INTERNATIONAL TELECOMMUNICATION UNION

ITU-T Y.1310

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(03/2004)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

Internet protocol aspects – Transport

Transport of IP over ATM in public networks

ITU-T Recommendation Y.1310

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	**Y.1300–Y.1399**
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Numbering, naming and addressing	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation Y.1310

Transport of IP over ATM in public networks

Summary

With the rapid growth of IP-based networks and applications in both private and public networks, it is necessary to consider arrangements to transport IP services over ATM in the public network environment.

For the private network environment, the ATM Forum has specified Multi-Protocol Over ATM (MPOA) [ATM_MPOA]. The Internet Engineering Task Force (IETF) has specified Classical IP Over ATM (C-IPOA) [CIP_ATM] and Next Hop Resolution Protocol (NHRP) [NHRP] and Multi-Protocol Label Switching (MPLS) [MPLS_ARCH]. To ensure that public networks will interwork with each other supporting a set of services defined in this Recommendation, and to ensure the interworking of public and private networks, it is necessary to recommend the preferred approach for transporting IP over ATM in public networks.

The approach adopted in this Recommendation is to identify generic requirements, key IP services and determine which IP over ATM approach is preferred for each service. It is preferable that the same approach is used for all services considered. This approach is recommended for all identified services using IP over ATM transport in public networks.

Source

ITU-T Recommendation Y.1310 was approved on 15 March 2004 by ITU-T Study Group 13 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation Y.1310

Transport of IP over ATM in public networks

1 Scope

This Recommendation addresses the transport of IP services over ATM. IP services in this Recommendation are defined as services provided at the IP layer. IP services in this Recommendation do not include those at the application layer (e.g., network banking).

This Recommendation identifies the IP over ATM approach for public networks adopting ATM technology, including service provider networks and carrier networks, but does not preclude the same approach where applicable in access networks, private networks and end systems. Approaches taken into account include classical IPOA, MPOA, and MPLS. These approaches are described briefly in Appendix I.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

2.1 Normative references

2.1.1 ITU-T

[I.321] ITU-T Recommendation I.321 (1991), *B-ISDN protocol reference model and its application*.

[I.326] ITU-T Recommendation I.326 (2003), *Functional architecture of transport networks based on ATM*.

[I.356] ITU-T Recommendation I.356 (2000), *B-ISDN ATM layer cell transfer performance*.

[I.361] ITU-T Recommendation I.361 (1999), *B-ISDN ATM layer specification*.

[I.364] ITU-T Recommendation I.364 (1999), *Support of the broadband connectionless data bearer service by the B-ISDN*.

[I.371] ITU-T Recommendation I.371 (2004), *Traffic control and congestion control in B-ISDN*.

[I.432] ITU-T Recommendations I.432.1 (1999), I.432.2 (1999), I.432.3 (1999) and I.432.4 (1999), *B-ISDN user-network interface – Physical layer specification*.

[Q.2931] ITU-T Recommendation Q.2931 (1995), *Digital Subscriber Signalling System No. 2 – User-Network Interface (UNI) layer 3 specification for basic call/connection control*.

[Q.2941] ITU-T Recommendation Q.2941.2 (1999), *Digital Subscriber Signalling System No. 2 – Generic identifier transport extensions*.

[Y.1311.1] ITU-T Recommendation Y.1311.1 (2001), *Network-based IP-VPN over MPLS architecture*.

2.1.2 ISOC/IETF

[ATM_MULTI] IETF RFC 2684 (1999), *Multiprotocol Encapsulation over ATM Adaptation Layer 5*.

[ATM_VCID] IETF RFC 3038 (2001), *VCID Notification over ATM Link for LDP*.

[CIP_ATM] IETF RFC 2225 (1998), *Classical IP and ARP over ATM*.

[CONTROL_SER] IETF RFC 2211 (1997), *Specification of the Controlled-Load Network Element Service*.

[CR_LDP] IETF RFC 3212 (2002), *Constraint-Based LSP Setup using LDP*.

[DIFF_AF] IETF RFC 2597 (1999), *Assured Forwarding PHB Group*.

[DIFF_ARCH] IETF RFC 2475 (1998), *An Architecture for Differentiated Services*.

[DIFF_HEADER] IETF RFC 2474 (1998), *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*.

[DIFF_EF] IETF RFC 2598 (1999), *An Expedited Forwarding PHB*.

[GUAR_SER] IETF RFC 2212 (1997), *Specification of Guaranteed Quality of Service*.

[IP_V4] IETF RFC 791 (1981), *Internet Protocol*.

[IP_V6] IETF RFC 2460 (1998), *Internet Protocol, Version 6 (IPv6) Specification*.

[LDP] IETF RFC 3036 (2001), *LDP Specification*.

[MPLS_ARCH] IETF RFC 3031 (2001), *Multiprotocol Label Switching Architecture*.

[MPLS_ATM] IETF RFC 3035 (2001), *MPLS using LDP and ATM VC Switching*.

[MPLS_DIFF] IETF RFC 3270 (2002), *Multi-Protocol Label Switching (MPLS) Support of Differentiated Services*.

[MPLS_ENCAPS] IETF RFC 3032 (2001), *MPLS Label Stack Encoding*.

[NHRP] IETF RFC 2332 (1998), *NBMA Next Hop Resolution Protocol (NHRP)*.

[RSVP_AGG] IETF RFC 3175 (2001), *Aggregation of RSVP for IPv4 and IPv6 Reservations*.

[RSVP_FUN] IETF RFC 2205 (1997), *Resource Reservation Protocol (RSVP) – Version 1 Functional Specification*.

[RSVP_REFR] IETF RFC 2961 (2001), *RSVP Refresh Overhead Reduction Extensions*.

[RSVP_TE] IETF RFC 3209 (2001), *RSVP-TE: Extensions to RSVP for LSP Tunnels*.

[TCP] IETF RFC 793 (1981), *Transmission Control Protocol*.

[UDP] IETF RFC 768 (1980), *User Datagram Protocol*.

2.1.3 ATM Forum

[ATM_MPOA] ATM Forum AF-MPOA-0087.000 (1997), *Multi-Protocol Over ATM Specification v1.0*.

2.2 Informative references

2.2.1 ISOC/IETF

[BGP_VPN] IETF RFC 2547 (1999), *BGP/MPLS VPNs*.

3 Terms and definitions

This clause lists alphabetically the acronyms of the key terms used in this Recommendation and the references to their sources of definitions. Refer to clause 4 for the acronyms and to clause 2 for the references:

CR-LDP [CR_LDP]

DS [DIFF_ARCH]

DSCP [DIFF_ARCH]

FEC [MPLS_ARCH]

LIB [MPLS_ARCH]

LSR [MPLS_ARCH]

MPLS [MPLS_ARCH]

PHB [DIFF_ARCH]

RSVP [RSVP_FUN]

RSVP-TE [RSVP_TE]

VPN [BGP_VPN]

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations:

AAL ATM Adaptation Layer
ABR Available Bit Rate
ABT ATM Block Transfer
AESA ATM End System Address
ARP Address Resolution Protocol
ATC ATM Transfer Capability
ATM Asynchronous Transfer Mode
ATMARP ATM Address Resolution Protocol
BGP Border Gateway Protocol
BUS Broadcast and Unknown Server
CE Customer Edge
CE Customer Equipment
C-IPOA Classical IP over ATM
CLP Cell Loss Priority
CLS Controlled Load Service
CoF Coordination Function
CR-LDP Constraint-based Routing LDP
DBR Deterministic Bit Rate
DS Differentiated Services
DSCP Differentiated Service Code Point
ER Explicit Routing
ES End System
FEC Forwarding Equivalence Class
FIB Forwarding Information Base

GFR	Guaranteed Frame Rate
GS	Guaranteed Service
ILMI	Integrated Local Management Interface
IP	Internet Protocol
IPOA	IP Over ATM
IPSF	IP Service Functions
IP-SSCS	IP-Service Specific Convergence Service
IS	Integrated Service
ISP	IP Service Provider
LANE	Local Area Network Emulation
LDP	Label Distribution Protocol
LEC	LANE Client
LECS	LANE Configuration Server
LER	Label Edge Router
LES	LANE Server
LIB	Label Information Base
LIS	Logical Internet Subnet
LLC	Logical Link Control
LSP	Label Switched Path
LSR	Label Switching Router
MAC	Medium Access Control
MBS	Maximum Burst Size
MCR	Minimum Cell Rate
MPC	MPOA Client
MPLS	Multi-Protocol Label Switch
MPOA	Multi-Protocol Over ATM
MPS	MPOA Server
NAT	Network Address Translation
NHC	NHRP Client
NHRP	Next Hop Resolution Protocol
NHS	NHRP Server
NNI	Network to Network Interface
OSPF	Open Shortest Path First
PCI	Protocol Control Information
PCR	Peak Cell Rate
PDR	Peak Data Rate
PE	Provider Edge
PHB	Per Hop Behaviour
PIM	Protocol Independent Multicasting
PPP	Point-to-Point Protocol
PSC	Per Hop Scheduling
QoS	Quality of Service
RSVP	Resource Reservation Protocol
RSVP-TE	Resource Reservation Protocol-Traffic Engineering
SBR	Statistical Bit Rate
SLA	Service Level Agreement

SNAP Subnet Attachment Point
SSCS Service Specific Control Service
TCP Transmission Control Protocol
TMN Telecommunications Management Network
UDP User Data Protocol
UNI User Network Interface
VPN Virtual Private Network
VPN-ID VPN Identifier
xDSL x-Digital Subscriber Loop

5 Generic requirements

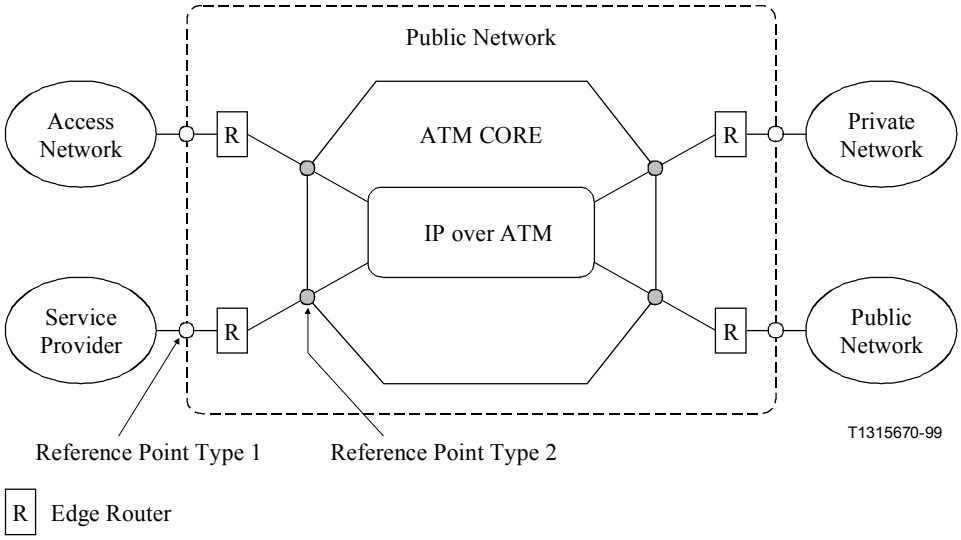
This Recommendation places a number of generic requirements on IP over ATM approaches. Such requirements are applicable to all identified IP services. The mandatory generic requirements are as follows:

- The recommended approach must be independent of the IP version supported.
- The recommended approach must have sufficient scalability to support large networks. Items to be taken into account regarding scalability include:
  - use of VCI and VPI values;
  - complexity of routing calculation at Layer 2 and Layer 3;
  - complexity of address resolution mechanism;
  - control messaging load (e.g., frequency of setup and cleardown of ATM connections, frequency of IP-related signalling messages);
  - complexity of the packet classification mechanism needed to support QoS. The less granularity in the QoS (e.g., from per IP flow, to per IP flow aggregation, to per service, as in Diffserv) the simpler the packet classification mechanism.
- The recommended approach must include the capability for efficient and scalable solutions to support IP multicasting over ATM networks.
- The recommended approach must have sufficient robustness to support large networks. Items to be taken into account include:
  - capability to support restoration systems.

6 Framework architecture

The framework architecture to support IP layer services over ATM is defined as comprising the network architecture and the protocol architecture to support the IP services required.

6.1 Network architecture



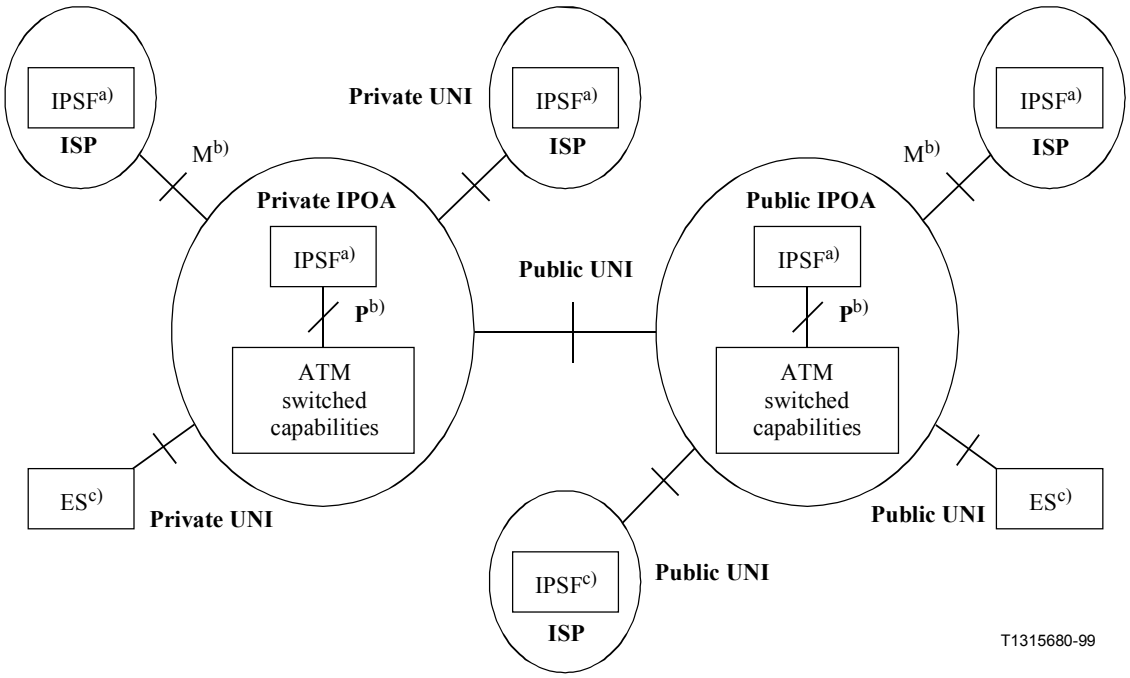
NOTE 1 – Internet service providers may also provide the ATM core.

NOTE 2 – Reference point types 1 and 2 can be a standardized reference point such as an ISDN S, T interface or a non-standardized interface.

Figure 6-1/Y.1310 – Reference network architecture for IP over ATM

The reference network architecture for IP over ATM is shown in Figure 6-1. This configuration illustrates possible scenarios to support IP services over ATM identified in this Recommendation. The dotted rectangle indicates a public network on which we focus. Note that the public network in this Recommendation is restricted to one which has an ATM core. The boxes inside the dotted rectangle describe the generic arrangement inside the public network. It includes an ATM core, IP over ATM capabilities, and edge routers. A number of different kinds of networks are described outside the dotted rectangle, with each identifying a scenario in which the public network provides a certain identified IP service to a certain type of a network. From the viewpoint of the public network, these networks are regarded as subscriber networks.

In this figure, two types of distinct reference points are described. The reference point type 1 is the boundary between the public network and subscriber networks, and the reference point type 2 is the interface to the IP over ATM network inside the public network. The arrangement of the reference point type 1 may depend on the facilities of subscriber networks, and the definition of IP services provided. Interworking functions and/or adaptation functions may be required in edge routers. This Recommendation primarily focuses on the reference point type 2, and on the approach adopted inside the public network.



a) IPSF: IP Service Functions.

b) P or M: on the basis of ITU-T Recommendation I.364.

c) ES: End System has a full IPOA protocol stack.

Figure 6-2/Y.1310 – Reference configuration for IP services over ATM

Figure 6-2 illustrates the reference configuration for IP services over public and private ATM networks. In the private and public IPOA networks, the provision of IP services is realized by means of ATM switched capabilities and IP Service Functions (IPSF). In this case, interfaces between ATM switched capability and IPSF shall be defined at the P or M reference points [I.364]. The IP Service Functions (IPSF) are those functions necessary to enable IP over ATM. A typical example of IPSF is address resolution service. As an end system, the IPSF is essentially a router with an ATM interface.

The IPSF function can be implemented in the same equipment together with the ATM switched capabilities. In this case, there is no need to define the interface at the P reference point. The IPSF function and ATM switched capabilities can be also implemented in separate equipment. In this case, interfaces shall be defined at the M or P reference points depending whether the IPSF is located outside or inside the core ATM network.

ISPs and End Systems (ES) outside the ATM networks may be connected to the private or public ATM networks. Each ES has a full IPOA protocol stack and is connected by way of private UNI for the private IPOA, or public UNI for the public IPOA.

6.1.1 Network and service interworking

In network interworking scenario the IP Protocol Control Information (PCI) and payload data are transferred transparently across the ATM network to another IP-based network by means of an interworking function (IWF) between the two networks. Typically the IWF simply encapsulates the IP packet by means of an adaptation function and transfers it transparently to the remote IWF. For current IP and ATM interworking, network interworking is the typical case, with ATM providing a backbone or core network for transport of Internet protocol. In this scenario, the ATM network may be viewed as an underlay transport for the Layer 3 (and above) protocols.

For the case of service interworking, the IWF terminates the IP protocol and translates the PCI to the ATM network PCI for transfer, control and management functions. Since in general not all functions may be supported in one or other of the networks, the service interworking scenario may only be able to provide a "best fit" conversion between the two different technologies. However, this should not result in any loss of user data since this is not affected by the PCI conversion at the service interworking IWF.

Figures 6-1 and 6-2 illustrate network interworking associated with IP over ATM.

6.2 Protocol architecture

6.2.1 General description of IPOA protocol reference model

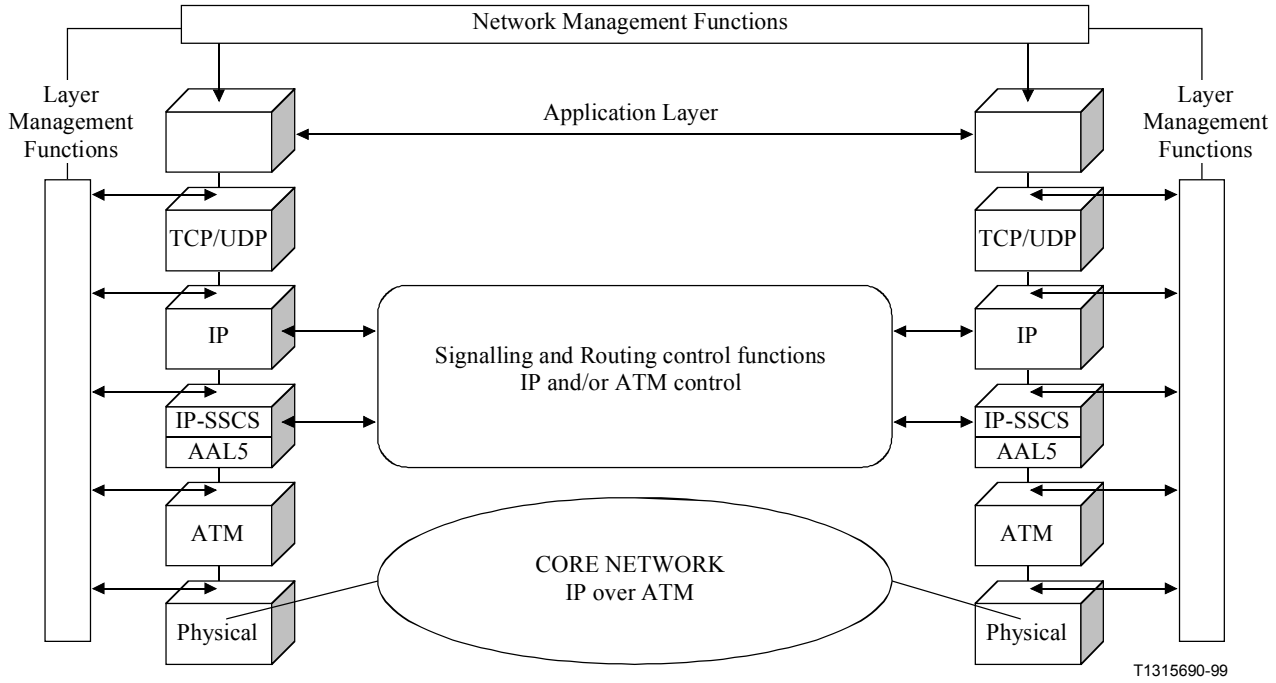


Figure 6-3/Y.1310 – Protocol reference model for IP over ATM

Figure 6-3 generalizes the protocol reference model for IP transport over ATM in public networks. It may be noted that the underlying protocol reference model concepts of layer management, network management and signalling and routing control are extended to include the Layer 3 and above functional blocks. It should be noted that the blocks shown in Figure 6-3 correspond to logical representations of the functions and therefore do not signify or constrain any particular network implementation.

The interfaces between the functional blocks may either be internal, non-standardized communication between the sub-layers, or external, standardized protocols. Each layer in the general model has its associated layer management functional block. The layer management blocks are responsible for processing of management and Protocol Control Information (PCI) for that layer only. Communication of information between the layers may only occur through the network management function. This is performed via the Coordination Function (CoF) of the network management.

All functional blocks do not need to be present in all network applications of IPOA. Thus, the blocks may be viewed as basic "building blocks" to enable any given network application of IPOA. However, the basic relationships and ordering between the different blocks must be maintained to ensure consistent interoperability.

6.2.2 Functional description of IPOA protocol reference model

This clause describes only the IPOA-related functional blocks, since detailed descriptions of the physical, ATM and ATM control layer functions are dealt with in other Recommendations [I.321], [I.326], [I.361], [I.432] and [Q.2931]. The application layer block is outside the scope of this Recommendation.

6.2.2.1 IP-SSCS/AAL5 functions

The IP-SSCS/AAL5 incorporates transfer functions required to map the IP packet onto the AAL5. The IP-SSCS/AAL5 functional block provides the encapsulation and multi-protocol multiplexing functions defined by the IEEE 802.2-based Link Layer Control/Subnetwork Attachment Point (LLC/SNAP) protocol as adopted by the IETF in [ATM_MULTI].

6.2.2.2 IP layer functions

IP layer functions provide IP forwarding (IP datagram delivery) from a source to a destination over an interconnected system. IP forwarding is the process of receiving a packet and using a very low overhead decision process determining how to handle the packet. The packet may be delivered locally or forwarded externally. For traffic that is forwarded externally, the IP forwarding process also determines which interface the packet should be sent out on, and if necessary, either removes one media layer encapsulation and replaces it with another, or modifies certain fields in the media layer encapsulation.

IPOA protocol architecture must be independent of IP version. Currently, there are two versions: IPv4 (IP version 4) and IPv6 (IP version 6). IP layer functions are equal to those defined by IETF in [IP_V4] and [IP_V6] according to IPv4 and IPv6, respectively.

The IP layer function does not provide a reliable communication facility. There are no acknowledgements either end-to-end or hop-by-hop.

Note that IP layer functions should not be changed to use IP-SSCS/AAL5 functions over ATM.

6.2.2.3 IP layer management functions

IP layer management function has two basic functions: addressing and fragmentation. IP layer functions use the addresses carried in the IP header to transmit IP datagrams toward their destinations. The selection of a path for transmission is resolved as using signalling and routing functions block. IP layer functions use fields in the IP header to fragment and reassemble IP datagrams when necessary for transmission.

IPv4 protocol uses four basic key mechanisms in providing its service: type of service, time to live, options, and header checksum. IPv6 is a new version of the Internet protocol, designed as the successor to IPv4. The changes from IPv4 to IPv6 fall primarily into the following categories: expanded addressing capabilities, header format simplification, improved support for extensions and options, flow labelling capability and authentication and privacy capabilities. The IP layer management function does not provide error control for data, only a header checksum. There are no retransmissions. There is no flow control.

6.2.2.4 Transport layer functions

Transport layer includes connection-oriented type TCP functions and connectionless type UDP functions respectively. These depend on application programme type.

TCP functions provide reliable connection service between pairs of processes. TCP functions are equal to those defined by IETF in [TCP]. TCP functions include the following facilities: basic data transfer, reliability, flow control, multiplexing, connections and precedence and security.

UDP functions provide datagram transfer. UDP functions are equal to those defined by IETF in [UDP]. The UDP is transaction oriented, and delivery and duplicate protection are not guaranteed.

Note that transport layer functions should not be changed to use IP layer functions over ATM.

6.2.2.5 Application layer functions

The application layer and its associated layer management functional blocks include user or network-specific applications such as HTTP, FTP, TELNET, etc. The description of application layer functions is outside the scope of this Recommendation.

Note that in the TCP/IP protocol architecture, the application layer function is generally taken to include the session and presentation layer functions.

6.2.2.6 Network management functions

The network management functions depend on the specific network application for IPOA. In general, they include the TMN (Telecommunications Management Network) functions associated with: fault management, performance management, configuration management, security management, etc.

6.2.2.7 Signalling and routing control functions

This includes the signalling and routing functional blocks in IP and/or ATM control. IP control and signalling encompasses various aspects of IP control including routing. ATM control includes ATM signalling and routing.

7 IP services

A range of IP services is included in this Recommendation as a means to determine the preferred IP over ATM approach in public networks. Initially, mapping of IP QoS with ATM and VPN services are addressed. Additional services are for further study.

7.1 Mapping of IP QoS with ATM

7.1.1 Introduction

Two major approaches for the support of QoS differentiation at the IP level are documented in the IETF: the Intserv paradigm, aimed at supporting per IP flow QoS differentiation, and the Diffserv paradigm, aimed at supporting "coarse" QoS differentiation for aggregations of IP flows.

7.1.1.1 The IP Intserv paradigm

The Intserv paradigm relies on explicit per IP flow QoS requests carried by the RSVP protocol and on flow admittance control at the RSVP capable routers along the path of the flow. In the Intserv paradigm, two services are defined: the Guaranteed Service – GS [GUAR_SER] and the Controlled Load Service – CLS [CONTROL_SER]. In GS the maximum queuing delay for the flow is controlled. To compute the maximum delay a datagram will experience, the latency of the path must be determined and added to the maximum queuing delay [GUAR_SER]. CLS does not provide firm delay guarantees, but the service provided to the flow should be comparable to what the flow would experience in a lightly loaded network, even when it is not the case [CONTROL_SER]. In practice, CLS requires long-term available bandwidth.

Both services require that the characteristics of the flow be specified by means of a token bucket specification [RSVP_FUN] and that excess traffic be treated as best effort.

7.1.1.2 The IP Diffserv paradigm

The IETF Diffserv model is based on the concept of Per Hop Behaviours (PHB) [DIFF_HEADER] and [DIFF_ARCH]. The Diffserv PHBs are defined by a set of forwarding behaviours that each local router along the path adheres to. IETF has identified two main PHBs so far:

- Expedited Forwarding (EF) PHB [DIFF_EF]:

  The EF-PHB is characterized by a configurable amount of bandwidth that is not impacted by the other traffic sharing the link. The EF-PHB can be used to build an end-to-end service that requires low loss, low delay and low delay variation through Diffserv domains.

- Assured Forwarding (AF) PHB group [DIFF_AF]:

  The AF-PHB group is characterized by four AF classes, and each AF class is allocated a certain amount of forwarding resources such as buffer and bandwidth in a Diffserv node. Within each AF class, IP packets are marked with one of three possible drop precedence values. In case of congestion, the drop precedence of a packet determines the relative importance of the packet within the AF class. However, there is no standardized relationship between the relative performance of the four AF classes. The AF-PHB group can be used to ensure that the subscribed information rate of a service is guaranteed with high probability.

7.1.2 Network model to support QoS-aware IP services

This clause describes a network model to support QoS-aware IP services in IPOA networks. In the IETF framework, the end-to-end QoS is provided by coupling Intserv regions at the edge of the network with Diffserv regions in the core of the network. The network model proposed here, however, considers additional possibilities. Moreover, the link layer in this case is always assumed to be ATM.

7.1.2.1 Model description

The possible network models to support QoS-aware IP services are illustrated in Figures 7-1, 7-2 and 7-3. In each case, the shaded area indicates the active function used.

Case 1 – Intserv over ATM networks

In this model, communication between two Intserv stub networks is supported over IPOA core networks. The IPOA devices in the core networks may provide both Intserv and Diffserv capabilities. However, only Intserv functionality of the IPOA devices will be activated to support end-to-end integrated services. Both service level agreements (SLA 1 and SLA 2) require that the requirements for Intserv service are satisfied.

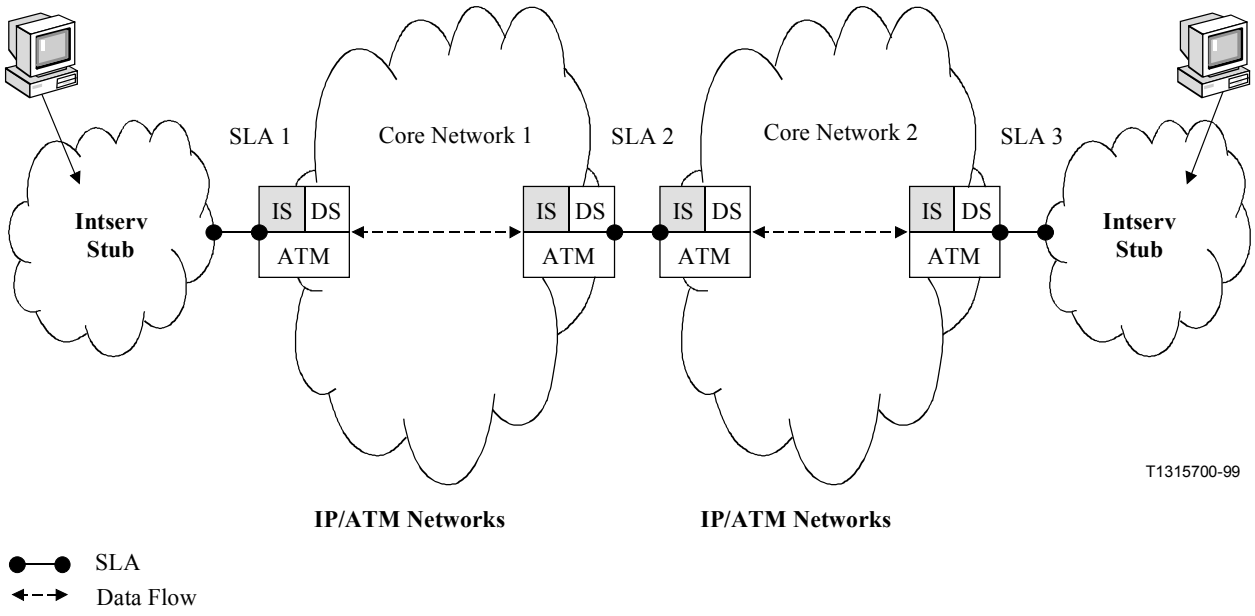


Figure 7-1/Y.1310 – Network model to support Intserv over ATM

Case 2 – Diffserv over ATM networks

In this model, communication between two Diffserv stub networks is supported over IPOA core networks. The IPOA devices in the core networks may provide both Intserv and Diffserv capabilities. However, only Diffserv functionality of the IPOA devices will be activated to support end-to-end differentiated services. Both service level agreements (SLA 1 and SLA 2) require that the requirements for Diffserv service are satisfied.

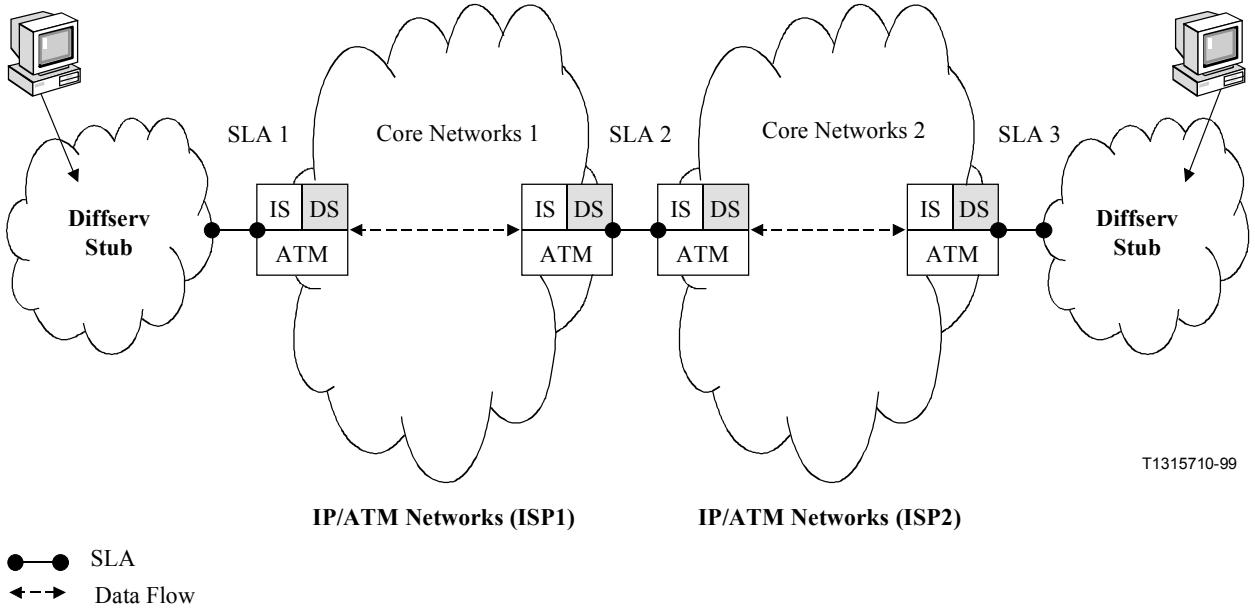


Figure 7-2/Y.1310 – Network model to support Diffserv over ATM

Case 3 – Intserv through Diffserv domains over ATM networks

In this model, communication between two Intserv stub networks is supported over IPOA core networks. In the IPOA core networks, some domains may provide only Diffserv, and the others may provide both Intserv and Diffserv capabilities. In this case, the Intserv may be transparently transported over Diffserv only domains. In this case there are two types of service level agreement.

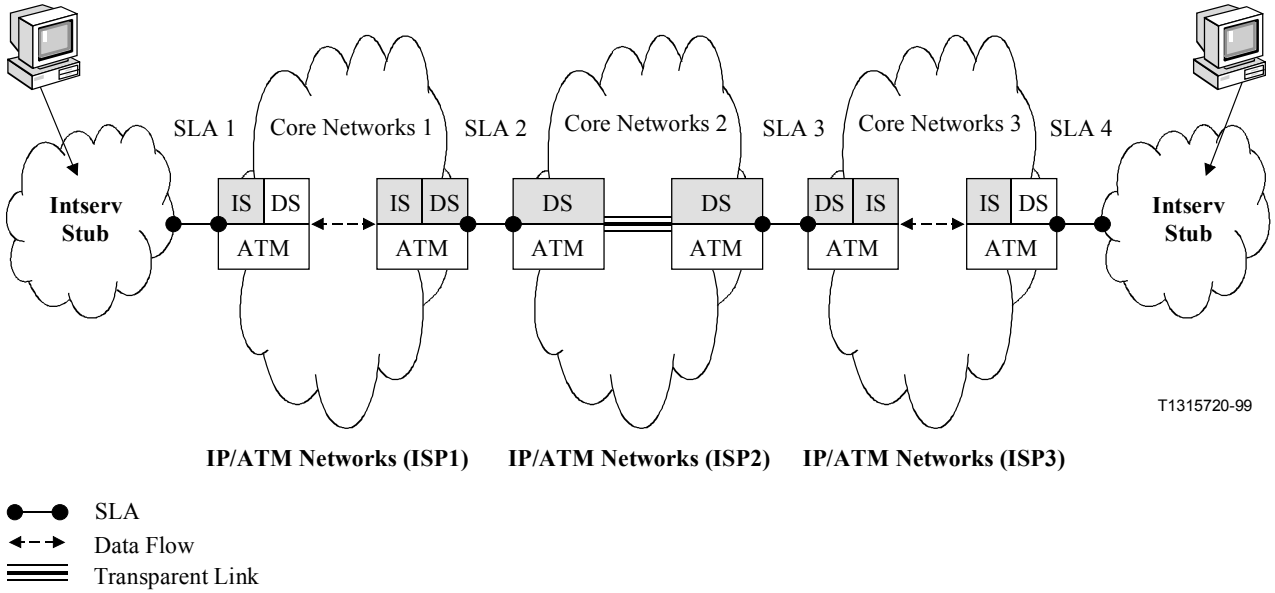


Figure 7-3/Y.1310 – Network model to support Intserv through Diffserv domains over ATM

7.1.3 List of service mapping functions

The service mapping functions do not depend on the architecture of the surrounding network but only on the way IP and ATM QoS is supported on both sides of the interface where the mapping is needed. Figure 7-4 thus displays the necessary set of possible IP services to ATM services mappings of the considered framework architecture (refer to clause 6).

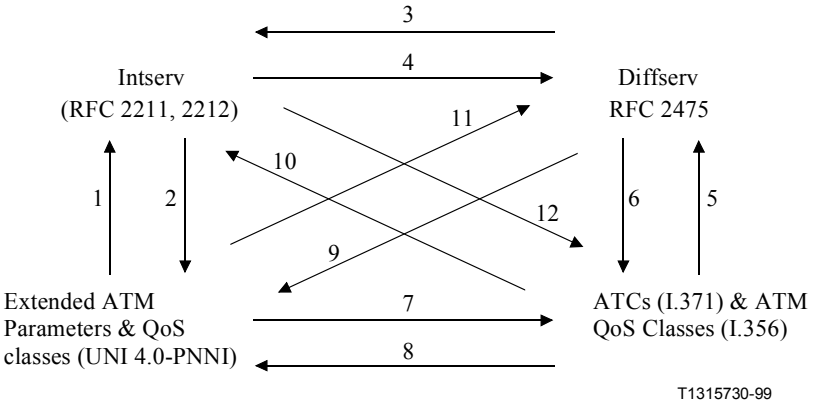


Figure 7-4/Y.1310 – List of service mapping functions

Among all these mappings, only mappings 6 and 12 are addressed in this Recommendation. Note that in this case at the egress of this ATM portion, there is no need of a mapping function of type 5 or 10, as in the destination IP network the support of QoS is totally based on the IP level information which is transparently carried by the ATM portion. Mappings 5 and 10 might be

required in the case of native ATM traffic having to cross or reach a pure IP network, and are for further study.

Mappings 3 and 4 pertain to the IP domain only and are part of activity within the IETF, whereas all the mapping originating/terminating into the extended ATM parameters and QoS classes (supported in private ATM networks) is part of ATM Forum's work.

7.1.4 Mapping of IP integrated services to ATM services

The issue of Intserv to ATM mapping arises whenever an IP flow requesting Guaranteed Services (GS) [GUAR_SER] or Controlled Load Services (CLS) [CONTROL_SER] must be supported by an ATM connection linking two Intserv capable routers, and it is independent from the specific approach to support IP over ATM.

Two different kinds of mapping are foreseen: one to one mapping and many to one mapping.

7.1.4.1 One to one mapping

One to one mapping occurs when a single ATM connection is entirely devoted to the support of a single IP flow. Specifically, the mapping process consists of the choice of an ATM service (i.e., ATC and associated QoS class) that can satisfy the QoS commitments of the IP service (GS or CLS), and in that view, several mappings are possible. More generally, however, the mapping process can be additionally thought as a way to communicate to the ATM level additional information regarding the characteristics of the carried flow, so that the downstream ATM network can make use of this information to efficiently carry the connection (e.g., multiplex it with others). In that view, one can rank all the possible mappings, and some mappings are better than others.

7.1.4.2 Many to one mapping

Many to one mapping occurs when a single ATM connection can carry more than one IP flow. In that case the mapping process consists in the choice of an ATM service that can satisfy the QoS commitments of a set of IP flows. As IP flows normally start and end asynchronously, this mapping can be viewed as an aggregation process which on the basis of the flow's IP level characteristics (e.g., token bucket and requested QoS), decides on the possibility of carrying the flow together with others on an already existing ATM connection (while still meeting the QoS constraints of the IP flow), or on the need of re-negotiating the connection's parameters.

Rules on how to take such a decision are outside the scope of this Recommendation.

7.1.4.3 Mapping of Guaranteed Service (GS) to ATM

ATM does not require any extensions to perform these mappings. However, the mapping scheme chosen should satisfy the following requirements:

- The chosen ATC must be one able to support delay requirements.
- The chosen ATC must be one able to reserve some bandwidth for the flow.

Appendix II provides suggestions on guidelines for implementation of the mappings.

7.1.4.4 Mapping of Controlled Load Service (CLS) to ATM

ATM does not require any extensions to perform these mappings. Appendix II provides suggestions on guidelines for implementation of the mappings. However, the mapping scheme chosen should satisfy the following requirement:

- The chosen ATC must be one able to reserve some bandwidth for the flow.

7.1.4.5 Impact on ATM traffic management

For further study.

7.1.4.6 Impact on ATM signalling

For further study.

7.1.4.7 Impact on ATM routing

For further study.

7.1.5 Mapping IP differentiated services to ATM services

The IP Differentiated Services (Diffserv) model uses the concept of Per Hop Behaviour (PHB) [DIFF_HEADER] and [DIFF_ARCH] to enable QoS-based IP services.

The PHBs may be used as an important factor to define an IP service in the Diffserv domain. However, PHB itself is not related to end-to-end IP QoS services. So, the mapping between Diffserv and ATM should be based on IP services and ATM services. Specifically, IP services can be defined by a combination of PHB implementations with traffic characteristics at the edges of Diffserv domains, and ATM services can be defined by a combination of ATM transfer capabilities [I.371] with QoS classes [I.356].

7.1.5.1 Service mapping

To provide services to the customers, Diffserv providers must combine PHB implementations with traffic conditioners and service provisioning strategies. The PHB concept is not addressed in ATM. Thus, the PHBs to ATM Transfer Capability mapping does not seem to be suitable. So, a service mapping may be considered from a particular differentiated service to an ATM service instead. The service mapping from Diffserv to ATM is clearly provided by the negotiation between two network providers, based on the definition of the IP services considered.

Therefore, the service mapping depends on the policy of service providers, and may vary among different service providers. Some examples of possible service mapping are given in Appendix II.

The following requirement applies to service mapping:

- No per connection minimum cell rate need be associated with the support of some Diffserv PHBs over ATM.

There is also a need to consider qualitative or relative service. The solutions are for further study.

7.1.5.2 Impact on ATM traffic management

For further study.

7.1.5.3 Impact on ATM signalling

For further study.

7.1.5.4 Impact on ATM routing

For further study.

7.2 IP Virtual Private Networks (IP-VPNs)

7.2.1 Scope of IP-VPN

The IP-VPN in this Recommendation is defined as the emulation of IP-based private wide area network facilities provided over a carrier-scale ATM transport network. Figure 7-5 shows the arrangements of the IP-VPN in this Recommendation. An example method to demonstrate IP-VPN support in an MPLS/ATM public network is provided in Appendix IV.

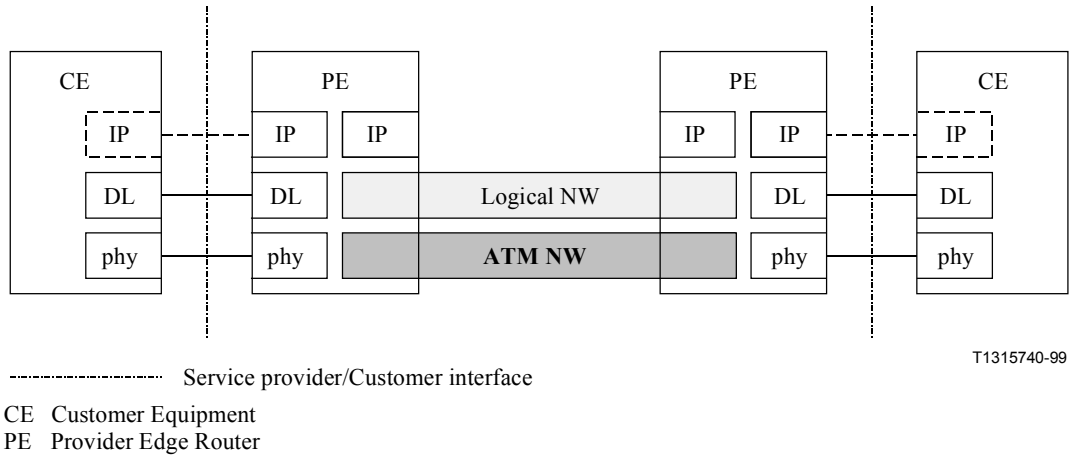


Figure 7-5/Y.1310 – Network model for IP-VPN

A customer site is connected to the service provider's network through a customer equipment device (CE). This can be a single host, a switch, or an IP router. On the other hand, the services provider's network connects the customer site with provider edge router (PE). When the CE is a router, we can configure in such way that it is a routing peer of the attached PE, but is not a routing peer of the CE at the other site. Routers at different sites do not directly exchange routing information with each other. This arrangement allows supporting very large VPNs easily, while the routing strategy for each individual site is greatly simplified. This capability is considered to be important for carrier-scale service providers who provide outsourced IP-VPN service.

7.2.2 Definition of IP-VPN service

The IP-VPN service enables customer sites to form groups: to and from which IP access is restricted. This group is called IP-VPN. A specific site may be a member of one or more IP-VPNs. Member sites of a specific IP-VPN can communicate among themselves using IP protocol. Specific sites can have additional capabilities that allow them to access to the sites outside the group, and/or to be accessed from the sites outside the group.

7.2.3 Requirements for IP-VPN service

7.2.3.1 Requirements for user-plane

7.2.3.1.1 Support for opaque packet transport

Opaque packet transport allows IP-VPN customers to use independent IP address inside their network. Service provider's network requires having the capability to route IP packets depending on VPN membership, even though they use overlapping address spaces. The functions to identify a VPN, (e.g., the use of VPN-ID) and/or the function to differentiate per-VPN packet forwarding may be required.

7.2.3.1.2 Support for data security

Data security provides IP-VPN customers with a certain level of secured communication among the member sites of an IP-VPN. Service provider's network requires ensuring that the snooping of data, misdirection or misinsertion of unrelated packets is avoided. Filtering functions, encryption functions and functions for authorization may be required.

7.2.3.1.3 Support for QoS

QoS allows IP-VPN customers to subscribe a certain level of assurances in the quality of communications among the member sites of an IP-VPN. Service provider's network requires having

the capabilities to support arbitrary categories of QoS services as it does to provide general IP services. Clause 7.1 describes the capabilities in detail.

7.2.3.2 Requirements for control-plane

7.2.3.2.1 Support for signalling logical network resources

Service provider should have the capability to signal its network resources to support the transport of IP packets for IP-VPN customers.

7.2.3.2.2 Support for transporting VPN-IDs

For further study.

7.2.3.2.3 Support for per-VPN routing

For further study.

7.2.3.3 Requirements for management-plane

7.2.3.3.1 Interoperable VPN-ID

If an IP-VPN spans over several service providers, each provider's network may require to distinguish the traffic of the IP-VPN correctly. In this case, the generally accepted VPN-ID definition is necessary to reduce the processing of edge routers.

7.2.3.3.2 Support for IP-VPN membership management

Service provider should have the capability to manage the VPN membership information, such as which customer site belongs to which IP-VPN. This capability should have sufficient interoperability to be used in different service providers in order to support IP-VPN spans over several service providers.

7.2.3.3.3 Support for configuration of logical network resource

Services provider's network should have the capability to configure its network resources to support the transport of IP packets for IP-VPN customers. This capability should have sufficient interoperability to be used in different service providers in order to support IP-VPN spanning several networks.

8 Preferred network solution

8.1 Recommended approach

Considering the generic requirements described in clause 5 as well as the services described in clause 7, it is recommended that MPLS [MPLS_ARCH] be adopted as the single preferred approach for public networks. MPLS supports all the services identified. It is recognized that MPLS does not provide significant benefits over properly engineered classical IP over ATM (as described in I.1.2) for the support of the Intserv service. However, MPLS does not offer less than classical IPOA for the support of Intserv while also providing support for all other services.

Additional motivations for the selection of MPLS as the single preferred approach include:

8.1.1 Small networks versus large networks

It is very well known that MPOA is very well suited for small networks but has limitations when applied to large networks. This Recommendation is dedicated to service providers and is therefore targeting large networks. MPLS has been designed to accommodate the requirements of large networks in terms of flexibility, scalability and manageability.

8.1.2 ATM versus non-ATM bearer

While the focus of this Recommendation is the transport of IP over ATM, it is useful to understand that large networks can use several distinct bearer technologies including ATM. In a wider scope, it is useful to choose a technology that is optimal for IP transport over ATM but at the same time optimal for IP transport over other link layer technologies. MPLS is probably the only possible strategy, which covers this wider scope.

8.1.3 Static versus dynamic control

From a routing perspective, the MPLS architecture gives the opportunity to have, and at the same time, the possibility to choose between provisioned routing and dynamic routing. It is the network operator's choice on what approach to select.

8.1.4 ATM versus non-ATM control in IPOA

It is preferable to have a generic control which is link layer independent. Also, ATM control can still be used on the same switches.

8.1.5 Traffic engineering of IP services

ATM has the most complete set of functionality for traffic engineering known to date. However, overlay IP over ATM models may not efficiently use all ATM capabilities and may tend to be limited in scalability due to the well-known "n-squared" problem when a full mesh of PVCs is provisioned. MPLS borrows some of ATM technology's capabilities in terms of QoS, routing, resource management and other aspects, adding the notion of explicit routing to help map traffic demand onto network topologies. Thus, use of MPLS offers new and more traffic management features than before.

8.1.6 Build on existing investments

Given the existing investment in ATM and other technologies, there is a clear need to carry IP traffic over ATM and other link layer protocols and therefore, a unifying switching technology is necessary. In today's carrier networks, ATM hardware is used in a provisioned mode to carry IP traffic, MPLS is seen as a logical evolution of C-IPOA in the near future, since explicit routing can build on the basis of existing provisioned PVCs and the architecture is flexible enough to accommodate potential network evolution.

8.1.7 VPN services support

The main advantage of MPLS is the ability to provide connection-oriented services over connectionless or explicit routing, which makes it ideal for dynamic tunnelling. There is no unique way to provide MPLS-based VPNs, which makes comparison more difficult with other IPOA technologies.

8.1.8 QoS aspects

There is a clear synergy between IP differentiated services and MPLS, since both evolved with service provider's requirements inherent in their design. The label, with its extended semantics can carry Diffserv related information, and end-to-end LSPs can guarantee consistency of QoS mechanisms within a specific MPLS domain through appropriate resource reservation mechanisms.

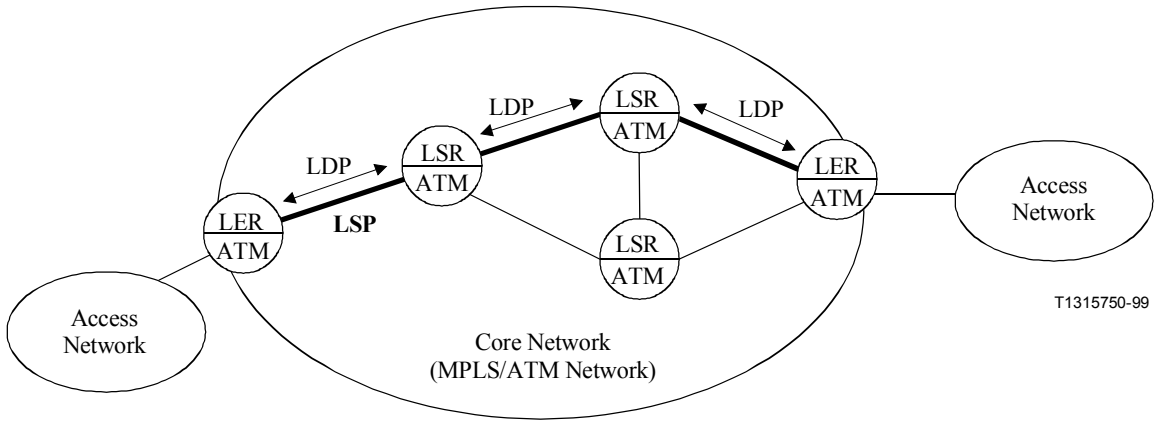
8.2 Framework for MPLS over ATM in public networks

8.2.1 Architectural model

Figure 8-1 illustrates the general network model of MPLS/ATM core network. The public network is implemented as MPLS with ATM networks that are composed of Label Edge Routers (LERs) and Label Switching Routers (LSRs). The LER is located at the edge of the MPLS network as a MPLS aware ingress/egress router. The edge of the MPLS network may or may not coincide with

the edge of the core ATM network. The LER performs full functions of Layer 3 and label binding based on LIB (Label Information Base) generated by running LDP. The LER is connected to interior LSRs. The LSR performs label swapping based on LIB. LSP (Label Switched Path) between LERs or LER and LSR is set up using LDP [LDP] and [MPLS_ENCAPS].

Based on this simple model, various IP services such as IP QoS (Intserv and Diffserv) and IP-VPN can be efficiently and flexibly provided for IP customers through different access networks (such as pure ATM, frame relay, xDSL, pure IP, etc. including non-MPLS domain).



LDP Label Distribution Protocol
LER Label Edge Router
LSP Label Switching Path
LSR Label Switch Router

Figure 8-1/Y.1310 – Network model of MPLS/ATM core network

8.2.2 Control protocol for MPLS over ATM

- *Label advertisement mode*

  In MPLS over ATM networks, ATM VCI and VPI are used as the label. The label can be advertised in the networks in the following two ways:

  – explicit label distribution protocol such as LDP;

  – piggybacking on other control messages such as RSVP, BGP, etc.

  Both explicit label distribution and piggybacking can be used in the network. In this Recommendation, however, the LDP is recommended for hop-by-hop label distribution.

- *Label allocation mode*

  If LDP is used, the label can be allocated among LSRs in the network in the following two ways:

  – unsolicited downstream mode;

  – downstream on demand mode.

  In this Recommendation, the downstream on demand mode is recommended as the label advertisement mode in MPLS over ATM networks for the following reasons:

  – In the downstream on demand mode, VPI/VCI values are only consumed when requested.

  – The downstream on demand mode is more similar to the conventional signalling such as ATM signalling, and thus can interoperate with the existing public networks.

- *Label Switching Path (LSP) control mode*

  The labels for an LSP are controlled in the following two ways:

  – Ordered LSP control, where an LSR only binds a label to a particular FEC if it is the egress LSR for that FEC, or if it has already received a label binding for that FEC from its next hop for that FEC.

  – Independent LSP control mode, where each LSR, upon noting that it recognizes a particular FEC, makes an independent decision to bind a label to that FEC and to distribute that binding to its label distribution peers.

  In this Recommendation, the ordered control mode is recommended for the following reasons:

  – In the independent control mode, since each LSR independently assigns labels to IP forwarding equivalence classes, it is possible that different LSRs may make inconsistent decisions. For the ordered control mode, this is not the case.

  – Compared to the independent control mode, in the ordered control mode, resources such as VCI/VPI can be utilized more efficiently.

To satisfy the service provider's traffic engineering requirements, two signalling approaches are possible:

1) MPLS/LDP with CR-LDP [CR_LDP].

2) MPLS/LDP with RSVP-TE extensions [RSVP_TE, RSVP_REFR, RSVP_AGG].

Service providers may choose which approach to use based on their own specific requirements, service needs, and deployment experience.

Criteria for this choice might include functional capabilities, interoperability issues, operational and management complexity level.

Note that setup protocols for complete support of provisioning end-to-end quality of service are not yet available.

Appendix I

Approaches for IP over ATM

I.1 Classical IP over ATM

Classical IP and ARP Over ATM (C-IPOA) is defined in [CIP_ATM]. Figure I.1 provides a functional description of classical IP over ATM.

C-IPOA defines a mechanism for ATM networks to carry multiple types of protocols including IP over an ATM transport using AAL5 adaptation. In this approach, one of the two types of encapsulations can be chosen when an ATM VC (PVC or SVC) is established. They are IEEE 802.2 Link Layer Control/Subnet Attachment Point Encapsulation (LLC/SNAP Encapsulation) or VC-based multiplexing. LLC/SNAP encapsulation is the default packet format for IP datagrams. In LLC/SNAP multiplexing, protocol distinctions are made through the use of an LLC/SNAP protocol ID in every Layer 3 message, IP in this case. To reduce the encapsulation overhead, the VC-based multiplexing mechanism can be used. The protocol to be used on a VC is defined during VC setup time and is maintained throughout the VC connection time. This mechanism, however, does not provide the multi-protocol encapsulation capability available with LLC/SNAP encapsulation.

Multi-protocol encapsulation alone, although necessary, is not sufficient to provide routing and forwarding of IP datagrams over ATM transports. Resolution of IP addresses to ATM native

addresses is required. [CIP_ATM] defines a classical IP over ATM model. Figure I.1 shows the functional blocks for building the signalling, management and user planes and the message flows between an IP host and an ATMARP server. An ATM network is partitioned into discrete administrative and functional domains called Logical IP Subnets (LISs). Each LIS functions independently of other LISs. All members (hosts and routers) within an LIS have the same IP network/subnet address prefix and address masks. In this model, the deployment of ATM is used as a direct replacement of wide area networks supporting IP. Thus, an Address Resolution Protocol (ARP) type of server, called ATMARP is needed for the resolution of target IP addresses to target ATM addresses within a single LIS. The ATM addresses can either be E.164 addresses or ATM End System Addresses (AESA). The ATMARP functions stay within a single LIS.

In the classical model, hosts communicate among themselves directly via ATM within the same LIS using the ATMARP service for target address resolution. Communication outside the local LIS is provided via an IP router. The use of Next Hop Resolution Protocol (NHRP) to communicate among LISs is an extension to the classical model (reference Figure I.1 and I.1.1). ATMARP is a query-response client-server protocol. ATMARP clients (ATM hosts) must be configured with or learned through ILMI of the ATM address of the ATMARP server before the query-response operation is possible. Prior to an ATMARP query-response operation, an ATMARP client needs the establishment of an SVC or uses a pre-configured PVC to register itself with the ATMARP server (Step 1 in Figure I.1). During an ATMARP operation, the client sends an ATMARP-Request message to the server over this VCC. The source IP and ATM addresses are included in the request message along with the target IP address. The server is expected to respond with the appropriate target ATM address in an ATMARP-Reply message if the IP address can be resolved. If otherwise, an ATMARP-NAK message will be returned (Steps 2 to 6 in Figure I.1). Once the target ATM address is resolved, communication between two hosts can commence by establishing an ATM VCC and performing data transfer (Steps 7 and 8 in Figure I.1). Each ATMARP client maintains a table keeping records of all the resolved address entries. A client must refresh this table with its server within the Aging period using the registration procedures. An inverse address resolution process (In ATMARP) is also provided by the classical model and is used to resolve the target IP address given the target ATM address of an LIS member.

I.1.1 Next Hop Resolution Protocol (NHRP)

NHRP, specified in [NHRP], extends the classical model by providing communication between multiple LISs. In NHRP, a source station (host or router), known as a source Next Hop Client (NHC), intending to communicate with a destination station known as the destination NHC (host or router), uses the NHRP request and response protocol to obtain the ATM address of the destination station. An NHRP request transverses through a series of NHRP servers (NHSs) following the path as defined by the routing protocol in use until it reaches the NHS serving the destination station, whereby an NHRP response is returned to the source station. A "shortcut" path is then established between the source and destination stations via a direct ATM virtual circuit. If the destination station is within the ATM network served by an NHS, it will be directly reached via this shortcut. If it is outside the network or for any policy constraint, an egress router "closest" to the destination will be connected via the NHS. If the destination host is not served by any NHSs, a negative NHRP response will be returned and routing to the destination will follow the normal routing protocol.

I.1.2 Use of local ATM shortcut

NHRP turns particularly useful when QoS demanding IP flows (e.g., GS or CLS Intserv flows) have to be supported, as it eliminates IP hopping at each LIS boundary. However, it requires the introduction in the network of dedicated servers and adds the complexity of another query-response protocol. Moreover, the setup delay of the shortcut connection supporting the flow may be significant.

Another way to avoid multiple IP hopping without requiring any additions to the classical model is to perform local ATM shortcuts at each LIS boundary [51]. This requires the LIS border routers to be hybrid IP/ATM devices, i.e., not only routers with ATM interfaces, but integrated IP/ATM switches capable of sharing some information between the two layers.

Specifically, the basic functions that these hybrid devices should perform are:

- the building and maintenance of an association map between IP flows and the ATM connections supporting them, both for the incoming and outgoing direction;
- the local ATM shortcut on the basis of this association map.

The effort of building such an association map, however, is only worthwhile for long lasting, QoS-demanding IP flows. GS or CLS Intserv flows requiring some QoS through RSVP signalling are the most straightforward example.

Note that if the hybrid device is in charge of setting up ATM connections in response to the reception of RSVP signalling messages, then the building of the association map for the outgoing side is straightforward and does not require the usage of any particular standard mechanism: all that is needed is an internal communication between the IP and ATM components of the device. For the incoming side, on the contrary, the hybrid device needs to exploit some information that can be carried in the ATM signalling messages. Specifically, ITU-T Rec. Q.2941.2 [Q.2941] defines DSS2 signalling capability to carry, among others, Internet-related identifiers (i.e., an IPv4 or IPv6 session identifier, which identifies an IP flow). This information of course is available if an upstream analogous hybrid device takes care of filling the above-mentioned fields in ATM signalling messages.

The table below lists if and how a hybrid device can exploit the association map and perform a local ATM shortcut, depending on the type of association between IP flows and ATM connections (one IP flow per ATM connection or many IP flows on a single ATM connection).

In case 1, the maximum advantages are obtained: the IP flow is supported across several LIS by a concatenation of ATM connections, but, thanks to the possibility of on-the-fly forwarding, the resulting QoS is identical to the one that would be obtained on a direct ATM connection.

In case 3 (VC merging case), there is only the advantage of avoiding IP level processing, but on-the-fly forwarding is unachievable. In the other two cases, IP processing is still needed and no performance gain is achievable on that hop.

	Incoming association	**Outgoing association**	**IP processing required**	**On-the-fly forwarding allowed**
1	One to one	One to one	No	Yes
2	Many to one	One to one	Yes	No
3	One to one	Many to one	No	No
4	Many to one	Many to one	Yes	No

In such a scenario, each hybrid device is responsible for the type of association for the outgoing side, according to a given policy. For example, it may decide to always set up a separate ATM connection for each GS IP flow, and thus have a one to one association, and to always merge CLS flows on a single ATM connection to save VCIs. A coordinated choice for the policies of the hybrid devices of a single administrative domain is of course highly desirable.

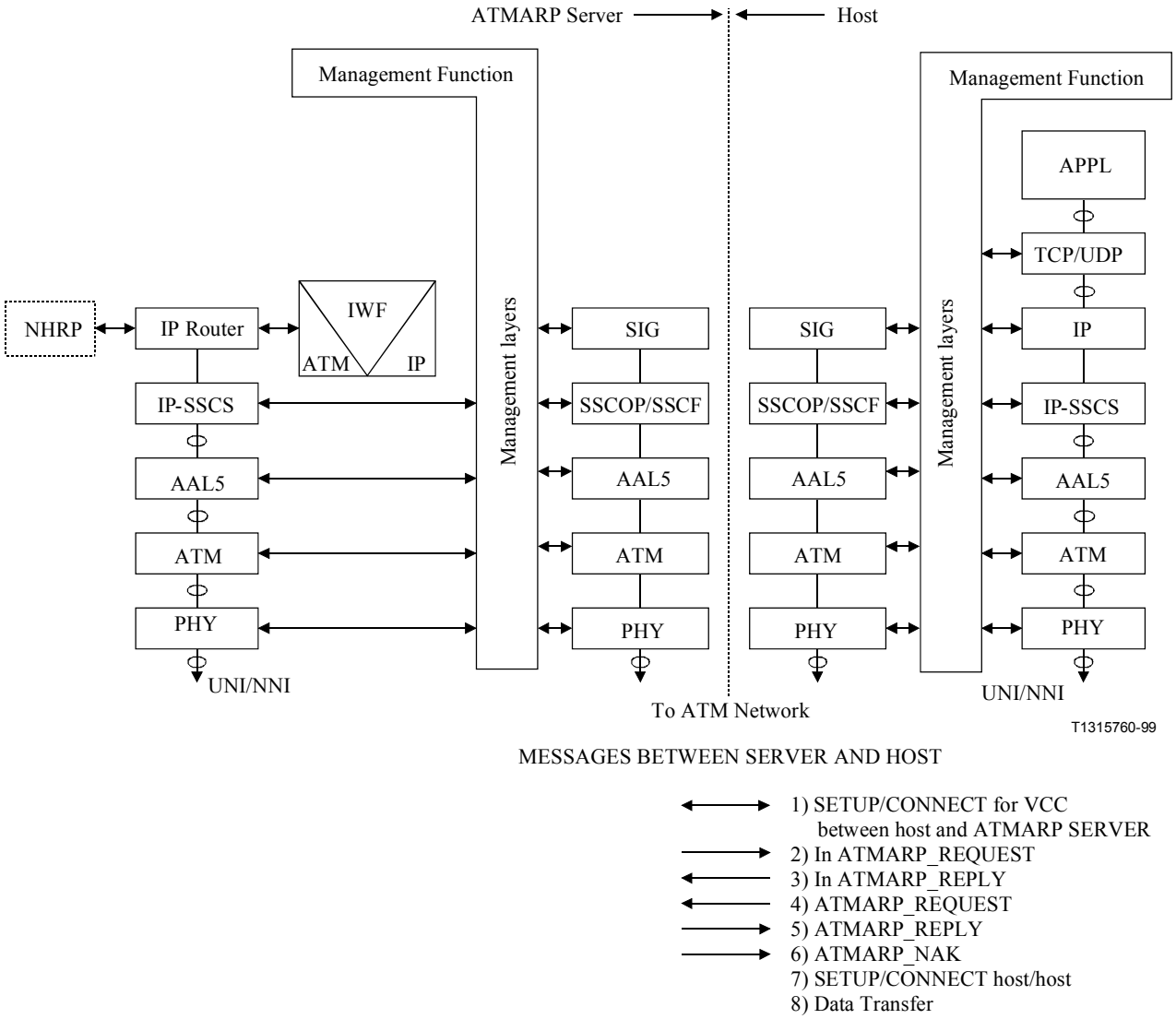


Figure I.1/Y.1310 – Functional description of classical IP and ARP over ATM

I.2 Multi-Protocol Over ATM (MPOA)

[ATM_MPOA] specifies a generic bridging and routing environment to transport multi-protocols (e.g., IP packets) over direct ATM VCCs. The technology combines the Local Area Network Emulation (LANE) technology with the Next Hop Resolution Protocol (NHRP) technology to provide an ATM shortcut paradigm. Figure I.2 illustrates the functional blocks of MPOA, showing the relationship between the control, management and data planes. The MPOA components in this figure are: NHS, NHC, MPC, MPS and LANE. Their functions are explained below.

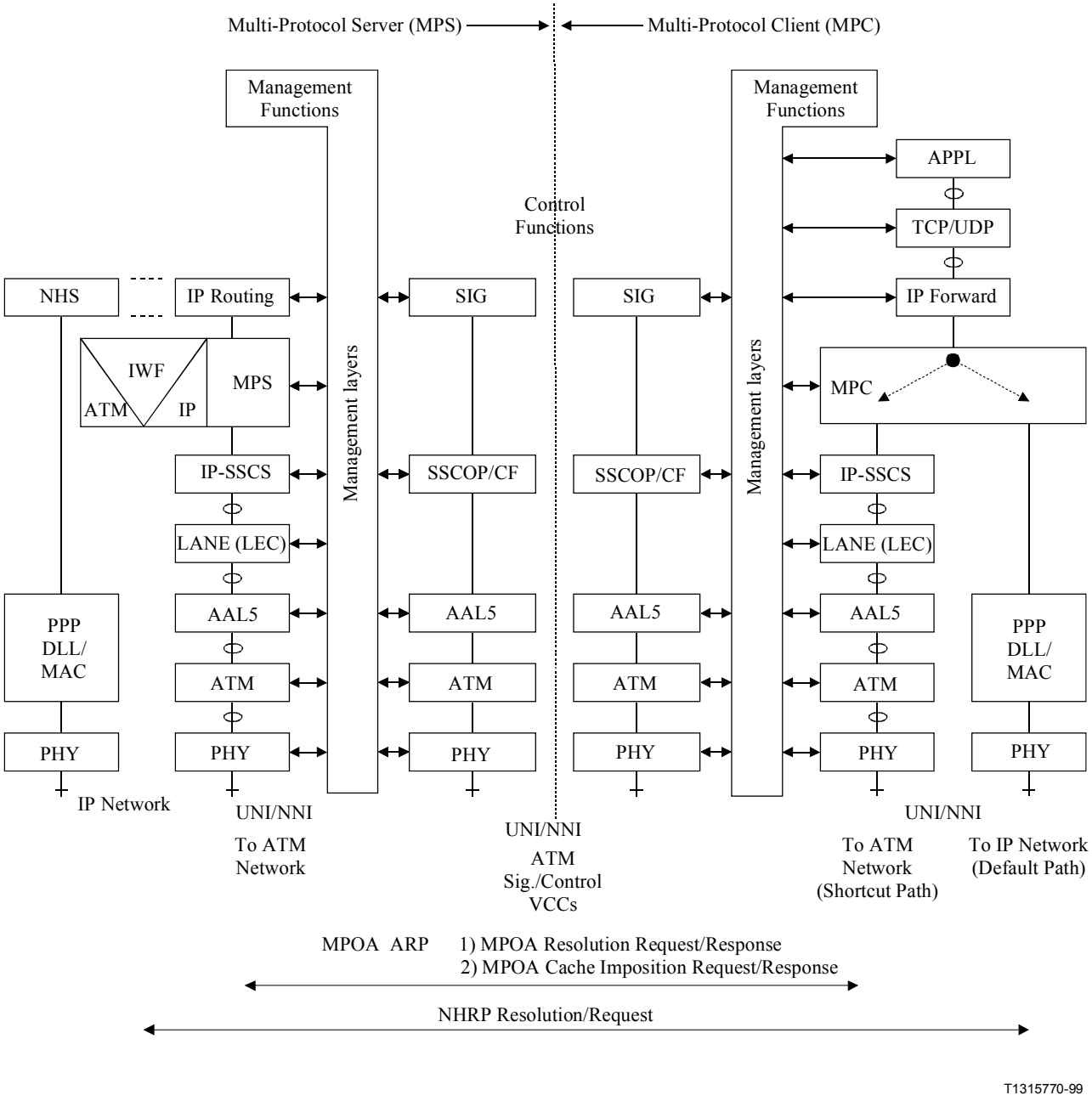


Figure I.2/Y.1310 – Functional description of MPOA

LANE forms an integral part of the MPOA protocol. LANE segregates a large ATM network into multiple domains, each of which can be emulated as a LAN segment. LANE specifies a set of protocols for LAN users to communicate among themselves within an ATM environment. These LANE users can be ATM-attached end systems or LAN-attached users. IP services are supported within this LAN environment. The LANE protocol operates between the ATM AAL5 and the network and LLC layers. LANE has four main LANE components: LANE Client (LEC), LANE Server (LES), Broadcast and Unknown Server (BUS) and LANE Configuration Server (LECS). A LEC (e.g., a LAN station) obtains configuration information from and registers with a LECS. A LES resolves MAC addresses of the LANE clients to their corresponding ATM addresses. The address resolution protocol (LE_ARP) functions similarly to that used by IP ARP. At the stable stage, direct data ATM VCs are used to connect these clients for data transfer. BUS distributes client data before address resolution is complete and data paths are established or when a client does

not know which direct data VC to use. IP packets are carried via LLC/SNAP encapsulation or via VC multiplexing as described earlier.

The other integral component of MPOA is NHRP, which is described in I.1. The classical IP over ATM model has a constraint of serving a single LIS. NHRP extends this capability by allowing "shortcuts" across multiple LISs within an ATM network.

I.3 Multi-Protocol Label Switching (MPLS)

MPLS was developed to achieve fast and efficient data forwarding for Internet routers [MPLS_ARCH]. Although architecturally targeted for multiple protocol applications, so far MPLS is primarily used for the IP protocol. In the connectionless IP environment, IP routers conventionally perform IP data forwarding on each datagram along a routed path to the destination based on a hop-by-hop routing decision. This next hop decision involves the examination of the IP packet header by the router to assign the packet to a Forwarding Equivalence Class (FEC) and the mapping of the FEC to a next hop thus determining the direction of the routing path. This process can be simplified and made more efficient by an MPLS process. With MPLS, the assignment of an IP packet to a FEC is done once by the Ingress Label Switching Router (LSR) and the FEC is represented and encoded by a fixed length label. The label is attached to the IP packet header. The header is no longer used by subsequent routers for forwarding the packet. LSRs, along the label switching path (LSP), use the label to index a table that specifies the next hop and a new label. Old labels are replaced with new ones as the packet transverses the LSRs along the LSP towards the destination. Labels are locally significant and their coding is specified in [MPLS_ENCAPS]. Labels represent the complete forwarding behaviour of a packet. It follows a hop-by-hop behaviour that includes choosing the next hop for the packet and the operation to be performed on the label, such as removal or replacement. Under normal situations, an LSP follows the same path as determined by normal IP routing protocols such as OSPF. MPLS can run over any link layer transport such as ATM, frame relay or Point-to-Point Protocol (PPP). Figure I.3 depicts the protocol structure of MPLS running over ATM. The main MPLS protocol elements in this figure are LDP, LIB and FIB. LDP is described in the next paragraph. Label Information Base (LIB) and Forwarding Information Base (FIB) are information databases that contain label binding information and forwarding information on the labels [MPLS_ARCH], [MPLS_ENCAPS] and [LDP].

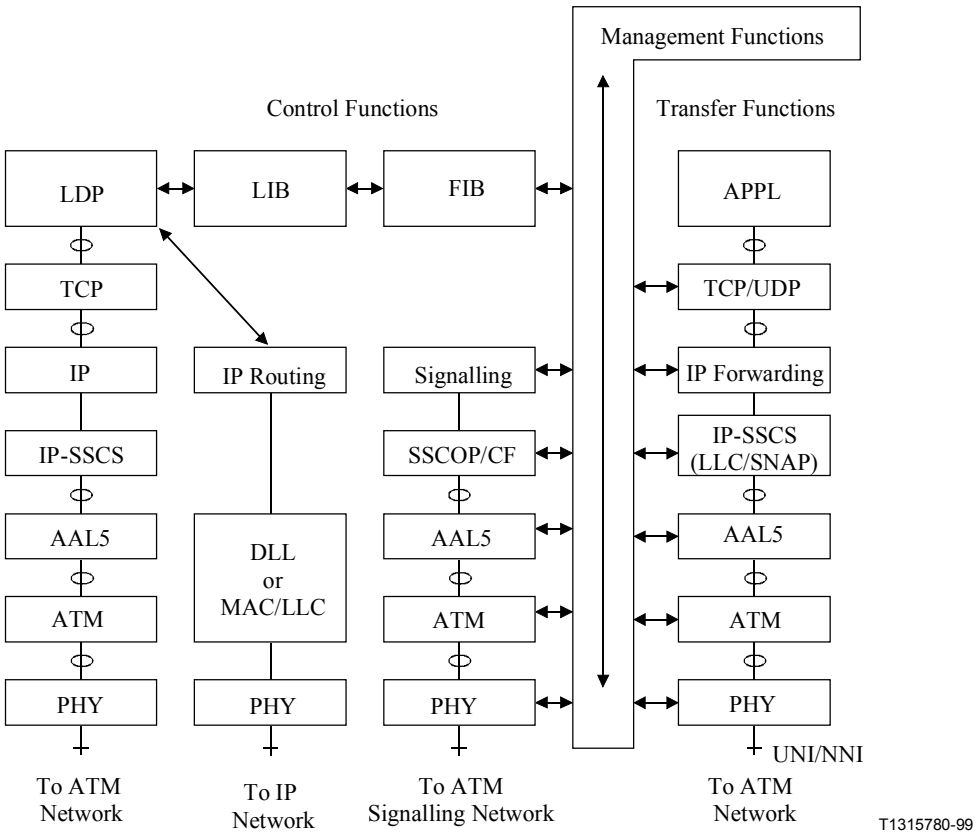


Figure I.3/Y.1310 – Functional description of MPLS

NOTE – The use of ATM signalling is only required for MPLS to B-ISDN interworking.

In order to provide a meaningful definition and common understanding of MPLS labels within an MPLS domain, an MPLS signalling protocol is required. This can be achieved by the use of the Label Distribution Protocol (LDP) [LDP], which provides a standardized MPLS signalling mechanism for allocating and distributing labels. As depicted in Figure I.3, MPLS can use LDP to build an LIB information base derived from the routing protocol in use and also sets up LSP connections between the corresponding ingress and egress LSR endpoints. LDP runs primarily over reliable TCP connections (except for the discovery process stated below which uses UDP). LDP has four phases of operation:

- Discovery: To announce and maintain the presence of LSRs in the network.
- Session: To establish and maintain sessions between LDP peers.
- Advertisement: To perform label allocation and distribution.
- Notification: For error reporting.

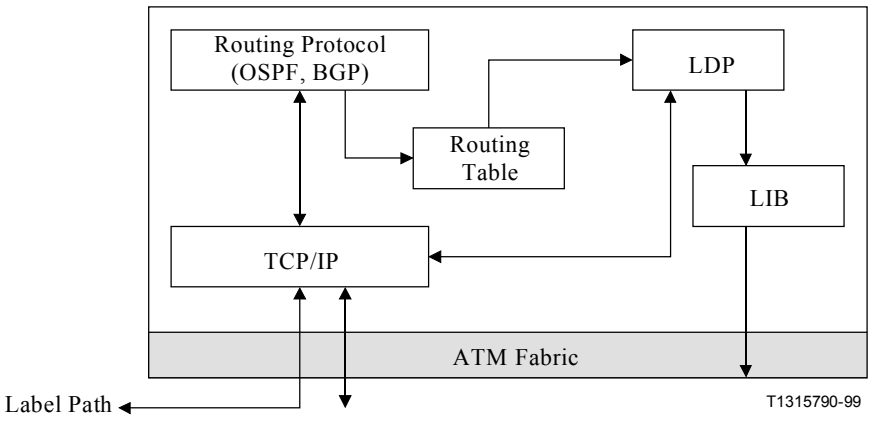
When labels are distributed, certain mechanisms or modes can be chosen. For example, a mechanism is the downstream-on-demand label distribution, where labels are distributed by a downstream LSR in response to an explicit request from its upstream LSR. Other distribution mechanisms and modes are detailed in [LDP]. Other than LDP, pre-configuration or existing IP protocols such as RSVP and BGP can be extended to handle label distribution [MPLS_ARCH].

Constraint-based Routing (CR) is a mechanism used to deliver Traffic Engineering (TE) capability and QoS performance characteristics within a network. These requirements can be met by extending the "conventional" LDP [LDP] or Resource Reservation Protocol (RSVP) [RSVP_FUN] for the support of Constraint-based Routed Label Switched Paths (CR-LSPs). Both extended protocols CR-LDP [CR_LDP], RSVP-TE [RSVP_TE] provide the following CR capabilities:

- *Explicit Routing (ER):* An explicit route can be defined as a list of nodes and established via signalling. This can divert from the conventional LSPs that are based on IP routing. Both strict and loose ERs are supported.
- *Traffic Characterizations:* the traffic characteristics of a CR-LSP can be defined using traffic parameters [CR_LDP, RSVP_TE].
- *Path Pre-emption:* During a path establishment, signalling provides this new path with the ability to pre-empt existing CR-LSPs, should such a need arise. Whether a new path can pre-empt an existing path depends on the setup priority of the new path and the holding priority of the existing path. This ability allows a network operator to satisfy network policy and engineering requirements within the available resources.
- *Route Pinning:* This option allows a segment of a loose ER to be fixed.
- *Resource Classes:* Network resources can be categorized by a network operator into "resource classes".

ATM switches can be used as label switching nodes. When an ATM switch is used as a label switching node or router (called ATM-LSR), the label on which forwarding decisions are made is carried in the VCI/VPI field of the ATM cell header. To support label switching, an ATM-LSR must support the control and signalling protocol for label switching such as LDP and participate in a network layer routing protocol such OSPF. ATM-specific routing and addressing are not needed. Between peer ATM-LSR, a dedicated ATM virtual connection (dedicated VPI/VCI) must be established for LDP control signalling. As in conventional LSRs, other methods such as OSPF, RSVP, PIM can be used for label distributions. An ATM-LSR can perform label switching on a VPI, VCI or VPI/VCI field, depending if VC or VP merging is used for flow aggregation. Peer ATM-LSRs can be connected directly over an ATM link or remotely through an ATM cloud over an ATM virtual connection. In the latter case, ATM signalling will have to carry the binding information.

Figure I.4 shows a protocol structure of ATM-based MPLS architecture. The MPLS/ATM architecture has two parts: one is MPLS routing module and the other is ATM forwarding module. The MPLS routing module includes IP routing protocol functional block supporting OSPF and BGP, TCP/IP protocol stack, and LDP and its running result, LIB used for label distribution and assignment. The ATM forwarding module is the ATM fabric.



LDP Label Distribution Protocol
LIB Label Information Base

Figure I.4/Y.1310 – An example of MPLS implementation

Appendix II

Guidelines for mapping services to ATM connections

II.1 Mapping Intserv services to ATM connections

II.1.1 Mapping of Guaranteed Service (GS) to ATM

II.1.1.1 Network model for GS

The network model assumed by [GUAR_SER] is reported in Figure II.1.

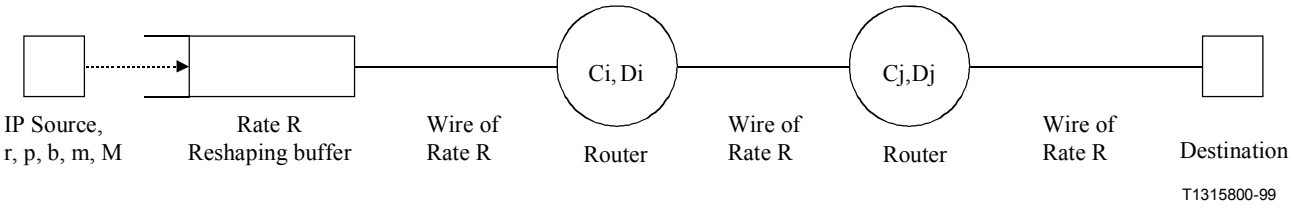


Figure II.1/Y.1310 – Network model for GS

The source of the IP flow requesting GS emits traffic according to its token bucket specification (r, p, b, m, M). After a transient phase, during which traffic may be carried as best effort, resources are allocated so that the network can be modelled as a sequence of "wires" of rate R. Just before the first wire, all traffic instantaneously exceeding rate R (even if conforming to the token bucket specification) is buffered and reshaped at the rate R. By assuming this ingress reshaping at the allocated rate R, the network model is considering the worst case-situation for what concerns delay variation. Wires are linked by devices (routers) that introduce a "distortion" from the ideal fluid model (a single wire of rate R). The distortion introduced by each router is taken into account by two terms called C (rate dependent) and D (rate independent) respectively. RSVP messages carry to the receiver the sum of all Ci and Di values, so that it can compute an upper bound for the variable part of the delay. If the rate of the wires is R this upper bound is:

$$\frac{b-M}{R}\cdot\frac{p-R}{p-r}+\frac{M+C_{tot}}{R}+D_{tot} \quad \text{for } r \le R < p \text{ [1]} \tag{II-1}$$

$$\frac{M+C_{tot}}{R}+D_{tot} \quad \text{for } r \le p \le R \tag{II-2}$$

So, the receiver requests that the bandwidth reserved on wires be R (by sending upstream an RSVP RESV message) in order to keep Equation II-1 or II-2 below a target value. Note that the delay in the reshaping buffer (see Figure II.1) is taken into account by the first term of Equation II-1, but if R > p (as in Equation II-2) traffic will never be delayed in the reshaping buffer.

II.1.1.2 Choice of the ATM service

The problem of mapping arises when wires have to be replaced by ATM connections. Strictly speaking, if the wire model has to be respected, no choice would be possible as ATM connections always introduce a CDV that reflects on a delay variation for packets, whereas wires do not. In practice, the choice should be limited to those ATCs that can be associated to a QoS class that ensures limited CDV (i.e., QoS Class 1 [I.356] or QoS Class 5 [I.356]). This CDV can then be taken into account on the D term of Equation II-1 or II-2.

The ATCs DBR [I.371] and SBR1 [I.371] may be associated with QoS Class 1 [I.356]. The ATCs SBR2 [I.371] and SBR3 [I.371] may be associated with QoS Class 5 [I.356]. So, what criteria can be used to choose the ATM service? A first one may be to consider whether the IP flow, before being sent to the ATM level, is really reshaped at the IP level at a rate R, as supposed by the model. If so, the most natural choice is to adopt a QoS class 1 DBR ATM connection with PCR = (ATM equivalent of *R*)[2]. Choosing an SBR connection with SCR = (ATM equivalent of *R*) would require uselessly more resources than the DBR one, except in the case PCR = SCR = (ATM equivalent of *R*) and MBS = 0, that means falling in the DBR case again.

On the contrary, if the ATM connection starting point is unaware of any IP level packet reshaping, the best mapping would be with a SBR connection with SCR = (ATM equivalent of *R*), PCR = (ATM equivalent of *p*), MBS = [ATM equivalent of $bp/(p - r)$]. DBR would require more resources to accommodate conforming bursts of traffic up to the rate *p*, and therefore it needs a PCR = (ATM equivalent of *p*) or MCR = (ATM equivalent of *p*).

Of the three SBR versions, the one that best matches the service model of GS is SBR3, which allows the tagging of non-conforming traffic. This allows to leave up to the ATM level all the complexity of treating the excess traffic as best effort, as requested by [GUAR_SER]. Associated QoS class is QoS class 5 [I.356].

---

1 This formula remains the same even if the reshaping does not occur at the ingress but at one or more than one of the crossed routers.

2 See II.1.1.3.

Table II.1/Y.1310 – Preferred GS to ATM mapping

	ATM level is aware of an IP level packet reshaping at rate *R* immediately before connection's starting point	ATM level is unaware of any IP level packet reshaping at rate *R* immediately before connection's starting point
Preferred ATC and QoS class	DBR class 1	SBR3 class 4
Mapping between ATM traffic descriptors and token bucket's parameters	PCR = (ATM equiv. of *R*)	Note 1 Note 2 PCR = (ATM equiv. of *p*) SCR = (ATM equiv. of *R*) MBS = (ATM equiv. of $bp/(p - r)$)
NOTE 1 – The parameter mapping to SBR is always valid when $R \leq p$. However, in the GS the *R* parameter can be set greater than *p* (see Equation II-2). As PCR cannot be set below *R*, it will result PCR = SCR = $R > p$ and there would be no reason to have an MBS > 0. So, in the case $R > p$, the preferred mapping is the DBR with QoS class 1, with PCR = (ATM equivalent of *R*). NOTE 2 – When SBR is used, there is an intrinsic inefficiency in the mapping scheme, due to the fact that the network model of GS considers "wires" of rate *R* while SBR connections are much "better" than wires with rate *R*, in the sense that they can absorb instantaneous bursts of traffic up to the rate *p*, without relying on buffering the traffic exceeding *R*. This inefficiency is reflected into an over-allocation for *R*, due to the first term of Equation II-1 which is assumed by the GS model but in the real network could be non-existent or much smaller.		

II.1.1.3 ATM equivalents of token bucket parameters

When translating token bucket's parameters into ATM traffic descriptors, one should recall that the former are expressed in bytes or in bytes/s, while the latter are in cells or cells/s. Furthermore, ATM and AAL overheads must be taken into account.

An upper bound for the number of cells needed to carry an IP packet of *B* bytes is:

$$C(B) = (H + B + T + 47)/48 \qquad \text{(II-3)}$$

where *H* and *T* are the AAL PDU header and trailer lengths and "47" accounts for the last cell which may be only partially filled.

The ATM equivalents for the terms appearing in Table II.1 are listed in Table II.2. The assumption is that the token bucket specification is (r, b, p, m, M).

Table II.2/Y.1310 – ATM equivalents for GS to ATM mapping

Mapping to DBR class 1	Mapping to SBR3 class 4
$PCR = \left\lfloor \frac{R}{m} \right\rfloor C(m)$	$PCR = \left\lfloor \frac{p}{m} \right\rfloor C(m)$ $SCR = \left\lfloor \frac{R}{m} \right\rfloor C(m)$ Note $MBS = \left\lfloor \frac{bp}{m(p-r)} \right\rfloor C(m)$
NOTE – For how much time can a source conforming to the token bucket specification send "bytes" at the peak rate p[bytes/s]? For $T = b/(p - r)$ seconds. How many bytes can it send, at the peak rate p, before becoming "non-conforming"? $bp/(p - r)$. How many packets at most? bp/[$m(p - r)$]. An SBR ATM connection carrying this traffic should therefore transmit this amount of packets "transparently" (i.e., at their peak rate), and therefore it should have the indicated MBS (in cells).	

These equivalencies are worst case, i.e., are calculated assuming all the packets having the minimum declared length, thus considering the maximum possible overhead impact. A more realistic estimate can be done replacing m in the above formulas with a value in the range [m, M], but this requires a detailed knowledge of the generating application's distribution of packet sizes.

II.1.1.4 Accounting for CDVT

After the packet segmentation is performed, cells belonging to a packet are simultaneously ready for transmission and may be sent at the line rate if no cell level reshaping is done. When no cell level shaping function to absorb the burst due to the packet segmentation is assumed to exist, it is necessary to account for this bursty behaviour by adding a proper value to the CDVT on the PCR parameter of the DBR or of the SBR connection. This value can be evaluated as:

$$(C(M)-1)\left(\frac{1}{PCR}-\frac{1}{LCR}\right) \qquad \text{(II-4)}$$

II.1.2 Mapping of Controlled Load Service (CLS) to ATM

Also for CLS the network model considered by [CONTROL_SER] consists of a source whose traffic is described by a token bucket specification (r, p, b, m, M) and a sequence of routers linked by "wires" (see Figure II.1), but as there is no explicit guarantee on the delays there is not any specific formulae as Equation II-1 or II-2. When wires have to be replaced by ATM connections, the choice is no more limited to ATCs that can have limited CDV. As the requirement of CLS is simply to have a "long term"[3] available bandwidth, and low losses, suitable ATM services may be:

- DBR with class 2;
- ABT with class 2;
- ABR with class 3;
- SBR1 with class 2;
- SBR2 with class 3;
- SBR3 with class 3;

[3] "Long term" means on timescales significantly larger than b/r, where b and r are part of the token bucket's parameter of the source.

- GFR1;
- GFR2.

Mapping with DBR with class 2 would require setting the PCR somewhere between (the ATM equivalent of *r* and *p*)[4], i.e., finding an equivalent bandwidth for the single flow, but besides being an implementation specific issue, this also risks to be inefficient. Moreover, class 2 does not allow the ATM layer to take care about the best effort treatment of the traffic exceeding the token bucket's specification, as requested by [CONTROL_SER].

Mapping with ABT/dt with class 2, even if potentially attractive, has the drawback of the re-negotiation overhead at the burst level, probably unbearable. With longer term re-negotiations, the same drawback outlined for DBR applies.

Mapping with ABR with class 3 could be performed by setting MCR = (ATM equivalent of *r*), but has the drawback that all the traffic instantaneously exceeding r is treated as best effort. Setting MCR somewhere between (the ATM equivalent of *r* and *p*) has the same drawback outlined for DBR without guaranteeing, anyway, that the ATM layer can make an exact discrimination between traffic conforming and non-conforming to the token bucket specification.

Mapping with SBR1 with class 2 or SBR2 with class 3 allows the ATM level to know the maximum amount of information to perform an efficient statistical multiplexing, but as regards the treatment of the non-conforming part of the traffic it does not meet the expectation that it is treated as best effort.

The remaining three mappings all meet both the goal of reflecting at the ATM level as closely as possible the traffic characteristics as specified by the token bucket specification and of providing a best effort treatment for exactly the portion of traffic exceeding it. The detailed ATM equivalents are reported in Table II.3.

Table II.3/Y.1310 – ATM equivalents for CLS to ATM mapping

Mapping to SBR3	Mapping to GFR1 or GFR2
$PCR = \left\lfloor \frac{p}{m} \right\rfloor C(m)$	$PCR = \left\lfloor \frac{p}{m} \right\rfloor C(m)$
$SCR = \left\lfloor \frac{r}{m} \right\rfloor C(m)$	$MCR = \left\lfloor \frac{r}{m} \right\rfloor C(m)$
$MBS = \left\lfloor \frac{bp}{m(p-r)} \right\rfloor C(m)$	$MFS = C(m)$
	$MBS = \max\left(\left\lfloor \frac{bp}{m(p-r)} \right\rfloor C(m), MFS \right)$

The same considerations about the need of replacing *m* with a value between [*m*, *M*] and about the value to add to CDVT expressed at the bottom of II.1.1.4 apply.

II.2 Mapping Diffserv services over ATM

This clause describes some possible examples of the differentiated service mapping to ATM services for information. IETF has just described application services that can be supported by each PHB or PHB group [DIFF_AF] and [DIFF_EF].

4 See also II.1.1.3.

- *Leased line emulation service*

  It is also called the "premium service". This service can be implemented using the EF-PHB. This kind of service usually requires stringent low loss and delay guarantees. This service is also characterized by its peak rate. Accordingly, this service could easily be mapped to the DBR ATC using QoS class 1 in order to meet such loss and delay requirements. The peak rate can be mapped to the PCR parameter PCR of DBR ATC in a straightforward fashion.

- *Quantitative assured service*

  It is also called the "assured rate service". This service can be implemented using one of four AF-PHB classes. This service is characterized by a minimum rate guaranteed on a statistical basis. This service offers looser assurances than the leased line emulation service, but is still considered a quantitative service. In particular, it promises to deliver traffic with a high degree of reliability and with bounded latency, up to a negotiated rate. Accordingly, it seems to be a perfect fit to map this service to ABR ATC using QoS class 3. In this case, the MCR can be set equal to the service minimum rate.

II.3 Intserv in MPLS over ATM

The traffic parameters of Intserv including *p, r, b* and *R* are defined in the RSVP objects such as *Tspec* and *Rspec*. If CR-LDP is used to support Intserv in MPLS over ATM networks, instead of RSVP, the following requirements must be considered:

– When the RSVP/Intserv flow, including guaranteed service and controlled-load service, enters an ingress LSR in MPLS networks, the RSVP *Tspec* parameters such as *p, r* and *b* must be reflected on the traffic parameters in the label request message of CR-LDP.

– To support the guaranteed service, the RSVP *Rspec* parameters such as *R* and *S* must be reflected on the traffic parameters in the label mapping message of CR-LDP.

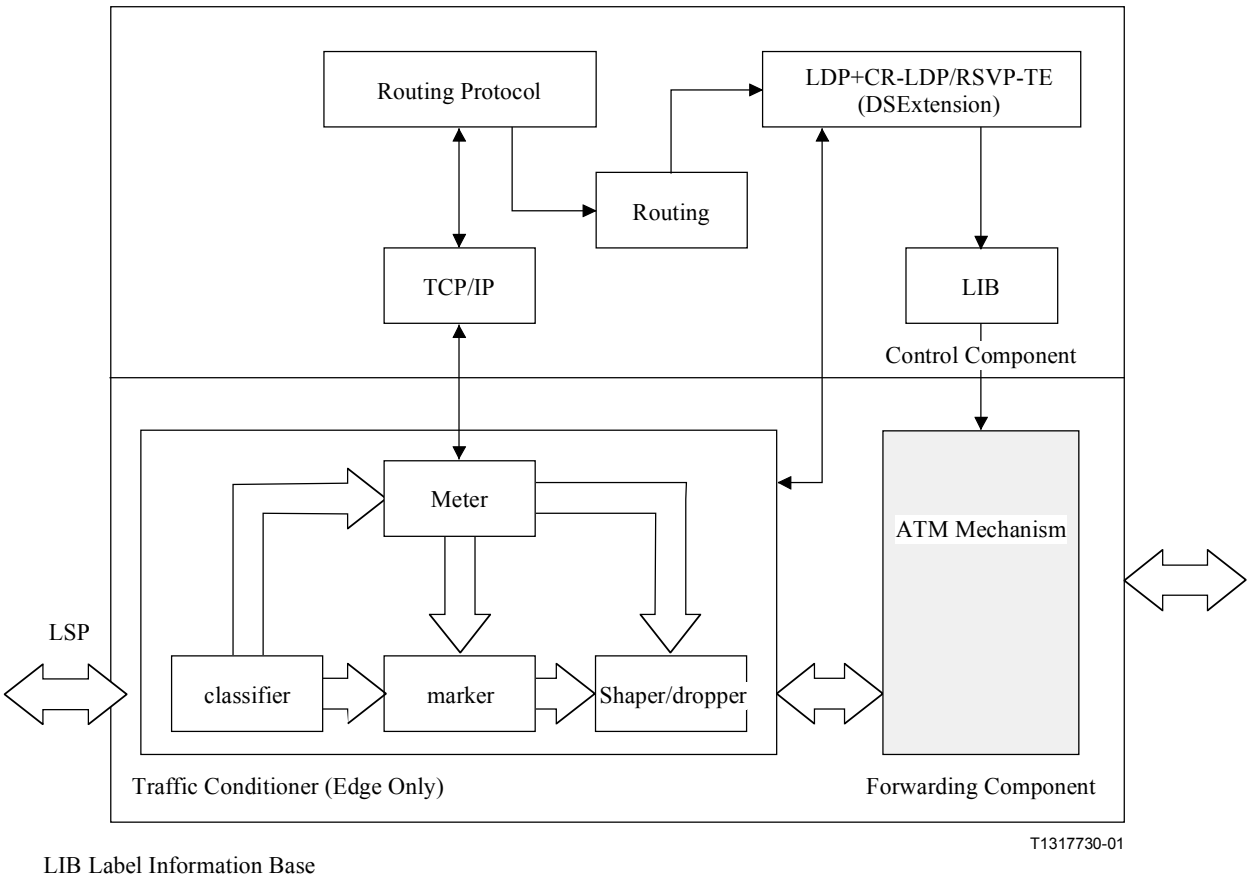
The traffic parameter mapping from Intserv to CR-LDP depends on the traffic conditioning policy at the Ingress LSR.

II.4 Diffserv in MPLS over ATM

This clause depicts an approach of supporting Diffserv in a MPLS ATM network. A Diffserv capable ATM-LSR should have the logical structure as specified in Figure II.2. It is worth noting that a transit ATM-LSR will usually not need the traffic conditioning element, but edge ATM-LSR must have this element to perform the packet classifying, marking, metering and shaping/dropping functions required by the Diffserv Architecture [DIFF_ARCH].

Either CR-LDP or RSVP-TE may be used as the signalling system.

For details about the Traffic Conditioner, refer to the reference [DIFF_ARCH].



LIB Label Information Base

Figure II.2/Y.1310 – Logical architecture of a Diffserv capable ATM-LSR

II.4.1 LSP setup procedures

The basic ATM MPLS Diffserv LSP setup procedure will include following actions:

- On the edge of the ATM MPLS Diffserv Domain, the ATM LERs will deal with the service requests and perform the service classification action. Then the LERs will determine the PHB (per hop behaviour) that will be used by the service. Then, the ATM LER should map the PHB into <PSC, CLP> (per hop scheduling, cell loss priority) pairs. The mapping relationships are specified in Table II.4.
- According to the Diffserv requirements, using the MPLS signalling system (for instances, we can use the CR-LDP with Diffserv support [MPLS_DIFF] as the signalling system) to perform the service provisioning process [DIFF_ARCH] and set up a QoS aware LSP for the service. At this point, the forwarding tables on the LSRs along the LSP will have a new column for incoming PSCs.

II.4.2 Label forwarding procedure

The basic label forwarding operation of ATM MPLS Diffserv scheme will include the following actions:

- On the ingress of the ATM MPLS Diffserv domain, through the checking of the IP address and DSCP (differentiated service code point) value carried by the IP packet, the LERs determine the packet's FEC and PHB class.
- The ingress ATM-LSR then performs the traffic conditioning, determine the outgoing CLP for the packet. When an ATM L-LSP has been set up, only the CLP field can be rewritten.

- The ingress ATM-LSR then determines the outgoing VCI and interface number, performs the outgoing PSC for the packet, encapsulates the IP packet into an ATM packet and send it onto the outgoing interface.
- In the MPLS Diffserv domain, the transit LSR checks the VCI and CLP field in the header of the ATM packet. By using the forwarding table, it determines the PSC required by the packet.
- By using the mapping table shown in Table II.4, the transit LSR then determines the PHB required by the packet. Usually, the transit LSR will not perform the traffic conditioning actions, it just realizes the PHB for the packet and uses the forwarding table to forward the packet to the downstream LSR. If the transit LSR does need to perform the traffic conditioning actions, it uses the results of the traffic conditioning procedures to change the outgoing CLP for an ATM packet.
- On the egress of the ATM MPLS Diffserv domain, the egress ATM LSR checks the VCI and CLP field in the header of the ATM packet. By using the forwarding table, it determines the PSC required by the packet. By using the mapping table shown in Table II.4, the transit LSR then determines the PHB required by the packet. Then it performs the traffic conditioning procedures and uses the results combined with the incoming PHB to determine the outgoing PHB and the outgoing DSCP value for the packet.
- Then, the ATM egress will convert the ATM packet back to IP packet, which carries the IP address and DSCP field. (This field must be replaced by the DSCP got from the action above.)

Table II.4/Y.1310 – Mapping between the Diffserv PHBs and ATM <PSC, CLP> pairs

Diffserv PHBs	PSC	ATM CLP
DF	DF	0
CSn	CSn (Note 1)	0
AFi1 (Note 2)	AFCi	0
AFi2	AFCi	1
AFi3	AFCi	1
EF	EF	0
NOTE 1 – "n" ($1 \leq n \leq 8$) refers to the number of IP precedence. NOTE 2 – "i" ($1 \leq i \leq 4$) refers to the AF PHB class, for instance, when i = 1, AFi1 will represent AF11 which belongs to the AF PHB class 1 and have the drop precedence 1.		

II.4.3 The mappings between <PSC, CLP> and PHB

The following PHBs have been defined:

II.4.3.1 DF (Default PHB): This PHB is used for the best effort packets or packets with unknown DSCP values.

II.4.3.2 CS (Class Selector PHB): This PHB is used for backward compatibility with the existing 8-level IP precedence system.

II.4.3.3 EF (Expedite Forwarding PHB): This PHB is used for the services that require low packet loss rate, low latency, low jitter and bandwidth guarantee. A packet with this PHB will get the highest service priority and premium service in the domain.

II.4.3.4 AF (Assured Forwarding PHB): This PHB is used to classify the packets in the same connection with different drop precedence. IETF defined four AF classes, and within each class

there are three PHBs with different drop precedence. Thus there are 12 AF PHBs in total. An application example of this PHB is when traffic exceeds a certain transmission rate; the excess packet will be assigned a PHB with a higher drop precedence. Another important requirement of this PHB is that the packets belonging to a single connection and the same PHB class cannot be reordered.

Table II.4 shows the mapping between the PHBs and <PSC, CLP> pairs. These mappings must be consistent on each LSR in an ATM Diffserv domain, and these mappings must be configurable.

II.4.4 Implementation Considerations

The ATM MPLS LSRs should support the PHBs and the traffic conditioning rules of IP services. However, the detail packet treatment and traffic conditioning rules on the ATM MPLS LSRs are implementation issues.

Appendix III

Possible evolution scenarios to MPLS for IP over ATM in public networks

III.1 Introduction

What are the potential routes to MPLS from current network infrastructure? This will be a function of the actual state as well as the services to be delivered by a specific carrier.

In this Recommendation, we assume that MPLS will be deployed into existing ATM backbones. For the purposes of examining the MPLS evolution solutions, carriers will be classified by whether they are New versus Established, and whether they are full-service (data, voice, video, leased line) versus IP-centric. This is not a universal classification; it is rather a convenient way to categorize service providers based on the actual state of their network and its anticipated service offering.

This appendix considers several types of existing infrastructure, and broad strategies for introducing MPLS into those network types. It then examines various technologies for operating MPLS over non-MPLS-capable ATM equipment, and makes recommendations about the use of those technologies.

III.2 Proposed scenarios

Different scenarios are presented and discussed as follows:

III.2.1 Established full-service carrier

We assume that an established carrier has a legacy of voice network and either transports the Data traffic over the TDM network or on a separate network. We also assume that such carrier is in the process of merging its data network and voice network over the same infrastructure.

The existing carrier likely has a legacy ATM infrastructure that is being used for data traffic (IP or Frame Relay) and may be used for voice and video traffic or any other native ATM services. In this case, ATM is used as a multi-service switching technology.

The existing carriage of IP in the carrier network is likely to be based on one of three cases:

– using point-to-point ATM PVCs with RFC 2684 encapsulation [ATM_MULTI];
– using Classical IP over ATM;
– using MPOA.

In any case, it will be necessary to introduce MPLS into a network which currently uses only PVCs, SPVCs, SVCs, PVP, and SPVPs, and does not currently use VCs with MPLS control. The MPLS VCs may be termed "Label VCs (LVCs)" to distinguish them from SVCs with PNNI or similar control.

III.2.2 Established voice service carrier

We assume that the carrier has a voice network only, traditional SS7/TDM without any substantial ATM investment and wants to carry both voice and data in the future using a cell-based MPLS infrastructure. What is the best evolution path?

The voice carrier would probably in a first phase decide to keep the SS7 control and transfer the voice traffic from a TDM network onto a packet network. Assuming the likelihood of choosing a cell-based MPLS, data as well as voice traffic will be transported onto this MPLS-based network. The service provider has either the choice to keep both networks separated or work towards their progressive integration.

III.2.3 IP-centric new carrier

The question here is whether or not it makes sense to do any ATM deployment. If the carrier chooses to deploy a cell-based MPLS, then there is little push towards having ATM control in the network. The main point would be to reuse ATM switching capabilities only.

III.2.4 Full-service new carrier

The full-service new carrier will be offering voice, video, and leased lines services along with IP-centric services. Due to the varying traffic types, we assume that the carrier might choose to deploy an ATM infrastructure to integrate their service offering onto one network.

Routers may be deployed at the edges of the network to support IP services, but the switched core will be ATM. Cell-based MPLS will be deployed in the core. Ships in the Night operation with ATM may be required to integrate MPLS and ATM based services. Explicitly routed MPLS will be required for traffic engineering, hop-by-hop MPLS would be required to handle traffic not carried within explicitly routed LSPs.

III.3 Hybrid ATM network

This clause discusses three possible ways for integrating MPLS equipment with non-MPLS equipment in an ATM network. This clause assumes support of ATM switching in a current network. Support for MPOA and C-IPOA is not addressed in this clause; however, the techniques discussed here may apply.

III.3.1 Technologies for hybrid ATM networks

During the introduction of ATM MPLS into an existing ATM network, it will sometimes be necessary to connect LSRs over traditional ATM equipment, forming a "hybrid" network. In hybrid networks, some switches and/or routers have MPLS capability and some do not. This clause discusses possible ways of implementing hybrid ATM networks: MPLS-over-PVCs, Virtual Trunks, and Virtual Connection Identifier Notification for LDP (VCID). These are illustrated in Figure III.1.

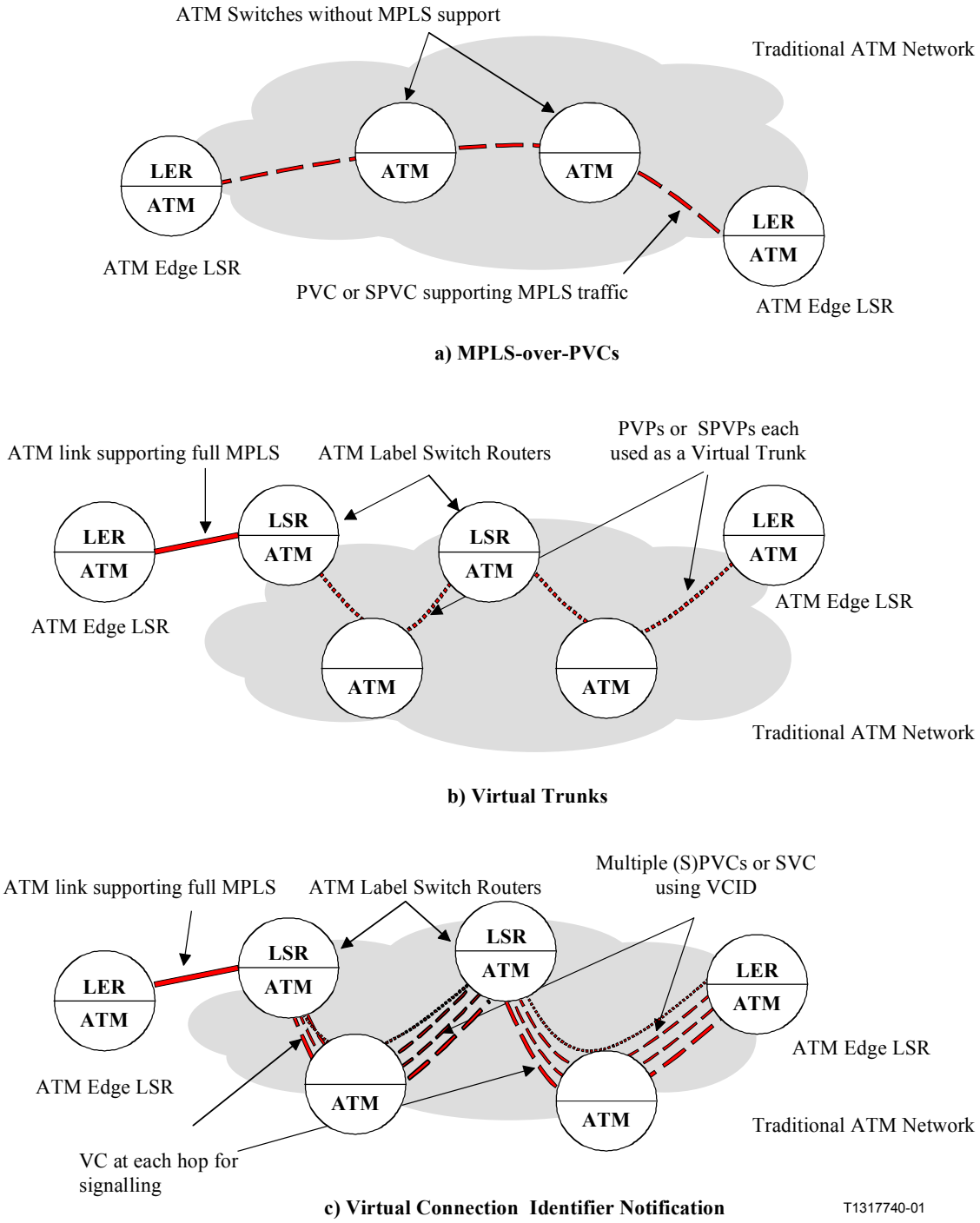


Figure III.1/Y.1310 – Technologies for hybrid networks

III.3.1.1 MPLS-over-PVCs

MPLS-over-PVCs is shown in Figure III.1 a). It can be used only to connect packet-based LSRs. It may not be used to connect ATM Label Switch Routers (ATM-LSRs) to each other. MPLS-over-PVCs connect packet-based Label Switch Routers (LSRs) by way of Permanent Virtual Circuit Connections (PVCs) over a traditional ATM network. Soft Permanent Virtual Circuit Connections (SPVCs) may also be used. (Any mention of a "PVC" with respect to MPLS-over-PVCs in this appendix equally applies to an SPVC.) The routers send MPLS packets to each other, with labels explicitly encapsulated along with the IP packet. This is called "packet-based labelling", as the MPLS label is applied to a whole packet, as opposed to individual cells. When packet-based labelling is used over PVCs, packets with many different labels are sent in the same PVC. This

differs from ATM MPLS, where each different label is represented by a different VC, known as a "Label VC" (LVC). Packet-based labelling over PVCs is virtually identical to the case where MPLS Label Switch Routers (LSRs) are connected by links such as Packet-over-SONET, Packet-over-SDH, or any other point-to-point links. Note that MPLS-over-PVC does not use ATM MPLS on the ATM switches supporting the PVCs. This means that service providers must continue to provision and manage PVCs on a scale equal to the traditional IP over ATM approach.

MPLS-over-PVCs uses the generic encapsulation that is described in the MPLS Label Stack Encodings specification [MPLS_ENCAPS]. Possible link-layer encapsulations for the PVC include Null encapsulation and LLC/SNAP encapsulation. If the PVCs are carrying only MPLS packets, then Null encapsulation is recommended. Otherwise LLC/SNAP should be used, with a SNAP header containing the ethertypes specified for MPLS over LAN media [MPLS_ENCAPS].

III.3.1.2 Virtual trunks

A different method of implementing hybrid ATM networks is the use of Virtual Trunks. Virtual trunks are based on Virtual Path (VP) connections. ATM MPLS normally involves labelling IP packets by putting them in different VCs in the same ATM trunk. Each different VC on the trunk represents a different label value. ATM LSRs treat virtual trunks almost identically to a physical trunk: each different VC within the VP represents a different label value. The difference is that the virtual trunk is not a physical trunk linking two adjacent LSRs. The virtual trunk is a Permanent Virtual Path Connection (PVP) or Soft Permanent Virtual Path Connection (SPVP) which connects ATM-LSRs by way of traditional ATM switches. Virtual trunks may also connect ATM edge LSRs to ATM-LSRs, or connect ATM edge LSRs to each other. Use of virtual trunks is illustrated in Figure III.1 b). Use of ATM MPLS with virtual trunks and VCI-based labels is described in the "MPLS using LDP and ATM VC Switching" document [MPLS_ATM], and is in most respects identical to use of MPLS over physical trunks with VPI/VCI-based labels. A VC must be assigned to carry the LDP control traffic and this VC must use the LLC/SNAP encapsulation.

III.3.1.3 Virtual Connection Identifier Notification for LDP (VCID)

VCID allows PVCs, SPVCs, and Switched Virtual Circuit Connections (SVCs) to be used in ATM MPLS [ATM_VCID]. (Here, "SVC" refers specifically to a dynamically-established VC in a traditional ATM network. VCs used directly by ATM MPLS are referred to here as "Label VCs", or "LVCs".) By contrast, MPLS-over-PVCs uses packet-based MPLS and not ATM MPLS, and Virtual Trunks uses PVP or SPVP connections and not PVCs, SPVCs or SVCs. VCID supports the use of PVCs, SPVCs and SVCs in similar network configurations to Virtual Trunks, as shown in Figure III.1 c). When VCID is used, a number of PVCs, SPVCs or SVCs are used to carry labelled packets between the ATM MPLS devices, with one VC per label. Because there is a distinct VC for each label, ATM MPLS packet forwarding can be used at the ATM MPLS devices using VCID.

There must be a default VC pre-established on each LSR-to-LSR "hop", in order to carry IP routing and LDP. This VC is in addition to the VCs used by VCID to correspond to labels.

III.3.2 Networks using MPLS-over-PVCs

III.3.2.1 Use of MPLS-over-PVC technology

The simplest network structure using MPLS-over-PVCs is the full mesh shown in Figure III.2 a). Operation of IP routing protocols in an MPLS network of this structure leads to the same scalability issues as for traditional IP-over-ATM networks of similar structure. One solution to this is to use a partial mesh between the routers, but this would lead to the use of inefficient, multi-hop routes. Another alternative is to add extra ATM edge LSRs, as shown in Figure III.2 b), or possibly a redundant pair of them. The extra ATM edge LSRs reduce the size of the meshes. Note that the performance requirements on the extra LSRs will be quite high, as they will carry a large part of the network traffic. There is no direct way of using ATM-LSRs in a network using MPLS-over-PVCs.

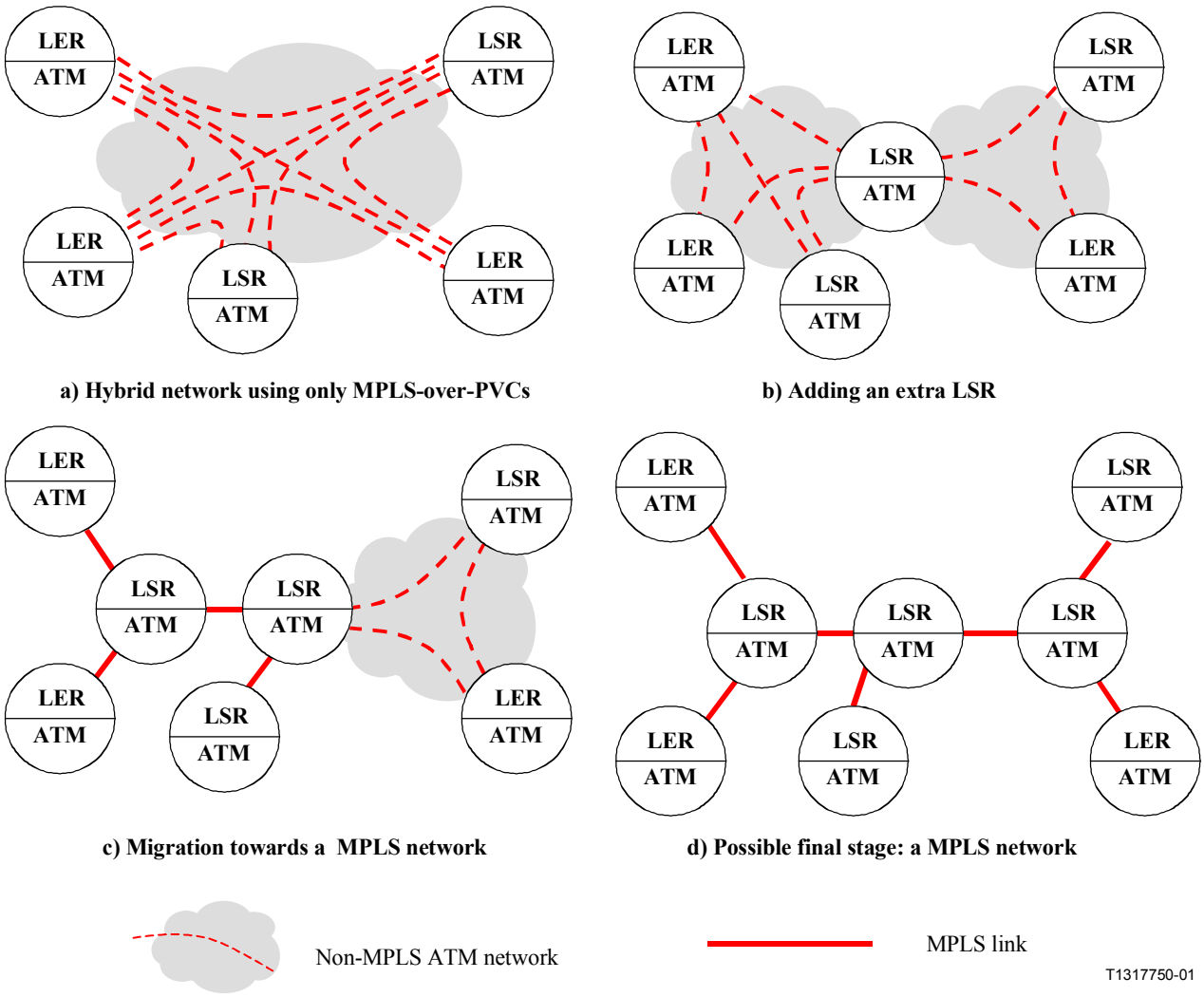


Figure III.2/Y.1310 – ATM MPLS networks using MPLS-over-PVCs

Some carriers might prefer to build an infrastructure for their MPLS traffic which is separate from their traditional ATM network. This MPLS network might use ATM MPLS. Alternatively, it might use packet-based MPLS, with packet-based LSRs and links such as PPP-over-SDH. Such a packet-based MPLS network might use MPLS-over-PVCs as a transitional stage, allowing a traditional ATM network to be used to carry MPLS traffic in the early stages of introduction of the packet-based MPLS network. In the course of growing this network, the MPLS-over-PVC links might be replaced with physical links. This possible future migration is shown in Figures III.2 c) and d).

III.3.2.2 Equipment for MPLS-over-PVCs

The core of an MPLS-over-PVCs network is a traditional ATM network which need only support PVCs or SPVCs. Virtually any ATM network can be used. The ATM edge LSRs should support the following:

- one or more ATM network interface cards;
- packet-based MPLS encapsulation over PVCs or SPVCs;
- traffic shaping to the PVC or SPVC parameters.

III.3.3 Networks using virtual trunks

- *Use of virtual trunks*

  A simple way of using virtual trunks is to use them to connect ATM edge LSRs without using any ATM-LSRs at all in the network, as illustrated in Figure III.3 a). This means that all MPLS packets are carried over virtual trunks, and no label switching actually occurs in the network. The ATM part of the network consists entirely of traditional ATM switches. More commonly, some switches in the ATM network will support the MPLS protocol stack, and some will not. Virtual trunks may be used to connect ATM-LSRs to ATM-LSRs, or to connect ATM edge LSRs to ATM-LSRs, as well as connecting ATM edge LSRs to each other. This is shown in Figure III.3 b).

- *Migration to full MPLS*

  Figures III.3 a), b) and c) show a possible migration process for introducing MPLS to a traditional ATM network:

  ATM edge LSRs are added around the edge of a traditional ATM network; alternatively, MPLS function may be added to existing routers. This enables MPLS VPNs as well as leading to the next steps.

  Next, MPLS function is added to some ATM switches, or extra ATM-LSRs are added to the network. This reduces the number of virtual trunks required, and starts to reduce some of the scalability problems of hybrid networks.

  More ATM-LSRs are added, which further reduces the number of virtual trunks, and starts to introduce native ATM MPLS links as shown in Figure III.3 c). This step naturally leads to the final one.

  Ultimately, all ATM switches are ATM-LSRs, and no virtual trunks are used at all. The full network runs ATM MPLS, and has none of the disadvantages of hybrid networks. This is illustrated in Figure III.3 d).

- *Other variations*

  LSRs and traditional ATM switches can be combined in many different ways. Figure III.4 shows some other hybrid network structures which may occur. Many other hybrid network structures are possible. An ATM MPLS network must include edge LSRs, but may use any nearly combination of zero or more ATM-LSRs and zero or more traditional ATM switches with virtual trunks.

- *Requirements to support virtual trunks*

  Virtual trunks are implemented using Permanent Virtual Paths (PVPs) or Soft Permanent Virtual Paths (SPVPs).

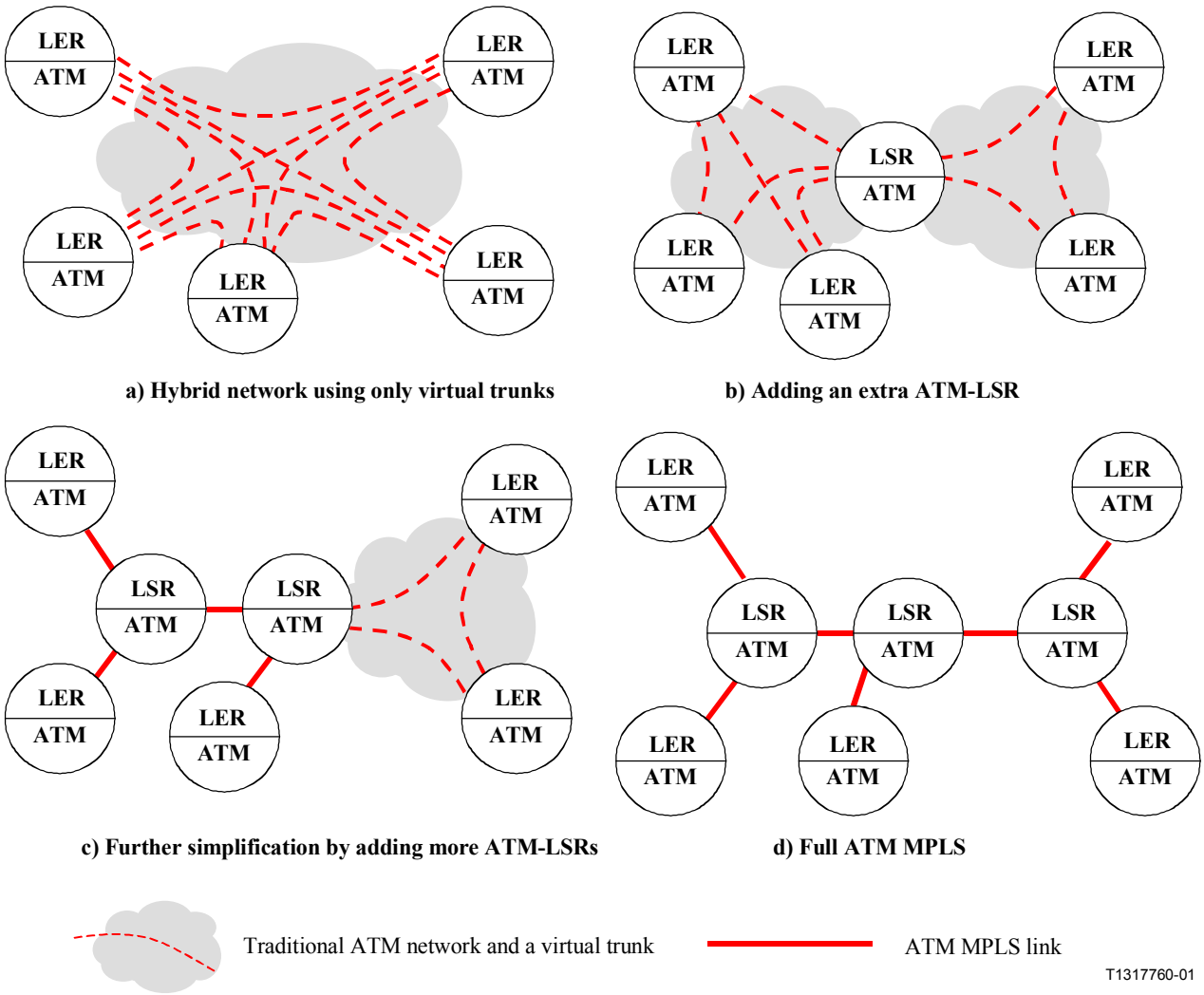


Figure III.3/Y.1310 – ATM MPLS networks using virtual trunks

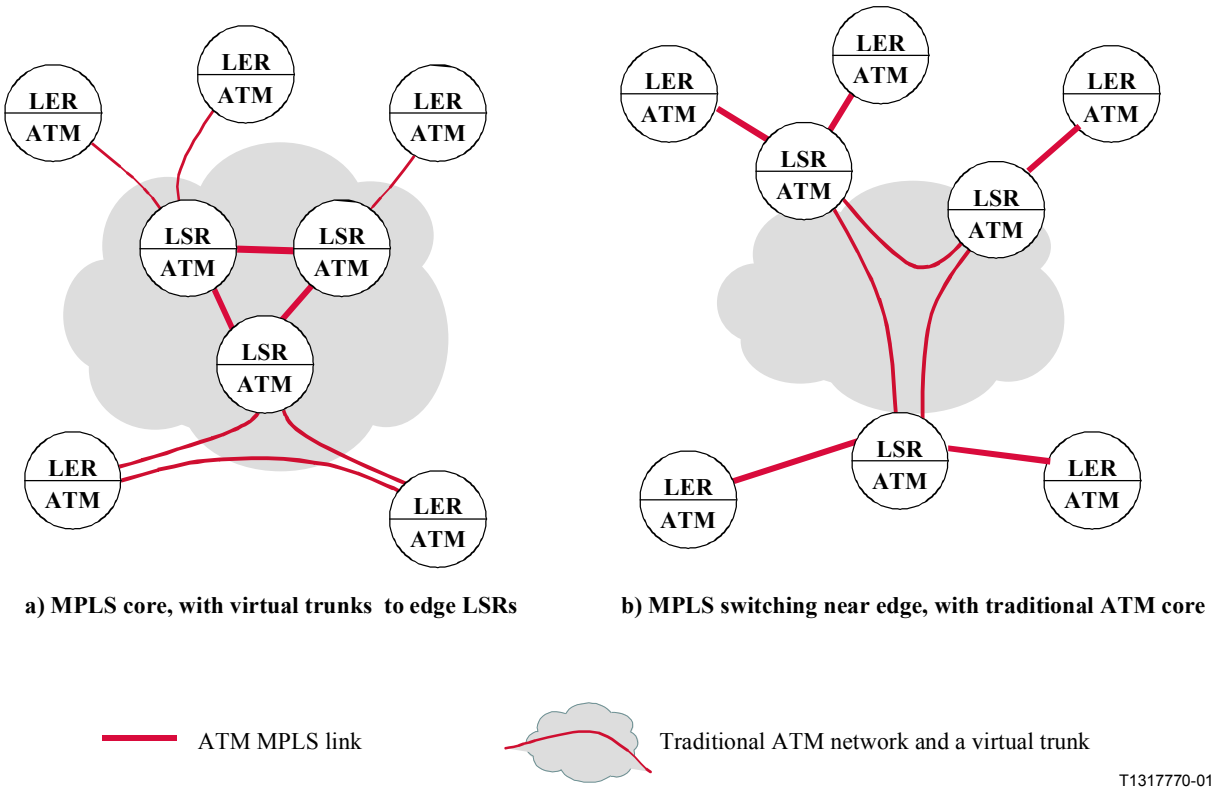


Figure III.4/Y.1310 – Other hybrid network examples using virtual trunks on ATM-LSRs

III.3.3.1 Supporting virtual trunks on traditional ATM switches

The switches in the traditional ATM networks must support PVP or SPVP connections with ATM Forum or ITU-T traffic management types which match those used on the edge LSRs. The switches are not required to support MPLS.

III.3.3.2 Supporting virtual trunks on edge LSRs

ATM edge LSRs must meet the following requirements in order to support virtual trunks:

- They must support one or more ATM network interface cards.
- If a particular virtual trunk uses a VPI x at the edge LSR, then the LDP signalling VC for the virtual trunk must be within x. It may have VPI = x, VCI = 32, instead of the normal default VPI = 0, VCI = 32 for LDP signalling [MPLS_ATM]. However, other VCI values can be configured through mutual bilateral agreements.

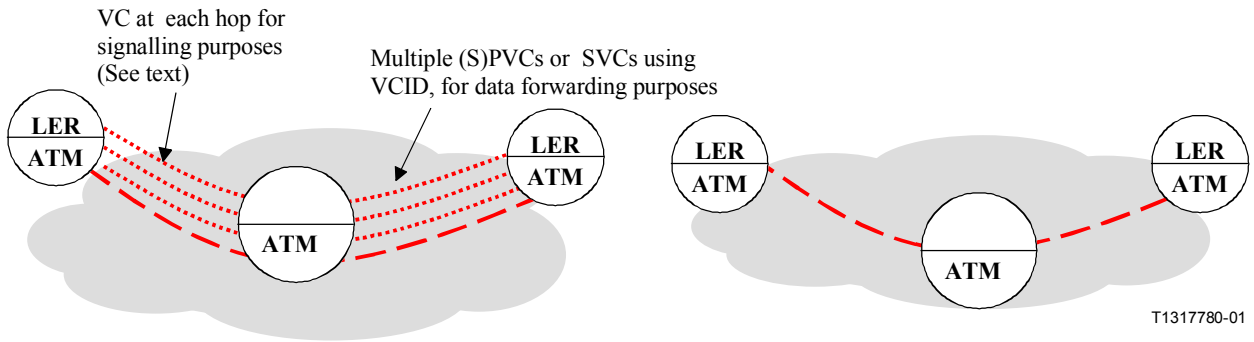
In order to support virtual trunks, ATM-LSRs must meet the same requirements as ATM edge LSRs:

- If a particular virtual trunk uses a VPI x at the ATM-LSR, then the LDP signalling VC for the virtual trunk must be within x. It may have VPI = x, VCI = 32, instead of the normal default VPI = 0, VCI = 32. However, other VCI values can be configured through mutual bilateral agreements.

III.3.4 Networks using VCID

III.3.4.1 The concept of a "Logical Link"

VCID uses multiple PVCs, SPVCs or SVCs to connect each pair of ATM MPLS devices across a traditional network [ATM_VCID]. Despite the differences between VCID and virtual trunks, VCID can be used in similar network configurations to virtual trunks. Figure III.1 illustrated this. In Figure III.5, a concept is introduced which allows VCID to be directly compared to virtual trunks.



a) Connections for VCID **b) Simplified representation as a VCID "logical link"**

Figure III.5/Y.1310 – Representing VCID connections as "logical links"

When connecting two ATM MPLS devices (ATM LSRs or ATM edge LSRs) with VCID, many PVCs, SPVCs or SVCs are required: one for signalling, and many for the MPLS labels. However, the group of PVCs, SPVCs or SVCs used by VCID between two ATM MPLS devices acts in place of a single ATM link in an ATM MPLS network. Consequently, it is useful to consider this group of PVCs, SPVCs or SVCs to be a single "logical link".

Figure III.3 showed how virtual trunks could be used when introducing MPLS to a traditional ATM network. VCID "logical links" can be used in an exactly analogous way, as shown in Figure III.6. The variant network structures shown in Figure III.4 also apply equally well with VCID.

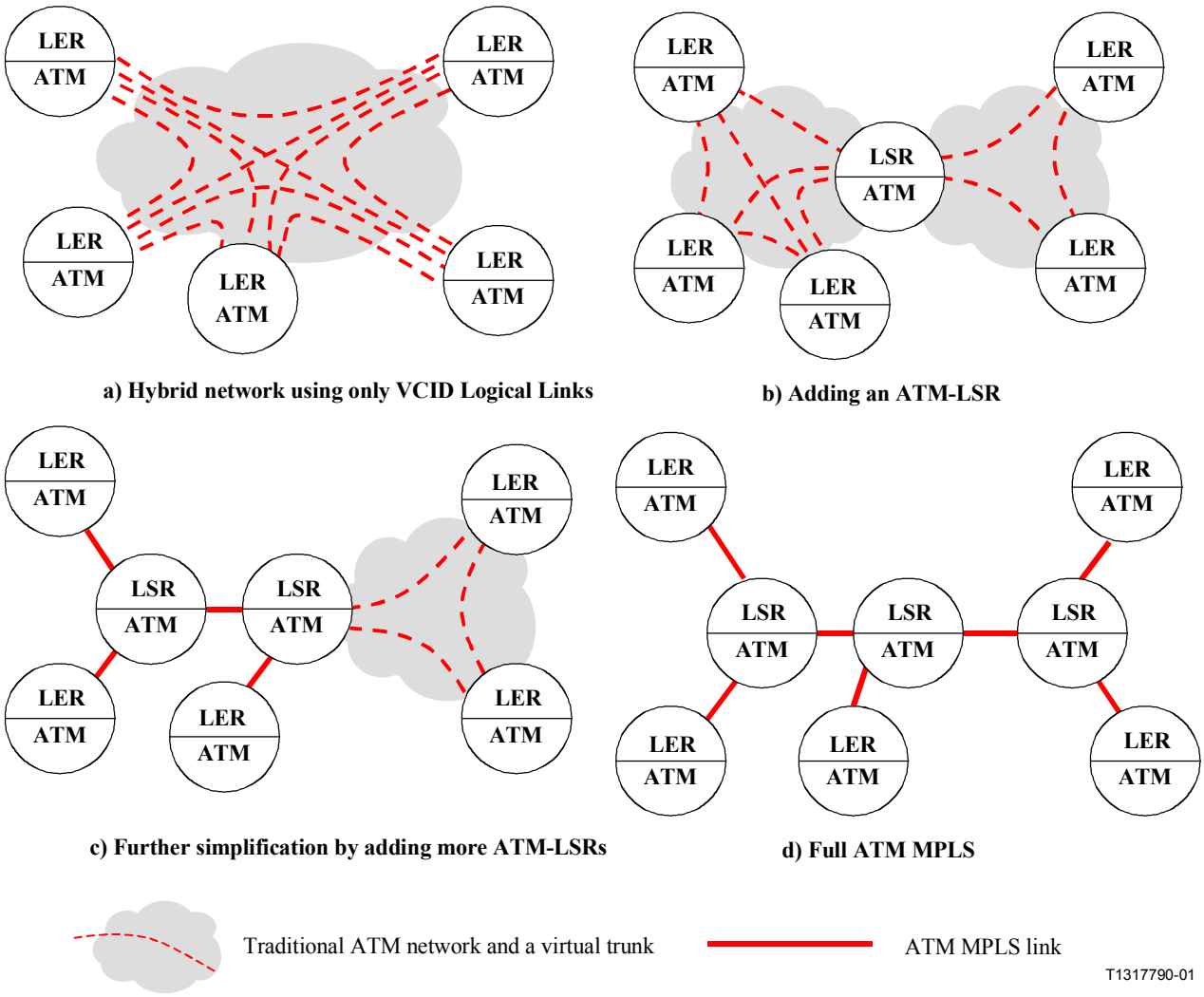


Figure III.6/Y.1310 – ATM MPLS networks using VCID "logical links"

III.3.4.2 Supporting VCID on traditional ATM switches

The switches in the traditional ATM networks must support PVC, SPVC or SVCs connections with ATM Forum or ITU-T traffic management types which match those used on the edge LSRs. They are not required to support VCID signalling or any MPLS functions.

III.3.4.3 Supporting VCID on ATM edge LSRs

The ATM edge LSRs must meet the following requirements:

- They must support one or more ATM network interface cards.
- They must support VCID in addition to the ATM MPLS protocols.

Appendix IV

Example methods for IP-VPN support in MPLS/ATM public network

IV.1 Introduction

This appendix describes example methods for using MPLS to provide IP Virtual Private Network services in a public network. MPLS provides a flexible and scalable basis for building IP-VPN services. Clause 7.2 defines the IP-VPN service and some requirements for the service.

[Y.1311.1] also provides requirements and architectural approach description for MPLS-based IP-VPN.

It is understood that Service Providers will make design decisions to support IP-VPNs based on their internal network and customer requirements. This appendix describes example methods and is not intended to constrain the deployment of VPNs inside a carrier's network.

Although the IP-VPN concept framed in the language of supporting enterprise customers by carriers, the same methods can be used by Service Providers to support other Service Providers (e.g., a carriers' carrier).

Figure IV.1 illustrates a generic scenario for IP-VPNs:

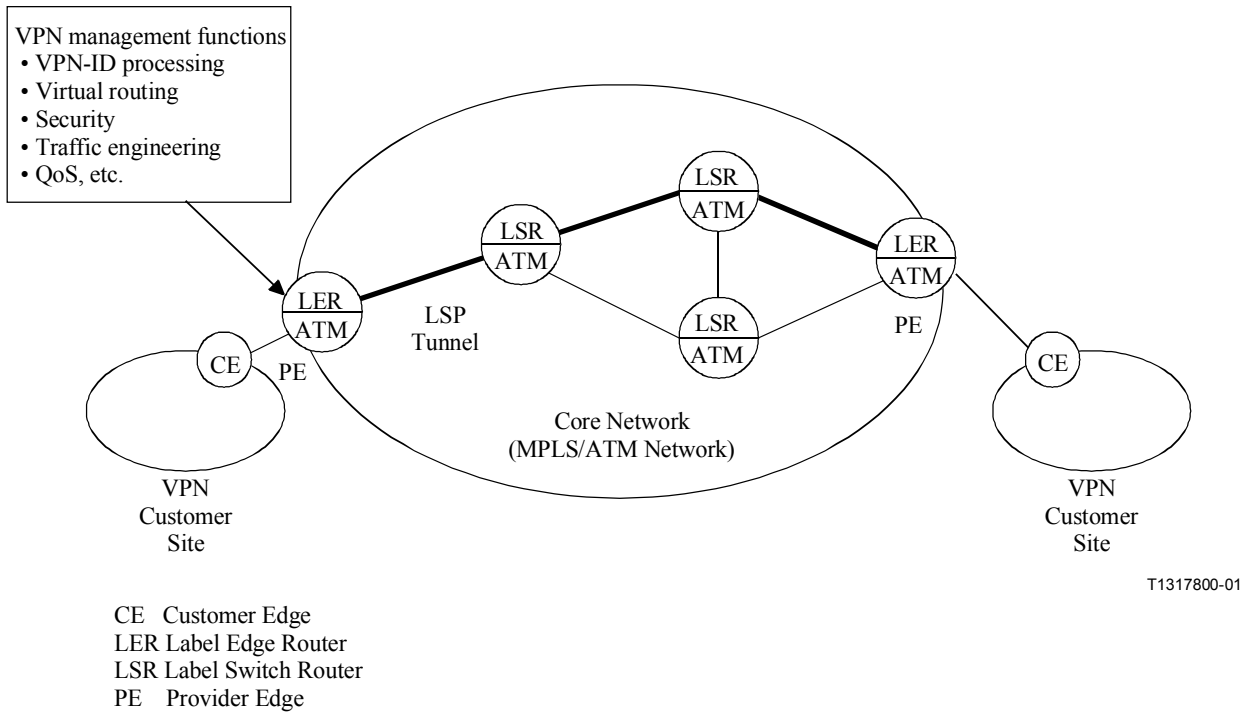


Figure IV.1/Y.1310 – Generic scenario for IP-VPNs using MPLS

The CE router is the customer edge router that interfaces a customer site with the service provider's network. The PE router is the service provider edge router that interfaces to customers' CEs.

A site is a set of (sub)networks that are part of the customer's network and is connected to the VPN through one or more PE/CE links. A VPN is a collection of sites sharing common routing information. A site can be part of different VPNs.

Figure IV.2 illustrates the case in which a Service Provider supports multiple VPNs. As shown, one site can belong to multiple VPNs. A site belonging to multiple VPNs may or may not provide transit between the two VPNs according to policy (how this is done is outside the scope of this Recommendation). If a site belongs to multiple VPNs, it must have an address space that is unique among the VPNs.

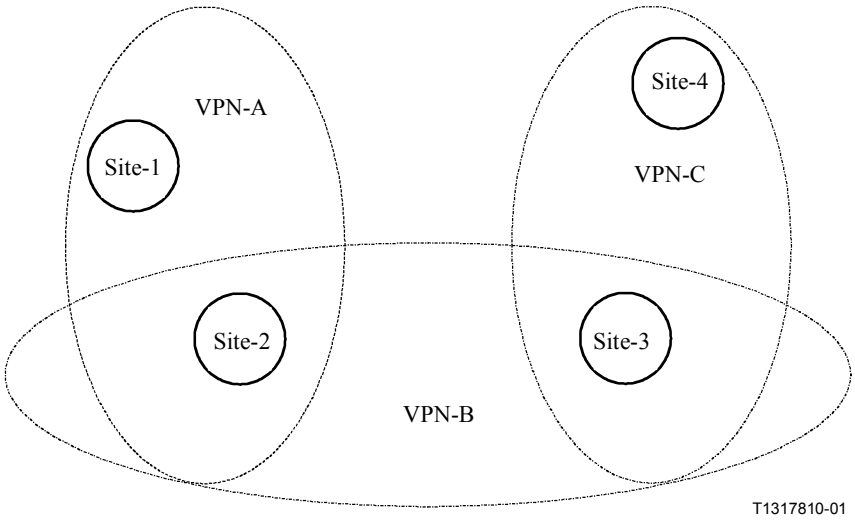


Figure IV.2/Y.1310 – Illustration of multiple VPNs

IV.2 Scenario 1

This clause describes an example method of using MPLS to provide IP-VPN services in a public network. The MPLS and its LDP provide a very flexible and powerful basis for building IP-VPN services. As a normal LDP operation, basic LSP setup is according to a topology-driven method. This is a base LSP setup using base label. In this case, two levels of LSP tunnelling (label stacking) for intra-VPN routing are used.

IV.2.1 Simple network configuration

IV.2.1.1 Architectural overview

Figure IV.3 illustrates an example of configuration composed of LER and LSR for IP-VPN services in MPLS/ATM core network.

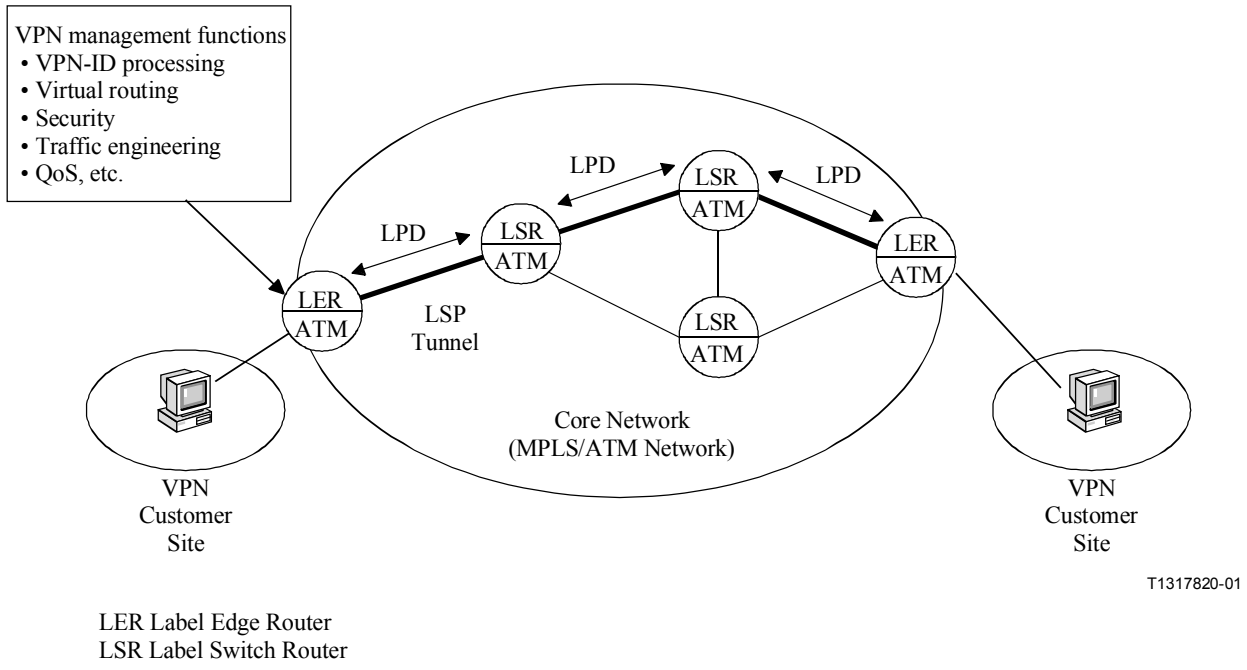


Figure IV.3/Y.1310 – Network model for IP-VPN support in MPLS/ATM public network

IV.2.2 Network components

IV.2.2.1 LER (Label Edge Router)

LER is a MPLS border router that is located at the edge of MPLS/ATM provider network. They serve as LSP tunnel ingress and egress points for IP traffic of VPN customers. If a LER is shared over many customers, it should perform virtual routing, which means a LER maintains separate forwarding tables for each VPN it serves, since their IP address spaces may not be distinct.

IV.2.2.2 LSR (Label Switched Router)

The MPLS/ATM core network is the provider's underlying network that is shared among customer IP-VPN services.

IV.2.2.3 Operations for establishing IP-VPN areas

The network provider wishing to offer IP-VPN service has first to configure the MPLS domain. An MPLS domain here means an IP-VPN area. An IP-VPN area consists of LERs and LSRs. As a normal LDP operation, basic LSP setup is according to a topology-driven method, which is defined as a base or level 1 LSP setup using base label. For intra-VPN routing, two levels of LSP tunneling (label stacking) are used.

IV.2.2.4 VPN membership discovery

Each LER discovers all other LERs in the VPN area that is serving the same IP-VPN. The LDP session initiation process is used as the method of LERs discovering their peers, since the ultimate intent of the scheme is to establish a second level of MPLS tunnels. Every LER sends an LDP hello message down every base network LSP that exits its LER. Hello messages are encapsulated with

the base MPLS label so that they are carried all the way to the destination LER. The LDP hello message is a form of query to determine if a LER for the same VPN (a peer) resides at the destination LER. When a hello adjacency is registered, the relevant LER proceeds to initiate an LDP session with its peer. One of the two LERs will initiate a TCP connection to the other. After the TCP connection is in place, and the necessary initiation messages have been exchanged, then an LDP session between the peer LERs exists. Immediately after the LDP session is established, each of the two LERs offers the other a label for a LSP tunnel to itself. If the LSP tunnel is a nested tunnel, its label is pushed onto a packets label stack before the base network LSP label.

IV.2.2.5 VPN membership and reachability information dissemination

The LER learns the IP address prefixes of the customer sites it is directly connected to as exchanging routing information. The LER needs to find its peer LERs. It has to discover which other LERs in the VPN area serve its VPN. The LER offers to establish a direct LDP session with every other LER in the VPN area. But only LERs serving the same VPN will discover each other, and go on to establish LDP sessions with each other. LDP sessions will only be successfully established between LERs that are supporting the same VPN.

IV.2.2.6 Intra VPN reachability

The first traffic that will flow over the nested tunnels is the exchange of routing information between LERs. It is assumed that when a LER is first configured for an IP-VPN, part of the configuration information is the routing protocol that it should use "intra VPN". It would also be given any security credentials that it needs in order to participate as a neighbour router to the other LERs. After any discovery phase of the "intra VPN" routing scheme, each LER will be advertising the VPN customer specific address prefixes reachable through it.

IV.2.2.7 IP packet forwarding

As a result of routing exchanges between LERs, each LER will build a forwarding table that relates VPN customer specific address prefixes (FEC: Forward Equivalency Classes) to the next hop. When IP packets arrive whose next hop is a LER, the forwarding process pushes first the label for the peer LER (the nested tunnel label). Then the base label, for the first hop of the base network LSP that leads to the LER, is pushed onto the packet. The doubly labelled packet is then forwarded to the next LSR in the base network LSP. When the packet arrives at the destination LER, the outermost label may have changed several times but the nested label has not changed. As the label stack is popped, the nested label is used to direct the packet to the correct LER. At a LER, the nested label space used by each VPN has to be disjoint from all other VPNs supported by the same LER.

IV.3 Scenario 2

This clause describes an example method of using MPLS and the Multiprotocol Border Gateway Protocol to provide IP-VPN services in a public network as defined in [BGP_VPN]. This clause provides an overview. Details can be found in [BGP_VPN].

IV.3.1 Architectural overview

Figure IV.1 illustrates an example of a configuration composed of LER and LSR for IP-VPN services in a MPLS/ATM core network.

Figure IV.4 illustrates the network model using [BGP_VPN].

IV.3.2 Network components

This clause introduces the network components for support of the IP-VPN and the terminology used.

IV.3.2.1 Provider Edge (PE) Router

The PE router is the service provider edge router that interfaces to customers' edge (CE) routers. For the purposes of this Recommendation, this router is an edge LSR (i.e., the interface between the customer and provider does not use MPLS).

IV.3.2.2 Customer Edge (CE) Router

The CE router is the customer edge router that interfaces to the service provider's edge (PE) routers. For the purposes of this scenario, the CE router does not implement MPLS and is an IP router. The CE does not have to support any VPN-specific routing protocols or signalling.

IV.3.2.3 Provider (P) Router

The P routers are the core Label Switch Routers.

IV.3.2.4 Site

A site is a set of (sub)networks that are part of the customer's network and is connected to the VPN through one or more PE/CE links. A site can be part of different VPNs.

IV.3.2.5 Route distinguisher

The provider assigns each VPN a unique identifier called a Route Distinguisher (RD) that is different for each Intranet or Extranet within the provider network. Forwarding tables in PE routers contain unique addresses, called VPN-IP addresses, constructed by concatenating the RD with the customer's IP addresses. VPN-IP addresses are unique for each endpoint in the service provider network, and entries are stored in forwarding tables for each node in the VPN (i.e., each PE router in the VPN).

IV.3.2.6 Connection model

Figure IV.4 below illustrates the connection model for the MPLS/BGP VPN.

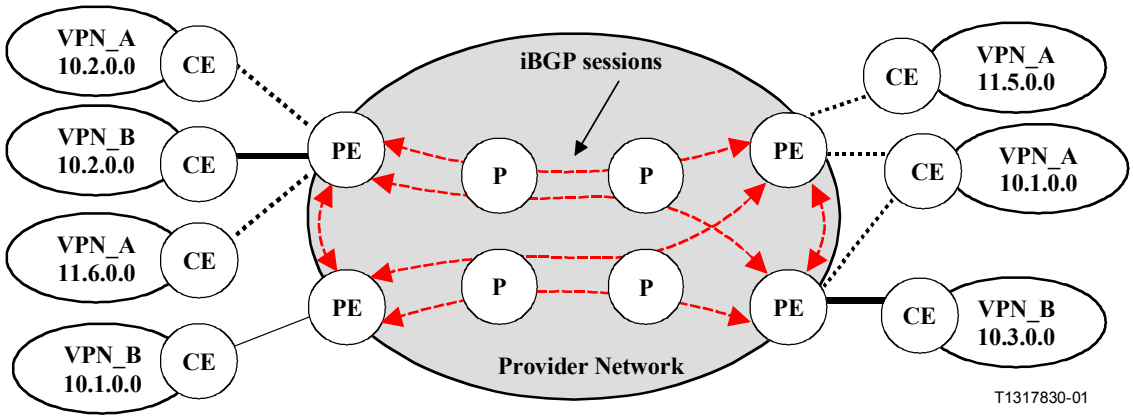


Figure IV.4/Y.1310 – Connection model for IP/VPNs using MPLS/BGP

P routers are in the core of the MPLS network. PE routers use MPLS to communicate with the core MPLS network and IP routing to communicate with the CE router. P and PE routers use an IP routing protocol (interior gateway protocol) for establishing IP routes through the MPLS core and LDP for label distribution between routers.

PE routers use multiprotocol BGP-4 to communicate with each other to exchange labels and policy for each VPN. The PEs are BGP fully meshed (unless a route reflector is used). Specifically, since the PEs are in the same Autonomous System, they use internal BGP (iBGP).

P routers do not run BGP and do not have any knowledge of the VPN. They use normal MPLS protocols and procedures.

PE routers may exchange IP routes with CE routers via an IP routing protocol. Static routes can also be used. Normal routing procedures are used between the CE and PE routers. The CE router does not have to implement MPLS or have any special knowledge of the VPN.

The PE routers distribute the customer routes to the other PE routers via iBGP. The VPN-IP address (constructed from the route distinguisher and IPv4 address) is used in BGP for distributing routes. Thus, different VPNs can use overlapping IPv4 address space without duplicate VPN-IP addresses.

The PE routers map the routes learned from BGP into their routing tables for forwarding packets received from the CE routers into the correct LSP.

Two levels of labels are used. The inside label is used to identify the correct VPN to the PE router. The outside label is used by the LSRs in the MPLS network to route the packets to the correct PE.

IV.3.2.7 Operations for establishing IP-VPN areas

The network provider wishing to offer IP-VPN service must design and provision the network according to the connectivity requirements.

Each PE must be configured for the VPNs which they must support and the VPNs to which each attached CE belongs. An iBGP peering relationship must be configured between PE routers in the MPLS network or a route reflector. Normal iBGP scaling capabilities can be used.

Normal routing protocol configuration must be done to communicate with the CE.

Normal MPLS configuration (LDP, IGP, etc.) must be performed to communicate with the MPLS core network.

The P routers should not have to be configured to support the VPN (beyond normal MPLS support).

IV.3.2.8 VPN membership and reachability information dissemination

The PE learns the IP address prefixes of the customer sites it is directly connected to by exchanging routing information via an IP routing protocol or via configuration (static routes).

The PE exchanges VPN-IP address prefixes with its BGP peers to learn routes to destination VPN sites. The PE also exchanges labels via BGP with its PE router peers in order to identify the LSP to use for connectivity between PE routers. These labels are used as second level labels and are not seen by the P routers.

PE routers maintain separate routing and forwarding tables for each VPN it supports. Each CE router attached to a PE router will use the appropriate routing table based on the interface to which it attaches.

IV.3.2.9 IP packet forwarding

As a result of routing exchanges between PEs, each PE will build a per-VPN forwarding table that relates VPN customer specific address prefixes to next hop PE routers.

When IP packets arrive from a CE router, the PE router will check the forwarding table for the VPN identified with that interface. If a match is found, the router will proceed as follows:

- If the next hop is a PE router, the forwarding process pushes first the label for the peer PE router (the nested tunnel label) identified by the routing table.
- The PE router then pushes the base label on to the packet, for the first hop of the base network LSP that leads to the destination PE router. The doubly labelled packet is then forwarded to the next LSR in the base network LSP.
- The P routers (LSRs) use the top-level labels and their routing tables to route the packet to the destination PE.

- When the packet arrives at the destination PE router the outermost label may have changed several times but the nested label has not changed.
- When the PE router receives a packet, it uses the embedded label to identify the VPN. At a PE router, the nested label space used by each VPN has to be disjoint from all other VPNs supported by the same PE router. The PE router checks the routing table associated with that VPN to determine over which interface to transmit the packet.

If a match is not found in the VPN routing table, the PE router checks the Internet routing table (if this capability is available from the provider) for routability. If no route is found, the packet is dropped.

VPN-IP forwarding tables contain labels that correspond to VPN-IP addresses. These labels route traffic to each site in a VPN. Because labels are used instead of IP addresses, customers can keep their private addressing schemes, within the corporate Internet, without requiring Network Address Translation (NAT) to pass traffic through the provider network. Traffic is separated between VPNs using a logically distinct forwarding table for each VPN. Based on the incoming interface, the switch selects a specific forwarding table, which lists only valid destinations in the VPN, thanks to BGP. To create Extranets, a provider explicitly configures reachability between VPNs. (NAT configurations may be required.).

IV.3.2.10 Security

Within the provider network, route distinguishers are associated with every packet by the PE router, so a user cannot "spoof" a flow or packet into another customer's VPN. Note that route distinguishers are not carried in user data packets. Users can participate in an Intranet or Extranet, only if they reside on the correct physical port and have the proper route distinguishers configured into the PE router. This setup makes virtually impossible to enter, and provides the same security levels users are accustomed to in a frame relay, leased-line, or ATM service.

BIBLIOGRAPHY

This bibliography includes explicit references embedded in the RFCs listed in 2.1.2. Embedded references which are already included as primary references are not repeated in this appendix.

[1] POSTEL (J.): DoD Standard – Internet Protocol, *RFC 760, USC/Information Sciences Institute*, January 1980.

[2] POSTEL (J.): DoD Standard – Transmission Control Protocol, *RFC 761, USC/Information Sciences Institute*, January 1980.

[3] POSTEL (J.): Internet Control Message Protocol – DARPA Internet Program Protocol Specification, *RFC 792, USC/Information Sciences Institute*, September 1981.

[4] POSTEL (J.): Service Mappings, *RFC 795, USC/Information Sciences Institute*, September 1981.

[5] POSTEL (J.): Address Mappings, *RFC 796, USC/Information Sciences Institute*, September 1981.

[6] BRADEN (R.): Requirements for Internet Hosts – Communication Layers, *STD 3, RFC 1122*, October 1989.

[7] RIVEST (R.): The MD5 Message-Digest Algorithm, *RFC 1321*, April 1992.

[8] ALMQUIST (P.): Type of Service in the Internet Protocol Suite, *RFC 1349*, July 1992.

[9] BRADLEY (T.), BROWN (C.), MALIS (A.): Multiprotocol Interconnect over Frame Relay, *RFC 1490*, July 1993.

[10] MOY (J.): OSPF Version 2, *RFC 1583, Proteon Inc*, March 1994.

[11] SIMPSON (W.): The Point-to-Point Protocol (PPP), *STD 51, RFC 1661*, July 1994.

[12] REYNOLDS (J.), POSTEL (J.): Assigned Numbers, *RFC 1700*, October 1994.

[13] REKHTER (Y.), LI (T.): A Border Gateway Protocol 4 (BGP-4), *RFC 1771*, March 1995.

[14] BAKER (F.), Editor: Requirements for IP Version 4 Routers, *RFC 1812*, June 1995.

[15] SCHULZRINNE (H.), CASNER (S.), FREDERICK (R.), JACOBSON (V.): RTP: A Transport Protocol for Real-Time Applications, *RFC 1889*, January 1996.

[16] McCANN (J.), MOGUL (J.), DEERING (S.): Path MTU Discovery for IP version 6, *RFC 1981*, August 1996.

[17] BRADNER (S.): Key words for use in RFCs to Indicate Requirement Levels, *BCP 14, RFC 2119*, March 1997.

[18] BRADEN (R.) et al.: Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, *RFC 2205*, September 1997.

[19] WROCLAWSKI (J.): The use of RSVP with IETF Integrated Services, *RFC 2210*, September 1997.

[20] WROCLAWSKI (J.): Specification of the Controlled-Load Network Element Service, *RFC 2211*, September 1997.

[21] SHENKER (S.), WROCLAWSKI (J.): General Characterization Parameters for Integrated Service Network Elements, *RFC 2215*, September 1997.

[22] SHENKER (S.), WROCLAWSKI (J.): Network Element Service Specification Template, *RFC 2216*, September 1997.

[23] HINDEN (R.), DEERING (S.): IP Version 6 Addressing Architecture, *RFC 2373*, July 1998.

[24] HEFFERNAN (A.): Protection of BGP Sessions via the TCP MD5 Signature Option, *RFC 2385*, August 1998.

[25] KENT (S.), ATKINSON (R.): Security Architecture for the Internet Protocol, *RFC 2401*, November 1998.

[26] KENT (S.), ATKINSON (R.): IP Authentication Header, *RFC 2402*, November 1998.

[27] KENT (S.), ATKINSON (R.): IP Encapsulating Security Protocol (ESP), *RFC 2406*, November 1998.

[28] NARTEN (T.), ALVESTRAND (H.): Guidelines for Writing an IANA Considerations Section in RFCs, *RFC 2434*, October 1998.

[29] DEERING (S.), HINDEN (R.): Internet Protocol, Version 6 (IPv6) Specification, *RFC 2460*, December 1998.

[30] CONTA (A.), DEERING (S.): Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification, *RFC 2463*, December 1998.

[31] AWDUCHE (D.) et al.: Requirements for Traffic Engineering Over MPLS, *RFC 2702*, September 1999.

[32] POSTEL (J.): Internet Name Server, *USC/Information Sciences Institute, IEN 116*, August 1979.

[33] SOLLINS (K.): The TFTP Protocol, *Massachusetts Institute of Technology, IEN 133*, January 1980.

[34] CERF (V.): The Catenet Model for Internetworking, *Information Processing Techniques Office, Defense Advanced Research Projects Agency, IEN 48*, July 1978.

[35] BBN Technical Report 1822, *Specification for the Interconnection of a Host and an IMP*, Bolt, Beranek, and Newman, Revised May 1978.

[36] SHOCH (J.): Inter-Network Naming, Addressing, and Routing, *COMPCON, IEEE Computer Society*, Fall 1978.

[37] SHOCH (J.): Packet Fragmentation in Inter-Network Protocols, *Computer Networks, Vol. 3*, *No. 1*, February 1979.

[38] STRAZISAR (V.): How to Build a Gateway, *IEN 109, Bolt, Beranek and Newman*, August 1979.

[39] CERF (V.), KAHN (R.): A Protocol for Packet Network Intercommunication, *IEEE Transactions on Communications*, Vol. COM-22, No. 5, pp. 637-648, May 1974.

[40] DALAL (Y.), SUNSHINE (C.): Connection Management in Transport Protocols, *Computer Networks*, Vol. 2, No. 6, pp. 454-473, December 1978.

[41] DEMERS (A.), KESHAV (S.), SHENKER (S.): Analysis and Simulation of a Fair Queueing Algorithm, in *Internetworking: Research and Experience*, Vol. 1, No. 1, pp. 3-26.

[42] ZHANG (L.): Virtual Clock: A New Traffic Control Algorithm for Packet Switching Networks, in *Proc. ACM SIGCOMM '90*, pp. 19-29.

[43] VERMA (D.), ZHANG (H.), FERRARI (D.): Guaranteeing Delay Jitter Bounds in Packet Switching Networks, in *Proc. Tricomm '91*.

[44] GEORGIADIS (L.), GUERIN (R.), PERIS (V.), SIVARAJAN (K.N.): Efficient Network QoS Provisioning Based on per Node Traffic Shaping, *IBM Research Report No. RC-20064*.

[45] GOYAL (P.), LAM (S.S.), VIN (H.M.): Determining End-to-End Delay Bounds in Heterogeneous Networks, *Proc. 5th Intl. Workshop on Network and Operating System Support for Digital Audio and Video*, April 1995.

[46] FLOYD (S.), JACOBSON (V.): Link-sharing and Resource Management Models for Packet Networks, *IEEE/ACM Transactions on Networking*, Vol. 3, No. 4, pp. 365-386, August 1995.

[47] SHREEDHAR (M.), VARGHESE (G.): Efficient Fair Queueing using Deficit Round Robin, *Proc. ACM SIGCOMM 95*, 1995.

[48] BENNETT (J.), ZHANG (Hui): Hierarchical Packet Fair. Queueing Algorithms, *Proc. ACM SIGCOMM 96*, August 1996.

[49] STILIADIS (D.), VARMA (A.): Rate-Proportional Servers: A Design Methodology for Fair Queueing Algorithms, *IEEE/ACM Trans. on Networking*, April 1998.

[50] CONTA (A.), DOOLAN (P.), MALIS (A.): Use of Label Switching on Frame Relay Networks, *RFC 3034*, January 2001.

[51] NIKOLAOU (N.), RIGOLIO (G.), CASACA (A.), CIULLI (N.), STASSINOPOULOS (G.): Integration of IP and ATM for QoS and Multimedia Support, *4th International Distributed Conference (ICD 1999)*, 22-23 September 1999, Madrid, Spain.

[52] IETF RFC 2208 (1997), *Resource ReSerVation Protocol (RSVP) – Version 1 Applicability Statement – Some Guidelines on Deployment*.

[53] IETF RFC 3260 (2002), *New Terminology and Clarifications for Diffserv*.

[54] IETF RFC 3496 (2003), *Protocol Extension for Support of Asynchronous Transfer Mode (ATM) Service Class-aware Multiprotocol Label Switching (MPLS) Traffic Engineering*.

[55] IETF RFC 2382 (1998), *A Framework for Integrated Services and RSVP over ATM*.

[56] IETF RFC 3215 (2002), *LDP State Machine*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	**Global information infrastructure, Internet protocol aspects and Next Generation Networks**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2004