of unicast packets whose destination addresses map to a given IP address prefix is an example of an FEC.

The label-swapping forwarding algorithm requires packet classification at the ingress edge of the network to assign an initial label to each packet. In Figure 3-2, the ingress label switch receives an unlabeled packet with a destination address of 192.4.2.1. The label switch performs a longest-match routing table lookup and maps the packet to an FEC—192.4/16. The ingress label switch then assigns a label (with a value of 5) to the packet and forwards it to the next hop in the label-switched path (LSP).

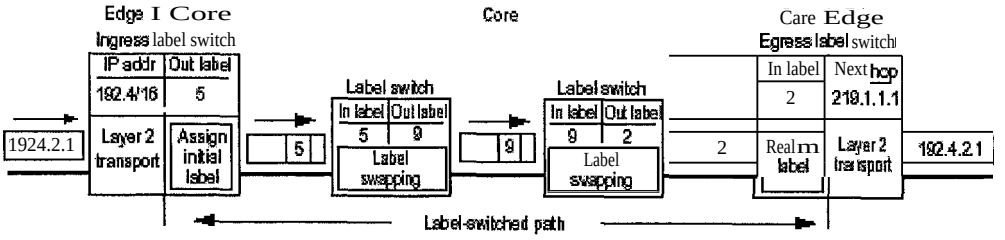


Figure 3-2: Packet Traversing a Label Switched Path

An LSP is functionally equivalent to a virtual circuit because it defines an ingress-to-egress path through a network that is followed by all packets assigned to a specific FEC. The first label switch in an LSP is called the ingress, or head-end, label switch. The last label switch in an LSP is called the egress, or tail-end, label switch.

In the core of the network, label switches ignore the packet's network layer header and simply forward the packet using the label-swapping algorithm. When a labeled packet arrives at a switch, the forwarding component uses the input port number and label to perform an exact match search of its forwarding table. When a match is found, the forwarding component retrieves the outgoing label, the outgoing interface, and the next-hop address from the forwarding table. The forwarding component then swaps (or replaces) the incoming label with the outgoing label and directs the packet to the outbound interface for transmission to the next hop in the LSP.

When the labeled packet arrives at the egress label switch, the forwarding component searches its forwarding table. If the next hop is not a label switch, the egress switch discards the label and forwards the packet using conventional longest-match IP forwarding.

Label swapping provides a significant number of operational benefits when compared to conventional hop-by-hop network layer routing:

Label swapping gives a service provider tremendous flexibility in the way that it assigns packets to FECs. For example, to simulate conventional IP forwarding, the ingress label switch can be configured to assign a packet to an FEC based on its destination address. However, packets can also be assigned to an FEC based on an unlimited number of policy-based considerations—the source address alone, the application type, the point of entry into the label-swapping network, the point of exit from the label-swapping network, the CoS conveyed in the packet header, or any combination of the above.

Service providers can construct customized LSPs that support specific application requirements. LSPs can be designed to minimize the number of hops, meet certain bandwidth requirements, support precise performance requirements, bypass potential points of congestion, direct traffic away from the default path selected by the IGP, or simply force traffic across certain links or nodes in the network.

The most important benefit of the label-swapping forwarding algorithm is its ability to take any type of user traffic, associate it with an FEC, and map the FEC to an LSP that has been specifically designed to satisfy the FEC's requirements. The deployment of technologies based on label-swapping forwarding techniques offer ISPs precise control over the flow of traffic in their networks. This unprecedented level of control results in a

network that operates more efficiently and provides more predictable service.

3.6 ISPs Migrate to the IP-over-ATM Model

In the mid-1990s, certain ISPs evolved their networks from router-based cores to the overlay model of running IP over ATM. ISPs undertook this migration because they needed greater bandwidth, deterministic forwarding performance, and traffic engineering to support the explosive growth occurring in their networks. One of the primary reasons that the IP-over-ATM overlay model was able to satisfy these operational requirements was ATM's use of a label-swapping forwarding algorithm.

The IP-over-ATM model (see Figure 3-3) was centered around ATM functionality, requiring ATM software controls (signaling and routing) and hardware forwarding (label swapping) on every system in the core of the network. The IP-over-ATM model met application requirements by using Layer 3 functionality at the edges of the network and maximized network throughput by relying on high-speed, label-swapping ATM switches and PVCs in the core. The role of IP routing was limited to the edges of the network because this model viewed software-based routers as the key source of poor network performance.

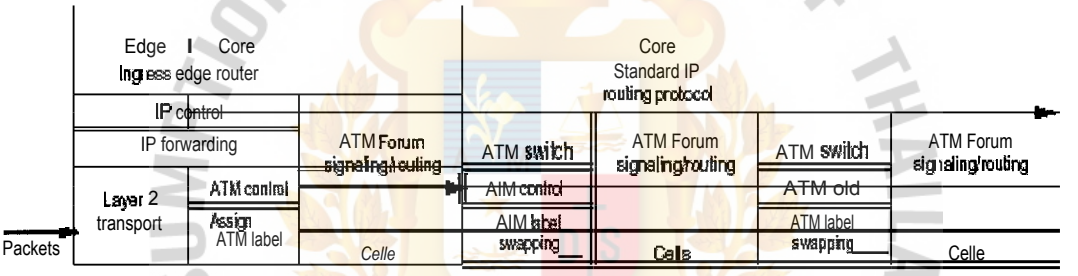


Figure 3-3: IP Over ATM Model

In the mid-1990s, networking equipment was not specifically designed for Internet backbone applications so ISPs were forced to adapt whatever equipment was commercially available to support their rapidly growing networks. Only the ATM switching infrastructure equipment provided the bandwidth and forwarding capacity to support their immediate requirements.

However, as ISPs continued their exponential growth and purpose-built equipment became available for Internet applications, continuing with the IP-over-ATM model and its inherent scalability problems made less sense. The scalability problems included the bandwidth limitations of ATM SAR interfaces, the 20 percent cell tax, the "n-squared" PVC problem, the Interior Gateway Protocol (IGP) stress, and the inability to operate over a non-ATM infrastructure.

Despite these scaling issues, the most challenging problem was the complexity of operating a network based on two disparate technologies that were independently designed and developed for entirely different tasks. IP and ATM are based on completely different protocol architectures (connectionless vs. connection-oriented), and they each have their own addressing models, routing protocols, signaling protocols, and resource allocation schemes.

While rapidly growing ISPs required the performance and control that ATM and label-swapping provided, they realized that in a packet-based network environment, it became more difficult to justify the complexity of the IP-over-ATM

model. Today, when high-performance Internet backbone routers are purpose-built for the core of the Internet, there are few good reasons to continue with an overly complex approach that requires two separate sets of equipment.

3.7 Multilayer Switching Alternatives to the IP-over-ATM Model

As ISPs continued migrating to the IP-over-ATM model, a number of technical, marketing, and financial trends began to influence the development of new technologies designed for the core of the Internet. The general public began to understand the Internet's prominent role in providing the foundation for a new global economy. It became clear that the Internet market was large enough to build equipment specifically designed for Internet backbone applications.

IP quickly became the only protocol that mattered, winning out over IPX, AppleTalk, OSI, and SNA. The notion of "IP convergence" provided venture capitalists and start-ups with a window of opportunity to compete with incumbent vendors. To be successful, a start-up needed to deliver a solution that provided the price and performance of an ATM switch and the control of an IP router, while eliminating the complex mapping required by the IP-over-ATM model. By late 1996, a number of vendors were promoting proprietary multilayer switching solutions that integrated ATM switching and IP routing, including:

IP Switching designed by Ipsilon/Nokia
Tag Switching developed by Cisco Systems
Aggregate Route-Based IP Switching (ARTS) designed by IBM Corporation
IP Navigator delivered by Cascade/Ascend/Lucent Technologies
Cell Switching Router (CSR) developed by Toshiba

Although these approaches had a number of characteristics in common, they were not interoperable because each relied on different technologies to combine IP routing and ATM switching into an integrated solution. However, by early 1997, many in the Internet community were impressed with the simplicity and elegance of these solutions that they began to view multilayer switching as the next logical evolutionary step for the design of large ISP backbone networks.

3.8 Similarities Among the Multilayer Switching Solutions

Each of the multilayer switching solutions sought to combine the best properties of IP routing and ATM switching, while still maintaining an IP focus. The fundamental by these strategies was to take the control software from an IP router, integrate it forwarding performance of a label-swapping ATM switch, and create an cost efficient IP router (see Figure 3-4).

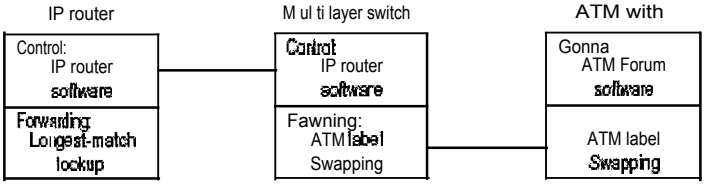


Figure 3-4: Multilayer Switch as a Fast IP Router

For the control component, each multilayer switch ran standard IP routing software (OSPF, IS-IS, and BGP-4) and a proprietary label-binding mechanism. The routing software permitted multilayer switches to exchange Layer 3 network reachability information. The label-binding mechanism mapped Layer 3 routes to labels (that is, to ATM VPI/VCIs) and distributed them to neighbors to establish LSPs across the core of the network. Running routing protocols on core systems rather than just edge systems provided a number of benefits that enhanced network operation:

- Eliminated the IP-over-ATM model's "n-squared" PVC scaling problem
  Reduced Interior Gateway Protocol stress by dramatically decreasing the number of peers that each router had to maintain
  Permitted information about the core's actual physical topology to be made available to Network layer routing procedures

For the forwarding component, multilayer switches used conventional ATM switching hardware and label swapping to forward cells across the core of the network (see Figure 3-5). However, the control procedures that assigned the labels to routes, distributed the labels among multilayer switches, and created the forwarding tables were managed by proprietary IP-based protocols, not ATM Forum protocols. ATM label swapping in the core of the network provided a number of benefits:

Label swapping optimized network performance by leveraging the benefits of hardware-based forwarding. Vendors believed that this would promote the creation of a new generation of products that offered superior price—performance ratios while dramatically reducing the time to market. While this was a critical issue in the mid-1990s, it is not one of the lasting benefits of multilayer switching because technological advances have provided better approaches to building Internet backbone routers.

Label swapping made explicit routing practical. An explicit route is a pre-configured sequence of hops that describes the path that traffic should take across a service provider's network, thus permitting the construction of a forwarding path that is different from the one typically created by destination-based routing. Explicit paths provide ISPs precise control over traffic flows, making it possible to support traffic engineering, QoS, and loop prevention.

Label swapping provided an instrument to extend control beyond the limitations of conventional destination-based routing. Multilayer switching's ability to provide enhanced forwarding control beyond that supported by traditional routing mechanisms is its lasting contribution to network design. Later in this paper we examine how multilayer switching facilitates the deployment of new types of routing functionality.

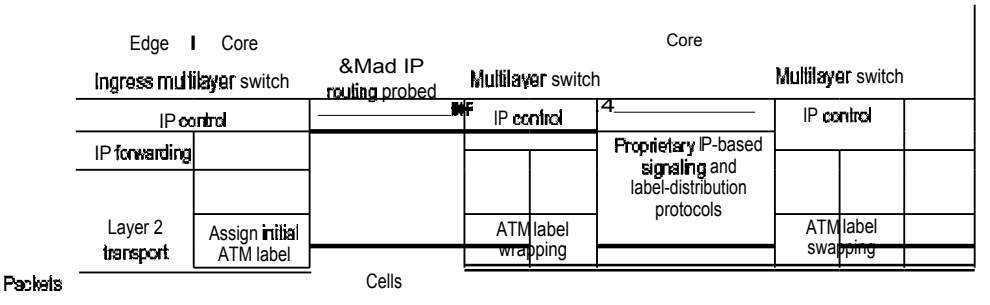


Figure 3-5: Multilayer Switching Solutions—Integrating IP Routing and ATM Switching

By excluding the ATM Forum's routing and signaling protocols, multilayer switching reduced operational complexity by eliminating the need to coordinate and map between two different protocol architectures—IP and ATM. Although multilayer switching still used standard ATM VPI/VCIs as labels, they were assigned and distributed using proprietary IP-based protocols rather than the standard ATM Forum protocols. This streamlined the integration of IP routing with ATM switching by eliminating the need to perform complex mapping between architectures. However, a critical limitation of the majority of multilayer switching solutions was that they were restricted to running over a cell-based ATM infrastructure, when the Internet was becoming increasingly packet-oriented.

3.9 Fundamental Differences Among the Multilayer Switching Solutions

While the various multilayer-switching solutions had numerous features in common, they relied on two fundamentally different approaches to initiate the assignment and distribution of label bindings to establish LSPs:

- Data-driven model
- Control-driven model

3.10 Data-Driven Model

In the data-driven model, label bindings are created when user data packets arrive. A flow is a sequence of packets that have the same source and destination IP addresses and TCP or UDP port numbers. A multilayer switch can either create a label binding as soon as it sees the first packet in traffic flow or wait until it has seen a number of packets in the flow. The benefit of waiting for a number of packets ensures that the flow is long enough to merit the overhead of assigning and distributing a label. Multilayer switching solutions that implemented the data-driven approach were IP Switching (Ipsilon) and the Cell Switching Router (Toshiba). Note that MPLS does not support the data-driven model.

The advantage of the data-driven model is that a label binding is created only when there is a traffic flow that uses the label binding. However, this model has a number of limitations for deployment in the core of a large ISP network, where there can be an enormous number of individual traffic flows:

Each multilayer switch must provide sophisticated and high-performance packet classification capabilities to identify traffic flows.

Typically, there is latency between the recognition of a flow and the assignment of a label to the flow. This means that each multilayer switch must also support longest-match IP forwarding during the setup phase so packets that have not been assigned to a flow can be forwarded and not dropped.

The amount of control traffic needed to distribute label bindings is directly proportional to the number of traffic flows.

The presence of a significant number of relatively short-lived flows can impose a heavy burden on network operations.

Conventional wisdom dictates that the data-driven model does not have the scaling properties required for application in the core of the Internet.

3.11 Control-Driven Model

In the control-driven model, label bindings are created when control information arrives. Labels are assigned in response to the normal processing of routing protocol traffic, control traffic such as RSVP traffic, or in response to static configuration. Multilayer switching solutions that implemented the control-driven model were Tag Switching (Cisco Systems), **IP** Navigator (Ascend/Lucent), and ARIS (IBM). In addition, MPLS uses the control-driven model.

The control-driven model has a number of benefits for deployment in the core of a large ISP network:

- Labels are assigned and distributed before the arrival of user data traffic. This means that if a route exists in the IP forwarding table, a label has already been allocated for the route, so traffic arriving at a multilayer switch can be label swapped immediately.
- Scalability is significantly better than in the data-driven model, because the number of label switched paths is proportional to the number of entries in the IP forwarding table, not to the number of individual traffic flows. For traffic engineering in large ISP networks, scaling could be even better—proportional to the number of exit points in the network. Label assignment based on prefixes, rather than individual flows, permits a single label to represent a highly aggregated FEC.

  In a stable topology, the label assignment and distribution overhead is lower than in the data-driven model because label-switched paths are established only after a topology change or the arrival of control traffic, not with the arrival of each "new" traffic flow.

  Every packet in a flow is label switched, not just the tail-end of the flow as in thedata-driven model.

3.12 Fundamental Problem with Multilayer Switching Solutions

Each multilayer switching solution maintained the IP control component and used ATM label swapping as the forwarding component. The challenge facing the ISP community was that each solution was proprietary and therefore not interoperable. Also, the majority of multiplayer switching solutions required an ATM transport because they could not operate over mixed media infrastructures (Frame Relay, PPP, SONET, and LANs). If multilayer switching was to be widely deployed by ISPs, there had to be a multivendor standard that could run over any link layer technology. In early 1997, the IETF established the MPLS working group to produce a unified and interoperable multilayer switching standard.

3.13 The Beginning of MPLS

The initial deployment of the Internet addressed the requirements of data transfer over the network. To meet the requirements a simple device based on router platform include the interface to support the existing connection at low speed was sufficient. Since the increasing demand of higher speed and the ability to support high-bandwidth of transmission, the technologies and devices with capabilities to switch at the layer 2 – Data Link layer and the layer 3 – Network layer had to be concerned.

The rapid growth of Internet and the rapid growth of Internet and increase in real-time and multimedia applications have created a need to improve Internet routing technology in terms of bandwidth, performance, scalability and delivery of new functionalities. There are several technologies involving applications of layer 2 switching technology to layer 3 routing have been made to counter above challenges. The efforts in these directions are Ipsilon's IP Switching, Cisco's Tag Switching, IBM's ARTS, Toshiba's CSR and MPLS.

3.13.1 IP Switching

Ipsilon Networks Inc has developed *IP Switching* technology. The goal of Ipsilon is to make IP faster and offer the quality of service support. The "IP over ATM" approach tries to hide the underlying network topology from IP layer by treating the datatlink layer as large, opaque network cloud. However, this leads to inefficiency, complexity and duplication of functionality in the resulting network Ipsilon's approach is to discard the connection-oriented ATM software and implement the connectionless IP routing directly on the top of ATM hardware. This approach takes the advantage of robustness and scalability of connectionless IP and speed, capacity and scalability of ATM switches.

IP switch is basically an IP router with attached switching hardware that has the ability to cache routing decisions in switching hardware. To construct an IP switch, a standard ATM switch is taken, the hardware is left untouched, but all the control software above AAL-5 is removed. It is replaced by standard IP routing software, a flow classifier to decide whether to switch a flow or not and a driver to control the switch hardware. At system startup a default virtual channel is established between the control software of the IP switch and its neighbors, which is then used for default hop-by-hop forwarding of IP datagrams. To gain the benefits of switching, a mechanism has been defined to associate IP flows with the ATM labels.

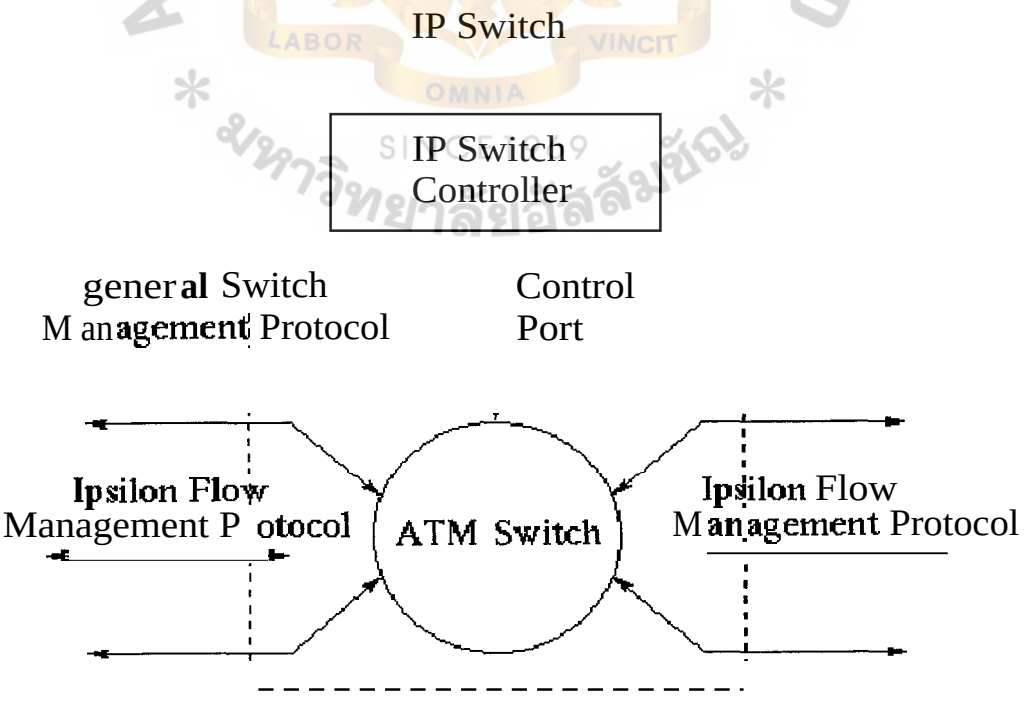


Figure 3-6: Structure of an IP Switch

3.13.2 Tag Switching

This technology has been developed by Cisco Systems Inc. to enhance routing in terms of bandwidth, scalability, support to newer routing functionalities, multicast, hierarchy of routing knowledge and flexible routing control. It uses the label (tag) switching technology for layer 3 packets forwarding. A Tag switching network consists of Tag Edge Routers and Tag Switching Routers, with packet tagging being the responsibility of the edge router. Standard **IP** routing protocols are used to determine the next hop for traffic. Tags are matched to routes in a routing table and distributed to destination via a Tag Distribution Protocol.

3.13.3 Cell Switched Router

Cell Switched Router (CSR) technology has been introduced by Toshiba Corporation, Japan and presented to the IETF in 1994. Toshiba has proposed CSR based Internetworking Architecture, which tries to merge the two approaches by extending the current routers to handle resource reservation and **IP** flows using ATM's cell switching capabilities. It was one of the earliest public proposals for using IP protocol to control an ATM switching fabric. CSR is designed to function as a router for connecting logical IP subnets in a classical IP over ATM environment. Label switching devices communicate over standard ATM virtual circuits. CSR labeling is data-driven.

The Flow Attribute Notification Protocol (FANP) is used to identify the dedicated VCs between CSR's and to establish the association between individual flows and individual dedicated VCs. The objective of the CSR is to allow cut through forwarding of flows to switch the ATM cell flow that constitutes the packet rather than reassembling it making an IP level forwarding decision on it.

3.13.4 Aggregate Route Based IP Switching (ARIS)

IBM has introduced this concept. The goal of ARIS is to improve aggregate throughput of IP and other network layer protocols by switching datagrams at the media speed. It is similar architecturally to Tag Switching. ARIS binds labels to aggregate routes rather than flow. Label binding and label switched paths are set up in response to control traffic rather than data flows, with the egress router generally the initiator. Routers that are ARIS capable are called Integrated Switch Routers (ISR).

ARIS was designed with a focus on ATM as the Data link layer. The ARIS Protocol is a peer-to-peer protocol that runs between ISRs directly over **IP** and provides a means to establish neighbors and to exchange label binding. A key concept of ARIS is the "egress identifier".

Since multiple proprietary solutions for label-based switching is clearly not an acceptable direction, it was recognized that standards were needed and that an IETF Working Group had to be formed. A charter was agreed to the IETF in early 1997 and the inaugural meeting of the working group was held in April 1997. Then the term "Multi-protocol Label Switching (MPLS)" was selected as the vendor independent name for the set of standards that will be produced

The Internet Draft MPLS Framework states that the goal of standardization is to "integrate the label swapping forwarding paradigm with network layer routing" with an initial focus on IP v4 and IP v6. MPLS provides the mechanism and these can be applied in various ways according to the network's needs.

Draft standards are not expected until the end of 1998, although vendors are already working on implementations. Those who building large MPLS-based IP networks and fully exploit the benefits of MPLS can be expected to become leaders in the next wave of inter-network expansion.

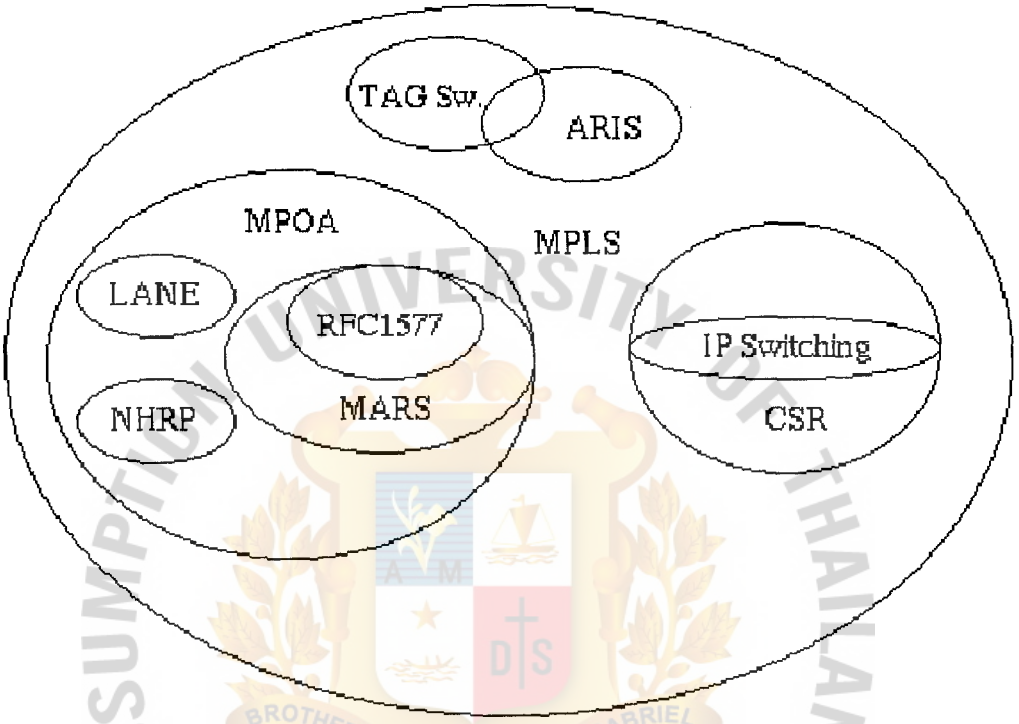


Figure 3-7: Approach in Layer 3 Switching

3.14 Benefits of MPLS

MPLS provides the following major benefits to service provider networks:

- Scalable support for SVirtual Private Networks (VPNs)—MPLS enables VPN services to be supported in service provider networks, thereby greatly accelerating Internet growth. The use of MPLS for VPNs provides an attractive alternative to the building of VPNs by means of either ATM or Frame Relay permanent virtual circuits (PVCs) or various forms of tunneling to interconnect routers at customer sites.

  Unlike the PVC VPN model, the MPLS VPN model is highly scalable and can accommodate increasing numbers of sites and customers. The MPLS VPN model also supports "any-to-any" communication among VPN sites without requiring a full mesh of PVCs or the backhauling (suboptimal routing) of traffic across the service provider network. For each MPLS VPN user, the network of the service provider appears to

function as a private IP backbone over which the user can reach other sites within the VPN organization, but not the sites of any other VPN organization. From a user perspective, the MPLS VPN model enables network routing to be dramatically simplified. For example, rather than needing to manage routing over a topologically complex virtual backbone composed of many PVCs, an MPLS VPN user can generally employ the backbone of the service provider as the default route in communicating with all of the other VPN sites.

• Explicit routing capabilities (also called constraint-based routing or traffic engineering)--Explicit routing employs "constraint-based routing," in which the path for a traffic flow is the shortest path that meets the resource requirements (constraints) of the traffic flow. In MPLS traffic engineering, factors such as bandwidth requirements, media requirements, and the priority of one-traffic flow versus another can be taken into account. These traffic-engineering capabilities enable the administrator of a service provider network to perform the following tasks:

– Control traffic flow in the network
– Reduce congestion in the network
– Make best use of network resources

Thus, the network administrator can specify the amount of traffic expected to flow between various points in the network (thereby establishing a traffic matrix), while relying on the routing system to perform the following tasks:

– Calculate the best paths for network traffic
– Set up the explicit paths to carry the traffic

• Support for IP routing on ATM switches (also called IP and ATM integration)—MPLS enables an ATM switch to perform virtually all of the functions of an IP router. This capability of an ATM switch stems from the fact that the MPLS forwarding paradigm (namely, label swapping) is exactly the same as the forwarding paradigm provided by ATM switch hardware.

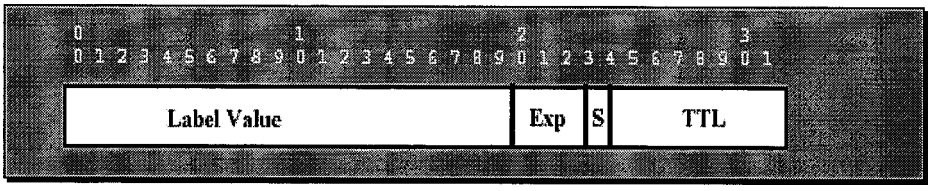
The key difference between a conventional ATM switch and an ATM label switch is the control software used by the latter to establish its virtual channel identifier (VCI) table entries. An ATM label switch uses IP routing protocols and the TDP to establish VCI table entries. An ATM label switch can function as a conventional ATM switch. In this dual mode, the ATM switch resources (such as VCI space and bandwidth) are partitioned between the MPLS control plane and the ATM control plane. The MPLS control plane provides IP-based services, while the ATM control plane supports ATM-oriented functions, such as circuit emulation or PVC services.

3.15 Label Switching Functions

In conventional Layer 3 forwarding mechanisms, as a packet traverses the network, each router extracts all the information relevant to forwarding the packet

from the Layer 3 header. This information is then used as an index for a routing table lookup to determine the next hop for the packet. In the most common case, the only relevant field in the header is the destination address field, but in some cases other header fields might also be relevant. As a result, the header analysis must be done independently at each router through which the packet passes. A complicated table lookup must also be done at each router.

In label switching, the analysis of the Layer 3 header is done only once. The Layer 3 header is then mapped into a fixed length, unstructured value called a *label.*



Label = 20 bits
Exp = Class of Service, 3 bits
S = Bottom of Stack, I bit
TTL = Time to Live, 8 bits

Figure 3-8: Label's header format

Many different headers can map to the same label, as long as those headers always result in the same choice of next hop. In effect, a label represents a forwarding equivalence class that is, a set of packets that, however different they may be, are indistinguishable by the forwarding function.

The initial choice of a label need not be based exclusively on the contents of the Layer 3 packet header; for example, forwarding decisions at subsequent hops can also be based on routing policy.

Once a label is assigned, a short label header is added at the front of the Layer 3 packet. This header is carried across the network as part of the packet. At subsequent hops through each MPLS router in the network, labels are swapped and forwarding decisions are made by means of MPLS forwarding table lookup for the label carried in the packet header. Hence, the packet header need not be reevaluated during packet transit through the network. Because the label is of fixed length and unstructured, the MPLS forwarding table lookup process is both straightforward and fast. Figure 2-9 shows packet forwarding of MPLS network

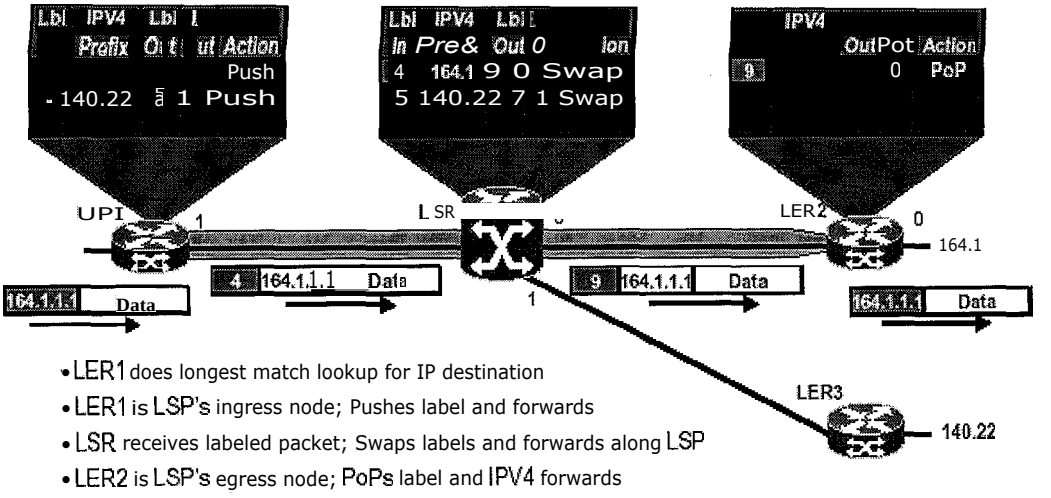


Figure 3-9: Packets forwarding

3.16 Distribution of Label Bindings

Each LSR in the network makes an independent, local decision as to which label value to use to represent a forwarding equivalence class. This association is known as a *label binding.* Each LSR informs its neighbors of the label bindings it has made- Figure 2-10. Neighboring routers facilitates this awareness of label bindings by neighboring routers by the following protocols:

- TDP—Used to support MPLS forwarding along normally routed paths
- Resource Reservation Protocol (RSVP)—Used to support MPLS traffic engineering
- Border Gateway Protocol (BGP)—Used to support MPLS VPNs

When a labeled packet is being sent from LSR A to the neighboring LSR B, the label value carried by the IP packet is the label value that LSR B assigned to represent the forwarding equivalence class of the packet. Thus, the label value changes as the IP packet traverses the network.

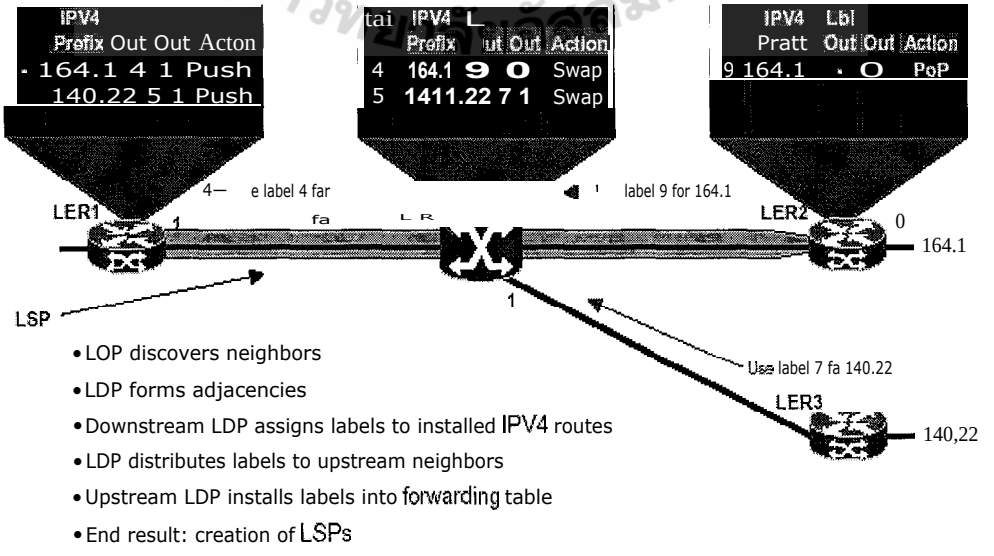


Figure 3-10: Label assignment

3.17 MPLS and Routing

A label represents a forwarding equivalence class, but it does not represent a particular path through the network. In general, the path through the network continues to be chosen by the existing Layer 3 routing algorithms such as OSPF, Enhanced IGRP, and BGP. That is, at each hop when a label is looked up, the dynamic routing algorithm determines the next hop chosen.

3.18 MPLS Operations

This section illustrates the passage of a frame through an MPLS system to highlight the function of several key MPLS components, Specifically, it illustrates MPLS through a frame-based infrastructure as opposed to a cell-based (ATM) system.

In Figure 2-11, a series of LSRs (edge and core) interconnection, forming a physical path between two elements, Station A and Station B.

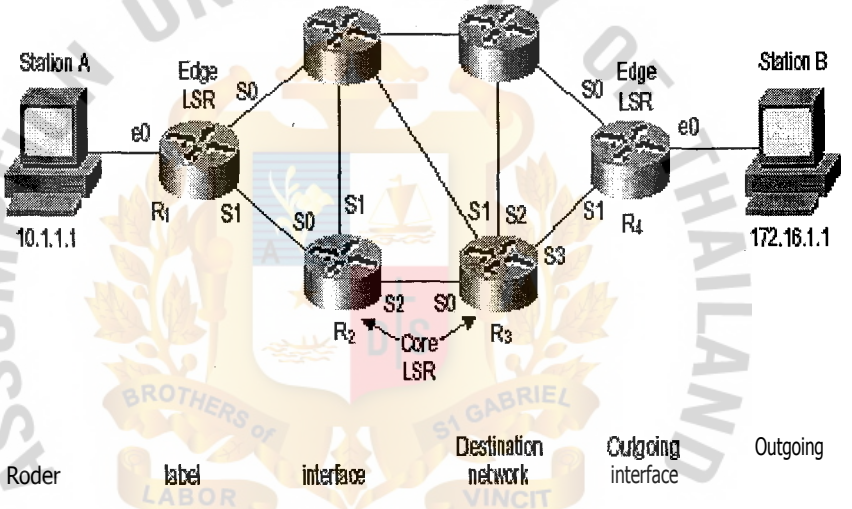


Figure 3-11, a series of LSRs interconnect

The frame generated by Station A follows the standard Ethernet format with a normal Layer 2 header followed by a Layer 3 header. Because the destination address resides in a different network, Station A targets the Layer 2 header to its default gateway. In this case, the default gateway also serves as the edge LSR (ingress side). The ingress LSR references its internal switch table (LFIB) and determines that it needs to forward the frame out port 2 toward the next LSR.

Furthermore, the ingress LSR must insert a label between the Layer 2 and Layer 3 headers to indicate what path the frame should travel on its way to Station B. Router 2 looks at the frame entering port 1 and determines that there is a label embedded between Layers 2 and 3. Therefore, the router treats the frame according to the configuration in its LFIB, which says to forward the frame out port 2 and replace the label with a new value. Each of the subsequent routers handles the frame in a similar manner until the frame reaches the egress LSR. The egress edge LSR strips off all label information and pass a standard frame to Station B. Because each of the routers between Stations A and B could switch the frame based upon content in the LFIB and did not need to perform usual routing operation, the frame was handled more quickly.

3.19 MPLS Architecture

MPLS relies on two principal components: forwarding and control. The forwarding component uses labels carried by packets and the label-forwarding information maintained by an LSR to perform packet forwarding. The control component is responsible for maintaining correct label-forwarding information among a group of interconnected label switches (LSRs). Details about MPLS's forwarding and control mechanisms follow.

3.19.1 Forwarding Component

The forwarding paradigm employed by MPLS is based on the notion of label swapping. When a packet with a label is received by an LSR, the switch uses the label as an index in its label information base (LFIB). Each entry in the LFIB consists of an incoming label and one or more subentries (of the form outgoing label, outgoing interface, outgoing link-level information). If the switch finds an entry with the incoming label equal to the label carried in the packet, then, for each component in the entry, the switch replaces the label in the packet with the outgoing label, replaces the link-level information (such as the MAC address) in the packet with the outgoing link-level information, and forwards the packet over the outgoing interface.

From the previous description of the forwarding component, we can make several observations. First, the forwarding decision is based on the exact-match algorithm using a fixed-length, fairly short label as an index. This enables a simplified forwarding procedure, relative to longest-match forwarding traditionally used at the network layer.

This, in turn, enables higher forwarding performance (higher packets per second). The forwarding procedure is simple enough to allow a straightforward hardware implementation. A second observation is that the forwarding decision is independent of the label's forwarding granularity. The same forwarding algorithm, for example, applies to both unicast and multicast: A unicast entry would have a single (outgoing label, outgoing interface, outgoing link-level information) subentry, while a multicast entry might have one or more subentries. This illustrates how the same forwarding paradigm can be used in label switching to support different routing functions.

The simple forwarding procedure is thus essentially decoupled from the control component of label switching. New routing (control) functions can readily be deployed without disturbing the forwarding paradigm. This means that it is not necessary to re-optimize forwarding performance (by modifying either hardware or software) as new routing functionality is added.

3.19.2 Label Encapsulation

Label information can be carried in a packet in a variety of ways:

- As a small, shim label header inserted between the Layer 2 and network layer headers
- As part of the Layer 2 header, if the Layer 2 header provides adequate semantics (such as ATM)
- As part of the network layer header (such as using the Flow Label field in IPv6 with appropriately modified semantics)

As a result, MPLS can be implemented over any media type, including point-to-point links, multi-access links, and ATM. The label-forwarding component is independent of the network layer protocol. Use of control component(s) specific to a particular network layer protocol enables the use of label switching with different network layer protocols.

3.19.3 Control Component

Essential to MPLS is the notion of binding between a label and network layer routes. MPLS supports a wide range of forwarding granularities to provide good scaling characteristics while also accommodating diverse routing functionality. At one extreme, a label could be associated (bound) to a group of routes (more specifically, to the network layer reach ability information of the routes in the group).

At the other extreme, a label could be bound to an individual application flow (such as an RSVP flow), or it could be bound to a multicast tree. The control component creates label bindings and then distributes the label-binding information among LSRs using a Label Distribution Protocol (LDP).

3.20 Label Distribution Protocols

With destination-based routing, a router makes a forwarding decision based on the Layer 3 destination address carried in a packet and the information stored in the forwarding information base (FIB) maintained by the router. A router constructs its FIB by using the information that the router receives from routing protocols, such as OSPF and BGP.

To support destination-based routing with MPLS, an LSR participates in routing protocols and constructs its LFIB by using the information that it receives from these protocols. In this way, it operates much like a router.

An LSR, however, must distribute and use allocated labels for LSR peers to correctly forward the frame. LSRs distribute labels using a label distribution protocol (LDP). A label binding associates a destination subnet to a locally significant label. (Labels are locally significant because they are replaced at each hop.) Whenever an LSR discovers a neighbor LSR, the two establish a TCP connection to transfer label bindings. LDP exchanges subnet/label bindings using one of two methods: downstream-unsolicited distribution or downstream-on-demand distribution. Both LSRs must agree as to which mode to use.

Downstream-unsolicited distribution disperses labels if a downstream LSR needs to establish a new binding with its neighboring upstream LSR. For example, an edge LSR may enable a new interface with another subnet. The LSR then announces to the upstream router a binding to reach this network.

In downstream-on-demand distribution, on the other hand, a downstream LSR sends a binding upstream only if the upstream LSR requests it. For each route in its route table, the LSR identifies the next hop for that route. It then issues a request (via LDP) to the next hop for a label binding for that route. When the next hop receives the request, it allocates a label, creates an entry in its LFIB with the incoming label set to the allocated label, and then returns the binding between the (incoming) label and the route to the LSR that sent the original request. When the LSR receives the binding information, the LSR creates an entry in its LFIB and sets the outgoing label in the entry to the value received from the next hop.

3.21 Hierarchical Routing

The IP routing architecture models a network as a collection of routing domains. Within a domain, routing is provided via interior routing (such as OSPF), while routing across domains is provided via exterior routing (such as BGP). All routers within domains that carry transit traffic, however (such as domains formed by Internet service providers), must maintain information provided by exterior routing, not just interior routing.

MPLS decouples interior and exterior routing so that only LSRs at the border of a domain are required to maintain routing information provided by exterior routing. All other switches within the domain maintain routing information provided by the domain's interior routing, which usually is smaller than the exterior routing information. This, in turn, reduces the routing load on non-border switches and shortens routing convergence time.

To support this functionality, MPLS allows a packet to carry not one, but a set of labels organized as a stack. An LSR can swap the label at the top of the stack, pop the stack, or swap the label and push one or more labels into the stack. When a packet is forwarded between two (border) LSRs in different domains, the label stack in the packet contains just one label.

When a packet is forwarded within a domain, however, the label stack in the packet contains not one, but two labels (the second label is pushed by the domain's ingress-border LSR). The label at the top of the stack provides packet forwarding to an appropriate egress-border label switch, while the next label in the stack provides correct packet forwarding at the egress switch. The stack is popped by either the egress switch or the penultimate switch.

3.22 Multicast Routing

In a multicast routing environment, multicast routing procedures (such as protocol-independent multicast [PIM]) are responsible for constructing spanning trees, with receivers as leaves. Multicast forwarding is responsible for forwarding multicast packets along these spanning trees.

Multicast in an MPLS environment is still under study by the IETF. However, MPLS supports multicast by utilizing data link layer multicast capabilities, such as those provided by Ethernet. Details are still in progress in the IETF committees.

3.23 Label Switching with ATM

Because the MPLS forwarding paradigm is based on label swapping, as is ATM forwarding, MPLS technology can be applied to ATM switches by implementing the control component. The label information needed for tag switching can be carried in the ATM VCI field. If two levels of labeling are needed, then the ATM VPI field could be used as well, although the size of the VPI field limits the size of networks in which this would be practical. The VCI field, however, is adequate for most applications of one level of labeling.

Implementing MPLS on an ATM switch would simplify integration of ATM switches and routers. An ATM switch capable of MPLS would appear as a router to an adjacent router. That would provide a scalable alternative to the overlay model and would remove the necessity for ATM addressing, routing, and signaling schemes. Because destination-based forwarding is topology-driven rather than traffic-driven,

application of this approach to ATM switches does not involve high call-setup rates, nor does it depend on the longevity of flows.

Implementing MPLS on an ATM switch does not preclude the capability to support a traditional ATM control plane (such as PNNI) on the same switch. The two components, MPLS and the ATM control plane, would operate independently with VPI/VCI space and other resources partitioned so that the components would not interact.

3.24 Quality of Service and Traffic Engineering

An important proposed MPLS capability is quality of service (QoS) support. Two mechanisms provide a range of QoS to packets passing through a router or a tag switch:

- Classification of packets into different classes
- Handling of packets via appropriate QoS characteristics (such as bandwidth and loss)

MPLS provides an easy way to mark packets as belonging to a particular class after they have been classified the first time. Initial classification uses information carried in the network layer or higher-layer headers. A label corresponding to the resultant class then would be applied to the packet. Labeled packets could be handled efficiently by LSRs in their path without needing to be reclassified. The actual packet scheduling and queuing is largely orthogonal: The key point here is that MPLS enables simple logic to be used to find the state that identifies how the packet should be scheduled.

The exact use of MPLS for QoS purposes depends a great deal on how QoS is deployed. If RSVP were used to request a certain QoS for a class of packets, then it would be necessary to allocate a label corresponding to each RSVP session for which state is installed at an LSR.

One of the fundamental properties of destination-based routing is that the only information from a packet that is used to forward the packet is the destination address. Although this property enables highly scalable routing, it also limits the capability to influence the actual paths taken by packets. This limits the capability to evenly distribute traffic among multiple links, taking the load off highly utilized links and shifting it toward less-utilized links.

For Internet service providers (ISPs) who support different classes of service, destination-based routing also limits their capability to segregate different classes with respect to the links used by these classes. Some of the ISPs today use Frame Relay or ATM to overcome the limitations imposed by destination-based routing. Because of the flexible granularity of labels, MPLS is capable of overcoming these limitations without using either Frame Relay or ATM. To provide forwarding along the paths that are different from the paths determined by the destination-based routing, the control component of MPLS allows installation of label bindings in LSRs that do not correspond to the destination-based routing paths. *Traffic engineering* allows a network administrator to make the path deterministic and bypass the normal routed hop-by-hop paths. An administrator may elect to explicitly define the path between stations to ensure QoS or have the traffic follow a specified path to reduce traffic loading across certain hops. In other words, the network administrator can reduce congestion by forcing the frame to travel around the overloaded segments. Traffic engineering, then, enables an administrator to define a policy for forwarding frames rather than depending upon dynamic routing protocols.

Traffic engineering is similar to source routing in that an explicit path is defined for the frame to travel. However, unlike source routing, the hop-by-hop definition is not carried with every frame. Rather, the hops are configured in the LSRs ahead of time along with the appropriate label values. Traffic engineering may be accomplished through *constraint-based routing*. Extensions to LDP allow traffic engineering to occur. Called *constraint-based LDP (CR-LDP)*, it enables a network engineer to establish and maintain explicitly routed LSPs called *constraint-based routed LSPs (CR-LSP)*.

3.24.1 MPLS Quality of Service

The quality of service (QoS) feature for MPLS enables network administrators to provide differentiated types of service across an MPLS network. Differentiated service satisfies a range of requirements by supplying for each packet transmitted the particular kind of service specified for that packet by its QoS. Service can be specified in different ways; for example, using the IP precedence bit settings in IP packets.

In supplying differentiated service, MPLS QoS offers packet classification, congestion avoidance, and congestion management. Table 3-1 lists these functions and their descriptions.

Service	QoS Function	Description
Packet classification	Committed access ale (CAR). Packets are classified at the edge of the network before labels are assigned.	Classifies packets according to input or output transmission rates. Allows you to set the MPLS experimental bits or the IP Precedence or DSCP bits (whichever is appropriate).
Congestion avoidance	Weighted Random Early Detection (WRED). Packet classes are differentiated based on drop probability.	Monitors network traffic to prevent congestion by dropping packets based on the IP Precedence or DSCP bits or the MPLS experimental field.
Congestion management	Class-based weighted fair queueing (CBWFQ). Packet classes are differentiated based on bandwidth and bounded delay.	An automated scheduling system that uses a queueing algorithm to ensure bandwidth allocation to different classes of network traffic.

Table 3-1: QoS Functions and their descriptions

3.24.2 Specifying the QoS in the IP Precedence Field

When you send IP packets from one site to another, the IP Precedence field (the first three bits of the DSCP field in the header of an IP packet) specifies the QoS. Based on the IP precedence marking, the packet is given the desired treatment such as the latency or the percent of bandwidth allowed for that quality of service. If the service provider network is an MPLS network, then the IP precedence bits are copied into the MPLS EXP field at the edge of the network. However, the service provider might want to set a QoS for a MPLS packet to a different value determined by the service offering.

This feature allows the service provider to set the MPLS experimental field instead of overwriting the value in the IP precedence field belonging to a customer. The IP header remains available for the customer's use; the QoS of an IP packet is not changed as the packet travels through the MPLS network.

. Figure 3-12 shows an MPLS network that connects two sites of a IP network belonging to a customer

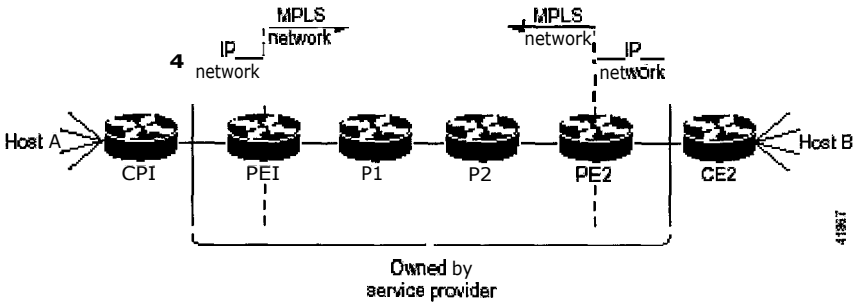


Figure 3-12: MPLS network that connects two sites of IP network belonging to a customer

In Figure 3-12, the symbols have the following meanings displayed in Table 3-2:

Symbol	Meaning
CHI	Customer equipment 1
PE1	Service provider edge router (ingress LSR)
PI	Service provider router within the core of the network of the service provider
P2	Service provider router within the core of the network of the service provider
PH2	Service provider edge router (egress LSR)
CH 2	Customer equipment 2

Table 3-2: Symbol of MPLS devices

In Figure 3-12, the following behavior occurs:

- Packets arrive as IP packets at PE1, the provider edge router (also known as the ingress label switching router).
- PE1 sends the packets as MPLS packets.
- Within the service provider network, there is *no IP Precedence field* for the queuing mechanism to look at because the packets are MPLS packets. The packets remain MPLS packets until they arrive at PE2, the provider edge router.
- PE2 removes the label from each packet and forwards the packets as IP packets.

This MPLS QoS enhancement allows service providers to classify packets according to their type, input interface, and other factors by setting (marking) each packet within the MPLS experimental field without changing the IP Precedence or DSCP field. For example, service providers can classify packets with or without considering the rate of the packets that PE1 receives. If the rate is a consideration, the service provider marks in-rate packets differently from out-of-rate packets.

3.24.3 MPLS Traffic Engineering

MPLS traffic engineering software enables an MPLS backbone to replicate and expand upon the traffic engineering capabilities of Layer 2 ATM and Frame Relay networks. MPLS is an integration of Layer 2 and Layer 3 technologies. By making traditional Layer 2 features available to Layer 3, MPLS enables traffic engineering. Thus, you can offer in a one-tier network what now can be achieved only by overlaying a Layer 3 network on a Layer 2 network. Traffic engineering is essential for service provider and Internet service provider (ISP) backbones. Such

backbones must support a high use of transmission capacity, and the networks must be very resilient so that they can withstand link or node failures.

MPLS traffic engineering provides an integrated approach to traffic engineering. With MPLS, traffic-engineering capabilities are integrated into Layer 3, which optimizes the routing of IP traffic, given the constraints imposed by backbone capacity and topology.

3.24.4 Reason of MPLS Traffic Engineering Using

WAN connections are an expensive item in an ISP budget. Traffic engineering enables ISPs to route network traffic to offer the best service to their users in terms of throughput and delay. By making the service provider more efficient, traffic engineering reduces the cost of the network.

Currently, some ISPs base their services on an overlay model. In the overlay model, transmission facilities are managed by Layer 2 switching. The routers see only a fully meshed virtual topology, making most destinations appear one hop away. If you use the explicit Layer 2 transit layer, you can precisely control how traffic uses available bandwidth.

However, the overlay model has numerous disadvantages. MPLS traffic engineering achieves the traffic engineering benefits of the overlay model without running a separate network. And without needing a non-scalable, full mesh of router interconnects.

3.24.5 How MPLS Traffic Engineering Works

MPLS traffic engineering automatically establishes and maintains LSPs across the backbone by using RSVP. The LSP resource requirements and network resources, such as bandwidth, determine the path that an LSP uses.

Available resources are flooded by means of extensions to a link-state-based Interior Gateway Protocol (IGP). Traffic engineering tunnels are calculated at the LSP head based on a fit between required and available resources (constraint-based routing). The IGP automatically routes the traffic onto these LSPs. Typically, a packet crossing the MPLS traffic engineering backbone travels on a single LSP that connects the ingress point to the egress point. MPLS traffic engineering is built on the following Cisco IOS mechanisms:

- IP tunnel interfaces—From a Layer 2 standpoint, an MPLS tunnel interface represents the head of an LSP. It is configured with a set of resource requirements, such as bandwidth and media requirements, and priority. From a Layer 3 standpoint, an LSP tunnel interface is the head-end of a unidirectional virtual link to the tunnel destination.
- MPLS traffic engineering path calculation module—This calculation module operates at the LSP head. The module determines a path to use for an LSP. The path calculation uses a link-state database containing flooded topology and resource information.
- RSVP with traffic engineering extensions—RSVP operates at each LSP hop and is used to signal and maintain LSPs based on the calculated path.
- MPLS traffic engineering link management module—This module operates at each LSP hop, does link call admission on the RSVP

signaling messages, and does bookkeeping of topology and resource information to be flooded.

• Link-state IGP (Intermediate System-to-Intermediate System (IS-IS) or OSPF—each with traffic engineering extensions)—These IGPs are used to globally flood topology and resource information from the link management module.

• Enhancements to the SPF calculation used by the link-state IGP (IS-IS or OSPF)—The IGP automatically routes traffic onto the appropriate LSP tunnel based on tunnel destination. Static routes can also be used to direct traffic onto LSP tunnels.

• Label switching forwarding—This forwarding mechanism provides routers with a Layer 2-like ability to direct traffic across multiple hops of the LSP established by RSVP signaling.

One approach to engineering a backbone is to define a mesh of tunnels from every ingress device to every egress device. The MPLS traffic engineering path calculation and signaling modules determine the path taken by the LSPs for these tunnels, subject to resource availability and the dynamic state of the network. The IGP, operating at an ingress device, determines which traffic should go to which egress device, and steers that traffic into the tunnel from ingress to egress. A flow from an ingress device to an egress device might be so large that it cannot fit over a single link, so a single tunnel cannot carry it. In this case, multiple tunnels between a given ingress and egress can be configured, and the flow is load-shared among them.

3.24.6 Mapping Traffic into Tunnels

This section describes how traffic is mapped into tunnels; that is, how conventional hop-by-hop link-state routing protocols interact with MPLS traffic engineering capabilities. In particular, this section describes how the shortest path first (SPF) algorithm, sometimes called a Dijkstra algorithm, has been enhanced so that a link-state IGP can automatically forward traffic over tunnels that MPLS traffic engineering establishes.

Link-state protocols, like integrated IS-IS or OSPF, use an SPF algorithm to compute a shortest path tree from the head end node to all nodes in the network. Routing tables are derived from this shortest path tree. The routing tables contain ordered sets of destination and first hop information. If a router does normal hop-by-hop routing, the first hop is over a physical interface attached to the router.

New traffic engineering algorithms calculate explicit routes to one or more nodes in the network. The originating router views these explicit routes as logical interfaces. In the context of this document, these explicit routes are represented by LSPs and referred to as traffic engineering tunnels.

The following sections describe how link-state IGPs can use these shortcuts, and how they can install routes in the routing table that point to these TE tunnels. These tunnels use explicit routes, and the router that is the head end of the tunnel controls the path taken by a TE tunnel. In the absence of errors, TE tunnels are guaranteed not to loop, but routers must agree on how to use the TE tunnels. Otherwise, traffic might loop through two or more tunnels.

3.24.7 Enhancement to the SPF Computation

During each step of the SPF computation, a router discovers the path to one node in the network, as follows:

- If that node is directly connected to the calculating router, the first hop information is derived from the adjacency database.
- If the node is not directly connected to the calculating router, the node inherits the first hop information from the parents of that node. Each node has one or more parents, and each node is the parent of zero or more downstream nodes.

For traffic engineering purposes, each router maintains a list of all TE tunnels that originate at this head end router. For each of those TE tunnels, the head end router knows the router at the tail end.

During the SPF computation, the TENT (tentative) list stores paths that are possibly the best paths and the PATH list stores paths that are definitely the best paths. When it is determined that a path is the best possible path, the node is moved from TENT to PATH. PATH is thus the set of nodes for which the best path from the computing router has been found. Each PATH entry consists of ID, path cost, and forwarding direction.

The router must determine the first hop information using one of the following methods:

- Examine the list of tail-end routers directly reachable by a TE tunnel. If there is a TE tunnel to this node, use the TE tunnel as the first hop.
- If there is no TE tunnel and the node is directly connected, use the first hop information from the adjacency database.
- If the node is not directly connected and is not directly reachable by a TE tunnel, copy the first hop information from the parent nodes to the new node.

As a result of this computation, traffic to nodes that are the tail end of TE tunnels flows over the TE tunnels. Traffic to nodes that are downstream of the tail-end nodes also flows over the TE tunnels. If there is more than one TE tunnel to different intermediate nodes on the path to destination node X, traffic flows over the TE tunnel whose tail-end node is closest to node X.

3.24.8 Special Cases and Exceptions

The SPF algorithm finds equal-cost parallel paths to destinations. The enhancement previously described does not change this behavior. Traffic can be forwarded over any of the following:

- One or more native IP paths
- One or more traffic engineering tunnels
- A combination of native IP paths and traffic engineering tunnels

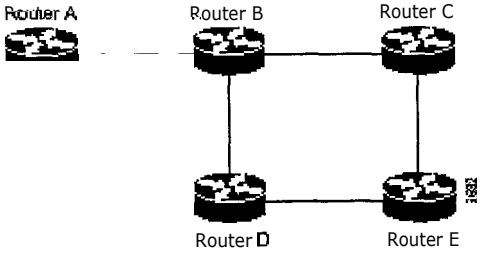


Figure 3-13: SPF algorithm in MPLS

If parallel native IP paths and paths over TE tunnels are available, the following implementations allow you to force traffic to flow over TE tunnels only or only over native IP paths. Assume that all links have the same cost and that a TE tunnel is set up from Router A to Router D.

- When the SPF calculation puts Router C on the TENT list; it realizes that Router C is not directly connected. It uses the first hop information from the parent, which is Router B.
- When the SPF calculation on Router A puts Router D on the TENT list; it realizes that Router D is the tail end of a TE tunnel. Thus Router A installs a route to Router D by the TE tunnel, and not by Router B.
- When Router A puts Router E on the TENT list, it realizes that Router E is not directly connected, and that Router E is not the tail end of a TE tunnel. Therefore Router A copies the first hop information from the parents (Router C and Router D) to the first-hop information of Router E.

Traffic to Router E now load balances over the following:

- The native IP path by Router A to Router B to Router C
- The TE tunnel Router A to Router D

—There are many neighbors or IP prefixes per router. A router that advertises substantial information causes the LSPs to be fragmented.

- Unpredictable results—In a large network, this solution can produce unpredictable results. A large network in transition pushes the limits regarding LSP flooding and SPF scaling. During the transition, the following behavior might occur:

—You can expect some extra network instability.

—Traffic engineering extensions might cause LSPs to be reflooded frequently.

- Ambiguity—If a router encounters different information in the old-style TLVs and the new-style TLVs, it may not be clear what the router should do.

These problems can be largely solved easily by using the following:

- All information in old-style and new style TLVs in an LSP
- The adjacency with the lowest link metric if an adjacency is advertised more than once

The main benefit to advertising the same information twice is that network administrators can use new-style TLVs before all routers in the network can understand them.

CHAPTER 4: Security of the MPLS Architecture

4.1 Introduction

Many enterprises are thinking of replacing traditional Layer 2 VPNs such as ATMor Frame Relay (FR) with MPLS-based services. As Multiprotocol Label Switching (MPLS) is becoming a more widespread technology for providing virtual private network (VPN) services, MPLS architecture security is of increasing concern to service providers (SPs) and VPN customers. This chapter gives an overview of MPLS architecture security for both SPs and MPLS users, and compares it with traditional Layer 2 services from a security perspective. This chapter also recommends how to secure an MPLS infrastructure. The focus is specifically on the MPLS/Border Gateway Protocol (BGP) VPN architecture.

MPLS is being used to achieve the following results: to engineer the core network more easily and efficiently (traditional MPLS and MPLS traffic engineering), to provide VPN services (MPLS-VPN), and to facilitate quality of service (QoS) across a network core (MPLS-DBP). In this chapter, the main emphasis is on security of the VPN provisioning aspect of MPLS, although most of it applies to other aspects of MPLS.

This chapter assumes that the MPLS core network is provided in a secure manner. Thus, it does not address basic security concerns such as securing the network elements against unauthorized access, mis-configurations of the core, internal (within the core) attacks, and so on. If a customer does not wish to assume the SP network is secure, it becomes necessary to run IP Security (IPSec) over the MPLS infrastructure.

Analysis of the security features of routing protocols is covered only to the extent that it influences MPLS. This chapter does not cover IPSec technology, except to highlight the combination of MPLS with IPSec.

Part A covers an analysis of the security that MPLS provides, compared to similar Layer 2 infrastructures. It targets the frequently asked question whether MPLS-based VPN services offer at least the same degree of security as ATMor Frame Relay-based VPNs.

Part B offers guidelines to secure an MPLS infrastructure. It discusses securing routing toward an MPLS core and interconnections between VPNs and Internet access. For additional security such as encryption, IPSec over an MPLS infrastructure is discussed, as well as remote access via IPSec into a specific VPN.

This chapter is targeted at technical staff of SPs and enterprises. Knowledge of the basic MPLS architecture is required to understand this chapter.

Part A: Analysis of the Security of the MPLS Architecture

This part answers the frequently asked question, whether MPLS provides the same level of security as traditional Layer 2 VPNs such as ATM and Frame Relay. Section 2 contains the requirements typically put forward by users of ATM or Frame Relay services, and Section 3 examines whether MPLS complies with these requirements.

4.2 Security Requirements of MPLS Networks

Both SPs offering MPLS services and customers using them have specific demands for the security of this special VPN solution. Mostly they compare MPLS-based solutions with traditional Layer 2-based VPN solutions such as Frame Relay and ATM, because these are widely deployed and accepted. This section outlines the security requirements typical in MPLS architectures. The next section discusses if and how MPLS addresses these requirements, for both the MPLS core and the connected VPNs.

4.2.1 Address Space and Routing Separation

Between two nonintersecting VPNs of an MPLS VPN service, it is assumed that the address space between different VPNs is entirely independent. This means, for example, that two nonintersecting VPNs must be able to use the 10/8 network without any interference. From a routing perspective, this means that each end system in a VPN has a unique address, and all routes to this address point to the same end system. Specifically:

- Any VPN must be able to use the same address space as any other VPN.
- Any VPN must be able to use the same address space as the MPLS core.
- Routing between any two VPNs must be independent.
- Routing between any VPN and the core must be independent.

From a security perspective, the basic requirement is to avoid the situation in which packets destined to a host a.b.c.d within a given VPN reach a host with the same address in another VPN or the core.

4.2.2 Hiding of the MPLS Core Structure

The internal structure of the MPLS core network (provider edge (PE) and provider (P) elements) should not be visible to outside networks (Internet or any connected VPN). Although a breach of this requirement does not lead to a security problem, many SPs feel this is advantageous if the internal addressing and network structure remains hidden to the outside world. A strong argument is that denial-of-service attacks against a core router, for example, are much easier to carry out if an attacker knows the address. Where addresses are not known, they can be guessed, but with this limited visibility, attacks become more difficult.

Ideally, the MPLS core should be as invisible to the outside world as a comparable Layer 2 (such as Frame Relay or ATM) infrastructure.

4.2.3 Resistance to Attacks

There are two basic types of attacks: denial-of-service (DoS) attacks, where resources become unavailable to authorized users, and intrusions, where the underlying goal is to gain unauthorized access to resources. Table 3-1 shows the two basic types of attack.

	Has Access	Has Ho Access
Authorized User	Normal	Denial at service
Unauthorized User	Intrusion	Normal

Table 4-1: Basic types pf attack

• For attacks that give unauthorized access to resources (intrusions), there are two basic ways to protect the network: first, to harden protocols that could be abused (such as Telnet to a router), and second, to make the network as inaccessible as possible. The latter is achieved by a combination of packet filtering or use of firewalls and address hiding, as discussed above.

• DoS attacks are easier to execute, because in the simplest case, a known IP address might be enough to attack a machine. The only way to be certain the network is invincible to this kind of attack is to make sure that machines are not reachable, again by packet filtering and address hiding.

MPLS networks must provide at least the same level of protection against both forms as current Layer 2 networks do. Note that this paper concentrates on protecting the core network against attacks from the "outside," or the Internet and connected VPNs. This chapter does not consider protection against attacks from the "inside," for example, if an attacker has logical or physical access to the core network, because any network can be attacked with access from the inside.

4.2.4 Impossibility of Label Spoofing

In a pure IP network, it is easy to spoof IP addresses, a key issue in Internet security. Because MPLS works internally with labels instead of IP addresses, the question arises whether these labels can be spoofed as easily as IP addresses.
Assuming the address and routing separation as discussed above, a potential attacker might try to gain access to other VPNs by inserting packets with a label that he doesn't "own." This could be done from the outside, for example, another customer edge (CE) router or from the Internet, or from within the MPLS core. This paper does not discuss the latter case (from within the core), because the assumption is that the core network is provided in a secure manner. If a network requires protection against an insecure core, it is necessary to run IPSec on top of the MPLS infrastructure.

It must be impossible to send packets with wrong labels from a CE router (the "outside") through a PE into the MPLS cloud, because this would make packet spoofing possible.

4.3 Analysis of MPLS Security

In this section the MPLS architecture is analyzed with respect to the security requirements listed above.

4.3.1 Address Space and Routing Separation

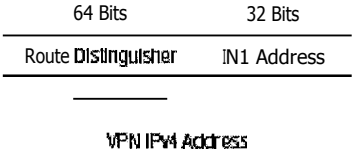


Figure 4-1: Format of VPN Ipv4 Address

MPLS allows distinct VPNs to use the same address space, which can also be private address space. This is achieved by adding a 64-bit route distinguisher (RD) to each IPv4 route, making VPN-unique addresses also unique in the MPLS core. This "extended" address is also called a "VPN-IPv4 address" and is shown in Figure 4-1. Thus, customers of an MPLS service do not need to change current addressing in their networks.

There is only one exception, which is the IP addresses of the PE routers the CE routers are peering with, in the case of using routing protocols between CE and PE routers (for static routing this is not an issue). Routing protocols on the CE routers need to have configured the address of the peer router in the core, to be able to "talk" to the PE router. This address must be unique from the perspective of the CE router and thus belongs logically to the address space of the VPN. In an environment where the SP also manages the CE routers as CPE, this setup can be made invisible to the customer.

Routing separation between the VPNs can also be achieved. Every PE router maintains a separate Virtual Routing and Forwarding instance (VRF) for each connected VPN. Each VRF on the PE router is populated with routes from one VPN, through statically configured routes or through routing protocols that run between the PE and the CE router. Because every VPN results in a separate VRF, there will be no interferences between the VPNs on the PE router.

Across the MPLS core to the other PE routers, this separation is maintained by adding unique VPN identifiers in multiprotocol BGP (MP BGP), such as the route distinguisher. VPN routes are exclusively exchanged by MP-BGP across the core, and this BGP information is not redistributed to the core network; it is redistributed only to the other PE routers, where the information is kept again in VPN-specific VRFs. Thus, routing across an MPLS network is separate per VPN.

Given the addressing and routing separation across an MPLS core network, we can assume that MPLS offers, in this respect, the same security as comparable Layer 2 VPNs such as ATM or Frame Relay. It is not possible to intrude into other VPNs through the MPLS cloud, unless this has been configured specifically.

4.3.2 Hiding of the MPLS Core Structure

For reasons of security, SPs and end customers do not normally want their network topologies revealed to the outside. This makes attacks more difficult. If an attacker does not know the target, he/she can only guess the IP addresses to attack or try to find out about addressing through a form of intelligence. Because most DoS attacks do not provide direct feedback to the attacker, a network attack is difficult.

With a known IP address, a potential attacker can launch a DoS attack against that device. So the ideal is to not reveal any information of the internal network to the outside. This applies equally to the customer networks as to the MPLS core. In

practice, numerous additional security measures have to be taken, primarily extensive packet filtering.

Figure 4-2 shows the visible address space of a given VPN. No P routers or other VPNs are visible to VPN1. The link between the CE and PE routers, which includes the interface address of the PE router, belongs to the VPN address space. All other addresses on the PE router, such as loopback interfaces, are not part of the VPN address space.

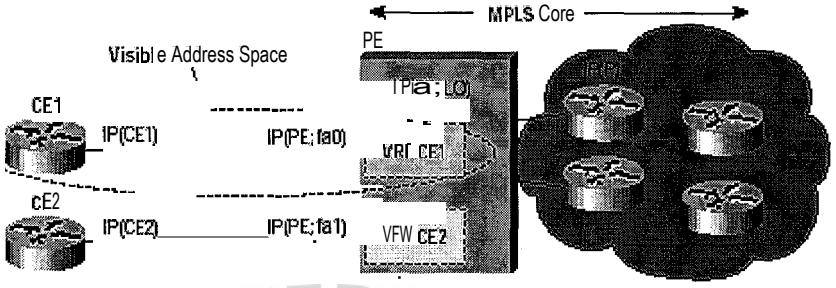


Figure 4-2: Hiding of the Core Infrastructure

MPLS does not reveal unnecessary information to the outside, not even to customer VPNs. Core addressing can be conducted with private addresses [RFC1918] or public addresses. Because the interface to the VPNs—and potentially the Internet—is BGP, there is no need to reveal any internal information. The only information required in the case of a routing protocol between PE and CE is the address of the PE router (IP PE in Figure 3-2). If this is not desired, static routing can be configured between the PE and CE. With this measure, the MPLS core can be kept completely hidden.

Customer VPNs will have to advertise their routes as a minimum to the MPLS core, to ensure reachability across the MPLS cloud. Although this could be seen as too "open," the following must be noted: First, the information known to the MPLS core is not about specific hosts, but networks (routes); this offers some degree of abstraction. Second, in a VPN-only MPLS network (such as one with no shared Internet access), this is equal to existing Layer 2 models in which the customer must trust an SP to some degree. Also, in a FR or ATM network, routing information about the VPNs can be seen on the core network.

In a VPN service with shared Internet access, an SP will typically announce the routes of customers who wish to use the Internet to upstream or peer providers. This can be done via a Network Address Translation (NAT) function to further obscure the addressing information of the customers' networks. In this case, the customer does not reveal more information to the general Internet than with a general Internet service. Core information will still not be revealed at all, except for the peering address(es) of the PE router(s) that hold(s) the peering with the Internet.

In summary, in a pure MPLS-VPN service, where no Internet access is provided, the information hiding is as good as on a comparable FR or ATM network; no addressing information is revealed to third parties or the Internet. If a customer chooses to access the Internet via the MPLS core, the customer must reveal the same addressing structure as for a normal Internet service. NAT can be used for further address hiding. If an MPLS network has no interconnections to the Internet, this is equal to FR or ATM networks. With an Internet access from the MPLS cloud, the SP has to reveal at least one IP address (of the peering PE router) to the next provider, and thus the outside world.

4.3.3 Resistance to Attacks

Section 4.2 shows that it is not possible to directly intrude into other VPNs. The only other possibility is to attack the MPLS core, and try to attack other VPNs from there. The MPLS core can be attacked in two basic ways:

- By attacking the PE routers directly
- By attacking the signaling mechanisms of MPLS (mostly routing)

To attack an element of an MPLS network, it is first necessary to know its address. As discussed in Section 4.2, it is possible to hide the addressing structure of the MPLS core to the outside world. Thus, an attacker does not know the IP address of any router in the core that he/she wants to attack. The attacker could now guess addresses and send packets to these addresses. However, because of the address separation of MPLS, each incoming packet will be treated as belonging to the address space of the customer. Thus it is impossible to reach an internal router, even through IP address guessing. This rule has only one exception, which is the peer interface of the PE router.

The routing between the VPN and the MPLS core can be configured two ways:

1. Static—In this case the PE routers are configured with static routes to the networks behind each CE, and the CEs are configured to statically point to the PE router for any network in other parts of the VPN (mostly a default route). There are now two subcases: The static route can point to the IP address of the PE router, or to an interface of the CE router.
2. Dynamic—Here a routing protocol (for example, Routing Information Protocol [RIP], Open Shortest Path First [OSPF], BGP) is used to exchange the routing information between the CE and the PE at each peering point.

In the case of a static route from the CE router to the PE router, which points to an interface, the CE router does not need to know any IP address of the core network, not even of the PE router. This has the disadvantage of a more extensive (static) configuration, but from a security point of view is preferable to the other cases.

In all other cases, each CE router needs to know at least the router ID (RID; peer IP address) of the PE router in the MPLS core, and thus has a potential destination for an attack. One could imagine various attacks on various services running on a router. In practice, access to the PE router over the CE/PE interface can be limited to the required routing protocol by using ACLs (access control lists). This limits the point of attack to one routing protocol, for example BGP. A potential attack could be to send an extensive number of routes, or to flood the PE router with routing updates. Both could lead to a DoS, however, not to unauthorized access.

To restrict this risk, it is necessary to configure the routing protocol on the PE router as securely as possible. This can be done in various ways:

- By ACL, allow the routing protocol only from the CE router, not from anywhere else—Furthermore, no access other than that should be allowed to the PE router in the inbound ACL on each CE interface.
- Where available, configure Message Digest 5 (MD5) authentication for routing protocols—This is available for BGP [RFC2385], OSPF [RFC2154], and RIP2 [RFC2082], for example. It prevents packets from being spoofed from parts of the customer network other than the CE router. Note that this requires that the SP and customer agree on a shared secret between all CE and PE routers. The problem here is

that it is necessary to do this for all VPN customers—it is not sufficient to do this for the customer with the highest security requirements.

• Configure, where available, parameters of the routing protocol, in order to further secure this communication—In BGP, for example, it is possible to configure dampening, which limits the number of routing interactions. Also, a maximum number of routes accepted per VRF should be configured where possible.

It should be noted that although in the static case the CE router does not know any IP address of the PE router, it is still attached to the PE router via some method; therefore, it could guess the address of the PE router and try to attack it with this address.

In summary, it is not possible to intrude from one VPN into other VPNs, or the core. However, it is theoretically possible to exploit the routing protocol to execute a DoS attack against the PE router. This in turn might have a negative impact on other VPNs. Therefore, PE routers must be extremely well secured, especially on their interfaces to the CE routers. ACLs must be configured to limit access only to the port(s) of the routing protocol, and only from the CE router. MD5 authentication in routing protocols should be used on all PE/CE peerings. It is easily possible to track the source of such a potential DoS attack.

4.3.4 Label Spoofing

Within the MPLS, network packets are not forwarded based on the IP destination address, but based on labels that are prepended by the PE routers. Similar to IP spoofing attacks, where an attacker replaces the source or destination IP address of a packet, it is also theoretically possible to spoof the label of an MPLS packet. In the first section, the assumption was made that the core network is secured by the SP.

Thus in this section the emphasis is on whether it is possible to insert packets with (wrong) labels into the MPLS network from the outside, that is, from a VPN (CE router) or from the Internet.

Principally, the interface between any CE router and its peering PE router is an IP interface (that is, without labels). The CE router is unaware of the MPLS core, and thinks it is sending IP packets to a simple router. The "intelligence" is done in the PE device, where based on the configuration, the label is chosen and pre ended to the packet. This is the case for all PE routers, toward CE routers as well as the upstream SP. All interfaces into the MPLS cloud require only IP packets, without labels.

For security reasons, a PE router should never accept a packet with a label from a CE router. In Cisco routers, the implementation is such that packets that arrive on a CE interface with a label will be dropped. Thus it is not possible to insert fake labels, because no labels at all are accepted.

There remains the possibility to spoof the IP address of a packet that is being sent to the MPLS core. However, because there is strict addressing separation within the PE router, and each VPN has its own VRF, this can harm only the VPN that the spoofed packet originated from; in other words, a VPN customer can attack himself/herself. MPLS does not add any security risk here.

4.3.5 Comparison with ATM/FR VPNs

ATM and FR VPN services often enjoy a very high reputation in terms of security. Although ATM and FR VPNs can also be provided in a secure manner, it has been reported that these technologies can also have severe security vulnerabilities.

Also, in ATM/FR the security depends on the configuration of the network being secure, and errors can also lead to security problems.

4.4 Options for Securing an MPLS Core

This part targets the SP: It tries to outline how MPLS-based VPN services can be secured, and what has to be addressed to implement network-based services such as Internet access, remote access to a VPN, or fire walling. It also explains what MPLS does not provide.

4.4.1 Securing the MPLS Core

This section is targeted toward the SP, to give guidelines about secure configuration of an MPLS core network. Many general security mechanisms, such as securing routers, are not discussed here. The following is a list of recommendations and considerations on configuring an MPLS network securely.

• Trusted devices—The PE and P devices, as well as remote-access servers and authentication, authorization, and accounting (AAA) servers, have to be treated as trusted systems. This requires strong security management, starting with physical building security and including issues such as access control, secure configuration management, and storage. Ample literature is available on how to secure network elements, so this topic is not treated here in more detail. CE routers are typically not under full control of the SP and, therefore, have been treated as untrusted.

• CE/PE interface—The interface between the PE and CE routers is crucial for a secure MPLS network. The PE router should be configured as close as possible. From a security point of view, the best option is to configure the interface to the CE router unnumbered, and route statically.

Packet filters (ACLs) should be configured to permit only one specific routing protocol to the peering interface of the PE router, and only from the CE router. All other traffic to the router and the internal SP network should be denied. This scenario prevents attack on the PE and P routers, because the PE router will drop all packets to the corresponding address range. The only exception is the peer interface on the PE router for routing purposes. This needs to be secured separately.

If private address space is used for the PE and P routers, the same rules with regard to packet filtering apply: All packets must be filtered to this range. However, because addresses of this range should not be routed over the Internet, attacks to adjacent networks are limited.

• Routing authentication—Routing is the signaling mechanism between the CEs and the PEs. To introduce bogus information into the core, routing protocols are the most obvious point for an attack. Thus it is essential that routing information is as secure as possible, and that it comes really from the router it is expected from, and not from a hacker's router. Toward this goal, all routing protocols should be configured with the corresponding authentication option toward the CEs and toward any Internet connection. All peering relationships in the network need to be secured this way: CE/PE (with BGP MD5 authentication), PE/P (with Label Distribution Protocol [LDP] MD5 authentication) and P/P. This setup prevents attackers from spoofing a peer router and introducing bogus routing information. Note specifically here the importance of secure management: Configuration files often contain shared secrets in cleartext (for example, for routing protocol authentication).

• Separation of CE/PE links—If several CEs share a common Layer 2 infrastructure to access the same PE router (for example, an Ethernet virtual LAN [VLAN]), a CE router can spoof packets as belonging to another VPN that also has a connection to this PE router. Securing the routing protocol as described above is not sufficient, because this does not affect normal packets. To avoid this problem, it is recommended to implement separate physical connections between CEs and PEs. The use of a switch between various CE routers and a PE router is also possible, but it is strongly recommended to put each CE/PE pair into a separate VLAN, to provide traffic separation. Note that although switches with VLANs increase security, they are not unbreakable. A switch in this environment must thus be treated as a trusted device, and configured with maximum security.

• LDP authentication—The LDP can also be secured with MD5 authentication across the MPLS cloud. This scenario prevents hackers from introducing bogus routers, which would participate in the LDP.

4.5 Interconnections between VPNs and Internet Access

4.5.1 Connectivity between VPNs

MPLS provides VPN services with address and routing separation between VPNs. In many environments, however, destinations outside the VPN must also be reachable. This could be for Internet access or for merging two VPNs, for example, in the case of two companies merging. MPLS not only provides full VPN separation, but also allows merging VPNs or access to the Internet.

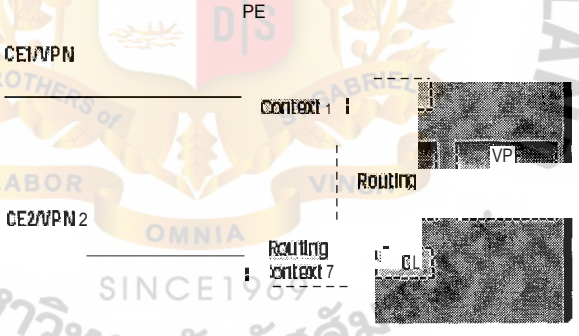


Figure 4-3: Connectivity between VPNs

To achieve this access, the PE routers maintain various tables: A routing context is specific to a CE router, and contains only routes from this particular VPN. From there, routes are propagated into the *VRF* (virtual routing and forwarding instance) *routing table,* from which a VRF forwarding table is calculated. For separated VPNs, the VRF routing table contains only routes from one routing context. To merge VPNs, different routing contexts (from different VPNs) are put into one single VRF routing table. This way, two or several VPNs can be merged to a single VPN. Note that in this case all merged VPNs must have mutually exclusive addressing spaces; in other words, the overall address space must be unique for all included VPNs. It is possible to control with ACLs which routes get redistributed into VRF tables.

For a VPN to have Internet connectivity, the same procedure is used: Routes from the Internet VRF routing table are propagated into the VRF routing table of the VPN that requires Internet access. Alternatively to propagating all Internet routes, a

default route can be propagated. In this case, the address space between the VPN and the Internet must be distinct. In other words, the VPN must use either publicly registered or private address space because all other addresses can occur in the Internet.

From a security point of view, the merged VPNs behave like one logical VPN, and the security mechanisms described above apply now between the merged VPN and other VPNs: The merged VPN must have unique address space internally, but further VPNs may use the same address space without interference. Packets from and to the merged VPNs cannot be routed to other VPNs, so all the separation functions of MPLS apply also for merged VPNs with respect to other VPNs.

If two VPNs are merged in this way, hosts from either part can reach the other part as if the two VPNs were a common VPN. This means that with the standard MPLS features, there is no separation or firewalling /packet filtering between the merged VPNs. Also, if a VPN receives Internet routes through MPLS/BGP VPN mechanisms, firewalling or packet filtering has to be engineered in addition to the MPLS features.

4.5.2 Firewalling Options

Two scenarios are examined in this section: securing VPNs against each other while maintaining inter-VPN connectivity, and securing Internet access.

Scenario 1: Firewalls between VPNs

One reason for merging two previously independent VPNs is two companies merging or interoperating over the network. In most of these cases, the companies want to maintain a logical separation from *other* companies, even if connectivity between the companies is required. Typically, firewalls are placed in such circumstances. As in traditional networks, the interconnection points between the two VPNs have to be secured with firewalls. However, whereas in traditional networks the border router is normally under the control of the company, in the MPLS/BGP VPN environment, the "peering point" between the VPNs is a PE router under the control of the SP.

Technically, the interconnection by announcing the routes of both VPNs to the other VPN as described above happens in one router. This way of interconnecting alone does not provide firewall capabilities. To position a firewall between two VPNs, the firewall must be provisioned as a separate entity in addition to the PE router. The PE router manages the two VPNs completely separate, as described above. This setup provides the required security between the two VPNs.

To interconnect the two VPNs via a firewall, an additional interface that leads to the firewall must be provisioned for each VPN. This way, packets from VPN A to VPN B would come from a router in VPN A, and they would be routed to the interconnecting PE router. The PE router has a route to VPN B, which points to the interface to which the firewall is connected. The packets traverse the firewall, and enter the PE router through another interface, which belongs to VPN B.

This way, it is also possible to use NAT on the firewall, with the effect that the merged VPNs do not have to have mutually exclusive address space.

The note on "Separation of CE/PE links" in Section 4 also applies here: Switches do not necessarily provide traffic separation. Thus if switches are used, it is strongly recommended not to put the interfaces of the firewall onto the same switch,

but to use separate switches. If this is not possible, different VLANs must be used for the two sides of the firewall. Because hubs do not provide any traffic separation, their use is strongly discouraged. There can be more than one interconnection point between VPNs. All interconnection points can be engineered this way.

Scenario 2: Firewalls to the Internet

The provisioning of a firewall for Internet access is similar: The PE router that connects to one or more other SPs will have to traverse a firewall before sending or receiving packets to/from the Internet. If one firewall can be applied to all VPNs equally (shared firewall), the setup consists of a PE router, which connects to a firewall before going to other SPs. The PE router maintains in the default VRF routing table the Internet routes or a default route to the Internet. The Internet routes (or the default) are propagated to the VRF routing tables of VPNs that require Internet connectivity. The routes from the VPNs are propagated to the default VRF routing table of the PE router, which announces them to the Internet over the firewall. Instead of dynamic routing, static routes can also be used. The addressing space of VPNs using Internet connectivity must be publicly registered address space.

Figure 4-4 shows one possible way to secure Internet access with a firewall of choice. The Internet routing table is treated as another VPN, and the connectivity to other VPNs is passing through an external firewall, providing all the features of this firewall, including NAT if required.

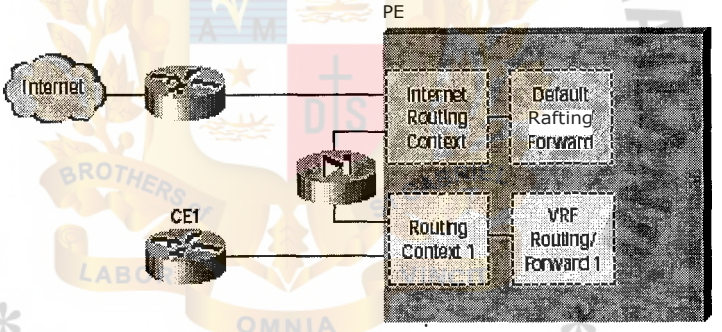


Figure 4-4: Example for a Firewall Installation to the Internet

This option is relatively easy to engineer, but has the disadvantage that all VPNs use the same firewall and are thus bound to one security policy at this point. To engineer an Internet firewall for each VPN separately, the above setup needs to be multiplied.

For separate firewalls per VPN, the PE router that connects to the Internet needs one interface per VPN, leading to one firewall per VPN. Beyond the firewalls, the connections can come together again in one router, which connects then to other providers. The advantage of this option is the capability to have a NAT function per customer, so that internally each VPN can use random address space, and on the firewalls this is mapped to publicly registered space.

Scenario 3: A Firewall per CE Router

Big networks tend to become unmanageable in terms of security, unless there is some form of separation between parts of the network. In a country-wide network that is internally completely open, a security incident such as a break-in in one office

might require all hosts of the entire network to be reinstalled, to ensure that the attacker has not left some Trojan horses somewhere. An increasing number of companies are securing their internal networks additionally by, for example, separating offices with firewalls. This is, in general, a good security practice. Given that Cisco routers can function as a firewall, the additional costs are normally manageable, because often only a software upgrade is required.

In the case where a company separates its network, implemented as a VPN on an MPLS service, putting a firewall on every CE router can make the overall network easier to manage and more secure. In this case, everything outside an office (connected with a CE router to the MPLS network) is treated as untrusted, and traffic from the same company is checked as well as traffic from other companies, which might be merged over the MPLS structure.

4.6 Incapability of MPLS

4.6.1 Protection against Misconfigurations of the Core and Attacks within the Core

The security mechanisms discussed here assume correct configuration of the involved network elements on the MPLS core network (PE and P routers). Deliberate or inadvertent misconfigurations from SP staff may result in undesired behavior, including severe security leaks. Note that this paragraph refers specifically to the core network; that is, the PE and P elements. Misconfiguration of any of the customer-side elements such as the CE router *is* covered by the security mechanisms above, meaning that a potential attacker must have access to either PE or P routers to gain advantage from misconfigurations. If an attacker has access to core elements or is able to insert additional equipment into the core, he/she will be able to attack both the core network and the connected VPNs. Thus the following is important:

- To avoid the risk of misconfigurations, it is important that the equipment is easy to configure, and that SP staff has the appropriate training and experience when configuring the network.
- To avoid the risk of "internal" attacks, the MPLS core network must be properly secured. This security includes network-element security, management security, physical security of the SP infrastructure, access control to SP installations, and other standard SP security mechanisms. MPLS can provide a secure service only if the core network is provided in a secure fashion. This paper assumes that it is.

4.6.2 Data Encryption, Integrity, and Origin Authentication

MPLS itself does not provide encryption, integrity, or authentication services. If these features are required, IPSec should be used over the MPLS infrastructure.

4.6.3 Customer Network Security

MPLS can be secured so that it is comparable with other VPN services. However, the security of the core network is only one factor for the overall security of a customer's network. Threats in today's networks come not only from the "outside" connection, but also from the "inside" and from other entry points (modems, for example). To reach a good security level for a customer network in an MPLS infrastructure, MPLS security is necessary but not sufficient.

4.7 Virtual Private Network (VPN) by MPLS

The IP-based virtual private network (VPN) is rapidly becoming the foundation for the delivery of New World services, and many service providers are offering value-added applications on top of their VPN transport networks. Emerging services such as e-commerce, application hosting, and multimedia applications will enable service providers to generate new incremental revenue and maintain long-term competitive advantage. Two unique and complementary VPN architectures based on IP Security (IPsec) and Multiprotocol Label Switching (MPLS) technologies are emerging to form the predominant foundations for delivery of New World services. This white paper examines these two VPN architectures, their similarities and differences, and the benefits they offer. It concludes with a unified view of IP VPNs that combine both IPsec and MPLS based on their respective strengths.

4.8 Characteristics of Two VPN Architectures

The service goal of VPNs is to provide customer connectivity over a shared infrastructure, with the same policies enjoyed in a private network. A VPN solution must therefore be secure from intrusion and tampering, deliver mission-critical data in a reliable and timely manner, and be manageable. The essential attributes of a VPN can be segmented into five broad categories (Table 4-2).

Salability	Must be scalable across VPN platforms ranging from a small office configuration through the largest enterpri - implementations ubiquitously on a global scale: the ability to adapt the VPN to meet changing bandwidth and connectivity needs is crucial in a VPN solution. Additionally, in the fiercely competitive and dynamic market environment, large orders can be won and must be provisioned rapidly, hence the VPN must be highly scalable in order to accommodate unplanned growth and changes driven by customer demand. A typical MPLS deployment must be designed for highly scalable solutions, enabling tens of thousands of VPNs over the same network for maximum revenue and profitability.
Security	Ensures business-critical traffic remains confidential via security mechanisms such as tunneling, encryption, traffic Sepug on, packet authentication, user authentication, and access control.
Quality of Service	Ensures prioritization of mission-critical or delay-sensitive traffic and manages congestion across varying bandwidth rates. Quality of service (QoS) functions such as queuing, network congestion avoidance, traffic shaping, and packet classification, as well as VPN routing services utilizing an optimal routing protocol.
Manageability	Essential for cost-effective provisioning to enforce security and QoS policies, management and billing, with advanced monitoring and automated flow-through systems to quickly roll out new services and support service-level agreements (SLA).
Reliability	For predictable and extremely high service avail a bility that business customers expect and require.

Table 4-2: Attributes of VPN

In recent years, two working groups from the Internet Engineering Task Force (IETF) have been focusing on mechanisms to address Internet security, label switching standardization, and QoS, which are closely aligned with the building blocks of a VPN. The IETF IPsec working group (under the Security Area) is concentrating on the protection of network layer by designing cryptographic security mechanisms that can flexibly support combinations of authentication, integrity, access control, and confidentiality. The IETF MPLS working group (under the Routing Area)

on the other hand is developing mechanisms to support higher layer resource reservation, QoS, and definition of host behaviors.

IETF has left the issue of integrating IPsec and MPLS at the discretion of the implementers. As a result, two VPN architectures have emerged, weighted heavily on IPsec or MPLS technologies respectively. Service providers today are deploying one or both of these VPN architectures primarily based on the customers they serve, and the New World value-added services they plan to offer.

4.9 Comparison Between IPsec and MPLS-Based VPN

Table 4-3 describes the characteristics, benefits, positioning, and the differentiation between the IPsec and MPLS-based VPN.

	IPsec-Based VPN	MP S-Based VPN
Service Model:	High-speed Internet services, business-quality IP VPN services, e-commerce and application-hosting services	High-speed Internet services, business-quality P Mg services, e-commerce and application-hosting services
Scalability	Large-scale deployment requires planning and coordination to address issues on key distribution, key management, and peering configuration	Highly scalable since no site-to-site peering is required. Atypical MPLS-based VPN deployment is capable of supporting tens of thousands VPN groups over the same network
Place in Ne Work IFigure 11	Best at the local Ioop, edge and off-net where there is a higher degree of exposure to data privacy and where IPsec security mechanisms such as tunneling and encryption can best be applied	Best within a service provider's core network where QoS, traffic engineering, and bandwidth utilisation can be fully controlled, especially if SLA or service-level guarantee (SLG) is to be offered as part of the VPN service

Table 4-3: IPSec and MPLS-Based VPN Comparison

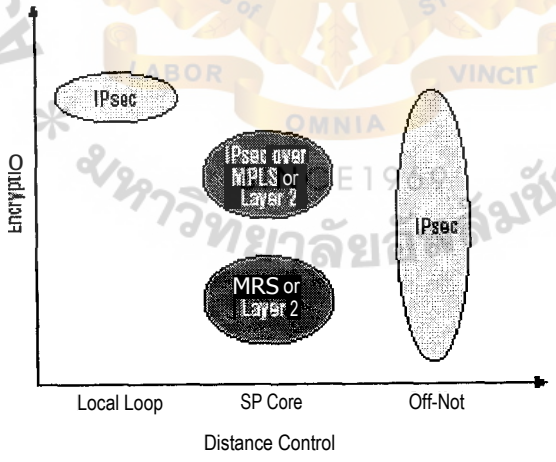


Figure 4-5: Network Placement Positioning

	IPsec-Based VPN	MPLS-Based VPN
Transparency	IPsec VPN resides at the network layer; it is transparent to the applications	MPLS VPN operates at the IP+ATM or IP environment; it is completely transparent to the applications
Provisioning	In general, no network level provisioning is required for managed CPE based service offering. When a networked based IPsec VPN service is deployed, service provider generally provides centralized provisioning and management support.	Because MPLS VPN site peers with a service provider network only, service activation requires just a one-time provisioning at the customer edge (CE) and provider edge (PE) devices to enable the site to become a member of a MPLS VPN group
Service Deployment	Fast time to market; can be deployed across any existing IP networks	Requires participating network elements at the core and edge to be MPLS capable, such as during a network upgrade or when a new MPLS network must be deployed
Session Authentication	Each IPsec session must be authenticated via digital certificate or preshared key; packets that do not conform to the security policy are dropped	VPN membership is determined by service providers—a provisioning function based on logical port and unique route descriptor; unauthorized access to a VPN group is denied by device configuration
Confidentiality (Figure 21)	IPSec VPN provides data privacy through a flexible suite of encryption and tunneling mechanisms at the IP network-layer	MPLS architecture separates traffic between customers offering security in a manner similar to a trusted Frame Relay or ATM network environment
Quality of Service, Service Level Agreement (Figure 21)	While the IPsec protocol does not address network reliability or QoS mechanisms, a Cisco IPsec VPN deployment can preserve packet classification for QoS with in art IPsec tunnel	A well-executed MPLS based VPN implementation provides scalable, robust QoS mechanism and traffic engineering capability enabling service providers to offer IP-based value-added services with guaranteed SLA compliance

Table 4-3 IPsec and MPLS-Based VPN Comparison (continued)

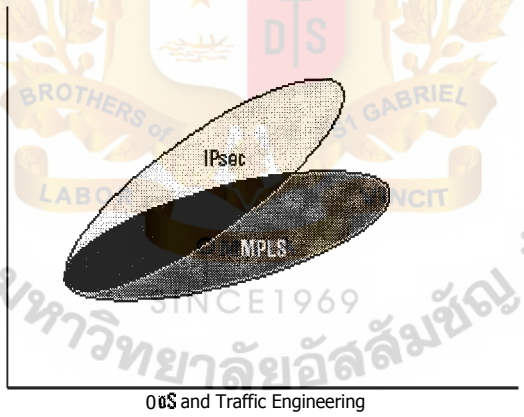


Figure 4-6: Confidentiality and QoS Positioning

	IPsec-Based VPN	MPI S-13asgd VPN
Client support	Required for client initiated IPsec VPN deployment; note that Cisco VPN client software is available for Microsoft Windows as well as nonWindows platforms such as Solaris, Linux, and Macintosh	Not applicable; MPLS VPN is a network-based VPN service
User Interaction	For client initiated IPsec VPN service offering, users need to interact with the IPsec client software	Not applicable; no user interaction is required

Table 4-3 IPsec and MPLS-Based VPN Comparison (continued)

4.10 IPsec and MPLS VPN Integration

While service providers can deploy one or the other of these VPN architectures to support New World value-added services, a greater benefit can be realized as they converge. Simply put, a well-executed, comprehensive VPN service offering may leverage both IPsec and MPLS. Service providers may choose IPsec for traffic that needs strong authentication and confidentiality and choose MPLS for its broader connectivity, traffic engineering, and QoS compared with traditional Layer 2 private data networking. With both architectures managed by the Cisco VPN Solution Center, this combination enables service providers to offer differentiated New World services that cover the spectrum of customer requirements for security, QoS, and traffic prioritization (Figure 4-7).

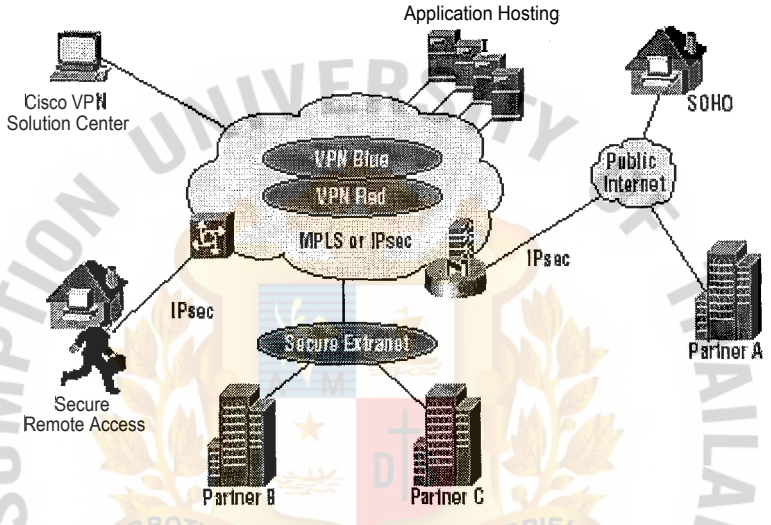


Figure 4-7: IPsec and MPLS Integrated VPN Architecture

Using MPLS VPNs provides the capability to deploy and administer scalable Layer 3 VPN backbone services including applications, data hosting network commerce, and telephony services to business customers. A VPN is a secure IP-based network that shares resources on one or more physical networks. A VPN contains geographically dispersed sites that can communicate securely over a shared backbone.

A one-to-one relationship does not necessarily exist between customer sites and VPNs; a given site can be a member of multiple VPNs. However, a site can associate with only one VPN routing and forwarding instance (VRF). Each VPN is associated with one or more VPN VRFs. A VRF includes routing and forwarding tables and rules that define the VPN membership of customer devices attached to CE routers. A VRF consists of the following:

- IP routing table
- CEF table
- Set of interfaces that use the CEF forwarding table
- Set of rules and routing protocol parameters to control the information in the routing tables.

VPN routing information is stored in the IP routing table and the CEF table for each VRF. A separate set of routing and CEF tables is maintained for each VRF. These tables prevent information from being forwarded outside a VPN and also

prevent packets that are outside a VPN from being forwarded to a router within the VPN.

4.11 Benefits

MPLS VPNs allow service providers to deploy scalable VPNs and build the foundation to deliver value-added services, including the following:

• Connectionless service—A significant technical advantage of MPLS VPNs is that they are connectionless. The Internet owes its success to its basic technology, TCP/IP. TCP/IP is built on packet-based, connectionless network paradigm. This means that no prior action is necessary to establish communication between hosts, making it easy for two parties to communicate. To establish privacy in a connectionless IP environment, current VPN solutions impose a connection-oriented, point-to-point overlay on the network. Even if it runs over a connectionless network, a VPN cannot take advantage of the ease of connectivity and multiple services available in connectionless networks. When you create a connectionless VPN, you do not need tunnels and encryption for network privacy, thus eliminating substantial complexity.

• Centralized service—Building VPNs in Layer 3 allows delivery of targeted services to a group of users represented by a VPN. A VPN must give service providers more than a mechanism for privately connecting users to intranet services. It must also provide a way to flexibly deliver value-added services to targeted customers. Scalability is critical, because customers want to use services privately in their intranets and extranets. Because MPLS VPNs are seen as private intranets, you may use IP services such as the following:

- Multicast
- Quality of service (QoS)
- Telephony support within a VPN
- Centralized services including content and web hosting to a VPN

You can customize several combinations of specialized services for individual customers. For example, a service that combines IP multicast with a low-latency service class enables videoconferencing within an intranet.

• Scalability—If you create a VPN using connection-oriented, point-to-point overlays, Frame Relay, or ATM virtual connections, the VPN's key deficiency of the VPN is scalability. Specifically, connection-oriented VPNs without fully meshed connections between customer sites are not optimal. MPLS-based VPNs instead use the peer model and Layer 3 connectionless architecture to leverage a highly scalable VPN solution. The peer model requires a customer site to only peer with one provider edge (PE) router as opposed to all other CPE or CE routers that are members of the VPN. The connectionless architecture allows the creation of VPNs in Layer 3, eliminating the need for tunnels or virtual connections.

The following are scalability issues of MPLS VPNs due to the partitioning of VPN routes between PE routers and the further

partitioning of VPN and IGP routes between PE routers and provider (P) routers in a core network:

– PE routers must maintain VPN routes for those VPNs that are members.

—P routers do not maintain any VPN routes.

This increases the scalability of the provider's core and ensures that no one device is a scalability bottleneck.

• Security—MPLS VPNs offer the same level of security, as connection-oriented VPNs. Packets from one VPN do not inadvertently go to another VPN. Security is provided

—At the edge of a provider network, ensuring that packets received from a customer are placed on the correct VPN.

— At the backbone, ensuring that VPN traffic is kept separate. Malicious spoofing (an attempt to gain access to a PE router) is nearly impossible because the packets received from customers are IP packets. These IP packets must be received on a particular interface or subinterface to be uniquely identified with a VPN label.

• Easy to create—To take full advantage of VPNs, it must be easy for you to create new VPNs and user communities. Because MPLS VPNs are connectionless, no specific point-to-point connection maps or topologies are required. You can add sites to intranets and extranets and form closed user groups. When you manage VPNs in this manner, it enables membership of any given site in multiple VPNs, maximizing flexibility in building intranets and extranets.

• Flexible addressing—To make a VPN service more accessible, customers of a service provider can design their own addressing plan, independent of addressing plans for other service provider customers. Many customers use private address spaces and do not want to invest the time and expense of converting to public IP addresses to enable intranet connectivity. MPLS VPNs allow customers to continue to use their present address spaces without Network Address Translation (NAT) by providing a public and private view of the address. A NAT is required only if two VPNs with overlapping address spaces want to communicate. This enables customers to use their own unregistered private addresses, and to communicate freely across a public IP network.

• Integrated Quality of Service (QoS) support—QoS is an important requirement for many IP VPN customers. It provides the ability to address two fundamental VPN requirements:

– Predictable performance and policy implementation

– Support for multiple levels of service in an MPLS VPN

Network traffic is classified and labeled at the edge of the network before traffic is aggregated according to policies defined by subscribers and implemented by the provider and transported across the provider

core. Traffic at the edge and core of the network can then be differentiated into different classes by drop probability or delay.

• Straight forward migration—For service providers to quickly deploy VPN services, use a straightforward migration path. MPLS VPNs are unique because you can build them over multiple network architectures, including IP, ATM, Frame Relay, and hybrid networks. Migration for the end customer is simplified because there is no requirement to support MPLS on the CE router and no modifications are required to intranet belonging to a customer.

Figure 4-8 shows an example of a VPN with a service provider (P) backbone network, service provider edge routers (PE), and customer edge routers (CE).

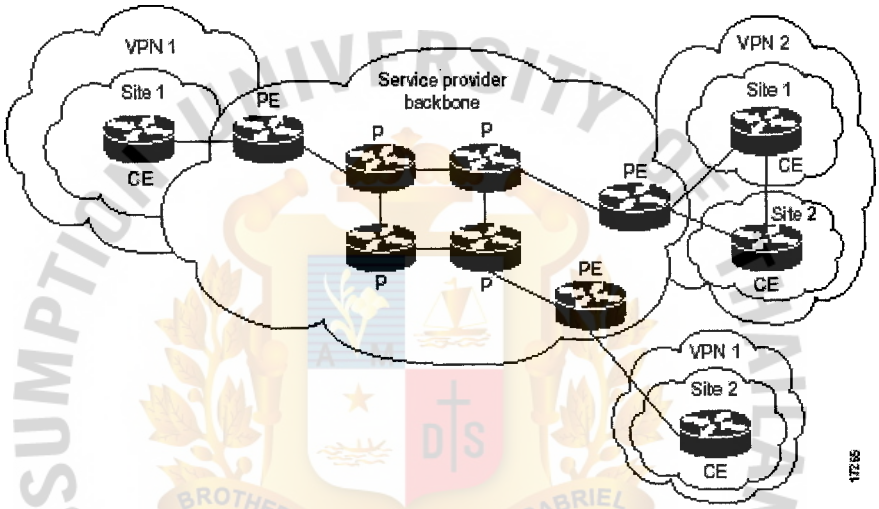


Figure 4-8: an example of a VPN

A VPN contains customer devices attached to the CE routers. These customer devices use VPNs to exchange information between devices. Only the PE routers are aware of the VPNs.

Figure 3-9 shows five customer sites communicating within three VPNs. The VPNs can communicate with the following sites:

- VPN1—Sites 2 and 4
- VPN2—Sites 1, 3, and 4
- VPN3—Sites 1,3, and 5

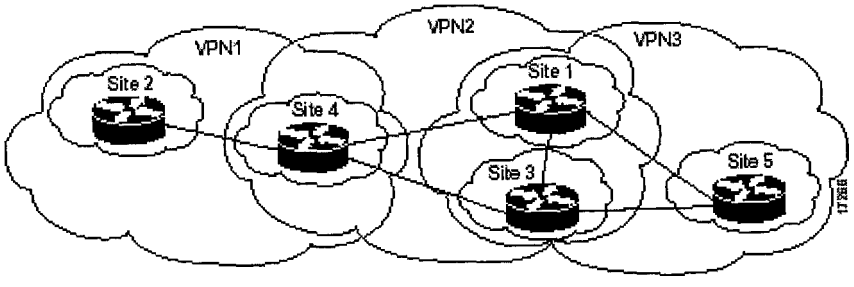


Figure 4-9: Five customer sites communicating within three VPNs

4.12 Increased BGP Functionality

The following is a list of increased BGP functionality:

- Configuring BGP hub and spoke connections—Configuring PE routers in a hub and spoke configuration allows a CE router to re-advertise all prefixes containing duplicate autonomous system numbers (ASNs) to neighboring PE routers. Using duplicate ASNs in a hub and spoke configuration provides faster convergence of routing information within geographically dispersed locations.
- Configuring faster convergence for BGP VRF routes—Configuring-scanning intervals of BGP routers decreases import processing time of VPNv4 routing information, thereby providing faster convergence of routing information. Routing tables are updated with routing information about VPNv4 routes learned from PE routers or route reflectors.
- Limiting VPN VRFs—Limiting the number of routes in a VRF prevents a PE router from importing too many routes, thus diminishing the performance of a router. This enhancement can also be used to enforce the maximum number of members that can join a VPN from a particular site. A threshold is set in the VRF routing table to limit the number of VRF routes imported.
- Reusing ASNs in an MPLS VPN environment—Configuring a PE router to reuse an existing ASN allows customers to configure BGP routes with the same ASNs in multiple geographically dispersed sites, providing better scalability between sites.
- Distributing BGP OSPF routing information—Setting a separate router ID for each interface or subinterface on a PE router attached to multiple CE routers within a VPN provides increased flexibility through OSPF when routers exchange routing information between sites.

4.13 VPN Operation

Each VPN is associated with one or more VRFs. A VRF defines the VPN membership of a customer site attached to a PE router. A VRF consists of an IP routing table, a derived CEF table, a set of interfaces that use the forwarding table, and a set of rules and routing protocol parameters that control the information that is included into the routing table.

A one-to-one relationship does not necessarily exist between customer sites and VPNs. A given site can be a member of multiple VPNs, as shown in Figure 3-9. However, a site can only associate with one (and only one) VRF. A customer's site VRF contains all the routes available to the site from the VPNs of which it is a member.

Packet forwarding information is stored in the IP routing table and the CEF table for each VRF. A separate set of routing and CEF tables is maintained for each VRF. These tables prevent information from being forwarded outside a VPN, and also prevent packets that are outside a VPN from being forwarded to a router within the VPN.

4.14 Distribution of VPN Routing Information

The distribution of VPN routing information is controlled through the use of VPN route target communities, implemented by BGP extended communities. Distribution of VPN routing information works as follows:

- When a VPN route learned from a CE router is injected into BGP, a list of VPN route target extended community attributes is associated with it. Typically the list of route target community extended values is set from an export list of route targets associated with the VRF from which the route was learned.
- An import list of route target extended communities is associated with each VRF. The import list defines route target extended community attributes that a route must have in order for the route to be imported into the VRF. For example, if the import list for a particular VRF includes route target extended communities A, B, and C, then any VPN route that carries any of those route target extended communities—A, **B,** *or* C—is imported into the VRF.

4.15 BGP Distribution of VPN Routing Information

A PE router can learn an IP prefix from a CE router by static configuration, through a BGP session with the CE router, or through the Routing Information Protocol (RIP) exchange with the CE router. The IP prefix is a member of the IPv4 address family. After it learns the IP prefix, the PE converts it into a VPN-IPv4 prefix by combining it with an 8-byte route distinguisher (RD). The generated prefix is a member of the VPN-IPv4 address family. It uniquely identifies the customer address, even if the customer site is using globally non-unique (unregistered private) IP addresses.

The RD used to generate the VPN-IPv4 prefix is specified by a configuration command associated with the VRF on the PE router.

BGP distributes reach ability information for VPN-IPv4 prefixes for each VPN. BGP communication takes place at two levels: within IP domains, known as autonomous systems (Interior BGP or IBGP) and between autonomous systems (Exterior BGP or EBGP). PE-PE or PE-RR (route reflector) sessions are IBGP sessions, and PE-CE sessions are EBGP sessions.

BGP propagates reach ability information for VPN-IPv4 prefixes among PE routers by means of the BGP multi-protocol extensions, which define support for address families other than IPv4. It does this in a way that ensures that only other members of that VPN, enabling members of the VPN to communicate, learn the routes for a given VPN.

4.16 MPLS Forwarding

Based on routing information stored in the VRF IP routing table and VRF CEF table, packets are forwarded to their destination using MPLS.

A PE router binds a label to each customer prefix learned from a CE router and includes the label in the network-layer reach ability information for the prefix that it advertises to other PE routers. When a PE router forwards a packet received from a CE router across the provider network, it labels the packet with the label learned from

the destination PE router. When the destination PE router receives the labeled packet, it pops the label and uses it to direct the packet to the correct CE router. Label forwarding across the provider backbone, is based on either dynamic label switching or traffic engineered paths. A customer data packet carries two levels of labels when traversing the backbone:

- The top label directs the packet to the correct PE router.
- The second label indicates how that PE router should forward the packet to the CE router.

4.17 Inter-autonomous Systems for MPLS VPNs

The inter-autonomous system for MPLS VPNs feature allows an MPLS VPN to span service providers and autonomous systems.

As VPNs grow, their requirements expand. In some cases, VPNs need to reside on different autonomous systems in different geographic areas. (An autonomous system is a single network or group of networks that is controlled by a common system administration group and that uses a single, clearly defined routing protocol.) Also, some VPNs need to extend across multiple service providers (overlapping VPNs). Regardless of the complexity and location of the VPNs, the connection between autonomous systems must be seamless to the customer.

The inter-autonomous system for MPLS VPNs feature provides seamless integration of autonomous systems and service providers. Separate autonomous systems from different service providers can communicate by exchanging IPv4 network layer reach ability information (NLRI) in the form of VPN-IPv4 addresses. The border edge routers of autonomous systems use the EBGP to exchange that information. Then, an IGP distributes the network layer information for VPN-IPv4 prefixes throughout each VPN and each autonomous system. Routing information uses the following protocols:

- Within an autonomous system, routing information is shared using an IGP.
- Between autonomous systems, routing information is shared using an EBGP. An EBGP allows a service provider to set up an inter domain routing system that guarantees the loop-free exchange of routing information between separate autonomous systems.

An MPLS VPN with inter-autonomous system support allows a service provider to provide to customers scalable Layer 3 VPN services, such as web hosting, application hosting, interactive learning, electronic commerce, and telephony service. A VPN service provider supplies a secure, IP-based network that shares resources on one or more physical networks.

The primary function of an EBGP is to exchange network reach ability information between autonomous systems, including information about the list of autonomous system routes. The autonomous systems use EGBP border edge routers to distribute the routes, which include label-switching information. Each border edge router rewrites the next hop and MPLS labels.

Inter-autonomous system configurations supported in an MPLS VPN can include the following:

- Inter-provider VPN—MPLS VPNs that include two or more autonomous systems, connected by separate border edge routers. The

autonomous systems exchange routes using EBGP. No IGP or routing information is exchanged between the autonomous systems.

• BGP confederations—MPLS VPNs that divide a single autonomous system into multiple sub-autonomous systems, and classify them as a single, designated confederation. The network recognizes the confederation as a single autonomous system. The peers in the different autonomous systems communicate over EBGP sessions; however, they can exchange route information as if they were IBGP peers.

Benefits of inter-autonomous Systems for MPLS VPNs are as follows:

• Allows a VPN to cross more than one service provider backbone—The inter-autonomous systems for MPLS VPNs feature allows service providers, running separate autonomous systems, to jointly offer MPLS VPN services to the same end customer. A VPN can begin at one customer site and traverse different VPN service provider backbones before arriving at another site of the same customer. Previous MPLS VPNs could only traverse a single BGP autonomous system service provider backbone. The inter-autonomous system feature allows multiple autonomous systems to form a continuous (and seamless) network between customer sites of a service provider.

• Allows a VPN to exist in different areas—The inter-autonomous systems for MPLS VPNs feature allows a service provider to create a VPN in different geographic areas. Having all VPN traffic flow through one point (between the areas) allows for better rate control of network traffic between the areas.

• Allows confederations to optimize IBGP meshing—The inter-autonomous systems for MPLS VPNs feature can make IBGP meshing in an autonomous system more organized and manageable. You can divide an autonomous system into multiple, separate sub-autonomous systems and then classify them into a single confederation (even though the entire VPN backbone appears as a single autonomous system). This capability allows a service provider to offer MPLS VPNs across the confederation because it supports the exchange of labeled VPN-IPv4 NLRI between the sub-autonomous systems that form the confederation.

4.18 Routing Between Autonomous Systems

Figure 4-10 illustrates one MPLS VPN consisting of two separate autonomous systems. Each autonomous system operates under different administrative control and runs a different IGP. Service providers exchange routing information through EBGP border edge routers (ASBR1 and ASBR2).

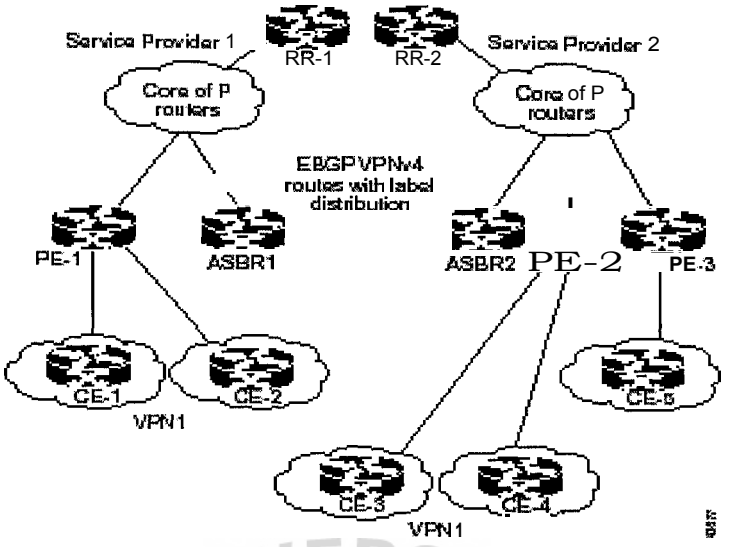


Figure 4-10: One MPLS VPN consisting of two separate autonomous systems

Step 1 The provider edge router (PE-1) assigns a label for a route before distributing that route. The PE router uses the multiprotocol extensions of a BGP to send label-mapping information. The PE router distributes the route as a VPN-IPv4 address. The address label and the VPN identifier are encoded as part of the NLRI.

Step 2 The two route reflectors (RR-1 and RR-2) reflect VPN-IPv4 internal routes within the autonomous system. The border edge routers of autonomous systems (ASBR1 and ASBR2) advertise the VPN-IPv4 external routes.

Step 3 The EBGP border edge router (ASBR1) redistributes the route to the next autonomous system (ASBR2). ASBR1 specifies its own address as the value of the EBGP next hop attribute and assigns a new label. The address ensures the following:

- That the next hop router is always reachable in the service provider (P) backbone network.
- That the label assigned by the distributing router is properly interpreted. (The corresponding next hop router must assign the label associated with a route.)

Step 4 The EBGP border edge router (ASBR2) redistributes the route by ASBR2 changes the next hop address of updates received from the EBGP peer, then forwards it or ASBR2 must propagate a host route for the EBGP peer through the IGP. The EBGP VPN-IPv4 neighbor host route is automatically installed in the routing table when the neighbor comes up. This is essential to establish the label-switched path between PE routers in different autonomous systems.

4.19 Summary of Configuration Options

4.19.1 Option 1: Dynamic versus Static Routing between CEs and Pes

Routing protocols between CEs and PEs must be secured with the appropriate authentication mechanisms to ensure that only the CE router can send routing updates. Furthermore, the routing protocols must be further secured from the SP side in order to not be vulnerable to routing attacks (malicious or inadvertent). For example, in

BGP, it is possible to configure dampening parameters, where only a limited number of routing updates are accepted in a period of time.

Even with these precautions, an attacker cannot be prevented from finding a way to flood the router with bogus routing messages. The existence of a peer IP address might be enough to prevent this type of attack. Flooding the PE router from a CE can not break security as far as the MPLS mechanisms are concerned, but might have an effect on router performance, which might in turn influence performance of other VPNs.

However, practically, it must be considered that in this scenario the potential attack can come only from a legal user of the VPN service. This type of attack is easily traceable to one port—and thus one VPN customer. It can easily be stopped by shutting down that particular interface.

If there is no routing protocol running between CEs and PEs, static routing is required. If this routing is configured on an unnumbered link, just pointing to an interface rather than to a peer IP address, the CE does not need to know any addressing information of the MPLS core. Sending any type of message to the PE router will not have an effect, because it will be treated within the VRF. An attacker could still guess the address space of the router (for example, the loopback address), but it can be protected with ACLs that do not permit any packet, because no communication is required with this address. In this scenario, the security is very high and fully comparable to similar Layer 2 services (FR, ATM).

4.19.2 Option 2: Internet Service

As long as MPLS/BGP VPNs are not connected to the Internet or other VPNs, MPLS provides a high level of security. In the case of Internet access through the MPLS network, all the rules of accessing the Internet in general apply. Most important, a firewall should be placed between the VPN and the Internet. The various options are described above. If configured correctly, Internet access over MPLS can be offered in a secure manner.

The same applies to various VPNs that are merged on the MPLS network. MPLS itself does not provide firewalling mechanisms, but an MPLS core can be engineered such that firewalls secure VPNs but allow connectivity.

4.19.3Option 3: Running IPSec over the MPLS Cloud

If the security of the SP MPLS network is considered insufficient, there is the additional option to run IPSec on the CE routers or behind or over the MPLS cloud, with encryption (Encapsulating Security Protocol [ESP]) and authentication (AH). In this case, even attacks within the MPLS cloud cannot break the security of the overall VPN. Traffic on the MPLS network can be traced back to only two routers; the content is not legible. Changes on existing packets as well as fake or spoofed packets will be detected by the IPSec AH mechanisms.

In the case of dynamic routing, the potential routing attacks as described above can still be carried out, so DoS from a neighbouring VPN might be possible. Static routing provides more security than dynamic routing, but in this case static routing can be easily configured also from the SP side, because the SP sees only packets between the CE routers. Thus on each PE router, one static route to the corresponding CE router is sufficient. Note, however, that security of a given VPN depends on the security of the overall MPLS service.

4.19.4 Option 4: Including the CE Router in the SP Management

All discussions so far have assumed that the interface between the customer and the SP is between the CE and PE routers. However, in reality, many existing service offerings include the CE router as SP-managed customer premises equipment (CPE). This setup has numerous consequences for security:

• The core network can now be completely hidden to the customer networks, because customers have no direct connection to the MPLS core. This setup improves security.

• In addition to the PE, the CE can now also be configured with strict ACLs that also control access to the PE router.

• The routing protocol between the CE and the PE is now under the control of the SP. Routing must also run between the customer's network and the CE, so there is still some potential for routing attacks. But because the CE can be secured with ACLs and additional routing security, the overall setup will be more secure.

• If IPSec is required, running IPSec from the CE routers means in this scenario giving control of encryption to the SP. Some organizations might prefer to keep control of the encryption, in which case the IPSec should be done on another device before the CE router.

4.20 Security comapration of MPLS and ATM/FR

MPLS provides full address and routing separation as in traditional Layer 2 VPN services. It hides addressing structures of the core and other VPNs, and it is in today's understanding not possible from the outside to intrude into the core or other VPNs abusing the MPLS mechanisms. It is also not possible to intrude into the MPLS core if it is properly secured. However, there is a significant difference between MPLS-based VPNs and, for example, FR- or ATM-based VPNs: The control structure of the core is on Layer 3 in the case of MPLS. This fact has caused significant scepticism in the industry toward MPLS, because this setup might open the architecture to DoS attacks from other VPNs or the Internet (if connected). Table 4-4 compares ATM/FR with MPLS.

	ATM/FR	MPLS
Address Space Separation	Yes	Yes
Routing Separation	Yes	Yes
Resistance to Attacks	Yes	Yes
Resistance to Label Spoofing	Yes	Yes

Table 4-4: Security comapration of MPLS and ATM/FR

As shown in this chapter, it is possible to secure an MPLS infrastructure to the same level of security as a comparable ATM or FR service. It is also possible to offer Internet connectivity to MPLS VPNs in a secure manner, and to interconnect different VPNs by firewalls. Although ATM and FR services have a strong reputation with regard to security, it has been shown that security problems can also exist in these networks.

With regard to attacks from within the MPLS core, all VPN classes (MPLS, FR, ATM) have the same problem: If an attacker can install a sniffer, he/she can read information in all VPNs, and if the attacker has access to the core devices, he/she can execute a large number of attacks, from packet spoofing to introducing a new peer

router. Numerous precaution measures that an SP can use to tighten security of the core are outlined above, but the security of the MPLS architecture depends on the security of the SP. If the SP is not trusted, the only way to fully secure a VPN against attacks from the "inside" of the VPN service is to run IP Sec on top, from the CE devices or beyond.

This chapter discusses many aspects of MPLS security. It should be noted explicitly that the overall security of MPLS architecture depends on all components, and is determined by the security of the weakest part of the solution. For example, a perfectly secured static MPLS network with secured Internet access and secure management is still open to many attacks if there is a weak remote-access solution in place.

CHAPTER 5: Migration to MPLS, Case study and Investment

5.1 Introduction

The business potential and operational effective of using new MPLS networks to carry IP traffic are quickly become apparent. It is only a matter of time until IP/MPLS becomes the industry's multiservice network technology platform of choice. However, incumbent service providers cannot ignore the fact that ATM, frame relay and private lines network are hard at work today generating the majority of their data service revenue.

5.2 Migration Consideration

Given that ATM, frame relay and private lines remain large revenue generators and that a tight economy is putting a squeeze on new capital investments, many service providers are torn as to whether they should risk embarking on migration to IP/MPLS network. Equipments vendors are urging them to do so. The promise is that IP/MPLS will open doors to new high-margin services that can be provisioned and managed in a simpler and more scalable manner than using virtual circuit-based ATM and frame relay technologies.

But MPLS is still in the early of adoption. It is about two years into a three- to five-year initial implementation period with a scant degree of network integration. As such, the technology does not yet have an established history of business success. So service providers are having a tough time justifying a major overhaul of their network infrastructure and the back-end support systems that accompany them.

At the same time, service providers don't want to be tardy to market with new generation IP service because they have postponed new infrastructure investments. To minimize risk while maximize revenue, service providers should consider a multiservice switching migration path that enables them to move stages to telephony-grade IP/MPLS network infrastructure. This way, they can be ready with MPLS when it becomes a competitive necessity. This phased migration must be conducted transparently to existing customers so as not to forfeit revenue.

There are two basic options for introduction IP/MPLS platforms. One is for network operators to run parallel networks- one supporting IP/MPLS and the other legacy ATM and frame relay traffic. Over time, the ATM-based network simply depreciates and eventually gets unplugged. The second, more cost-effective and operationally efficient approach is to begin integrating new equipment with old, switch by switch. The new equipment should accommodate all of the legacy networks in place today in addition to IP/MPLS networks.

To this end, service providers are starting to look to a new generation of multiservice switching equipment that can support legacy, revenue-generating ATM, frame relay and private lines network alongside IP/MPLS connections. Such equipments, in addition to being telephony-grade in terms of uptime and reliability, should support open interface to back-end management and operations support systems. It should also support unprecedented throughput capacity – scaling up into the hundreds of gigabits per second – for accommodating increased traffic loads with no performance degradation.

Use of equipment with these characteristics will enable service providers to begin installing highly reliable, multiservice IP/MPLS networks at a pace with which they are comfortable and without impacting the services already enjoyed by their

paying customers. This equipment can be installing in the conjunction with the existing ATM and frame relay switches and, going forward, can eventually replace legacy switch both at the network edge and in the core.

A graceful strategy like this represents the migratory path of least disruption and expense for service providers. It protects the existing, profitable network infrastructure while preventing carriers from having to continue investing in outdated multiservice switching equipment.

5.3 Releases for Incumbent Provider Pain

Since the vast majority of service provider data revenue is currently coming from virtual circuit and private line networks, it is advisable for incumbent operators to use existing revenue to fund the expansion and profitable of existing services without mortgaging their MPLS potential. Through Internet access services are rampant, few highly profitable IP/MPLS services are being sold yet, largely because if the technology investment and migration issues that have been raised in this paper. Most carriers are averse to being early MPLS adopters under current market conditions.

Still, it is expensive to continue investment in ATM and frame relay switches that will not reliably and efficiently accommodate IP/MPLS service going forward. The expense is not only the capital and operation costs. There are also significant costs associated with the service opportunities lost by not having an IP/MPLS platform in place when customer demand for new applications drives the need for an upgraded architecture.

Existing multiservice switching will still have a place in certain segments of the network for many years to come. But in certain places in the network- at the aggregation edge, which does not have the benefit of a fully redundant core.

Scalability, footprint, hitless software upgrades and downgrades, and common back-end interfaces join network availability as key requirement. Each contributes to a service provider's ability to deploy and manage services quickly with minimum investments in new capital, real estate, and sales and support staff retraining.

5.3.1 Network availability

Most service level agreement (SLAs) associated with today's IP virtual private network (VPN) service promise customers between 99.8% and 99.9% uptime. A few IP VPN service providers offer SLAs for 99.99% uptime. Virtual circuit service,ATM and frame relay generally carry 99.99% uptime SLAs.

5.3.2 Scalability

This factor applied to service provisioning and management, network capacity, and number of calls per second supported.

The first item- scalable service management is perhaps MPLS's greatest value-add to service provider. MPLS eliminate use of the virtual circuit identifier/virtual path identifier (VCI/VPI) in ATM networks and data link connection identifier (DLCIs) in frame relay networks for addressing. Service provider can add customers and new network sites in a meshed configuration without having to set up VPI/VCI or DLCI for each pair of communications sites.

5.3.3 Footprint

Along with capital equipments investments, the cost of real estate to new house switches also comes at a premium. In addition, service providers do not always have a choice as to where they run a point of presence. For example, they might need to collocate equipment with another provider where there is limited space available. It is thus financially advantageous to have as much functionality and capacity as possible packed into a small form factor with the lowest possible power consumption.

5.3.4 No interruptive upgrades

Not only must a system's software architecture be resilient to software faults, it must have a built-in mechanism that allows for hides system upgrades and downgrades with no disruption in functionality. This capacity has long been solved in existing switches on hardware side with is known as hot swap capability.

The same capacities are necessary on the software side. Just as hardware components can be automatically fail over to backup resources.

5.3.5 Common management protocols and operation support system (OSS) APIs

For multiservice switches to scale while supporting interoperability among various legacy and next generation networks, an integrated network management system is imperative.

Interoperability with existing OSSs is a key migration component. Incumbent service providers have spent years building up their backend provisioning, billing, and customer service management systems and developing the related staffs experience required for running them. Facing an overhaul of this portion of the network represents more of hardship to many service providers than finding capital for new equipment investments.

5.4 Case study of MPLS

5.4.1 Completel: National Metro Service Deployment

Established in France, Germany and the United Kingdom, Completel has deployed fiber optic Metropolitan Area Networks (MANs) in 13 cities to deliver Ethernet LAN-to-LAN services to its business clients. Completel offers companies a simple and fast way to meet its service needs, including telecom, high-speed data, Internet access, hosting, email, and associated services. Completel's services include higher performance, more reliability, more competitive pricing than existing services, and a more attractive alternative to the services provided by traditional telecom operators.

When Completel wanted to upgrade its basic metro LAN-to-LAN service with a more scalable and service-rich offering, it chose Riverstone. Completel is now using Riverstone's metro routers to offer premium services based on Gigabit Ethernet to its customers in its MAN deployments throughout France and Germany. A key deciding factor was Riverstone's support for industry-leading Multi-Protocol Label Switching (MPLS) to allow the creation of services within and between MANs.

Completel's MAN Service Challenges

Completel built an optical backbone network to link the 13 MAN and 5 hosting centers in less than 3 years, deploying around 2,350 km of fiber that is now fully operational. The 13 MANs are in Paris, Lille, Lyon, Marseilles, Grenoble, Toulouse, Nice/Sophia-Antipolis, Nantes, and Strasbourg in France, and Berlin, Essen, Munich, and Nuremberg in Germany.

Completel's existing LAN-to-LAN metro service offers basic point-to-point LAN connections to customers using Ethernet and VLAN technology. Customers profit from a high performance and reliable service that is optimized for LAN-to-LAN traffic. However, the service is limited to within each city MAN; to provide LAN-to-LAN services between cities, Completel has to route the traffic between MANs. This results in a considerable impact on the performance of its existing metro equipment.

Completel wanted to offer new premium services, such as high-speed Internet access, Transparent LAN Services, and VPN services, to its enterprise and SME customer base. Completel realized that it would be unable to change its business from a connection and bandwidth business to a value-added service oriented business unless it replaced its existing metro equipment. So it started to evaluate and test metro router solutions from a number of vendors.

Riverstone's Solution

The solution proposed by Riverstone was based on the RS 38000 metro backbone router to provide the key services both within and between the MANs.

In each city, Riverstone proposed a ring of RS 38000s built using parallel Gigabit Ethernet connections to provide a high-speed metro backbone ring. In some cities, multiple parallel Gigabit Ethernet connections are used to create very high capacity metro backbone rings. The RS 38000s have both 10/100 and Gigabit Ethernet interfaces for customer service delivery, with inter-metro connections using SDH STM-4.

Riverstone proposed a two-phase approach for the introduction of new services.

- For the initial national services, Riverstone proposed an IP VPN solution based on Layer 2 MPLS tunnels to allow Completel to offer Extranet services to corporate customers, with Internet services provided using Riverstone's proven implementation of BGP, scalable Network Address Translation (NAT), and Access Control List capability.
- For the follow-on LAN-to-LAN service, Riverstone proposed a solution that uses Transparent LAN Services based on mapping of 802.1Q VLAN tags and 802.1p Quality of Service (QoS) to Layer 2 MPLS tunnels, with rate limiting to control customer bandwidth in 1 Mbps increments. Completel will use Riverstone Stackable VLAN to provide scalability and Rapid Spanning Tree Protocol to ensure fast reconfiguration in the event of a failure.

As Riverstone has a number of network management technology partners, Riverstone was able to propose a Network Management System (NMS) to meet Completel's immediate needs but which also provides a migration path to a more

complete Operation and System Support (OSS) solution in the future. The NMS includes provisioning and service assurance tools to allow Completel to provision new services for customers quickly and to monitor those services accurately to ensure that Completel can meet the specified Service Level Agreement (SLA) for each customer.

*Why **Riverstone?***

Riverstone responded to Completel's initial request very quickly, providing a full response that immediately addressed Completel's commercial and technical issues.

Completel was impressed by the strength of Riverstone's MPLS implementation, which was evaluated against competitive offerings during extensive benchmark testing. Riverstone's compliance to standards, leading role in the MPLS Forum, demonstration of future MPLS capability, and commitment to implement MPLS in hardware also impressed Completel.

The Riverstone solution met Completel's benchmark test criteria thanks to Riverstone's high-quality technical support throughout the duration of the tests, fast turnaround of software code, and a commitment to provide features specifically to meet Completel's requirements. When compared to the competition, Completel found that the Riverstone solution is carrier class, integrates Ethernet connectivity with IP routing and switching, and has the most features implemented in hardware ASICs.

The rich network management features provided by Riverstone was important to Completel, including the support for SNMP, RMON1 and RMON2 in hardware, and the Lightweight Flow Accounting Protocol (LFAP) which allows accounting information to be collected accurately and reliably. Completel was also impressed by the NMS solution that Riverstone proposed to meet Completel's current needs, while allowing a future migration to a more complete OSS solution as and when required.

5.4.2 Cable & Wireless Further Extends OC-192 Using MPLS Across Global IP Network

Cable & Wireless (NYSE: CWP; LSE: CW), the global telecommunications group, today announced it has upgraded the US portion of its global IP network to OC-192 running MPLS from coast-to-coast. The network upgrade, which delivers OC-192 network speeds from the west coast of the United States across the Atlantic Ocean and into Europe, further reinforces Cable & Wireless' leadership position for providing superior network performance, quality and reach, consistently around the globe. Through the combination of its global high performance network and IP services with Exodus' hosting and content delivery services, Cable & Wireless is now the premier choice for eBusiness infrastructure solutions in the US, Europe and Asia-Pacific.

The upgrade follows the OC-192 multi-city links brought on-line in the fourth quarter of 2001 connecting Washington DC and New York with London, Paris, Brussels, Amsterdam and Frankfurt. The current upgrade adds Anaheim, Atlanta, Chicago, Dallas and Santa Clara. The high speed connections support the unique network performance requirements of large scale Internet content providers in particular, as well as providing sufficient scale to keep Cable & Wireless ahead of increasing demand for IP services.

"Cable & Wireless continues to define what is truly a global IP network," said Andy MacLeod, chief operating officer, Cable & Wireless. "Cable & Wireless is committed to providing customers, whether large global enterprises or content providers, with an industrial strength IP infrastructure that provides the highest quality of service and performance on a global scale."

Multi-protocol label switching (MPLS) technology allows Cable & Wireless to provide enterprises and service providers with an Internet infrastructure to support all their applications, connectivity and content needs - even those time critical and mission critical services that they would typically not consider transmitting over other best effort Internet backbones. Cable & Wireless' global IP network provides:

* Large-scale Internet content providers with the high quality network connections required to efficiently and securely link hosted and cached content with their target market users even when the geographical span is several continents.
* Global ISPs with a quality IP backbone to deliver traffic-engineered services to their own customers. The additional capacity and predictability offered by the upgrade is crucial for maintaining this ability.
* Large multinational enterprise customers with the extra capacity, performance and predictability consistently around the globe to keep them ahead of the curve in terms of offering services to their customers and linking their own operations.

Network Architecture

Cable & Wireless' global IP network is based on core traffic engineering platform using intelligent MPLS routing and switching capability at 10 Gbps. This leading platform provides a scalable, reliable and more economic transport system for individual services, such as IP transit, hosting, and content delivery services.

With this network architecture, Cable & Wireless is able to:

* Offer high capacity OC-48/STM-16 IP access services to carriers, content providers, ISPs and large enterprises.
* Scale the network to handle the anticipated increase in Internet traffic in and between the US and Europe.
* Optimally restore network traffic and minimize service disruptions in the event of major network disruptions.

About Cable & Wireless

Cable & Wireless is a major global telecommunications business with revenue of over 5.9 billion pounds sterling (US$8.6 billion) in the year to 31 March 2002 and customers in 70 countries. The company consists of two core and complementary divisions: Cable & Wireless Regional and Cable & Wireless Global. Cable & Wireless Regional offers a full range of telecommunications services in 33 countries around the world. Cable & Wireless Global's focus for future growth is on IP (Internet protocol) and data services and solutions for business customers. It has developed advanced IP networks and value-added services in the US, Europe and the Asia-Pacific region in support of this strategy. With its financial strength and the capability of its global IP infrastructure, Cable & Wireless holds a unique position in

terms of global coverage and services to business customers. For more information about Cable & Wireless, go to http://www.cw.com.

5.4.3 Heidelberg Chooses AT&T For Global Networking Solution In $17m Agreement

Heidelberg, Germany, -- Heidelberger Druckmaschinen AG (Heidelberg), the world's leading solution providers for the print media industry has chosen AT&T to consolidate its telecommunications services currently provided by more than 30 other providers. AT&T will become Heidelberg's preferred supplier of business data communications and networking worldwide.

AT&T will provide a fully managed global wide-area network (WAN) service, connecting 150 Heidelberg locations in Europe, Asia Pacific and the Americas, supporting the growing globalisation of the Heidelberg group. AT&T will also provide managed Internet services. Under the terms of an initial three-year agreement, worth $12m AT&T will migrate Heidelberg's current network, based on global Frame Relay technology, to the latest IP based Multi Protocol Label Switching (MPLS) infrastructure. In a further $5m agreement, AT&T will provide a fully managed hosting solution for Heidelberg. Heidelberg also uses AT&T local service, long distance, teleconferencing, video-conferencing, frame relay and private line services in the US.

This new structure is expected to result in substantial cost savings for Heidelberg, estimated at up to 40% versus its current solution and emphasizes Heidelberg's long term commitment to further cost reductions in its communication and IT expenditure. AT&T will deliver one seamless network, with standard network services, pro-active monitoring and network management. This will facilitate the ready introduction of global remote access services for Heidelberg employees worldwide, supported by a global helpdesk for network services.

Heidelberg will also cooperate with AT&T Labs on future developments, ranging from new e-commerce services to web-based remote maintenance applications.

Commenting on the agreement, Dr. Herbert Meyer, Heidelberg's CFO said: "We recognize that if we are to maintain our global leadership in our industry we need to grasp opportunities for change and innovation ahead of our competitors. This new network solution enables us to introduce new, innovative networking technologies and cost effective management processes while considerably reducing our budget. We chose AT&T as a networking partner as it is able to provide us with a global network that can deliver optimal performance for a wide mix of enterprise applications."

Michael Neff, CIO of Heidelberg commented: "Heidelberg looked for a networking partner with "Thought Leadership" expertise, capable of contributing to the evaluation of new e-business opportunities. It also clearly has the resources to help us realize such projects. We have entrusted our operations to a stable carrier with a network capable of providing us with consistent performance and reliability."

Ken Sichau, President AT&T Business Sales, said, "Today's most critical business applications ride on networking. In this environment networking is not a commodity, it is a source of competitive advantage. We are proud to help Heidelberg maintain and extend its industry leadership position. This

announcement demonstrates that European companies trust in AT&T as a truly global service provider."

Heidelberg made the decision to appoint AT&T after the development of its new strategic plan, "Heidelberg Global Network", which defines Heidelberg's demands for networking capabilities based on future business developments and processes. Over the last few years, the group has made enormous investments in enterprise resource planning, supply chain management, and customer relationship management systems. It therefore recognized the need for a new, highly flexible networking solution, which would enable it to rapidly introduce new applications, connect new sites and have immediate access to additional bandwidth.

Managed services are central to AT&T's overall business strategy for enterprise customers. AT&T Managed Services consist of integrated offer portfolios — including, Enterprise VPNs, Hosting, Business Continuity (High Availability and Security Services) – that are combined with the seamless support of AT&T's integrated Global Enterprise Management System (iGEMS), which provides proactive and predictive networking-management capabilities.

AT&T Managed Services address the interdependency of networking, hosting and businesses' needs for reliability, security and recovery. They simplify the network complexity and operational challenges faced by companies while allowing them to easily extend their networks to customers, suppliers and partners. AT&T Managed Services can encompass access service options; hosting; virtual private networks (VPNs); content distribution; managed wide-area networks (WAN); enhanced transport maintenance; networked computing services; business continuity and security services.

5.5 Investment of MPLS, core network model

This chapter describes a business investment model of the IP packets network service of provider. It provides details of two components of the provision of the IP network services (routers based) and new scenario of implementation of MPLS and value add network model. By using estimated Thailand service provider's data and the IP network architectures that currently exist.

5.6 The Model Aim

The objective of the models is to calculate the costs of migration of IP Routing based network to MPLS network and the example of operation costs of providing services on MPLS based network. As mentioned, IP network service is the main revenue of service providers. The model anticipates VPN services over MPLS network. These costs are calculated yearly and then discounted at a cost of capital of 16% based on industry in year 1996. Summing the yearly costs provides the total discounted costs per year. The model is base on the premise routers and VPN implementation over the MPLS network. The model results that are presented in this section are given in terms of costs of implementation and migration MPLS network incurred by the service provider.

5.7 Existing IP networks of service providers

There are many kinds of devices and components to build up an IP core network of service providers. For traditional routing based network, routers are one device of core components in network implementation. Cisco is the one of router provider who has the largest numbers of world market share. Most service provider used Cisco's router as their core router in the network's backbone. In the table 5-1 show the model of Cisco routers and its' market price for Thailand customer.

Cisco 12000	7,200,000	450,000	7,650,000
Cisco 7000	714,000	240,000	954,000
Cisco 4000	380,000	190,000	570,000
Cisco 3600	180,000	9,500	189,500
Cisco 2600	119,000	38,500	157,500
Cisco 1700	72,000	24,000	96,000

Table 5-1 : Cisco's router Market price

The designing assumption in this paper will use Cisco Router as base product in the example of Model.

Another components, leased lines, interconnection between routers are the one of important components in IP network backbone. Table 5-2 show the market pricing of domestic leased lines in Thailand.

Table 5-2: Leased line market prince

Speed	Installation Charge	Within province	Connected Province	Unconnected Provinces (Distance/Km.)			
				0-200	201-400	401-600	>600
19.2Kbps.	10,000		8,000	12,000	18,000	21,000	30,000
[illegible]	10,000	8,000	11,000	15,000	24,000	28,000	40,000
128Kbps.	10,000	12,000	17,000	25,000	36,000	45,000	60,000
192Kbps.	10,000	15,000	21,000	30,000	45,000	54,000	75,000
256Kbps.	10,000	17,000	26,000	36,000	54,000	64,000	90,000
384Kbps.	20,000	23,000	34,000	45,000	63,000	80,000	113,000
512Kbps.	20,000	28,000	38,000	55,000	81,000	96,000	135,000
768Kbps.	20,000	36,000	51,000	69,000	104,000	122,000	172,000
1.024Mbps.	30,000	44,000	60,000	82,000	123,0001	145,000	210,000
1.536Mbps.	30,000	57,000	77,000	101,000	160,000	185,000	261,000

5.8 Model Designing Assumption

The assumption of model design based on the existing core network and its components. Figure 5-1 show the existing devices of Model's IP core network — Routers based network.

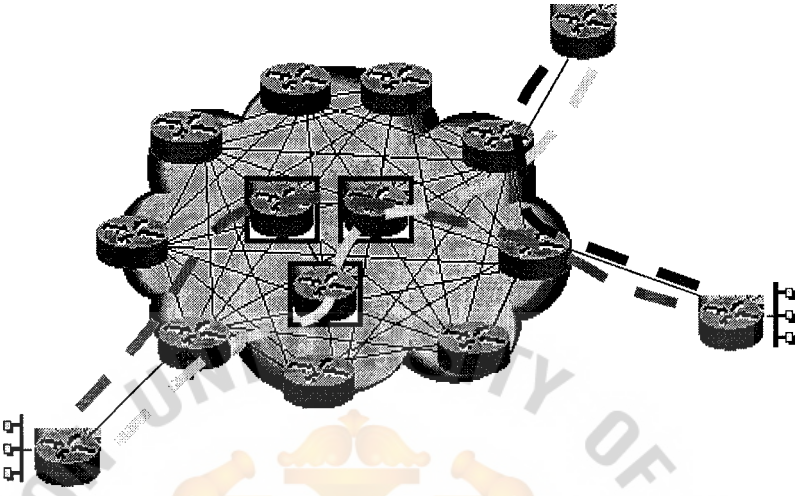


Figure 5-1: Model Designing Assumption

In case of migration from traditional IP network to MPLS in this paper applies to router-only networks, which might use MPLS for supporting VPN services or IP traffic engineering. In this structure, customer sites are connected directly to router-based edge LSRs. The edge LSRs are connected to other LSRs, which are also based on router platforms. The routers are interconnected by virtually any sort of link: serial, Ethernet, Packet over SONET (PoS), and so on, and packets are sent, with MPLS headers, over these links. The routers involved will typically be Cisco 7200, 7500, or 12000 series Gigabit Switch Routers (GSR). Midrange routers (Cisco 2600, 3600, and 4700 series) might be used in lower-bandwidth applications and as the CPE. Figure 5-2 shown the simple packet-based MPLS network.

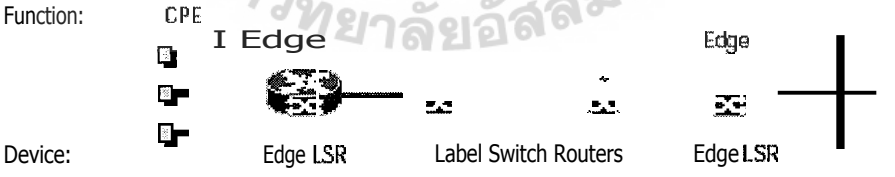


Figure 5-2: Simple packet-based MPLS network

5.9 Architecture of Model Network

The model network in this paper is designed based on replacing the old devices with new model of routers with MPLS supported.

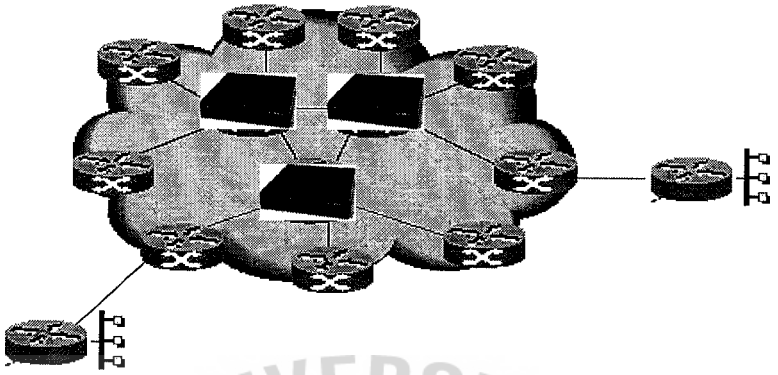


Figure 5-3: Migration to MPLS network

3 Backbone routers are Cisco4000 with serial (WAN) ports module for interconnection between edge routers.
9 Edge routers are cisco2600, which connected to customer routers.
Leased line connection between backbone routers, there are 3 circuits as the backbone, leased lines at speed 1.5Mbps (T1) connect each Cisco4000.
Leased line connection between backbone routers and edge routers, Cisco4000 and Cisco2600. Speed 512Kbps for each. There are 9 circuits, which are PPP topology.

From the network diagram, one Cisco2600 can support maximum 8 customers per one box so there will be 72 customers, maximum of the model designing. In this paper will assume that there are 36 customers existing in backlog of service provider.

So cost of investment Model network, which are assumed as the initial cost shown in the table 5-3 (routers components)

Components	Unit(s)	Cost/unit (Bht)	Total (Bht)
Cisco4000	3	570,000	1,710,000
Cisco2600	9	157,500	1,417,500
Leasedline-1.5 Mbps	3	30,000	90,000
Leasedline-512 Mbps	9	20,000	180,000
Installation charge			240,000
Total cost			3.63 [illegible]

Table 5-3: Devices pricing (router components)

Note that use installation charge of leased lines as the initial cost of implementations

5.10 Revenue Forecast

Revenue estimates this model is derived by apply number of customers in each year, tariffs, subscription and value added service (VAS) charges to the customers.

5.10.1 Pricing Assumption

Pricing assumption base on license of Internet from CAT. The market prices of Internet service are show in the following table 5-4.

Speed (kbps)	Market Price (Baht)	
	Monthly	Installation
64	16,000	N/A
128	30,000	N/A
256	55,000	N/A
512	100,000	N/A
1,000	180,000	N/A

Table 5-4: Inter service Market pricing

From the characteristic of MPLS, customer has invisibles the MPLS core network. Customer might not see any changing in Internet service except new value-add services from MPLS such as VPN, SLAB or QoS. There will be the assumption that 30-50% markup of MPLS value-add service from current Internet access charge.

Speed (kbps)	Market Price (Baht)	
	Monthly (+ 20%)	Installation
64	20,800	N/A
128	39,000	N/A
256	71,500	N/A
512	130,000	N/A
1,000	234,000	N/A

Table 5-5: MPLS network service pricing

From model network, there are 36 customers in backlog for the first year of investment and every customer access to core model network at speed 64kbps. The first year revenue is 8,985,000 Baht (20,800 bht x 12 months x 36 customers).

Revenue forecast of the years after, assume that there will be average 6 customers per year then 6 years afterward the model network capacity will be reach maximum number of customers, 72 customers.

Revenue/ No. of customers	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7
36 Customers	8,985,000	8,985,000	8,985,000	8,985,000	8,985,000	8,985,000	8,985,000
6 Customers		1,497,600	1,497,600	1,497,600	1,497,600	1,497,600	1,497,600
6 Customers			1,497,600	1,497,600	1,497,600	1,497,600	1,497,600
6 Customers				1,497,600	1,497,600	1,497,600	1,497,600
6 Customers					1,497,600	1,497,600	1,497,600
6 Customers						1,497,600	1,497,600
6 Customers							1,497,600
Total	8,985 000	10,483,200	11,980,800	13,478400	14,976 000	16,473 000	17.971,200

Table 5-6: Revenue forecast

In customer perspective, the pricing in the table are only the yearly market prices of Internet access service only. They are not including leased line for corporate connection. For all pricing structure, it will not include leased line price into Internet access charge. Customer has to take the responsibility to absorb the leased line cost. So leased line cost, which are the interconnection between customer and ISP are not shown in the total yearly service charge from any service providers.

5.11 Operational Cost Assumption

There are yearly operational costs to manage and maintenance the network availability. These costs include maintenance cost, supplied cost (electricity), leased lines monthly charge and salaries of administrators and/or engineer who take care network operation.

Maintenance cost in this model assume of 5% of yearly revenue. It also includes Maintenance Agreement (MA) of routers supplier and network management expenditure.

Supplies cost or electricity usage cost for each year. This part is calculated under the model network architecture only.

Leased lines cost, according to 3 circuits of T1 and 9 512kbps for interconnection between nodes of model network. This part is assumed as fixed rate for every year because there is not the upgrading of backbone circuits.

For employee salaries, there are 6 technicians and administrators and salaries will increase 10% every year according to profit of model network.

Tables 5-7 show the conclusion of operation cost of Model network in this paper.

Operation Cost	Year 1	Year 2	Year 3	Year 4	Year 5	bear 6	Year 7
Maintenance	449280	524160	599040	673920	748800	823680	898560
Supplies	718848	838656	958464	1078272	1198080	1317888	1437696
Leased Line	5356800	5356800	5356800	5356800	5356800	5356800	5356800
Technician Salary	720000	792000	871200	958320	1054152	1159567.2	1275523.92
Admin Cost	720000	741600	763848	786763.44	810366.34	834677.33	859717.65

Table 5-7: Cost assumption

5.12 Depreciation Calculation

This model selected straight-line depreciation method for calculation. The Straight-line depreciation is the simplest, and best known of the various depreciation methods. It is a method that a constant depreciation charge is made. To obtain the annual depreciation charge, the total amount to be depreciated, Basis - Salvage value, is divided by the useful life, in years. The useful life of all network equipments and software is assumed for 7 years. So in this model, depreciation is 541,500 Baht each year (3,637,500/7)

5.13 Statement of Cash Flows Output

The statement of cash flows is used to help answer the project generating enough cash to purchase the assets required for expansion.

Table 5-8: Statement of Cash Flow

			Y2		Y4	Y5	Y6	
Revenue		8 985 60C	10,483,200	1 980.8C 0	13,478,400	14,976,000 I	16,473,600	17 .71 200
Cost								
Maintenance		449,280	524,160	599,040	673,920	48 00	823,680	8 8,56C
ε ppl'e		718,848	83 656	958,464	1,078,27	1 198 080	1,317,888	, 37 696
Leased Line		5,356,800	5,356,800	5,356 8 0	5,356,800	5,356,800	5,356,800	4,190,400
Technician Salary		720,000	792 000	871,200	958,320	1 0 4 152	1 159,56 '	1,275,524
Admin Cost		720,000	741.600	763,848	786 763	10,36	834,677	859,718
Depreciation		519.843	51 84	519,643	519,643	519,643	519,643	519,643
Gross profit		501,029	41	2,911.805	4,104.682	5 288 1.9	6.461,345	8,789,660
Tax		150.309	2	873.542	1,231 405	1 58 48	1 938,403	2,636.898
Net profit after tax		350,720	1,197,239	2,038,264	2,873,277	3,701 711	4,522,941	6,152,762
'Depreciation		519,643	51 643	519,643	519,643	519,643	519,643	519,643
Add change in WC		(898560)	(149,760)	(149,76	(149,76	(149,760)	(149,760)	(149,760)
Investment	(3 637500)							
FCF	(3,637,500)	28.197	1 567 122	2 408 6	3 243 16	4 0 5	4 892 824	6 522 8

5.14 Summary - NPV, IRR, and Payback Period Output

Based on parameters and data from prior step. The model presents the net present value (NPV), Internal rate of return (IRR) and payback period for evaluating the project and for decision making in investment of MPLS. The results are presented as below table.

IRR	47%
NPV	6,113,319
Discount Rate	16%
Pay back	4

Table 5-9: IPP, NPV and Payback period

From the table 5-9, we found that pay back period is 4 years after investment for new model network, MPLS. it is depend on the revenue in each year after the investment . The average target is only six customers per year (6 corporate sites). Several strategies in marketing should launch after implementation in case of to increase the trusting of customers if they want to invest with our new service, MPLS.

CHAPTER 6: Recommendation and Conclusion

6.1 Introduction

Service providers need to offer MPLS services instead of IP services. They need to in cent their customers to fully implement the capabilities of MPLS, especially the setting of the prioritization. Service providers can still keep the "flat-rate" approach to pricing (versus usage-based pricing), but perhaps modify it to provide either credits or premium charges depending on the mix of traffic sent across the network. So, specifically, service providers could count the number of packets in each prioritization level and calculate a weighted average of the prioritization mix. Based on that, the monthly charges may vary anywhere from 80% to 120% of the standard monthly rate. All this requires systems changes and changes to service level agreements, none of which is trivial; however, this may become a survival issue.

Since most IP traffic today is highly delay-tolerant, enterprises will financially benefit from marking this traffic as low priority. Similarly, since traffic marked high priority will receive preferential treatment and therefore higher performance, there is strong incentive for enterprises to mark that traffic accordingly and for the service provider to receive a premium price for those packets, whether or not they arrive at the peak or off-peak.

6.2 The financial impact on the service provider

Worst case, revenues fall by 20% (although they could be smarter about setting the base price). Costs can easily fall by 30% or more (probably much more if customers really mark all their Web and email traffic as low priority).

Furthermore, overall traffic will likely increase. Prioritization will likely increase confidence in using IP for performance-sensitive applications (e.g., voice and video)— without significantly increasing costs. Since these high-priority flows are priced at a premium, the revenues increase faster than the capacity required to carry the traffic.

6.3 Mending the Peak Hour

There is another piece that is independent of MPLS and traditional traffic shifting, and that is shifting traffic off the peak-hour. The network needs to be built with both the peak-instant and peak-hour in mind. During the peak-hour, the peak-instants are relatively close together, so the threshold point needs to be relatively high so that you can de-buffer the lower-priority flows between the peak-instants. Shifting any traffic away from this peak-hour creates further benefits.

Even better is identifying new applications that use traffic during the non-peak-hour. These new applications would add virtually no cost to your backbone since they use capacity otherwise sitting idle; therefore, you have the opportunity to set bandwidth pricing for these applications quite low.

A number of examples of new services can be created around this concept. One example would be network-based PC backup. By setting the price of network capacity for this application reasonably low and automating the process, the economics for this type of solution can be incredible compared to either the cost of current solutions or the cost/risk of not doing backups at all.

Of course, pricing and management of these types of new offers are critical. The service provider must fully understand the incremental cost of delivering the capacity during the off-peak and then must carefully balance the opportunity for high margins with the desire to in cent deep market penetration through disruptive pricing.

New vendors are emerging to control and manage the pricing complexity of these types of off-peak applications. An example of technology that can enable this kind of new offer is Merkato from InvisibleHand Networks. This product creates a "market" where users connected to the network have a dynamic price set depending on the current supply/demand dynamic on the network. Pricing is set on a five-minute rolling window, which is well suited to the peak-hour situation. In effect, all network users "bid" on the market for capacity from the service provider.

6.4 Building for Better Profits

Broadband has certainly raised the stakes for everyone in the telecommunications industry. Predicting the take up of broadband has become an industry in itself—and one in which most people have been wrong. Major industry players like Global Crossing, Level 3, and most metro players are under intense pressure due to the debt incurred to build out massive fiber networks that have so far only been fractionally lit. Video file sharing can suddenly seize local networks; gigabit Ethernet streams can totally change the business case for metro and long haul capacity; new bandwidth hungry applications can outrun idle capacity in a matter of a few short years.

The trick is to continually reassess the large number of factors that can change on any day and to flow that through to the bottom line.

6.5 Conclusion

MPLS, was specifically designed to address the most significant issues facing service providers today—the need for a highly scalable foundation to deliver value-added IP business services. The innovative label-based forwarding mechanism of MPLS both simplifies IP traffic routing in complex networks and enables a plethora of very scalable value-added IP services. Service providers can now solve the three most challenging business IP service issues they face today:

- Provisioning connectionless IP VPNs that have same privacy as Frame Relay without tunneling or encryption
- Supporting multiple classes of service in an IP VPN to enable business policies on a per-customer basis
- Expanding market share with low-cost, managed IP services to capture new customers that need a lower cost, simpler alternative for intranet and extranet connectivity MPLS is the key to enabling IP+ATM solutions, allowing providers to build brand recognition and provide one-stop shopping to their subscribers while expanding their revenue and profits.

To do this quickly, service providers must retain existing subscribers by creating and provisioning new services while attracting new customers. MPLS gives service providers the means to offer an affordable, diversified IP+ATM services portfolio, with no migration risks. By reducing complexity and speeding up provisioning, MPLS enables low-cost operations. This opens the door for selling low-cost managed services to new markets that previously could not afford them. Through

MPLS, IP+ATM solutions give service providers the ability to continue offering their revenue-generating transport services today, while enabling their infrastructure for tomorrow's profit-generating value-added business IP services. It's the best of both worlds.

References

- Telechoice Perspective, Super-Broadband Deployment Initiatives, March 18, 2002, www.TeleChoice.com
- Internet Service Provider Traffic Aggregation, White paper, Cisco Systems www.cisco.com
- Enabling Business IP Services with Multiprotocol Label Switching" (White Paper), Rob Redford, Cisco Systems www.cisco.com
- A Comparison Between IPSec and Multiprotocol Label Switching Virtual Private Networks", White Paper, Cisco Systems www.cisco.com
- Positioning and Developing a Migration Strategy to Offer Advanced IP Service Based on MPLS, White Paper, Darryl Wortham,Jeremy Lawrence,Robert Redford, Cisco Systems www.cisco.com
- Internetwork Technology Handbook, Cisco Systems www.cisco.com
- A Comparison Between IPsec and Multiprotocol Label Switching Virtual Private Networks, White Paper, Cisco Systems www.cisco.com
- Multiprotocol Label switching-Enhance Routing in the New Public Network, White Paper, Chuck Semeria, Cisco Systems www.cisco.com
- Building IP Networks-Connection-Oriented Networking Solutions, www.marconi.com
- MPLS Migration Strategy for Multi-service Providers, www.marconi.com
- Technology guide Series- Multiprotocol Label Switching (MPLS), www.technology.com
- Financial Management Theory and Practice", Ninth Edition, Dryden 1999, Eugene F. Brigham Louis C. Gapenski, Michael C. Ehrhardt
- Managerial Accounting" Third Edition, 1997, Ronald W. Hilton
- MPLS Resource Center, www.mplsrc.com
- MPLS Forum, www.mplsforum.org
- AT&T, www.worldclasscampus.com
- IETF MPLS Frame work, RFC sources
- RFC3031 - MPLS Architecture, RFC sources
- RFC3032 - MPLS Label Encoding, RFC sources
- RFC3036 - MPLS LDP, RFC sources
- RFC2283 is MP-iBGP, RFC sources
- Draft-ramachandra-bgp-ext-communities-09, RFC sources
- RFC2547bis - BGP / MPLS VPNs, RFC sources
- draft-behringer-mpls-security-00, RFC sources
- Path to MPLS, White Paper, WaveSmith Networks, Inc. (wavesmith.com)
- NECTEC, www.nectec.or.th

Appendix

Service Providers Checklist:

Below are important equipment characteristics to look for when considering a paced migration to a next generation multiservice switching infrastructure:

- One platform that supports interoperable ATM, Frame relay, Private-leased line and MPLS interfaces under a common management umbrella.
- A protocol-independent switching architecture.
- Telephony-grade reliability.
- Modular software architecture.
- Fully redundant hardware components.
- Support for industry-standard APIs and managements protocols that integrate new network capabilities with existing back-end support and management systems.
- A compact form factor that allows for physical scalability as traffic volumes grow and reduces real estate requirements and costs.
- System throughput that can scale to hundreds of gigabits per second.
- Uplinks connections that can scale to OC-48/STM-16 (2.5G bps) and Gigabit Ethernet speeds.