Designing ATM MPLS Networks

Version 0.51. Cisco Internal Use Only.

Jeremy Lawrence, Multiservice Switching Business Unit

Note on Versions

This document is for Cisco Internal Use Only. An shorter Cisco Confidential version with some sensitive content removed is (or will soon be) available from "IP+ATM Solutions" page at http://wwwin.cisco.com/cmc/cc/cisco/mkt/wan/ipatm/index.htm. The Cisco Confidential version (but not this version) may be released to Service Provider customers under a Non-Disclosure Agreement.

Outline

Multiprotocol Label Switching (MPLS) extends the capabilities of IP routers and ATM Switches in several key ways:

- It fully integrates IP Routing control with ATM switches, providing native support of IP services such as IP Class of Service and IP Multicast on ATM switches as well as routers.
- It provides support for scalable and flexible IP Virtual Private Network (VPN) services on both switches and routers.
- It provides support for IP Traffic Engineering, which is the fine-grained adjustment of IP traffic flows according to the resources in a network.

This version of *Designing ATM MPLS Networks* concentrates on the fundamentals of ATM MPLS network design. These fundamentals apply to all ATM MPLS networks, irrespective of the services being offered on them:

- Choosing MPLS equipment
- IP+ATM concepts and equipment
- Designing and dimensioning ATM MPLS points of presence
- Dimensioning MPLS network links
- IP routing in ATM MPLS networks
- Dimensioning MPLS Label VC requirements
- Ongoing network design
- Migrating ATM networks to MPLS

Open Issues

Specific Comments & Errata

- 3.1: In Figure 4c it's not clear that the data flow is CPE -- MGX -- eLSR -- MPLS cloud. A revised version of the diagram would make this clearer.
- 3.4: The equations are either correct or "safe" over-estimates. However some slightly tighter limits apply to the edge LSRs in networks with switches which use unpaired cross-connects, namely the LS1010, 6400 and 8540 MSR. These should be mentioned.

- 3.x: Add note on this:
  "Actually, there is one thing - if the 7200 is acting as an LSC, then the maximum throughput is limited to full-duplex OC3, as that is currently the fastest TSC-BPX link. There is no point loading the box up with six OC3 interfaces, unless you actually want five of them to run at 20% utilization. Unless you're really loading up the box with _heavy_ edge functions (e.g. CAR + MPLS VPN Edge + WRED) in addition to LSC function, there won't be much point having an NPE 300 in it."
- 4.1: Either here or in an earlier section, it might be useful to illustrate the different MPLS label encapsulations (e.g. shim, VPI/VCI).
- 4.5: Add a diagram showing network with ordinary MPLS links plus IMA links
- 4.5: Add a diagram showing where BNI cards may be used in a network

Desirable Minor Changes and Additions

- Add more details on redundancy options and configuration, including dual-controller cooperative redundancy.
- Add more detail on MPLS-over-Frame Relay
- Add more detail on dimensioning the number of edge routers and RPMs, with reference to benchmarking of MPLS Edge + CAR + WFQ, etc, is done. Include example calculations.
- LDP modes: some discussion on upstream, downstream, conservative, optimistic, etc modes would be useful.
- 3.x: Expand BPX, MGX issues coverage:
  - LVC allocation
  - "Port" and "trunk" issues, including feeders. (Material from recent training slides.)
  - Bandwidth, and what happens when zero bandwidth is assigned to MPLS
  - Dimensioning for LSC, mentioning the 1xOC3 throughput limit

Desirable Major New Content

- VPNs, with pointer to Packet Magazine article as well as more detail on network-level design issues. For background, refer to the VPN design guide & RFC 2457 for more details.
- MPLS Quality of Service:
  - MPLS CoS configuration
  - MPLS CoS Philosophy & network configuration steps: refer to white paper
  - IP+ATM aspects (service template tradeoffs in BPX)
- IP Traffic Engineering
- Tools for managing IP VPN networks.

Table of Contents

1. Introduction 4
  1.1 MPLS Fundamentals 4
  1.2 Note on Tag Switching and Terminology 4
  1.3 MPLS Network Structure 5
  1.4 Applications for MPLS 6
  1.5 MPLS and Other IP-over-ATM Schemes 7
  1.6 Steps in Designing MPLS Networks 7
2. Choosing MPLS Equipment 9
  2.1 Structures for MPLS Networks 9
  2.2 Choosing Cisco MPLS Equipment 12
  2.3 IP+ATM 14
3. Designing MPLS Networks 17
  3.1 Points of Presence Structures 17
  3.2 Dimensioning an MPLS Network's Links 20
  3.3 IP Routing in an MPLS Network 27
  3.4 Dimensioning MPLS Label VC space 30
  3.5 On-going Network Design 38
4. Migration of MPLS into Traditional ATM Networks 40
  4.1 Hybrid ATM Networks 40
  4.2 Networks Using MPLS-over-PVCs 41
  4.3 Networks Using VP Tunnels 43
  4.4 Requirements to Support VP Tunnels 44
  4.5 Examples of Hybrid ATM Network Equipment 52
  4.6 VP Tunnels and LVC Usage 53

1. Introduction

1.1 MPLS Fundamentals

The essentials of Multiprotocol Label Switching (MPLS) operation have been described in other white papers and presentations. Some important points will be emphasized here, but for more background see the following:

- The “IP+ATM Solutions” page at http://wwwin.cisco.com/cmc/cc/cisco/mkt/wan/ipatm/index.htm is a good resource. See particularly the “MPLS and IP Quality of Service in Service Provider ATM Networks” white paper under “Internal Information”.
- The MSSBU SE Toolbox http://wwwin.cisco.com/WANBU/Tech/index.html contains much information. The links under Technical Information -> IP+ATM will be helpful http://wwwin.cisco.com/WANBU/Tech/ipatm/index.html
- Online MPLS video training is at http://showtime.cisco.com/index.html
- The “MPLS VPN Configuration and Design Guide” is at http://wwwin.cisco.com/WWSales/SE/design/uth/q100/index.html
- The OSPF version 2 specification is http://www.ietf.org/rfc/rfc2328.txt
- The “IS-IS for Routing in TCP/IP and Dual Environments” specification is http://www.ietf.org/rfc/rfc1195.txt
- IETF documents on MPLS are at http://www.ietf.org/html.charters/mpls-charter.html. The most important documents are:
  - “MPLS Architecture”, draft-ietf-mpls-arch-05.txt
  - “MPLS Label Stack Encodings”, draft-ietf-mpls-label-encaps-04.txt
  - “MPLS using LDP and ATM VC Switching”, draft-ietf-mpls-atm-02.txt
  - “LDP Specification”, draft-ietf-mpls-ldp-05.txt
  - “MPLS Support of Differentiated Services by ATM LSRs and Frame Relay LSRs,” draft-ietf-mpls-diff-ext-01.txt
- Other IETF documents on Differentiated Services are at http://www.ietf.org/html.charters/diffserv-charter.html
- The most important IETF documents on the Border Gateway Protocol are:
  - “A Border Gateway Protocol 4 (BGP-4),” http://www.ietf.org/rfc/rfc1771.txt
  - “Multiprotocol Extensions for BGP-4,” http://www.ietf.org/rfc/rfc2283.txt

  A further informational document shows how BGP can be used to support VPNs:
  - “BGP/MPLS VPNs,” RFC 2457, http://www.ietf.org/rfc/rfc2547.txt
- The following books on routing, MPLS and related topics are very useful:
  - Halabi, B., “Internet Routing Architectures, “Cisco Press, 1997.
  - Metz, C., “IP Switching Protocols and Architectures”, McGraw-Hill, 1999
  - Rekhter, et al., “Switching in IP Networks”, Morgan Kaufmann, 1998
- Useful magazine articles are:
  - Feldman, et al., “Evolution of Multiprotocol Label Switching”, IEEE Communications Magazine, Vol. 36, No. 5, May 1998
  - Metz, C., “Ingredients for Better Routing: Read the Label”, IEEE Internet Computing, Sept/Oct. 1998
- Some archives on MPLS and related technologies are:
  - http://infonet.aist-nara.ac.jp/member/nori-d/mlr/
  - http://dcn.soongsil.ac.kr/~jinsuh/home-mpls.html

(Thanks to Chris Metz for pointing out many of these resources.)

The rest of this design guide assumes that the reader has knowledge of MPLS fundamentals.

1.2 Note on Tag Switching and Terminology

MPLS is a standardized version of Cisco’s original Tag Switching proposal. MPLS and Tag Switching are identical in principle, and nearly identical in operation. This document uses MPLS terminology rather than Tag Switching terminology, as shown in Table 1. One exception is the term “Tag Distribution Protocol” (TDP). TDP and the MPLS Label Distribution Protocol (LDP) are nearly identical in general function, but use different message formats and some different procedures. The term “TDP” will be used in this design guide only when it is important to distinguish TDP from LDP. Otherwise, any reference to “LDP” in this design guide also applies to TDP.

Note in Table 1 that the term “Label Edge Router” is not used. This is because the equivalent term “Edge LSR” is used, and having two different terms meaning the same thing will lead to confusion. “Edge LSR” is technically the more correct term[1].

Some other terms used are as follows:

- “ATM MPLS” is the form of MPLS that runs in networks with ATM switches that do MPLS switching. More specifically, it is the form of MPLS where each different label on a link is represented by a different VC.
- “Packet-based MPLS” means the form of MPLS that runs in networks which do not use ATM MPLS. More specifically, it is the form of MPLS where labels are carried as an extra header attached to each packet. Packet-based MPLS is also known as “Non-ATM MPLS”, “Frame-based MPLS” and “Router-based MPLS”. The term “Frame-based MPLS” is not used in this document, as it seems to imply “Frame Relay,” but Packet-based MPLS does not necessarily have anything to do with Frame Relay.
- A “Packet-based LSR” is a device which manipulates whole packets rather than cells. A router running packet-based MPLS is a packet-based LSR. An ATM Edge LSR is also a type of packet-based LSR.
- “Traditional ATM” switches and networks are those which do not use ATM MPLS. Traditional ATM networks may support Packet-based MPLS traffic within Permanent Virtual Circuits (PVCs). A traditional ATM switch can support ATM MPLS within a Permanent Virtual Path (PVP) which acts a ‘Virtual Trunk’. In any case, the traditional ATM switches do not actually perform Multiprotocol Label Switching—they merely support ‘tunnels’ through which MPLS packets are carried. This is discussed further in “4. Migration of MPLS into Traditional ATM Networks” on page 40.

Table 1 Tag Switching terminology equivalents

Old Tag Switching Term	MPLS Term
Tag Switching	MPLS; Multiprotocol Label Switching
Tag Switch/ed	Label Switch/ed
Tag (short for “Tag Switching”)	MPLS
Tag (thing applied to a packet)	Label
Tag Core Router; TCR	Label Switch Router; LSR
Tag Distribution Protocol; TDP	Label Distribution Protocol; LDP (but see text)
Tag Edge Router; TER	Edge Label Switch Router; Edge LSR (see text)
Tag Switch Controller; TSC	Label Switch Controller; LSC
Tag Switch/ing Router; TSR	Label Switch Router; LSR
Tag VC; TVC	Label VC; LVC
ATM-TSR	ATM-LSR
TFIB	LFIB
Tag Switched Path; TSP	Label Switched Path; LSP

1.3 MPLS Network Structure

A typical structure for MPLS networks in providers, i.e. carriers or ISPs, is shown in Figure 1. An MPLS network consists of Edge Label Switch Routers (Edge LSRs) around a core of Label Switch Routers (LSRs). Customer sites are connected to the provider MPLS network. Figure 1 shows 9 customer sites and 6 Edge LSRs, but more typically there will be several hundred customer sites per edge LSR. The Customer Premises Equipment (CPE) runs ordinary IP forwarding. It typically will not run MPLS. If the CPE does run MPLS, it will typically use it independently of the provider. It is important to note that the Edge LSRs are part of the

---

[1] In packet-based MPLS, particularly in Cisco's implementation, there is nothing special about Edge LSR function compared to LSR function. An Edge LSR is just a router-based LSR which happens to have some ordinary IP interfaces. This is often true in ATM MPLS networks too. The BPX 8650, for example, has Edge LSR capability because of two things: the Label Switch Controller (LSC) runs IP routing and LDP, and the LSC is based on a 7200 or 7500 router, which can support ordinary ATM interfaces.

provider network, and are controlled by the provider. The edge LSRs are critical to network operation, and are not intended to be CPE under any circumstances. The provider might locate and manage routers on at the customer sites, but these will be running ordinary IP and will be outside the MPLS network itself[2].

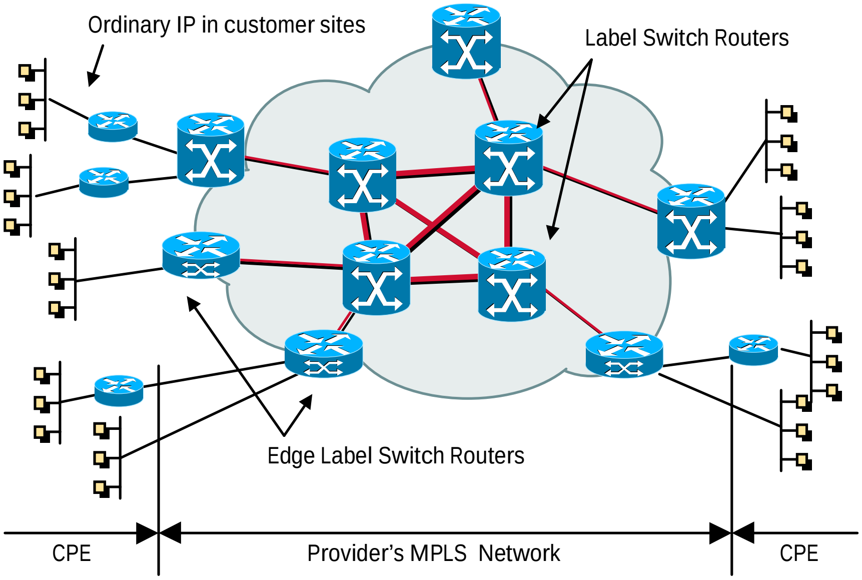


Figure 1 Typical MPLS network structure. Note that there will typically be many more customer sites than shown in this diagram.

1.4 Applications for MPLS

MPLS networks as shown in Figure 1 have three main applications. Typically, two or all three of these capabilities would be used simultaneously:

- IP+ATM integration. Because 'label switching' can be done by ATM switches, MPLS is a way of integrating IP services directly on to ATM switches. This involves putting IP routing and LDP software directly on ATM switches. Because it fully integrates IP onto ATM switches, MPLS allows ATM switches to optimally support IP services such as IP multicast, IP class of service, RSVP and Virtual Private Networks (see below). Optimal integration of IP+ATM means that MPLS does not have the scalability and complexity disadvantages of overlay schemes like MPOA, CSI, and IP Navigator.
- IP Virtual Private Network (VPN) services. VPNs form the infrastructure for Intranets and Extranets, which are IP networks on which corporations will base their whole business structures. A VPN service is a managed Intranet or Extranet service offered by a provider to many corporate customers. MPLS, in combination with the Border Gateway Protocol (BGP), allows one provider network to support thousands of customers' VPNs. So, MPLS with BGP offers a very flexible, scalable, and manageable way of providing VPN services on both ATM and packet-based equipment. Even on quite small providers' networks, the flexibility and manageability of MPLS+BGP VPN services are a major benefit.
- IP Explicit Routing and Traffic Engineering. An important problem in current IP networks is the lack of ability to finely adjust IP traffic flows to make best use of available network bandwidth. Also absent are related capabilities to send selected flows down selected paths, e.g. to select protected trunks for particular classes of traffic. MPLS uses Label Switched Paths (LSPs), which are a type of 'lightweight VC' which can be set up on both ATM and packet-based equipment. The IP traffic engineering capability of MPLS uses specially set up LSPs to finely adjust IP traffic flows.

---

[2] An edge LSR might be Customer Located Equipment, e.g. in a building basement, but only in circumstances where the location is secure and accessible only by the provider.

This design guide concentrates on MPLS in ATM networks, although some of the content also applies to packet-based networks. This version concentrates on fundamentals of MPLS network design which apply to all ATM MPLS networks, including those supporting VPNs and traffic engineering.

1.5 MPLS and Other IP-over-ATM Schemes

In ATM networks, MPLS allows ATM switches to directly support IP services, giving maximum efficiency compared to other approaches. Traditional IP over ATM connects routers over Permanent Virtual Circuits (PVCs). Multiprotocol over ATM (MPOA) and similar proprietary approaches carry IP traffic over Switched Virtual Circuits (SVCs). Traditional IP over ATM, MPOA, and proprietary approaches all have similar disadvantages:

- It is difficult to offer some types of IP services on the networks. For example, IP Class of Service cannot be offered natively by traditional ATM switches, and must be offered by translation to quite different ATM Forum Quality of Service concepts.
- Where IP services are offered, they are difficult to administer. Two levels of routing must be administered: IP routing (via OSPF or EIGRP or similar) and PNNI or similar routing for ATM. MPOA requires additional administration. Service translations, for example IP Class of Service to ATM Quality of Service, also require administration.
- IP services may be quite inefficient over ATM networks. For example, IP Multicast over ATM networks is difficult to achieve on a large scale due to the interaction of multicast routing, multicast group membership processing and ATM VC maintenance.
- There may be scaling limitations and/or dangerous interactions between IP routing (OSPF, etc.) and the ATM network, leading to unstable networks. Traditional IP over ATM can lead to storms of IP routing updates and subsequent network meltdown, if more than about 30 OSPF routers are connected in a full mesh over PVCs[3]. MPOA is unsafe when connecting routers to each other, and is intended only to connect hosts to routers or hosts to hosts[4].
- IP services require a substantial implementation and management effort. For example, an MPOA implementation requires PNNI, SVC signalling, ATM ARP, an ATMARP server, NHRP, and a NHRP server, in addition to AAL5, IP routing (OSPF, etc.) and an IPv4 stack.

MPLS in ATM networks avoids all of these disadvantages.

1.6 Steps in Designing MPLS Networks

The process of designing an MPLS network involves the following steps, which are described in the next sections. These will not necessarily be carried out in the following order, but they all need to be covered.

- Choose equipment types
- Design points of presence
- Design the backbone
- Design the IP routing
- Check MPLS-specific dimensioning issues

---

[3] If N number of routers are running OSPF and are connected in a full mesh over ATM PVCs, a single physical ATM link failure may result in ATM-layer re-routing of a large number of PVCs. If this takes too long, or if the ATM network can't re-route PVCs at all, this results in the effective failure of a large number of PVCs. The number of PVCs involved may be of the same order magnitude as N, and even N^2 in some cases. In any case, it is likely to be seen by $O(N)$ routers, where "$O(N)$" means "a number proportional to N". So, a single ATM link failure will cause each of $O(N)$ routers to send a link state advertisement (LSA) of size (at least) $O(N)$ to (N-1) neighbors. So, a single event in the ATM network results in $O(N^3)$ to $O(N^4)$ traffic. When a router receives an LSA, it must immediately re-calculate its routing table, as it must not forward packets based on old routing information. The processor load caused by a such a "storm" of routing updates may cause the routers to drop or not send keepalive packets, which appears to the neighboring routers as further link failures. These lead to further LSAs being sent, which perpetuates the problem. The net result is that a full mesh network can go persistently unstable after a single network event.

This is a fundamental problem, caused by the fact that the routers do not see the state of the ATM links and switches directly. IS-IS has somewhat better performance than OSPF in full mesh conditions because IS-IS has more sophisticated flooding capabilities (these capabilities, specifically the ability to pace flooding and block flooding on some interfaces, are also becoming available on OSPF). However this does not address the underlying problem. The solution to this problem is to let IP routing directly see the state of ATM links, which is what is done by ATM MPLS.

[4] This is not a simple issue to solve. It is a fundamental issue arising from conflict between routing protocols: PNNI routing at the ATM layer can make decisions which conflict with OSPF or similar routing at the IP layer. These conflicting decisions can lead to persistent loops. (See the NHRP Protocol Applicability Statement, RFC2333, for more discussion on this.) There has been further investigation on router-to-router NHRP at the IETF, but this revealed that router-to-router NHRP was not practical. The only reliable solution to this problem is to use the same routing protocol at the IP layer and ATM layer. This is exactly what MPLS does in ATM networks.

Class-of-service, MPLS VPNs, traffic engineering and other IP services will introduce additional design steps. Many of these are outlined in the "MPLS and IP Quality of Service in Service Provider ATM Networks" white paper. (See http://wwwin.cisco.com/cmc/cc/cisco/mkt/wan/ipatm/index.htm) See also the "MPLS VPN Design & Configuration guide" http://wwwin.cisco.com/WWSales/SE/design/uth/q100/index.html

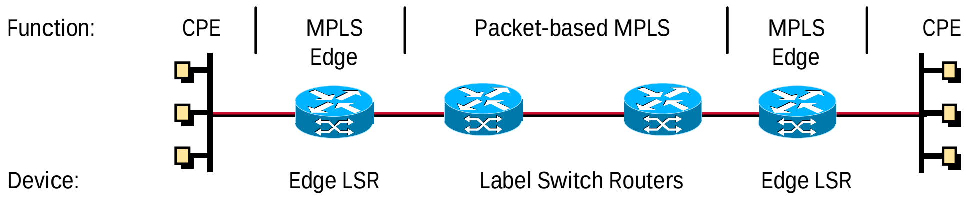
2. Choosing MPLS Equipment

2.1 Structures for MPLS Networks

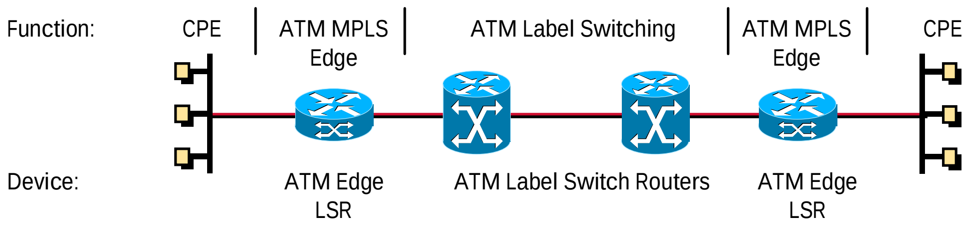
MPLS networks have Edge label switch router (LSR) and core LSR functions as shown in Figure 1. However these functions can be implemented on various types of equipment, and combined with access equipment there are various ways. These are shown in Figure 2.

Simple Packet-based MPLS

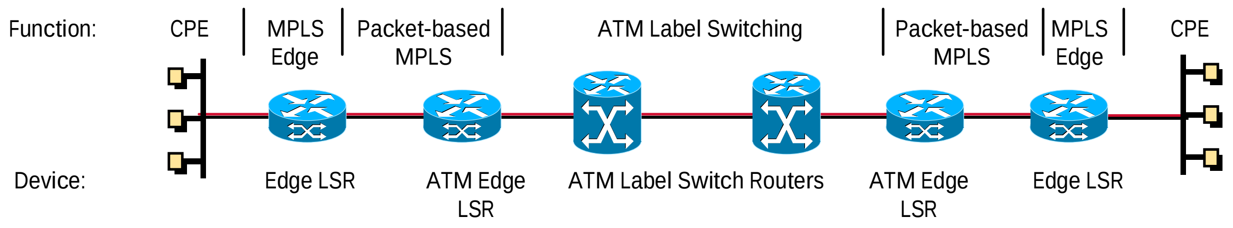
The simplest MPLS network structure is shown in Figure 2(a). This structure applies to router-only networks, which might use MPLS for supporting VPN services or IP Traffic Engineering. In this structure, customer sites are connected directly to router-based edge LSRs. The edge LSRs are connected to other LSRs, which are also based on router platforms. The routers are interconnected



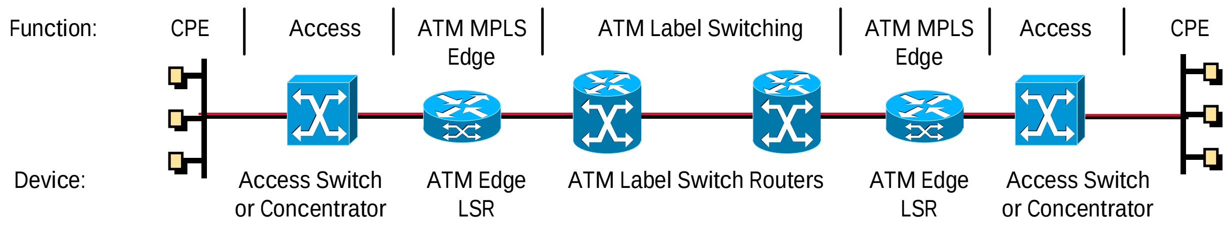
a. Simple packet-based MPLS



b. ATM MPLS with router-based Edge LSRs



c. Mixed ATM and packet-based MPLS



d. ATM MPLS with access concentrators

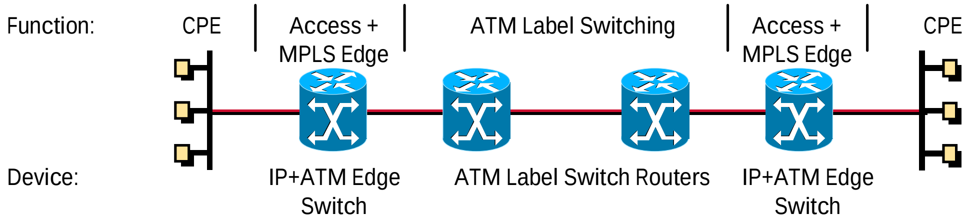
Figure 2 Devices in MPLS networks

by virtually any sort of link: serial, ethernet, packet-over-SONET, etc., and packets are sent, with MPLS headers, over these links. The routers involved will typically be 7200, 7500, or 12000-series Gigabit Switch Routers. Mid-range routers (3600 and 4700 series) might be used in lower-bandwidth applications.

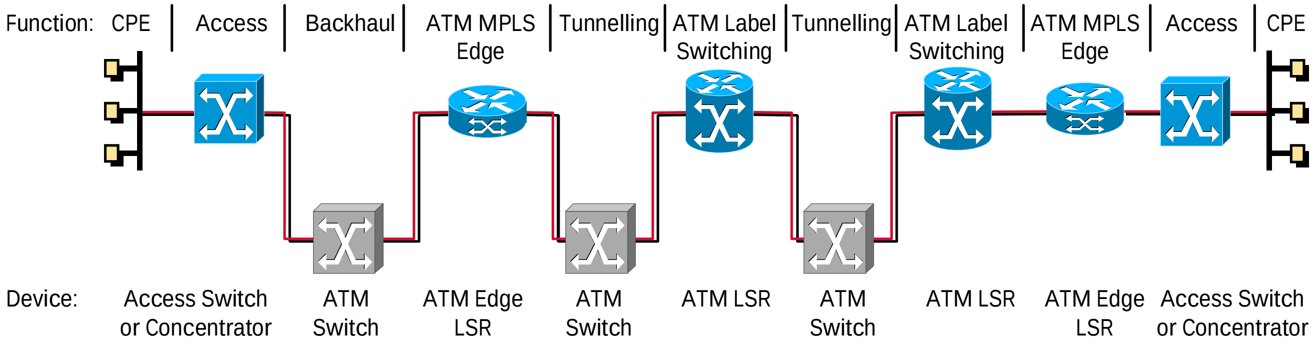
A variant of this structure is where the point-to-point links between the routers are actually ATM PVCs. These may be used during migration to ATM MPLS. This is discussed further in “4. Migration of MPLS into Traditional ATM Networks” on page 40.

ATM MPLS With Router-based Edge LSRs

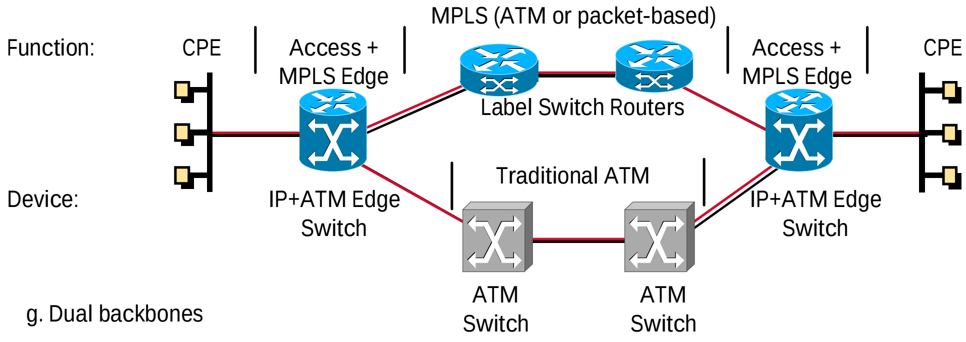
The simplest ATM MPLS network structure is shown in Figure 2(b). As with the previous case, customer sites are connected directly to router-based edge LSRs, typically 7200 or 7500 series routers. The Edge LSRs are connected by ATM links to the core devices which are ATM-LSRs. The ATM LSRs may be BPX 8650 IP+ATM Switches, LS1010, 8500 MSR, and later other ATM switches such as the MGX 8800 with PXM-45.



e. ATM MPLS with integrated ATM edge device



f. ATM MPLS with backhaul & tunnelling



g. Dual backbones

Figure 2   Devices in MPLS networks (continued)

The ATM switches carry packets with ATM MPLS labels; this means that, on each ATM link, there is a different MPLS Label VC (LVC) for each label.

Mixed ATM and Packet-based MPLS

It is possible to have a network with a mixture of ATM MPLS and packet-based MPLS. A simple example of this is shown in Figure 2(c). In a network such as this, some links run packet-based MPLS, and some links run ATM MPLS. The devices which interface between packet-based MPLS and ATM MPLS are the same routers which act as ATM Edge LSRs: anything from a3600 up to a 12000.

ATM MPLS With Separate Access Devices

ATM MPLS networks with router-based Edge LSRs may also use separate access devices, as shown in Figure 2(d). This will occur when access is required through a device which does not support MPLS services. There three common situations where this will be required:

- Access is required to both IP services, and ATM PVC services, through an access device which does not support MPLS. The most common example of this is the MGX 8220.
- The access device does not yet have software which supports MPLS. This will occur with the AS 5x00, uBR 7246 and IGX 8400 until they have MPLS support.
- Higher densities of low-bandwidth access lines can be supported by way of a separate access device, than simply by using an edge LSR.

Customer traffic is carried through the access device to the edge LSR. Between the access device and the edge LSR, there is a different logical link for each customer. This may be a Frame Relay or ATM PVC, or a PPP link.

ATM MPLS With Integrated IP+ATM Access Devices

The previous type of network can be simplified if the access device supports edge LSR function as well as Frame Relay, ATM, or other access services. This is shown in Figure 2(e). In the case of IP+ATM edge switches, a single device gives access to both MPLS services, and PVC or SVC services. IP+ATM edge switches include the BPX 8680, MGX 8850, 6400 universal access concentrator, and later the IGX 8450.

ATM MPLS Using Traditional ATM Switches

MPLS networks can use traditional ATM equipment. This will typically be done as a migration step in introducing MPLS to an existing ATM network. Traditional ATM switches can be used in three ways, as shown in Figure 2(f).

- Backhauling, when the access device is remote from the Edge LSR. The access device is connected to the Edge LSR by PVCs switched through an ATM network.
- Tunnelling through ATM switches between an Edge LSR and an ATM LSR. In the case, the edge LSR does not need to be adjacent to an ATM-LSR, but can be connected through an ATM network.
- Tunelling through ATM switches between ATM LSRs. In this case, the core network uses traditional ATM switches as well as ATM switches.

These uses of traditional ATM equipment have disadvantages, and must be used with care. These issues are discussed further in “4. Migration of MPLS into Traditional ATM Networks” on page 40.

Dual Backbones

Providers may sometimes want to keep an existing ATM infrastructure while building a new MPLS infrastructure (either ATM MPLS or packet based MPLS) alongside the old infrastructure. Cisco IP+ATM edge devices support this well, allowing customers to access both the MPLS network and services, and the old ATM network, even from a single access link. This is shown in Figure 2(g). The IP+ATM access devices can be any of those which can be used in Figure 2(e). The network in Figure 2(g) supports the same functions and services as Figure 2(e), but the Figure 2(g) network requires more equipment.

2.2 Choosing Cisco MPLS Equipment

Choosing ATM MPLS Edge Equipment

There are four main considerations when choosing ATM MPLS edge equipment:

- Type of services to be offered: IP+ATM, i.e. end-to-end PVC and SVC services as well as IP services, or just IP
- Type of access lines
- Number of access lines
- Requirements for redundancy and reliability. Key issues are whether the equipment can minimise (warm standby) or completely prevent (hot standby) disruption to data flow in the case of software or hardware failure, and whether individual components, e.g. port cards, can be hot-swapped. Hot standby means zero or almost zero (under 1 second) interruption to end-to-end data flows in the case of equipment failure, with no re-routing beyond the failed equipment.

Redundancy levels for MPLS edge devices can be classified as follows:

- None: The edge device has no redundancy features. The network must rely on re-routing for reliability. Customer sites must be dual-homed to two different access devices to ensure reliable service.
- Processor redundancy: The edge device has a redundant pair of processors and backplanes with warm or hot standby. Since port cards are not in redundant pairs, the network must still rely on re-routing for reliability. Customer sites must still be dual-homed to ensure reliable service in case of port card failure, but they may be dual-homed to two different cards on the same access device.
- Full redundancy. The edge device has redundant processors and port cards. Customer sites will receive reliable service from the edge equipment even if they are single-homed.

Equipment recommendations based on these requirements are shown in Table 2. In future, it is likely that Cisco's will support MPLS on most or all mid-range and high-end routers, most or all ATM switches, and most or all access products with routing capability. Consequently, more MPLS edge devices will become available in the future. Details of all Cisco products are available from http://www.cisco.com

There are some practical issues to bear in mind when examining the product specifications for products to be used as ATM Edge LSRs:

- When considering the number of access lines supported, take into account the card slots used by the ATM MPLS interfaces.
  For example, a Cisco 7206 router http://www.cisco.com/warp/public/cc/cisco/mkt/core/7200/prodlit/c7200_ds.htm has 6 card slots, and can nominally support 48 ethernet ports (or 8 per slot). However, when using the Cisco 7206 router as an ATM Edge LSR, at least one slot must be used for an ATM interface. So, the actual ethernet port capacity of a Cisco 7206 ATM Edge LSR is 40 ethernet ports.
- Some edge LSRs use card slots for cards performing the Edge LSR function, and these need to be taken into account.
  For example, the Cisco 6400 Universal Access Concentrator http://www.cisco.com/warp/public/cc/cisco/mkt/access/dslaggr/prodlit/6400_ds.htm has 8 card slots, but when it acts as an Edge LSR, at least one of these slots must be used for a Node Route Processor card, and not a line card.
- Some edge LSRs can deal with a throughput of MPLS traffic which varies with the number of processing cards.
  For example, at the time of writing the current Node Route Processor (NRP) card in a Cisco 6400 Universal Access Concentrator can handle roughly 150Mb/s full-duplex of MPLS edge traffic. This means that, assuming a typical activity fraction of 25%, a Cisco 6400 with one NRP can handle four OC3/STM-1 access lines; whereas a Cisco 6400 with two NRPs can handle eight OC3/STM-1 lines.
  The number of slots taken by processor cards also must be taken into account when calculating the number of access lines supported for any particular configuration.

Choosing ATM Label Switch Routers

There are five main considerations when choosing ATM LSRs:

- Type of trunks
- Number of trunks.

Table 2 Choosing MPLS Edge Equipment for ATM MPLS networks

Equipment	Type of Services	Access Lines[a]	Redundancy Support	Comments
3600 router	IP only	Relatively small numbers of async, modem, serial/frame relay, 10mb/s ethernet, ISDN BRI & PRI, HSSI, E1/T1 serial, fast ethernet, OC3/STM-1 ATM, voice interfaces, and others.	None	Small number of LVCs supported on ATM cards will lead to limitations on MPLS network size. Not recommended for provider ATM MPLS networks.
4700 router	IP only	Relatively small numbers of serial/frame relay, 10Mb/s ethernet, ISDN BRI, E1/T1 serial, fast ethernet, E3, T3 or OC3/STM-1 ATM, and others.	None	Small number of LVCs supported on ATM cards will lead to limitations on MPLS network size. Not recommended for provider ATM MPLS networks.
7200 router	IP only	Serial/frame relay up to E1/T1, 10Mb/s & fast ethernet, ISDN BRI, HSSI, high-speed serial, E3, T3 or OC3/STM-1 ATM, packet-over-SONET/SDH and others.	None	Minimum recommended for provider networks. PA-A2 CES-ATM port adaptors do not currently support MPLS.
7505 router, or 7507, 7513, or 7576.	IP only	Serial/frame relay or ISDN up to E1/T1, 10Mb/s, fast, & gigabit ethernet, HSSI, high-speed serial, ATM, packet-over-SONET/SDH and others.	Warm-Standby Processor Redundancy with dual RSPs.	
12008/12012	IP only	POSIP and ATM at OC3 to OC48 rates, and gigabit ethernet.[b]	Warm-Standby Processor Redundancy	Suitable for high-speed peerings between providers.
Catalyst 5500 with Route Switch Modules	IP+ATM	10Mb/s & fast ethernet, E3, T3, OC3/STM-1 ad OC12/STM-4 ATM, and others.	None	The Cat 5500 is primarily a LAN switch, but also has limited Edge LSR capability. The Cat 5500 may only be connected to an ATM MPLS network by tunnelling, as shown in Figure 2(f). This is discussed further in section 4. on page 40
6400	IP + ATM	ATM at E3/T3 to STM-4 rates, also ethernet and fast ethernet.	Warm-Standby Processor Redundancy	MPLS support is not yet shipping.
MGX 8850	IP + ATM	High numbers of 56k/64k Frame Relay, T1/E1 Frame Relay, channelized, and ATM, and higher-speed Frame Relay, serial and channelized T3.	Full Warm-to-Hot Standby, (but with FCS limitations)[c]	
BPX 8650	IP + ATM	High numbers of 56k/64k Frame Relay, T1/E1 Frame Relay, channelized, and ATM, ATM at E3/T3 to STM-4 rates, and others.	Excellent redundancy in general, but there is a single point of failure for Edge LSR function.	(See BPX 8680.)
BPX 8680	IP + ATM	High numbers of 56k/64k Frame Relay, T1/E1 ATM, Frame Relay and channelized. Also ATM at E3/T3 to STM-4 rates, and others[d].	Full Warm-to-Hot Standby (but with FCS limitations)[e]	BXM trunk cards must be used. BXM cards are required. MPLS is not supported on BNI cards, except if the BNI cards are used as Feeder Trunks. BCC cards must be BCC3-64 or later. BCC4 cards are strongly recommended.

[a] For more details see http://www.cisco.com.

[b] Note that the highest ATM bandwidth density supported by the 12000 series port cards is 1 x OC12 per slot. Since all traffic in an ATM edge LSR must go through an ATM interface into the ATM MPLS network, this relatively low ATM bandwidth density of the 12000 limits its capacity as an ATM Edge LSR.

[c] At FCS, the MGX 8800 will have hot-standby1:N redundancy capability for customer access lines, hot standby control for PVCs and hot-standby trunks. However it will not yet support redundant Edge LSRs processors (Route Processor Modules), so it will be effectively non-redundant as an edge LSR device. 1:N warm standby redundancy for RPMs is scheduled for release in CY2000,

[d] Extra 7200 or 7500 routers (or "Label Switch Controller" packages) may be required to act as Edge LSRs E3/T3 or faster ATM access lines are used. If IP service is to be supported for large numbers of ATM links at T3/E3 rates and above, its is more cost-effective to use separate, stand-alone, routers.

[e] .The BPX 8680 can include up to 16 MGX8850 shelves, with redundancy features as described above. Full redundancy for the combined device relies on redundancy for the label switch controller for the BPX 8600 shelf. A re dun and ant configuration using two simultaneously active controllers will be supported in BPX software release 9.3.

- Number of connections supported
- Whether VC Merge is required
- Requirements for redundancy and reliability, as discussed above.

Equipment recommendations based on these requirements are shown in Table 3. In future, it is Cisco's intention to support MPLS on all ATM switches. Consequently, more ATM-LSRs will become available in the future. In addition, any traditional ATM switch can be used in a Cisco MPLS network if tunnelling is used, subject to significant limitations. This is discussed further in "4. Migration of MPLS into Traditional ATM Networks" on page 40.

Table 3 Choosing ATM-LSRs

Equipment	Type and Number of ATM Trunks[a]	Number of Connections Supported	VC Merge Supported?	Redundancy Support	Comments
MGX 8850 with PXM1	4xOC3/STM1	8k[b]	No	Full Warm-to-Hot Standby	The MGX 8850 is intended primarily to be an edge LSR, but will also has limited ATM-LSR capability. This capability may not be available until late in CY 2000.
LS1010	32 x T1/E1 with Inverse Multiplexing over ATM (IMA), 32 x T3/E3, 32 x OC3/STM-1, 8 x OC12/STM-4	64k	Yes	None	
6400	16 x T3/E3, 16 x OC3/ATM-1, 8x OC12/STM-4	64k	Yes	Warm-Standby Processor Redundancy	ATM-LSR support on this switch is not yet shipping.
BPX 8650	144 x T3/E3, 96 x OC3/STM-1, 24 x OC12/STM-4	192k	CY2000[c]	Some redundancy features at FCS, full redundancy is possible later.[d]	All MPLS interfaces must be on BXM cards. BCC cards must be BCC3-64 or later. BCC4 cards are strongly recommended.
8540 MSR	64 x T1/E1 with IMA, 64 x T3/E3, 64 x OC3/STM-1, 32 x OC12/STM-4, 8 x OC48/STM-16	256k	Yes	Warm-Standby Processor Redundancy	Some trunk cards are not yet shipping.
MGX 8800 with PXM-45 card(s)	192 x T3/E3, 144 x OC3/STM-1, 48 x OC12/STM-4, 12 x OC48/STM-16	384k	Yes	Full Warm-to-Hot Standby	PXM-45 cards are not yet shipping

[a] If redundant pairs of trunks are used, the number of trunks supported will by half the numbers shown.

[b] 16k full-duplex connection legs are supported on the PXM card of the MGX 8850. If both legs of all connections are on the PXM card, then 8k connections are supported.

[c] VC Merge will be supported in BPX software release 9.3, with Enhanced BXM cards.

[d] The BPX8650 has supports hot standby trunks and switching fabrics. Full redundancy relies on redundancy for the label switch controller for the BPX 8600 shelf. A redundant configuration using two simultaneously active controllers will be supported in BPX software release 9.3.

Label Switch Routers Not Based On ATM Switches

LSRs other than ATM switches may be also be used. In particular, the following routers may be used as LSRs:

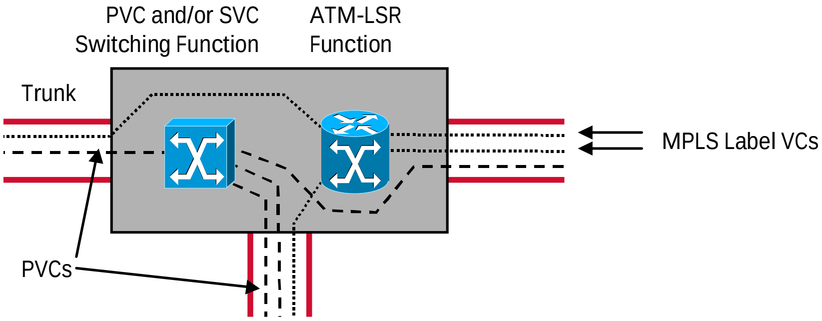
- 3600 & 4700 series routers (low-bandwidth applications only)
- 7200 & 7500 series routers
- 12000 series Gigabit Switch Routers

Using these routers, MPLS may be supported over virtually any link type: ATM, Packet-Over-SONET, Ethernet, etc. Router-based LSRs do not support native ATM Virtual Circuit connections, and all except the 12000 series have relatively low capacity compared to Cisco ATM-LSRs.

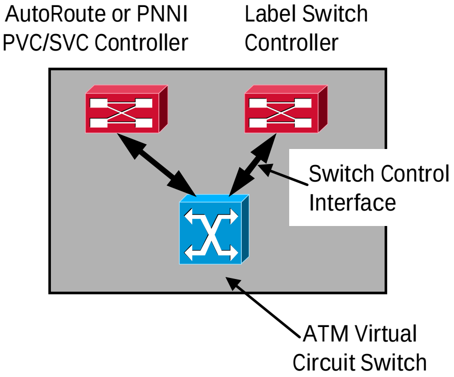
2.3 IP+ATM

IP+ATM capability is a major strength of Cisco ATM switches, and is one of the most important selling advantages of them when competing against other vendors. “IP+ATM” is Cisco’s name for equipment which simultaneously supports traditional ATM services (PVCs, SVCs, SPVCs, PVPs, etc.), as well as optimized IP transport using MPLS. The concept of an IP+ATM switch is shown in Figure 3(a). A single switch contains two logically separate switches: an MPLS ATM Label Switch Router (LSR) optimized for IP transport, and a traditional ATM PVC/SVC switch. Each trunk can support both PVCs (or SVCs, etc.), and MPLS Label VCs (LVCs).

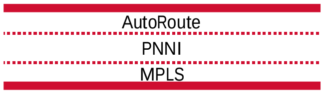
Although an IP+ATM switch contains logically separate switches, it physically contains only one switch. However it contains two (or more) separate sets of control software. One set of software controls PVCs, SVCs, etc., and the other control software controls MPLS. These controllers act independently, allowing the single physical switch to act as two (or more) virtual switches. In switches such as the BPX 8650 and MGX 8850, this independent control is implemented using a switch control interface (SCI). (The switch control interface used is the Virtual Switch Interface (VSI).) The SCI allows two (or more) separate controllers to independently control a single switch, as shown in Figure 3(b). The MPLS control software is physically located in a Label Switch



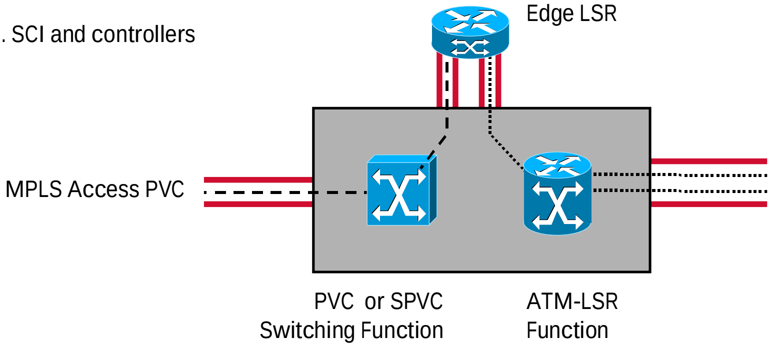
a. Logical view of an IP+ATM switch



b. SCI and controllers



c.Partitioning a trunk



d. Both functions in the data path

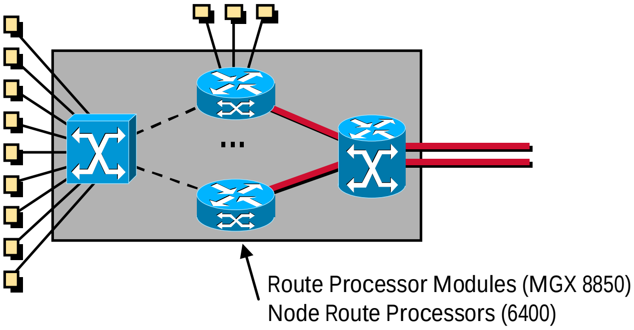


Figure 3 IP+ATM

Controller (LSC). In the BPX 8650, the LSC is a device separate from the main switch shelf. In the MGX 8850, the LSC is based on a Route Processor Module (RPM) in the switch shelf itself. Other SCI controllers may be software running on the switch control card[5]. The LS1010 and 8540 MSR implement similar functionality to the SCI using internal software interfaces.

In order that the control planes can act independently, the SCI "slave" process(es) in the switch must allocate resources two the different "control planes" (MPLS, PNNI, etc.). In the BPX 8650, resources for AutoRoute PVCs are reserved in a similar way. The resources partitioned in this way are:

- VPI/VCI space on trunks. Each control plane gets a range of VPIs to use.
- Bandwidth.Each control plane is guaranteed a certain bandwidth for Connection Admission Control (CAC) purposes. With "soft partitioning," there can be a pool of bandwidth which is shared between control planes for CAC purposes. Even with "hard" partitioning, spare bandwidth unused by a control plane is available on a cell-by-cell basis to other control planes.
- Traffic queues. One of the keys to IP+ATM is that MPLS traffic gets different traffic queues on the switch than the PVC and SVC traffic. This means that MPLS traffic can be handled by queues which directly support the MPLS "Class of Service" concept. The alternative is manually-configured translations to ATM Forum Service Types. The need for these translations is one of the main disadvantages of IP-over-ATM schemes apart from MPLS, and IP+ATM avoids this disadvantage.

Part of the configuration process for IP+ATM switches is the assignment of these resources to the different control planes. this involves creating different "partitions" of link resources for the different control planes, as shown in Figure 3(c).

Use of IP+ATM

IP+ATM can be used to offer MPLS services, and PVC, SVC, etc. services, on the same network. This means that all (or many) switches in the network act as both ATM-LSRs and traditional ATM switches, as in Figure 3(a). The traditional ATM services can also be used in conjunction with an MPLS service. Figure 3(d) shows the use of a PVC to connect ordinary IP traffic from customer site through to an ATM Edge LSR. A PVC used in this fashion is called an "MPLS Access PVC". Other PVCs are "traditional PVCs" as part of a traditional end-to-end PVC service. The traffic from the Edge LSR can then be fed back through the ATM-LSR function in the same switch that supports the MPLS Access PVC, or alternatively through a different switch. In any case, the end-to-end data path for customers' IP traffic can include both MPLS Access PVCs and MPLS Label VCs.

An integrated IP+ATM Edge Switch, such as the MGX 8850 or Cisco 6400, contains ATM-LSR function, as well as traditional access switch and PVC switching function. In addition, the Edge LSR function is also integrated into the device. This shown in Figure 3(e). In the MGX 8850, routing function is supported by way of Route Processor Modules (RPMs), and Node Route Processor (NRP) modules are used in the Cisco 6400. Each RPM or NRP acts as an Edge LSR. In the MGX 8850, one of the RPMs will simultaneously act as an LSC and an Edge LSR. The implications of this can been seen more clearly in Figure 7 on page 26.

---

[5] In the case of the BPX 8650 and MGX 8850, AutoRoute software, which controls PVCs, runs on the switch control card. PNNI control will be added to the BPX 8650 as a separate controller. Later MGX 8850 releases will have PNNI software running on the switch control card.

3. Designing MPLS Networks

The goal of designing an MPLS network prior to installation is to produce a network which will operate correctly, and which will come at least reasonably close to meeting performance goals. Because of the inherently connectionless nature of IP traffic, customers will not be able to tell a carrier exactly what traffic they want to send where. Because of this, it is not possible to perfectly design a network ahead of time. The initial design steps result in a working network:

- Design Points of Presence
- Dimension backbone links in the network
- Design IP routing
- Dimension MPLS Label VC space

The final design step is an on-going process of optimizing the network design:

- Refine the design once the network is operational

These steps are considered in the following subsections.

3.1 Points of Presence Structures

The design of Points of Presence (PoPs) for an ATM MPLS network is constrained by:

- The choice of access line types and equipment for a network, as discussed previously.
- Location of PoPs, which is largely determined by where the cities are.
- The population of user sites surrounding each location.

Some typical PoP designs are shown in Figure 4.

Single ATM Edge LSR

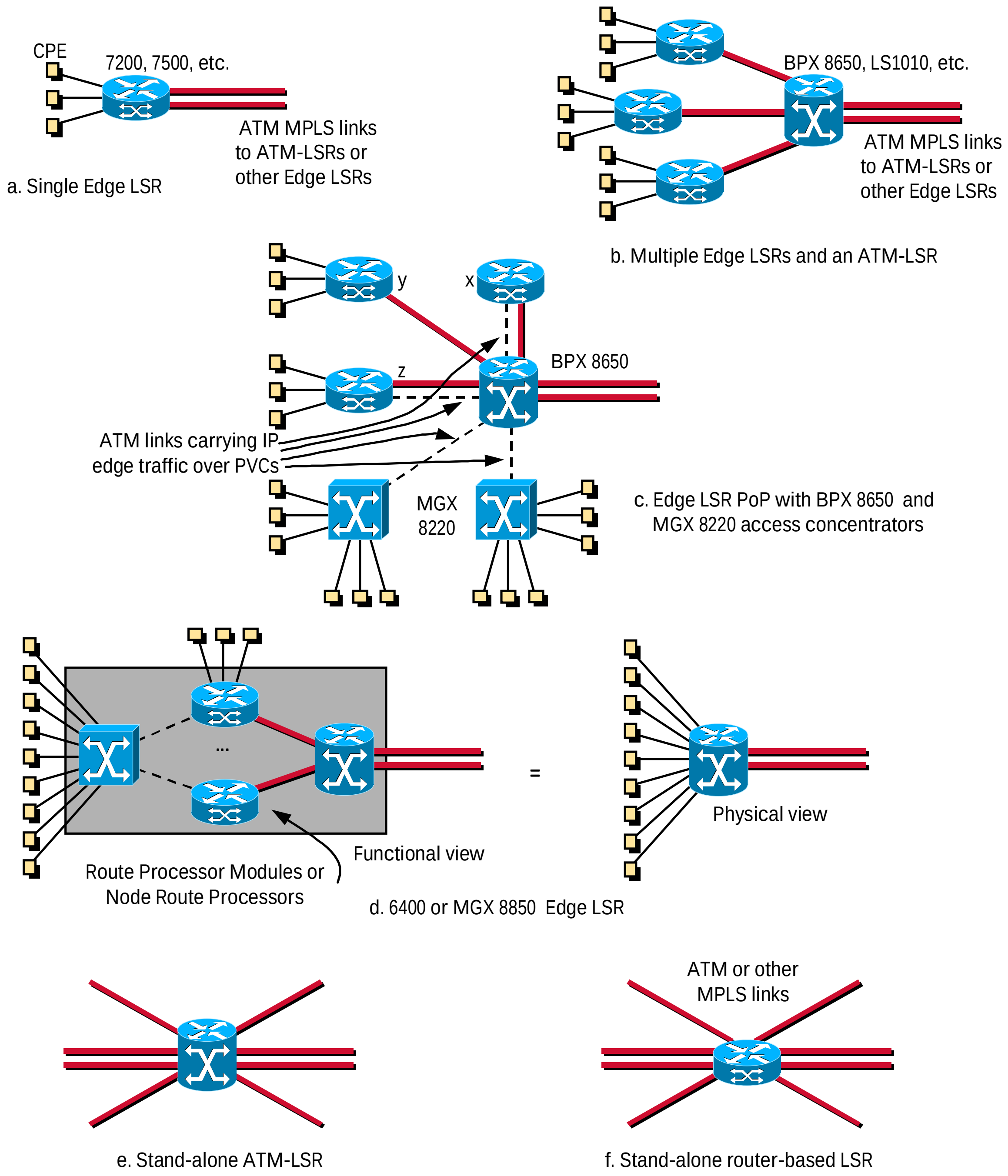
Where a single Edge LSR device is sufficient for supporting the number and types of access lines in a Point of Presence (PoP) location, then the simple structure shown in Figure 4(a) is sufficient. Numerous access lines (typically tens or hundreds) are brought into a single Edge LSR, which is connected to the rest of the ATM MPLS network. Numbers and types of access lines supported by single Edge LSRs are described in Table 2.

Multiple Edge LSRs and an ATM-LSR

A PoP may require more than one edge LSR because of a large number of access lines to be supported at that location. Alternatively, different types of edge LSR might be required because of different types of access lines to be supported. Where there are several edge LSRs in a PoP, it makes sense to also include an ATM-LSR. This is shown in Figure 4(b). The ATM-LSR:

- Locally switches traffic going between different edge LSRs in the PoP
- Concentrates traffic going from the PoP onto a single set of ATM MPLS links. The alternative would be either separate sets of links to all edge LSRs, or using one edge LSR to carry traffic to the others.
- Improves scalability of routing. Only one set of IP routing protocol (e.g. OSPF) peerings are required from the ATM-LSR to other points in the MPLS network. Without the ATM LSR, separate peerings would be required from all edge LSRs.

Depending on reliability requirements, redundant pairs of links would be used between the edge LSRs and the ATM-LSR.



a. Single Edge LSR

b. Multiple Edge LSRs and an ATM-LSR

c. Edge LSR PoP with BPX 8650 and MGX 8220 access concentrators

d. 6400 or MGX 8850 Edge LSR

e. Stand-alone ATM-LSR

f. Stand-alone router-based LSR

Figure 4 Point of Presence Structures for ATM MPLS networks

Edge LSR PoP With BPX 8650 and MGX 8220 Access Concentrators

An extension to the previous model is to use traditional access concentrators in addition to edge LSRs and an ATM LSRs. Circumstances where this is appropriate are discussed in “2.1 Structures for MPLS Networks” on page 9. One example of this type of PoP uses MGX 8220 access shelves, 7200 or 7500 edge LSRs, and BPX 8650. This is shown in Figure 4(c). IP traffic from access concentrators is carried in ATM PVCs to a edge LSRs. These may be carried through the same BPX 8650 that acts as an ATM-LSR, as this is an IP+ATM switch. Edge LSRs in such a PoP may have three different types of configuration:

- Typically, one or more edge LSRs will be dedicated to dealing with IP traffic on the edge PVCs. The router labelled ‘x’ in Figure 4(c) has this configuration. It must have at least two ATM interfaces: one for access PVCs (from the access concentrator) and one for ATM MPLS traffic.
- An edge LSR dealing with access PVCs may also have customer access lines directly terminating on it. Router ‘z’ in Figure 4(c) does this.
- There may also be edge LSRs in the PoP which don’t deal with access PVCs at all, and have only directly-connected access lines. Router ‘y’ has this configuration.

Note that the Label Switch Controller (LSC) in the BPX 8650 can act as an Edge LSR simultaneously with performing its LSC function. However, use of an LSC as an Edge LSR is not recommended for providers who consider the separation of MPLS control functions from data forwarding functions to be important. As an edge LSR, an LSC can perform any of the three functions discussed above. The number of Edge LSRs required in the PoP depends on:

- The total number of access lines
- The total bandwidth of the access lines, calculated from the average utilization. For example, the sum of the access lines bandwidths might be (say) 1Gb/s, the utilization might not exceed 500Mb/s.

The capacity of a 7200 or 7500 router running MPLS edge function is roughly the same as its ordinary IP capacity using Cisco Enhanced Forwarding (CEF). For example, a 7200 router with an NPE 200 processor can support close to 200 Mb/s of MPLS edge traffic, at normal IP packet sizes[6]. It can support MPLS edge function for about 700 access lines. [These figures are preliminary and need confirmation.]

Cisco 6400 and MGX 8850 Edge LSRs

The MGX 8850 and Cisco 6400 integrate the functions described in the previous example into a single device, illustrated in Figure 4(d). It consists of:

- A multiservice access concentrator with various types of Frame Relay and ATM access lines, as well as Circuit Emulation lines. Voice access capability and other types of access lines will be added later.
- One or more edge LSRs. Each edge LSR is a Route Processor Module (RPM) card in the case of the MGX 8850, or a Node Route processor (NRP) card in the case of the 6400. The number of RPMs or NRPs required to act as Edge LSRs depends on:
  - The total number of access lines
  - The total bandwidth of the access lines, downrated according to the utilization. For example, the sum of the access lines bandwidths might be (say) 1Gb/s, the utilization might not exceed 500Mb/s.

  An RPM with an NPE150 processor can support MPLS edge function for 700 access lines. It will support close to 150Mb/s of MPLS edge traffic, at normal IP packet sizes[7]. The NRP is similar. These limits will be raised in future due to software and hardware improvements.

---

[6] If additional edge functions such as CAR and WRED are used, then performance may be affected. In these circumstances, the performance of the routers needs to be verified for the particular combination of features which will be used. Note also that voice-over-IP packets are very short, and will have significantly lower throughput than indicated here.

[7] Again, the performance needs to be verified with the particular combination of edge functions to be used in addition to MPLS edge function.

- An ATM-LSR. In the MGX 8850, one of the RPM cards acts as a Label Switch Controller. It may perform both LSC function and edge LSR function simultaneously, if desired. Use of an RPM for simultaneous LSC and Edge LSR is not recommended for providers who consider the separation of MPLS control functions from data forwarding functions to be important. In the Cisco 6400, the main Node Switch Processor also acts an LSC.

The 6400 and MGX 8850 also have IP+ATM capability. Note that these functions are being phased in over time. In particular, LSC function will not be available until later. This has implications to network design which are discussed in "4.5 Examples of Hybrid ATM Network Equipment" on page 52.

Stand-Alone ATM-LSR s

Some sites in a network may have purely a switching role. In an ATM MPLS network, these sites will consist of a single ATM-LSR, as shown in Figure 4(e), or possibly a redundant pair of ATM-LSRs. The ATM-LSR would typically be a BPX 8650, 8540 MSR, or later an MGX 8800 with PXM-45. In some networks, it may make sense to use a router-based LSR instead, as shown in Figure 4(f). A 7500 or 12000-series router may be suitable for this application. Core LSRs usually have Edge LSR capability as well. For example, the BPX 8650 has limited edge LSR capability as part of its Label Switch Controller.

3.2 Dimensioning an MPLS Network's Links

This section is intended to be a guideline to the steps of dimensioning the links in a small MPLS network, but it is not the only way of doing this. Different providers will have their own unique procedures for designing and running networks, but they all will be roughly similar to this. Recall that it is not the purpose of initial design to produce a perfect network. Several approximations are made in this process which lead to a network which works, but is not necessarily optimal. The last step in the design process is to optimize the network, and this is discussed in "3.5 On-going Network Design" on page 38.

1. Design edge points of presence and their layout.

   The first step in MPLS network design will to be to choose the size, type, and layout of the Points of Presence according to the considerations described above. An example is shown in Figure 5. This example is based on a network in Australia. Australia was chosen for this example because it is small enough to make a suitable realistic example. The edge PoPs shown in Figure 5(a) are chosen based on the estimated customer link demand shown in Figure 5(b). BPX 8600-based edge LSR PoPs (which consist of several MGX 8800 shelves and a BPX 8650 for aggregation) are used in Sydney and Melbourne, which have the largest link bandwidths and number of links in this example. An MGX 8800 is used in Brisbane. Adelaide and Perth are smaller centres which can, in this example, be adequately served by router-based PoPs.
2. Estimate traffic from each point of presence

   Based on the total access line bandwidths, an estimate on the total traffic sent from customers into each PoP can be made. A 'busy-period' estimate should be used, e.g. of the rate during the busiest minute of the day. This is to ensure adequate dimensioning. A conservative estimate would be the total of the access line bandwidths at the PoP, as in Figure 5(b). However it will often be reasonable to take a somewhat lower estimate, e.g. 50% of the total access bandwidth, as shown in Figure 5(c)
3. Estimate the traffic matrix

   The exact process for this step will vary from network to network. In Australia, for example, the two main business centres are Sydney and Melbourne, with Sydney being slightly larger. In a large MPLS network for interstate business IP traffic, a reasonable first approximation may be that 50% of traffic will go to Sydney, 40%to Melbourne, 5% to Brisbane, and 2.5% to Adelaide and Perth respectively. An existing service provider would probably already have estimates for traffic patterns for their region. Based on the estimated traffic distribution percentages, and the total PoP traffic from Step 2, a traffic matrix can be estimated. The traffic matrix for this example is in Table 4. In a typical network, this matrix will be very roughly symmetrical. For example, in Table 4, the traffic from Sydney to Adelaide is 12.5Mb/s, but the traffic from Adelaide to Sydney is 25Mb/s. If the traffic were more asymmetrical than about 2:1 or 3:1, then there may be an error in traffic estimates or modelling.
4. In IP networks, traffic from x to y will often flows along the same path (but in the reverse direction) as traffic from y to x. Although this can be overridden by numerous routing protocol features, it maybe useful to assume that this will happen, particularly in small networks. Because of this, it may be easier to use bidirectional traffic flows rather than unidirectional flows

Table 4 Network example: uni-directional traffic matrix

Traffic Destination	Distribution Percentage	Traffic Source				
		Adelaide	Brisbane	Melbourne	Perth	Sydney
Adelaide	2.5%	1.25	2.5	10	1.25	12.5
Brisbane	5%	2.5	5	20	2.5	25
Melbourne	40%	20	40	160	20	200
Perth	2.5%	1.25	2.5	10	1.25	12.5
Sydney	50%	25	50	200	25	250
Total	100%	50Mb/s	100Mb/s	400Mb/s	50Mb/s	500Mb/s

in an initial network design. The estimated bidirectional flows for the example network are shown in Table 5. The bidirectional traffic bandwidth between Adelaide and Sydney, for example, is taken to be 25Mb/s, which is the maximum of the unidirectional bandwidth from Sydney to Adelaide (12.5Mb/s) and the bandwidth from Adelaide to Sydney (25Mb/s). Forming bidirectional flows in this way will tend to slightly overestimate the traffic in the network. This is useful as a conservative first approximation.

Table 5 Network example: approximate bidirectional traffic flows

	Adelaide	Brisbane	Melbourne	Perth	Sydney
Adelaide	1.25				
Brisbane	2.5	5			
Melbourne	20	40	160		
Perth	1.25	2.5	20	1.25	
Sydney	25	50	200	25	250

5. Design the layout of the backbone network

The layout of the backbone will involve consideration of

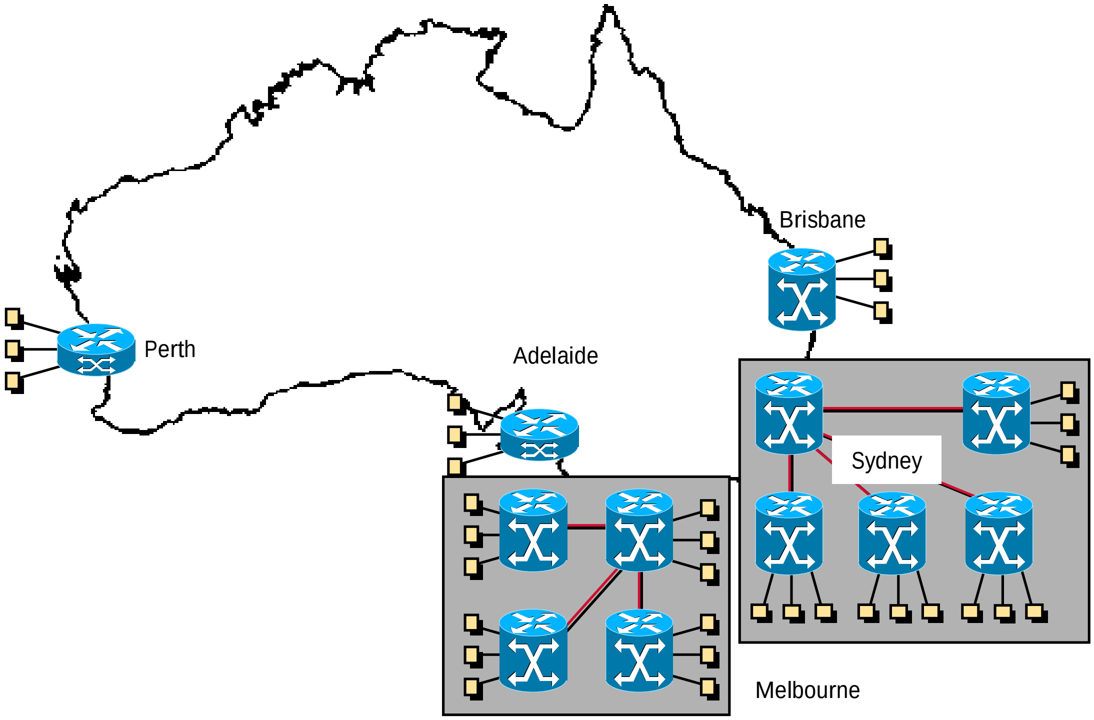
– Geographic layout, i.e. good locations for nodes

– Network-level redundancy, i.e. having multiple paths to each destination

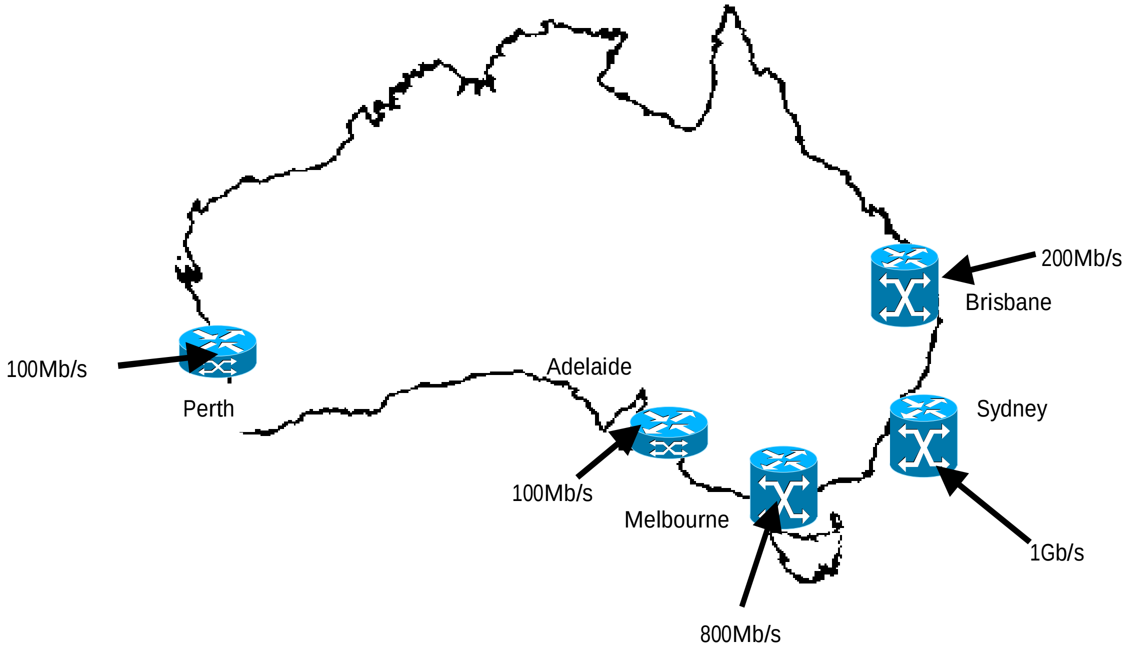
– Redundancy of trunks

The network layout chosen in this example is shown in Figure 5(d). Many layouts are possible. The one chosen consists of a combination of a partial ring, linking adjacent nodes at the outside of the network, and a star, connecting nodes back to an extra ATM-LSR in the core of the network. This provides a good degree of network-level redundancy, with at least two paths between each pair of nodes. The extra ATM-LSR node is placed in Bourke, a town which is roughly equidistant from four of the five customer PoPs. With a good degree of network-level redundancy, it is not essential to have redundant trunks, because it is possible to re-route MPLS Label VCs[8]. In this example, most trunks are chosen to be non-redundant for economy. A redundant pair of trunks is used for the link which is expected to have the heaviest utilization, namely between Sydney and Melbourne.

[8] In traditional connection-oriented networks, re-routing of virtual circuits is a last resort to be used only when all other redundancy mechanisms have failed. This is because it inevitably involves disruption of customer traffic for many seconds or minutes as all circuits are re-routed. In traditional IP networks, re-routing is a much less severe issue, as packet flows can be switched from one link to another almost instantaneously, once the IP routing protocol has converged. MPLS networks lie between theses two extremes. Re-routing in MPLS networks is particularly feasible if VC Merge is used, for two reasons: VC merge reduces the number of VCs which are used in the network, and it reduces the scope of changes required in connections when re-routing does occur.

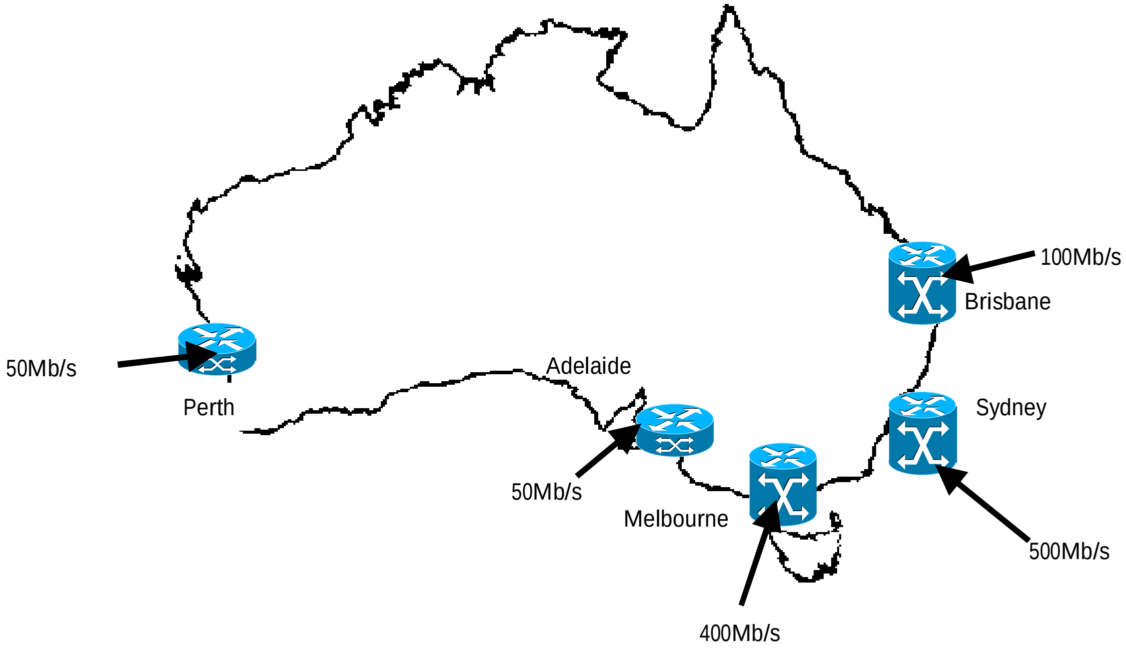


a. Point-of-Presence layout

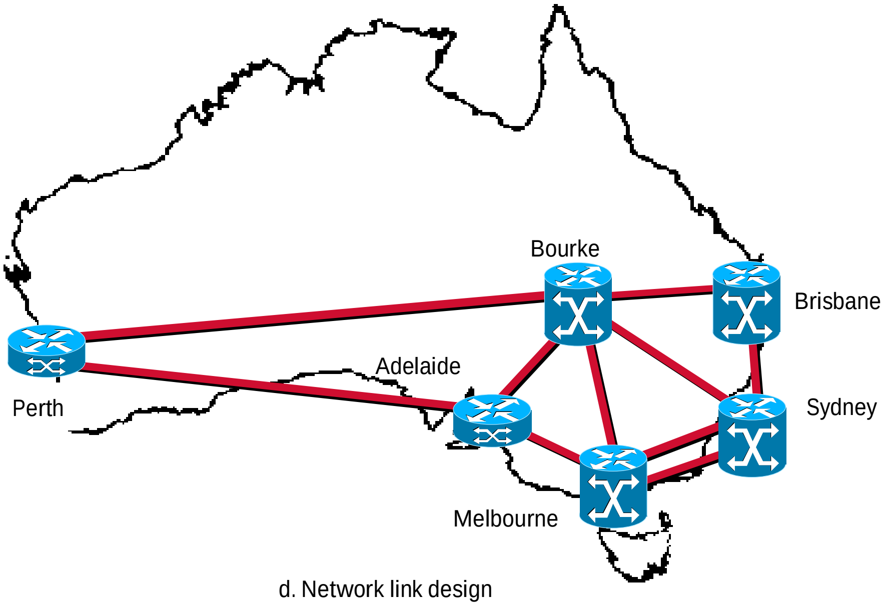


b. Total access line bandwidths

Figure 5 Network design example - geographic layout



c. Peak traffic sent from each PoP



d. Network link design

Figure 5 Network design example - geographic layout (continued)

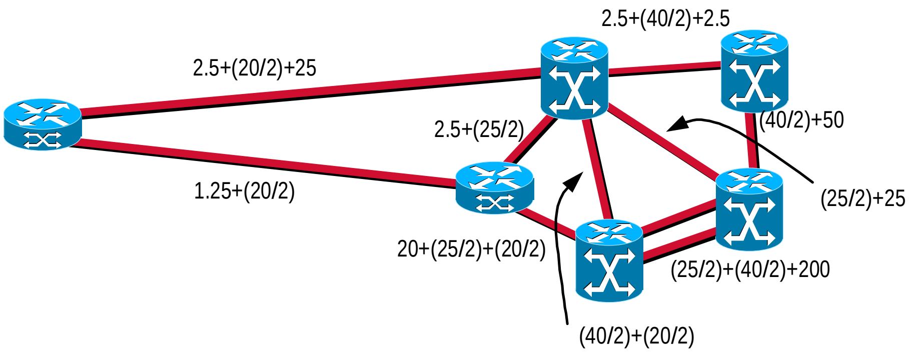
6. Estimate link flows based on the estimated traffic
   This involves calculating the paths taken by traffic through the network backbone. With IP routing protocols such as OSPF, this is a straightforward procedure, as the IP traffic will follow the minimum-hop path unless administrative costs are used. Where there are two or more minimum-hop paths, the traffic will be approximately balanced across them. The process of calculation of link flows for the traffic in Table 5 is shown in Figure 6(a), and the totals in Figure 6(b)
7. Assign link capacities
   Based on the estimated link flows, link capacities can be assigned to the links in the network. This would generally involve choosing the next standard link size (T3/E3, STM-1, etc.) larger than the link flows just calculated. This is illustrated in Figure 6(c).
8. Adjust for redundancy
   Where redundant trunks are not used, and the network relies on re-routing for reliability, it may be necessary to adjust link bandwidths to ensure that there is sufficient capacity to deal with link failures. For example, if the link labelled ‘r’ in Figure 6(c) fails, then links ‘s’ and ‘t’ will need to carry some or all of its traffic. The load on these links would then exceed E3 rates, so an STM-1 link (or multiple E3 links) would be required for each of links ‘s’ and ‘t’. Similarly, if link ‘u’ failed, then link ‘v’ would require more than E3 capacity. Also, if link ‘y’ failed then the offered load for ‘w’ would exceed E3. So, the final allocation of link bandwidths is as shown in Figure 6(d).
9. Check whether the selected equipment is adequate
   This involves checking in Table 2 and Table 3 whether the selected PoP equipment can support the number and size of links chosen in the network design. The network in this case would pass this check But for example, assume the Melbourne PoP had used an MGX8850 instead of a BPX8680. This PoP needs two STM-4 links and two STM-1 links, which is not yet supported on an MGX 8850. So, in this hypothetical example, the PoP would need to be re-designed, e.g. by using a BPX 8680 instead of an MGX8850.

   Note that any such re-designs, if required, are a relatively minor issue. Many different types of Cisco equipment can be used in ATM MPLS PoPs, and it will usually be found that a PoP can be built to meet the requirements of one location, simply by using combinations of equipment used at other locations. A BPX 8680, for example, combines several MGX 8850 shelves.

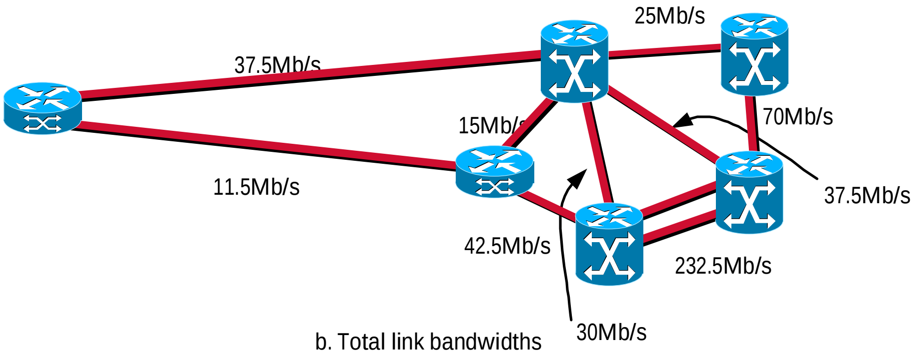
Note on Redundant Pairs of ATM Links

There are three main ways of achieving changeover for a redundant pair of ATM links:

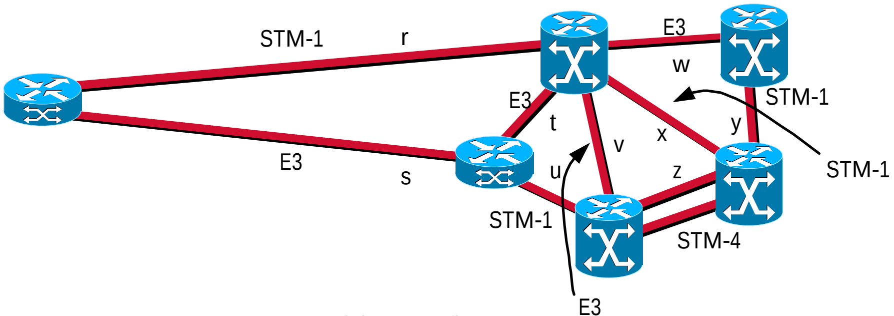
- Data link-level changeover. This is the normal link redundancy mechanism in traditional ATM networks. Changeover occurs because of physical and data link-layer monitoring in the ATM switch or ATM-LSR hardware. The switch hardware also typically sets up a copy of all virtual circuit state on the backup link. In a switch with data link-level redundancy, any single link failure will typically result in close to zero data loss on any virtual circuits. In addition, the network layers and routing (IP or PNNI, etc.) will not be affected by the link failure, or even be aware that it has occurred. However with data link redundancy, the backup link is not available to carry data except in the case of failure of the main link. Depending on how it is implemented, SONET Automatic protection Switching (APS) may be a form of data link redundancy. On the MGX 8850 and BPX 8650, SONET APS changeovers result in no change to the interfaces as seen by connection routing, and no loss of connection state. This does not necessarily apply to APS on other Cisco equipment.
- Inverse multiplexing over ATM (IMA). IMA carries distributes data over a group of links by distributing cells across the links in round-robin fashion. It offers both data-link level load sharing across links, and redundancy. If one of the links in a group fails, cells are no longer sent on that one link, but the others are still used. IMA is available only for low-speed links—groups of T1 or E1 links.
- Parallel links with network-layer changeover. In this case, a redundant pair of trunks is used, but data-link layer protection is not used at all, and all connection changeover takes place at the network layer. IP or PNNI routing is aware of all link failures and reacts to them. This is not particularly good for connection-oriented traffic, but works well with IP routing and MPLS. With OSPF equal-cost-multipath or similar, OSPF will choose to balance traffic for every route across both links in a pair of links. This causes a pair of MPLS Label LVCs to be set up for each destination, one per link. If one of the links fails, IP routing will simply divert



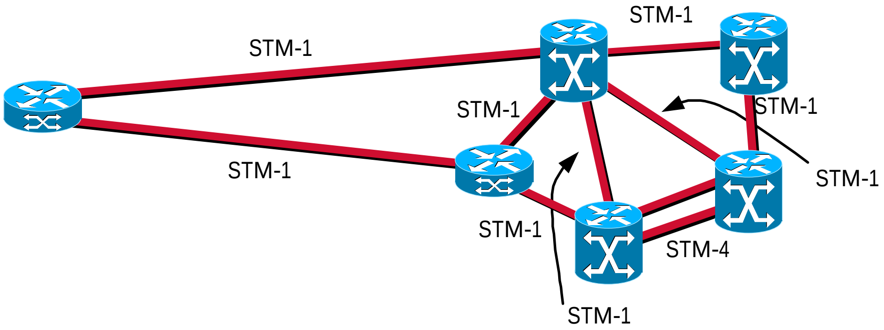
a. Calculating traffic flows according to IP routing



b. Total link bandwidths

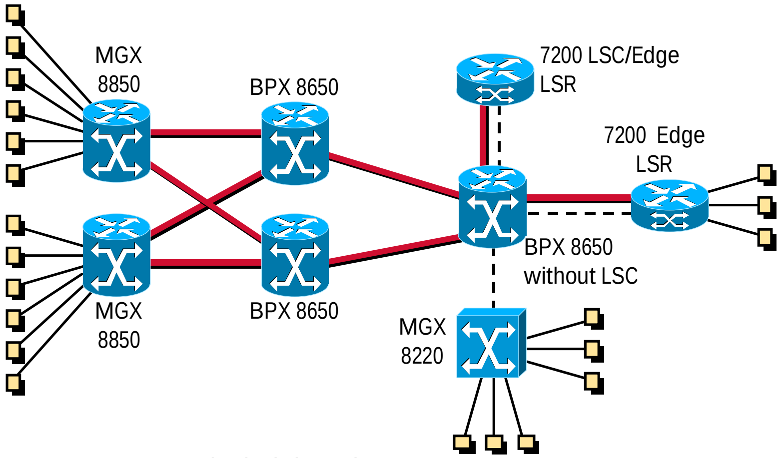


c. Link types - first pass

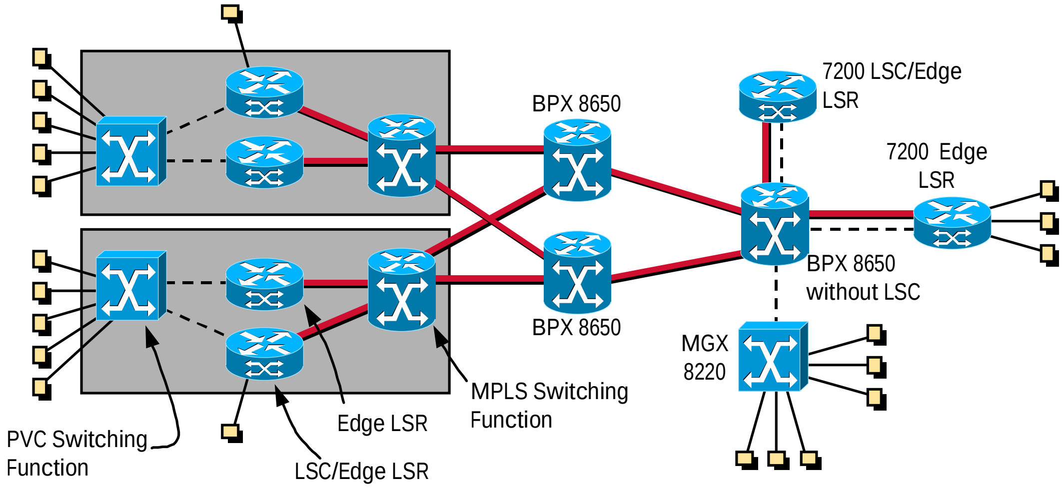


d. Link types, allowing for network-level redundancy

Figure 6 Network design example - calculating link bandwidths



a. Physical viewpoint



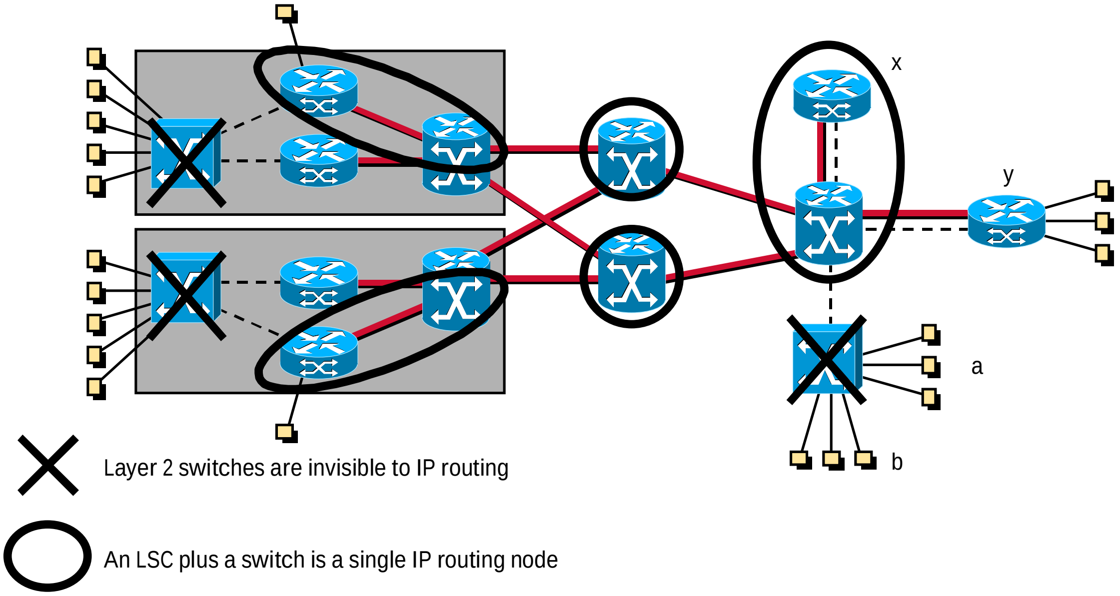
b. Functional viewpoint

Figure 7 Viewpoints of an ATM MPLS network

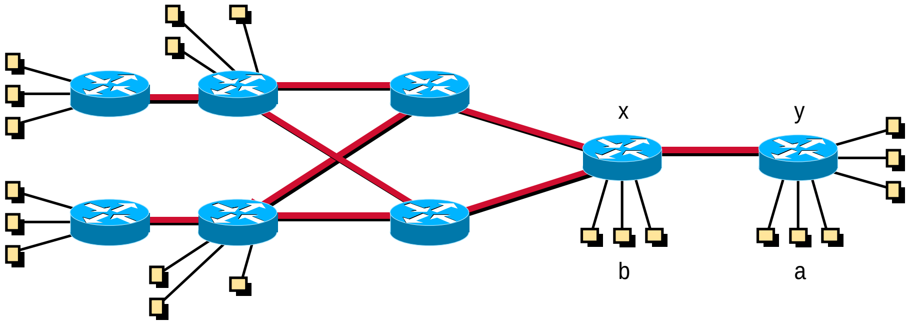
traffic onto the other, already-established, LVCs. If VC merge is used, this will require no more MPLS signalling, and could be achieved in a second or so. The advantage of this method is that it allows the bandwidth in the ‘backup’ trunk to be used, allowing more best-effort traffic to be carried in the network. SONET APS changeovers, as implemented on the non-MSSBU Cisco equipment, are a form of parallel link redundancy, but without the capability of equal-cost multipath to set up back-up links.

Where a redundant link is required, recommendations for use of these modes are as follows:

- IP+ATM networks: these should use Inverse Multiplexing for redundancy for low-speed trunks, and otherwise use data link-layer redundancy. This avoids costly re-routes of the connection-oriented traffic.
- Pure ATM MPLS networks: again, these should use Inverse Multiplexing for redundancy for low-speed trunks. Otherwise, if the network uses VC merge, parallel links with network-layer changeover should be used, in order to make the full network capacity available for use. Finally, if VC merge is not available, data link-layer redundancy should be used.



c. Deriving the routing viewpoint



d. Routing viewpoint

Figure 7 Viewpoints of an ATM MPLS network (continued)

3.3 IP Routing in an MPLS Network

MPLS uses ordinary IP routing protocols—OSPF, IS-IS, etc.—to determine the routes for IP traffic and LVCs. Every LSR runs ordinary IP routing protocols in the same way that ordinary IP routers do. An important implication of this is that OSPF (or IS-IS, etc.) “sees” an MPLS network as being exactly like an ordinary router network. This is illustrated in Figure 7. It is possible to have various viewpoints of an ATM MPLS network:

- Physical viewpoint: This viewpoint represents the physical devices and links in a network. An example is shown in Figure 7(a).
- Functional viewpoint: Where a product has several functions, these can be shown separately. For example, the MGX 8850s in Figure 7 each include two separate Edge LSRs, which are shown separately. In addition, it is useful to think of the PVC switching function of an MGX 8850 to be separate from the MPLS switching function. It is sometimes useful to consider the Label Switch Controller (LSC) function in an ATM-LSR as being separate from the switching function. This is particularly true if the LSC is also acting as an Edge LSR.

- Routing viewpoint: This viewpoint shows the network as it is seen by an IP routing protocol. An example of deriving it is shown in Figure 7(c) and (d).
  - Layer 2 PVC switches and PVC switching functions are invisible to IP routing. If a customer site is connected to a router by a PVC, then the PVC is a 1-hop direct connection from an IP routing perspective. See for example the sites labelled ‘a’ Figure 7(c), and assume that these are connected by PVCs to Edge LSR ‘y’. Then, in the routing viewpoint, the sites are directly adjacent to router ‘y’.
  - A Label Switch Controller and a switch together form a single routing node.

  Using these rules, the routing viewpoint of an MPLS network can be derived. This is shown in Figure 7(d).

Designing IP routing in an MPLS network is almost exactly the same process as designing IP routing for an ordinary IP network. By looking at the routing viewpoint, a network can be divided into Areas, route summarization can be designed, and so on. There are several design guides for IP routing on http://wwwin-cons.cisco.com/~dblack/design-guides/. Also, see the book “Internet Routing Architectures,” mentioned in the Introduction. There are a small number of routing issues specific to MPLS networks, and these are considered next.

MPLS-Specific IP Routing Issues

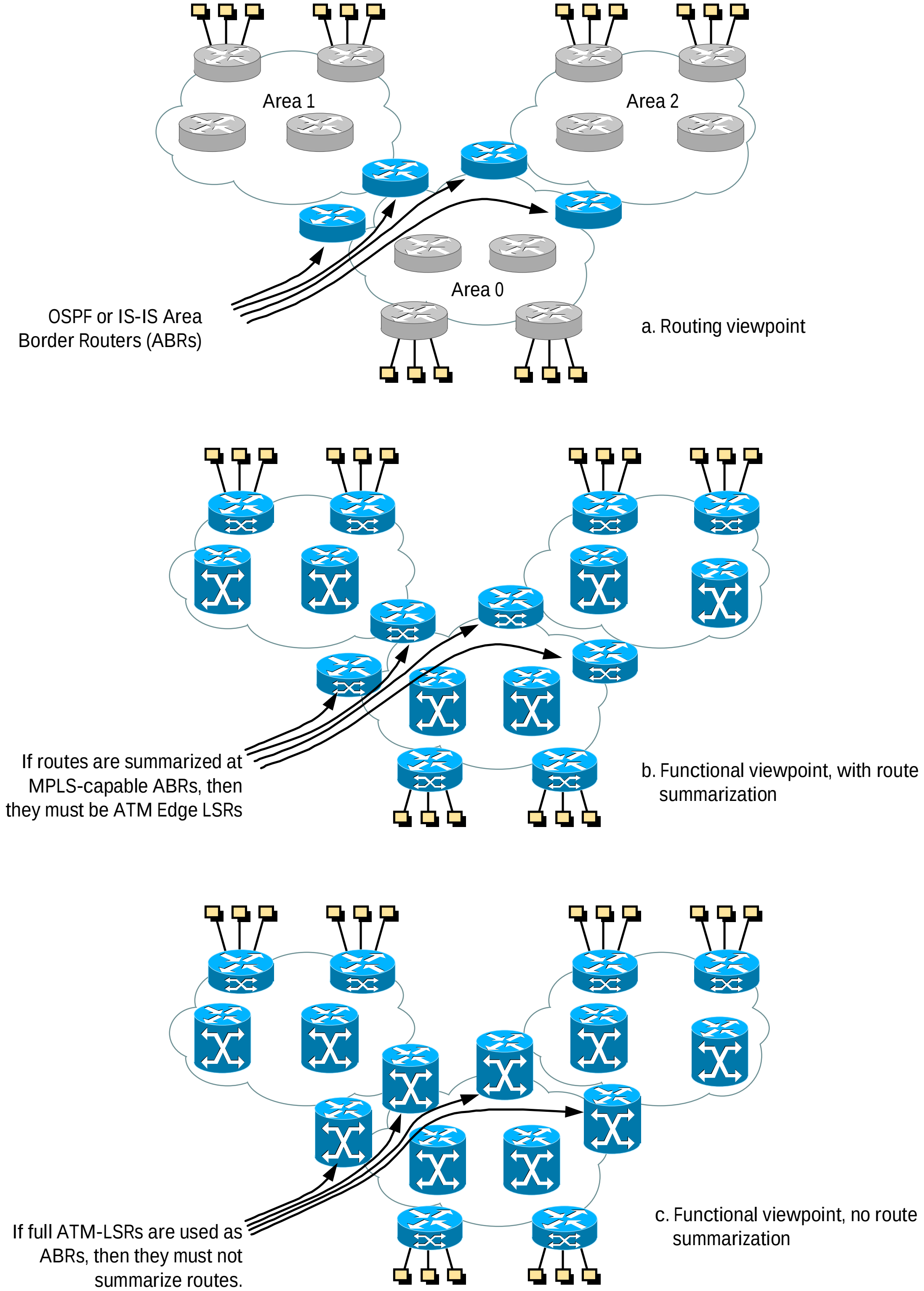
- The interior routing protocol used in MPLS backbones should be either OSPF or IS-IS[9]. EIGRP can also be used, but it will not work with an advanced MPLS-based IP traffic engineering feature called Routing with Resource Reservations (RRR). RRR requires a link-state routing protocol, i.e. OSPF or IS-IS[10]. Since EIGRP is a distance-vector routing protocol, it will not work with RRR. IGRP or RIP also will work with MPLS but not RRR, and are not recommended. Note that RRR is sometimes referred to loosely as “MPLS Traffic Engineering,” but is actually a specific type of MPLS traffic engineering.
- Use unnumbered IP links where possible. This reduces the number of IP destinations known to the routers, and hence reduces the number of LVCs used in the network. This is also discussed under “3.4 Dimensioning MPLS Label VC space” on page 30.
- Route summarization must not be done at an ATM-LSR. Multiple OSPF or IS-IS areas can be used in an ATM-MPLS network, as shown in Figure 8. An ATM-LSR may be used an OSPF or IS-IS Area Border Router (ABR), but only if no summarization is done at the Area Border routers. In Figure 8(c), this means that the address prefixes known in all the areas must be the same. An ABR in Figure 8(c) may not, for example, summarize reachability for 1.1.1.0/24, 1.1.2.0/24 and 1.1.3.0/24 with a single route for 1.1.0.0./16. If route summarization is required in an ATM MPLS network, it must be done at an ATM Edge LSR, as shown in Figure 8(b).
- The previous rule also applies to Autonomous Systems and BGP 4. An ATM-LSR may not be a BGP Autonomous System Boundary Router, but an ATM Edge LSR may be one.
- Routing with Resource Reservations (RRR) works best in backbones which contain a single OSPF or IS-IS Area. Currently, RRR may be not be used in multi-Area networks where the Area Border Routes are ATM-LSRs. This restriction will be eased in a later version of RRR.[11]
- Route summarization may not be done in the interior of an MPLS VPN network. MPLS VPN networks will be discussed in detail in a later version of this design guide. The interior of a MPLS network supporting VPNs may have multiple OSPF or IS-IS areas, but summarization should not be used.

---

[9] IS-IS is currently supported on most Cisco MPLS equipment, but not on the LS1010 and 8540 MSR.

[10] Both the OSPF and IS-IS Working Groups at the IETF are working on extensions to support IP Traffic Engineering, and these capabilities will be used by RRR.

[11] One specific requirement is that a Traffic Engineering (TE) or RRR tunnel must start and end at a device which can “push” and “pop” labels on a packet’s label stack. Currently this requires that the tunnels start and end at ATM edge LSRs or other packet-based LSRs. It is theoretically possible to let ATM-LSRs push and pop an extra label by carrying the additional label in a range of VPIs. However this has not yet been implemented. Another requirement is that TE/RRR tunnels may not pass through Area Border Routers; they must terminate and re-start. Together, these requirements currently mean that it is currently impossible to use RRR or TE in a network where the ABRs are ATM-LSRs. A future RRR release will allow TE/RRR tunnels to pass through Area Border Routers, and this will permit RRR to be used in networks where the ABRs are ATM-LSRs.



OSPF or IS-IS Area Border Routers (ABRs)

a. Routing viewpoint

If routes are summarized at MPLS-capable ABRs, then they must be ATM Edge LSRs

b. Functional viewpoint, with route summarization

If full ATM-LSRs are used as ABRs, then they must not summarize routes.

c. Functional viewpoint, no route summarization

Figure 8 Multiple routing areas and summarization in an ATM MPLS network

The restrictions on summarization exist because summarization stops some types of label-switched paths being set up end-to-end. For example, assume that an ABR summarizes reachability for 1.1.1.0/24, 1.1.2.0/24 and 1.1.3.0/24 with a single route for 1.1.0.0/16. Now assume that a packet with IP address 1.1.1.23 arrives with a label for 1.1.0.0/16. The ABR cannot label-switch the packet. It must look past the label and examine the IP address to find that the packet should go on to 1.1.1.0/24. Since ATM-LSRs cannot examine IP addresses, they may not do IP route summarization[12].

3.4 Dimensioning MPLS Label VC space

This design guide has shown how many of the issues of designing MPLS networks are similar to those of designing ordinary IP networks. One important exception to this is the dimensioning of MPLS LVC requirements on each link. This design problem is illustrated in Figure 9. In order to complete the design of an ATM MPLS network, a sufficient number of VCs must be reserved for use as LVCs on each link. This can be a problem, as any ATM switch will support only a certain number of active VCs. This is particularly important if there are multiple ATM services—MPLS, PNNI, etc.—sharing the resources of the links in an IP+ATM network. The design problem is to determine what number of LVCs is required.

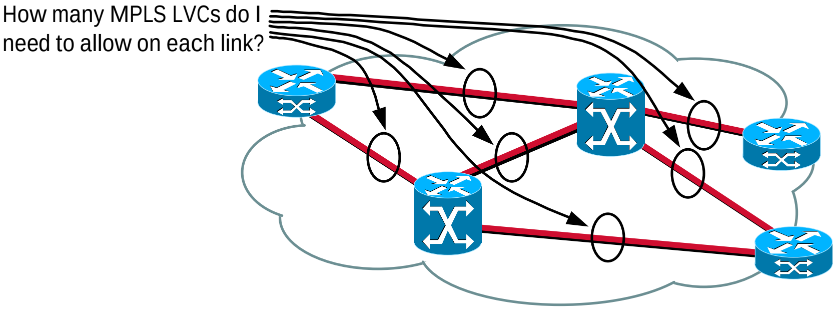


Figure 9 Label VC Requirements

The required number of LVCs depends on:
- The number of IP destinations in the network
- The relationship between destinations and LVCs
- Whether VC merge is used
- The paths chosen by IP routing

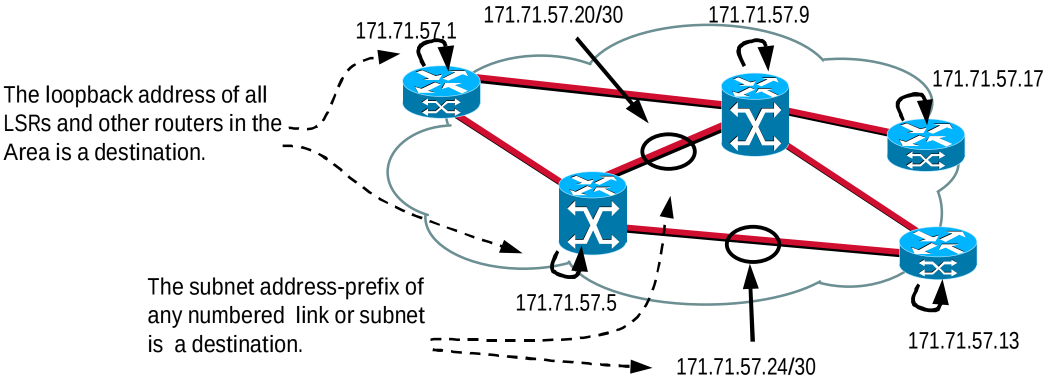
Destinations

The number of LVCs used in a particular Area of a network depends on the number of IP destination-prefixes known in that area. This follows the normal rules for an IP network:
- The loopback address of all LSRs and other routers in the area is a destination-prefix.
- The subnet address-prefix of any numbered point-to-point link, or any other subnet, is a destination-prefix. Because of this, it is best to use unnumbered links in MPLS networks.
- Any other address prefixes advertised into the area need to be counted as well[13]. If many addresses are summarized into a single address at the Area Border Router (or Autonomous System Border Router), then this counts as a single destination-prefix.

These rules are shown in Figure 10.

[12] Some ATM-LSRs, e.g. the BPX 8650, have a limited ability to examine IP addresses by sending the packets to the Edge LSR function in their Label Switch Controller. However this may be done only for a small minority of the traffic flowing through the ATM-LSR

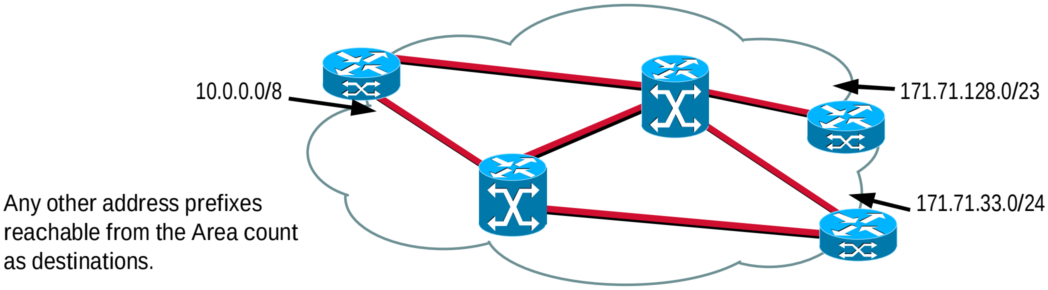
[13] Note that in MPLS VPN networks, this does not apply to VPN customers' addresses. VPN customers' destination-prefixes are not advertised into the core of the network. This is one of the keys to scalability of MPLS VPNs.



The loopback address of all LSRs and other routers in the Area is a destination.

The subnet address-prefix of any numbered link or subnet is a destination.

a. Destinations within an area



Any other address prefixes reachable from the Area count as destinations.

b. Destinations outside the area

Figure 10 Destination-prefixes in an MPLS network (or any other IP network)

LVC's Used Per Link & VC Merge

Each ATM Edge LSR and each Label Switch Controller will ask a neighbouring MPLS node for LVCs for the destination-prefixes it knows about. If MPLS Class of Service is used, it may ask for up to four LVCs for each destination-prefix. The requests for LVCs flow through the network according to the paths chosen by IP routing. With VC merge, the LVCs to each destination will be merged at each ATM-LSR. This means that on each link, there is at most one LVC per destination in the Area. This is shown in Figure 11(a). If MPLS Class-of-Service is used, then this is multiplied by the number of classes. If VC merge is not used, there may be many more LVCs: this is discussed later.

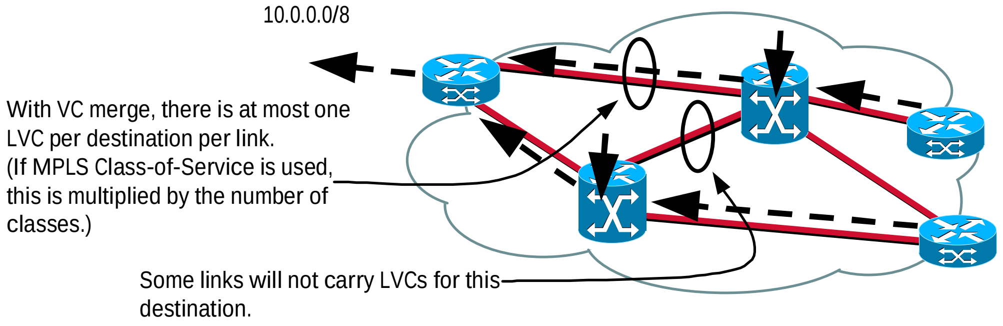
Design Calculations: Edge LSRs

For ATM Edge LSRs, the number of LVCs used per link depends on whether VC merge is being used in the network. Let d be the number of destination-prefixes known in an area, and c be the number of classes-of-service used in the network. If VC merge is used, then the number of LVCs used per link, l, satisfies

$$l \le cd\,. \tag{1}$$

If VC merge is not being used in the network, there is also a dependency on the number of LSCs and Edge LSRs in the area. There is also a dependency on how many destinations are directly reached through the edge LSR in question. If d_e is the number of destinations reachable through a particular ATM edge LSR (this will often equal 1, due to summarization) and the total number of ATM Edge LSRs and LSCs in the area is n, then the number of LVCs used per link satisfies

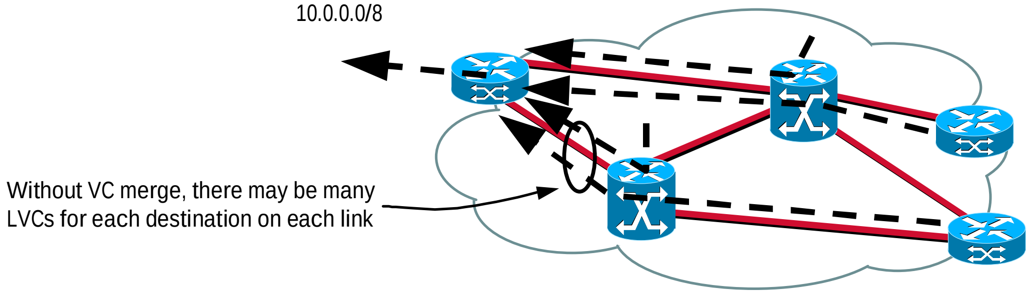
$$l \le c(d - d_e) + cnd_e\,. \tag{2}$$



With VC merge, there is at most one LVC per destination per link. (If MPLS Class-of-Service is used, this is multiplied by the number of classes.)

Some links will not carry LVCs for this destination.

a. LVCs to a destination with VC merge



Without VC merge, there may be many LVCs for each destination on each link

b. LVCs to a destination without VC merge

Figure 11 LVCs to each destination

In the particular case where VC merge is not used, and there is one destination-prefix per edge LSR or LSC, and all links are unnumbered, and there are no address prefixes from outside the area, a simpler equation applies. These conditions will often apply in the core of MPLS networks supporting VPNs, but not using VC Merge. The number of LVCs used per link on the ATM Edge LSR in this case is given by:

$$l < 2cn\,. \tag{3}$$

One of these equations is then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 6. Table 7 shows the LVC capacity of Cisco ATM Edge LSR interfaces.

Table 6 Checking the LVC limits of Edge LSR

Device	Situation	Key Parameter	Check Against
Edge LSR	The network uses VC merge	Number of active VCs supported per ATM link.	Equation (1)
Edge LSR	The network does not use VC merge, there is one destination-prefix per LSR or edge LSR, all links are unnumbered, and there are no out-of-area routes.	Number of active VCs supported per ATM link.	Equation (3)
Edge LSR	The network does not use VC merge, all other situations.	Number of active VCs supported per ATM link.	Equation (2)

Table 7 Cisco ATM Edge LSRs and LVC capacity

Device	Interface hardware	Number of active LVCs supported	Notes
3600	NM-1A ATM Network Modules	1024	
4700	NP-1A ATM Network Processor Module	1023	
7200, 7500	PA-A1 or standard ATM port adaptor	2048	
Catalyst 5500, 7200, 7500	PA-A3 ATM port adaptor.	4096	
6400	Node Route Processor (NRP)	2048	Capacity is reduced by 1 LVC for each active PVC which terminates on the NRP.
MGX 8850 IP+ATM switch	Route Processor Module (RPM)	4096	Capacity is reduced by 1 LVC for each active PVC which terminates on the RPM. In addition, the PXM is limited to 16k LVCs. This is unlikely to be a problem unless more than 3 RPMs are used in an MGX 8850 shelf.
12000 series routers	4xOC3 ATM Line Card	2047 [Needs confirmation]	The 2047 active VCs are shared between all four ports. Network capacity is reduced by 1 destination-prefix for every second and subsequent route chosen for each destination according to equal-cost multipath routing, if the extra route(s) are on the same card.
12000 series routers	1xOC12 ATM line card	2047 [Needs confirmation]	

Edge LSRs: Worked Examples

Q. Consider a network where VC merge is being used and one class of service is being used. If the edge LSRs are all 7200-series routers with PA-A3 port adaptors, then how many IP destination prefixes can safely be supported in the area?

A. VC merge is being used, so Table 6 indicates that Equation (1) should be used. One class of service is being used, so $c = 1$. Table 7 states that the PA-A3 port adaptor supports 4096 LVCs, so $l = 4096$. Substituting these into Equation (1) gives

$$4096 \le 1d$$

Or $d \ge 4096$. This means that 4096 destination-prefixes are guaranteed to be supported within the area, provided that the ATM-LSRs do not impose a tighter limit (this discussion has considered only the Edge LSRs).

Q. Consider a network where VC merge is not used, and 4 classes of service are being used. This network is the core of an MPLS VPN service, and there is one destination-prefix per LSR or Edge LSR. All links are unnumbered. No out-of-area routes are injected into the interior routing protocol. The edge LSRs are 7200 and 7500-series routers with a mixture of PA-A1 and PA-A3 ATM port adaptors. What is the largest number of LSRs which can be used, if the network consists of a single Area? Assume that the ATM-LSRs support a sufficiently large number of LVCs.

A. According to the conditions given, Table 6 indicates that Equation (3) should be used. Four classes of service are used, so $c = 4$. The Table 7 shows ATM interfaces each support 2048 or 4096 LVCs. The interfaces with 2048 LVCs give a tighter restriction, so use $l = 2048$. Substituting these into Equation (3) gives

$$2048 < 2(4)n$$

Or $n > 256$. This means that a maximum of 256 LSRs (edge LSRs or ATM-LSRs) may be used in the area, provided that the IP routing protocol supports that many routers in an area.

Q. Consider a network where VC merge is not used, and 4 classes of service are being used. This network is the core of an MPLS VPN service, and there is one destination-prefix per ATM-LSR or Edge LSR. All links are unnumbered. The network has multiple areas, and there are at most 100 ATM-LSRs and LSRs in each area. All the edge LSRs' ATM port adaptors are PA-A3. Assume that the ATM-LSRs support a sufficiently large number of LVCs. How many LSRs can be used in the entire network?

A. There are multiple areas, and so there will be out-of-area routes in each area. Table 6 indicates that Equation (2) should be used in this case. Four classes of service are used, so $c = 4$. Table 7 gives $l = 4096$. There are at most 100 ATM-LSRs and LSRs in each area, so use $n = 100$.

Observe that there is one route per edge LSR. In the worse case, all out-of-area routes are accessed through a single LSR—this concentrates the LVC requirements on the link(s) to that single LSR. In this case, $d_e = (d - 100)$. Substituting these into Equation (2) gives

$$4096 \le 4(100) + 4(100)(d - 100)$$

$$(d - 100) \ge 3696 / 400$$

or $d \ge 109$. This means that only 109 LSRs can be used in the network. By comparison with the previous example, we can see that using multiple areas can have major disadvantages in ATM MPLS networks without VC merge[14].

Design Calculations: ATM-LSRs With VC Merge

With VC merge, the LVCs to each destination will be merged at each ATM-LSR. This means that there is at most one LVC per destination on each link, as shown in Figure 11(a). If MPLS Class-of-Service is used, then this is multiplied by the number of classes. If d is the number of destination-prefixes known in an area, and c is the number of classes-of-service used, then the number of LVCs used per link, l , satisfies

$$l < cd . \tag{4}$$

Another important issue in switches which support VC merge is the number of LVCs which must be merged together in the switch, say m . This depends on the number of links into the switch, say k . The limit is:

$$m < cd(k - 1) . \tag{5}$$

These equations is then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 8. Both equations will need to be checked. Table 9 shows the limits of Cisco ATM-LSRs with VC merge.

Table 8 Checking the LVC limits of ATM-LSRs with VC Merge

Device	Key Parameters	Check Against
ATM-LSR with VC Merge	1. Number of active VCs supported per ATM link. 2. Number of merging LVCs supported per switch, or per port card, whichever is applicable to the switch architecture.[a]	Equation (4) Equation (5)

[a] In general, a per-switch limit will apply to shared-memory switches such as the LS1010 or 8540 MSR, and a per port card limit will apply to crossbar switches such as the BPX 8650 or MGX 8800 with PXM-45.

ATM-LSRs With VC Merge: Worked Examples

Q. A network uses BPX 8650 ATM-LSRs with VC merge. Two classes of service are used. Each BPX 8650 has 4x1-port OC12/STM-4 BXM cards, with each port used to link to another ATM-LSR or edge LSR. What limit do these ATM-LSRs put on the number of IP-destination-prefixes which can be supported inside an area?

---

[14] These examples indicated that ATM-LSRs without VC merge typically cannot be used in networks of larger than a few hundred nodes. An alternative which works around these limitations is to use the same switches, but to use MPLS-over-PVCs instead of ATM MPLS. This is discussed further in "4.2 Networks Using MPLS-over-PVCs" on page 41.

Table 9 Cisco ATM-LSRs and LVC capacity, if VC merge is used

Device	Interface Hardware	Number of active LVCs supported[a]	Number of active merging LVCs supported[a]
LS1010	Any ATM port hardware.	4096 per OC3 port, 16k per OC12 port, 16k? per OC48 port [needs confirmation]	64k per switch,
6400	Any ATM port hardware	4096 per OC3 port, 16k per OC12 port, 16k? per OC48 port [needs confirmation]	256k per switch
8540 MSR	Any ATM port adaptors	4096 per OC3 port, 16k per OC12 port, 16k? per OC48 port [needs confirmation]	256k per switch
BPX 8650 or 8680	BXM-E cards	32k per BXM, shared amongst up to 12 interfaces	32k per BXM, with a maximum of 16k per port on OC3 BXM cards and 2xOC12 BXM cards. T3/E3 BXM cards and 1xOC12 BXM cards have a limit of 32k per port.
MGX 8800 with PXM-45	AXSM cards	128k per AXSM, shared amongst up to 16 interfaces	128k per AXSM [Needs confirmation.]

a. These numbers are maximums, and the actual limits will be dependent on configurations. In a BPX 8650, for example, the actual number of active LVCs supported per link must be down-rated by a minimum of 270 lines per interface if AutoRoute is enabled on that interface. On all switches, the VC space reserved for PVCs, SVCs, etc. must be subtracted from the available VC space.

A. Table 8 shows that both Equation (4) and Equation (5) need to be checked. Two classes of service are used, so $c = 2$. Each switch has four ports, so $k = 4$. Looking up the BPX 8650 in Table 9 shows that BXM cards support 32k active LVCs. In this case, each BXM card has one port, so each link supports 32k LVCs, or $l = 32768$. Table 9 shows that 32k LVCs can be merged into a 1-port OC12 BXM card, so $m = 32768$.

Substituting these parameters into Equation (4) gives

$$32768 < 2d$$

or $d > 16k$. Substituting the parameters into Equation (5) gives

$$32768 < 2d(4 - 1)$$

or $d > 5461$. The limit from Equation (5) is tighter, which means that the limit imposed by the ATM-LSRs is 5461 destination-prefixes in the area. (The edge LSRs might impose a tighter limit.)

Q. A network uses 8540 MSR ATM-LSRs with VC merge. Four classes of service are used. Each 8540 MSR has 8xOC3/STM-1 with each port used to link to another ATM-LSR or edge LSR. What limit do these ATM-LSRs put on the number of IP-destination-prefixes which can be supported inside an area?

A. Table 8 shows that both Equation (4) and Equation (5) need to be checked. Four classes of service are used, so $c = 4$. Each switch has eight ports, so $k = 8$. Looking up the 8540 MSR in Table 9 shows that OC3 port cards support 4096 LVCs, or $l = 4096$. Similarly, Table 9 shows that the 8540 MSR supports 256k merging VCs, so $m = 262144$.

Substituting these parameters into Equation (4) gives

$$4096 < 4d$$

or $d > 1024$. Substituting the parameters into Equation (5) gives

$$262144 < 4d(8 - 1)$$

or $d > 9362$. The limit from Equation (4) is tighter, which means that the limit imposed by the ATM-LSRs is 1024 destination-prefixes in the area. (The edge LSRs might impose a tighter limit.)

Q. A network uses BPX 8650 ATM-LSRs with VC merge. Four classes of service are used. Each BPX 8650 has 8 ports, on 2x4-port OC3/STM-1 BXM cards, with each port used to link to another ATM-LSR or edge LSR. What limit do these ATM-LSRs put on the number of IP-destination-prefixes which can be supported inside an area?

A. Table 8 shows that both Equation (4) and Equation (5) need to be checked. Four classes of service are used, so $c = 4$. Each switch has eight ports, so $k = 8$. Looking up the BPX 8650 in Table 9 shows that BXM cards support 32k active LVCs. In this case, each BXM card has four ports, so we can assume that each link supports 32k/4 LVCs, or $l = 8192$. Similarly, Table 9 shows that 4xOC3 BXM cards can support 32k merging VCs, with a maximum of 16k per port. The worst case is when all LVCs try to merge into the same port, so use so $m = 16384$.

Substituting these parameters into Equation (4) gives

$$8192 < 4d$$

or $d > 2048$. Substituting the parameters into Equation (5) gives

$$16384 < 4d(8-1)$$

or $d > 585$. The limit from Equation (5) is tighter, which means that the limit imposed by the ATM-LSRs is 585 destination-prefixes in the area. (The edge LSRs might impose a tighter limit.)

Design Calculations: ATM-LSRs Without VC Merge

Without VC merge, there may be many VCs per destination on each link, as shown in Figure 11(b). If the total number of ATM Edge LSRs and LSCs in the area is n, then there may be up to $c(n-1)$ LVCs per destination on each link[15]. The number of LVCs used per link will then satisfy

$$l < cd(n-1)\,. \qquad \textbf{(6)}$$

A tighter limit applies in the particular case where VC merge is not used, and there is one destination-prefix per edge LSR or LSC, and all links are unnumbered, and there are no address prefixes from outside the area. These conditions will often apply in the core of MPLS networks supporting VPNs, but not using VC Merge. The number of LVCs used in this case is given by:

$$l \leq c(n^2/2)\,. \qquad \textbf{(7)}$$

One of these equations is then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 10. Table 11 shows the limits of Cisco ATM-LSRs without VC merge capability.

Table 10 Checking the LVC limits of ATM-LSRs without VC Merge

Device	Situation	Key Parameter	Check Against
ATM-LSR without VC Merge	There is one destination-prefix per LSR or edge LSR, all links are unnumbered, and there are no out-of-area routes.	Number of active VCs supported per ATM link.	Equation (7)
ATM-LSR without VC Merge	All other situations.	Number of active VCs supported per ATM link.	Equation (6)

Table 11 Cisco ATM-LSRs and LVC Capacity, if VC merge is not used

Device	Interface Hardware	Number of active LVCs supported per link[a]
BPX 8650	Older BXM cards (or pre-9.3.x software)	16k per BXM, shared amongst up to 12 interfaces

[a] These numbers are maximums, and the actual limits will be dependent on configurations. In a BPX 8650, for example, the actual number of active LVCs supported per link must be down-rated by a minimum of 270 lines per interface if AutoRoute is enabled on that interface. On all switches, the VC space reserved for PVCs, SVCs, etc. must be subtracted from the available VC space.

---

[15] This assumes that all LSRs in the area which are performing label switching (as opposed to label edge) function are ATM-LSRs. In other words, it assumes that the only router-based LSRs are at the edge of the area.

ATM-LSRs With VC Merge: Worked Examples

Q. A network uses BPX 8650 ATM-LSRs without VC merge. One class of service is used. Each BPX 8650 has 2x4-port OC3/STM-1 BXM cards, with each port used to link to another ATM-LSR or edge LSR. There is one destination-prefix per LSR or edge LSR. All links in the area are unnumbered, and there are no out-of-area routes known. What limit do these ATM-LSRs put on the number of LSRs or edge LSRs which can be supported inside an area?

A. Table 10 shows that Equation (7) should be checked in this case. One class of service is used, so $c = 1$. Looking up the BPX 8650 in Table 9 shows that BXM cards support 16k active LVCs. In this case, each BXM card has four ports, so each link supports 16k/4 LVCs, or $l = 4096$.

Substituting these parameters into Equation (7) gives

$$4096 \le 1(n^2/2)$$

$$(n^2 \ge 8192)$$

or $n > 90$. This means that the limit imposed by the ATM-LSRs is 90 LSRs or edge LSRs[16].

Q. A network uses BPX 8650 ATM-LSRs without VC merge. Two classes of service are used. Each BPX 8650 has 4x1-port OC12/STM-4 BXM cards, with each port used to link to another ATM-LSR or edge LSR. What limit do these ATM-LSRs put on the number of IP-destination-prefixes which can be supported inside an area?

A. No information is given on the relationship between devices and routes, so Table 10 shows that both Equation (6) should be checked. Two classes of service are used, so $c = 2$. Looking up the BPX 8650 in Table 9 shows that BXM cards support 16k active LVCs. In this case, each BXM card has one port, so each link supports 16k LVCs, or $l = 16384$.

Substituting these parameters into Equation (6) gives

$$16384 < 2d(n-1)$$

$$d(n-1) > 8192$$

n, the number of LSRs in the area, has not been given, so this question cannot be explicitly answered as stated. However, if a it is assumed that (say) $n = 50$, this would give and indicative value $d > 167$. In other words, the number of destination prefixes which may be supported depends on the number of LSRs in the area with, for example, a limit of 167 destination-prefixes if there are 50 LSRs in the area.

Note on Internet Routing Tables

The limits on destination-prefixes indicated in the previous examples are much smaller than the size of the Internet backbone routing table, which is about 70,000 routes. Despite this, ATM MPLS can still be used in networks with full Internet routing, by use of an MPLS feature known as BGP Next-Hop Labelling. This feature allows BGP Autonomous System Boundary Routers (ASBRs) to exchange the full Internet routing table with each other by way of BGP, while readvertising only a limited subset of these addresses (or none at all) into the Interior routing protocol (OSPF or ISIS) Area(s) through which they are connected. Since only a limited set of destination-prefixes is known on OSPF or IS-IS in the MPLS network, the limits discussed here are sufficient even though they are much smaller than the Internet routing table.

Cisco MPLS Virtual Private Networks (VPNs) extend the BGP Next-Hop Labelling technique to deal with address from many different customers' VPNs.

---

[16] These examples indicated that ATM-LSRs without VC merge typically cannot be used in networks of larger than a few hundred nodes. An alternative which works around these limitations is to use the same switches, but to use MPLS-over-PVCs instead of ATM MPLS. This is discussed further in "4.2 Networks Using MPLS-over-PVCs" on page 41.

Traffic Engineering

The limits shown in Equations (4) to (7) apply when MPLS Traffic Engineering is not being used. If Traffic Engineering is being used, then one LVC will be used for each Traffic Engineering Tunnel on each link, in addition to the limits shown above.

VP Tunnels

VP Tunnels are described in “4. Migration of MPLS into Traditional ATM Networks” on page 40. They involve several logical links terminating on a single physical interface on an LSR or ATM-LSR. When VP Tunnels are terminate on an interface, the LVCs on all VP Tunnels must be taken into account. For example, if 4 VP Tunnels terminate on a logical interface which supports 4000 LVCs, then an average of only 1000 LVCs will be available per VP Tunnel. This issue is considered in more detail in “4.6 VP Tunnels and LVC Usage” on page 53.

Alternative Calculations

The limits shown in Equations (4) to (7) can be quite ‘loose’. The number of LVCs actually used on a particular link may be much less than these limits suggest, particularly if VC merge is not being used. However it is difficult to calculate exactly how many will be required. This depends on the exact shape and state of the network, and the exact paths chosen by IP routing. If this can be analyzed, taking account of such things as failed links and multipath routing, then lesser numbers of LVCs could be safely reserved on each link. This would be a quite complex process. In any case, the limits shown above will be safe.

The Last Word

Josh Gahm, who wrote much of the LSC code, said this about running out of space for LVCs:

> “a network that has run out of cross-connects is basically non-functioning. There really isn't any way to control which [LVCs] get created and which don't. This may result in permanent black holes and/or [overloaded LSCs]. The network needs to be designed from the start such that it won't run out of connections.[17]”

In other words, it is very important to design the network and allocate switch resources so that a sufficient number of LVCs are available on each link.

3.5 On-going Network Design

Network design is an on-going process. Once an ATM-MPLS network is deployed, on-going design activities are required to:

- Verify the assumptions used in the initial design
- Adjust the network as new customers and PoPs are added.

The on-going process will involve the following steps:

1. Measure actual PoP and link traffic, and compare against
   - The predicted traffic
   - The link capacities
2. Based on the comparison between predicted and actual traffic, some modelling assumptions might be changed. For example, the traffic distribution across nodes might be different to that which was initially predicted. Review the initial design and dimensioning if the modelling assumptions are changed. Lower and higher-bandwidth links might be required.
3. As customers are added, and as traffic increases, review the initial design to:
   - Add new edge LSRs to PoPs
   - Add new links to the network

---

[17] Two different behaviors may occur when a link runs out of LVCs, depending on the type of traffic. The first case applies to ordinary IP traffic which would otherwise be carried with a single MPLS label. In this case, an edge LSR will deal with LVC set-up failure by sending the ordinary IP traffic on the default (0,32) LVCs, which are otherwise used for TDP/LDP signalling. This traffic will then be forwarded around the resource-starved links by the LSC processors, and not the ATM switch fabrics. This will not affect network stability, provided that MPLS Class of Service is used to give precedence to routing and TDP/LDP traffic. However if large amounts of user traffic are sent on the (0,32) LVCs, then this user traffic will experience poor performance when it exceeds the LSCs’ packet forwarding capacity.

The second case applies to advanced MPLS services such as VPNs. In this case, the packets which would otherwise use the failed paths are discarded. For these services, it is impossible to correctly deliver a packet unless it is fully labelled. This discarding of packets to certain destinations results in routing “black holes”.

In case of LVC starvation, it is usually impossible to predict which LVCs get created and which don’t. This is because LVCs are created roughly in the order in which routing converges. This is, in turn, dependent on random factors such as which links fail, and the exact timing of link failures compared to routing protocol update timers.

- Adjust routing
- Check for sufficient LVC allocation on links

4. Migration of MPLS into Traditional ATM Networks

4.1 Hybrid ATM Networks

Sometimes it is necessary to connect ATM MPLS devices over traditional ATM equipment, in a 'hybrid ATM network'. This is undesirable, as many of the disadvantages of traditional IP-over-ATM schemes, as discussed in "1.5 MPLS and Other IP-over-ATM Schemes" on page 7, are re-introduced in hybrid ATM networks. In particular:

- Hybrid ATM networks are inefficient at providing IP services such as multicast.
- Hybrid ATM networks have routing scaling problems similar to those of IP-over-PVC networks.
- Hybrid ATM networks are more difficult to manage than pure MPLS networks.

However hybrid ATM networks can support MPLS services such as Virtual Private Networks, and may be necessary as a migration step while introducing a full MPLS network.

Hybrid networks are not the same as IP+ATM networks. In IP+ATM networks, all ATM switches have both MPLS capability and traditional ATM switch capability. In hybrid networks, some of the switches do not have MPLS capability. Some of the switches in hybrid networks may have MPLS capability, and these particular switches may be IP+ATM switches.

There are two broad ways of implementing hybrid ATM networks: MPLS-over-PVCs and VP Tunnels. These are illustrated in Figure 12. MPLS-over-PVCs can be used only to connect packet-based MPLS devices. It may not be used to connect ATM Label Switch Routers (ATM-LSRs). MPLS-over-PVCs connects packet-based Label Switch Routers (LSRs) by way of PVCs over a traditional ATM network. 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 a 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. MPLS-over-PVCs does not use ATM MPLS at all. This means that service providers must continue to provision and manage PVCs on a scale equal to the traditional IP over ATM approach. However MPLS-over-PVCs does have the advantages of simplicity, and of using more stable IOS code than VP shaping.

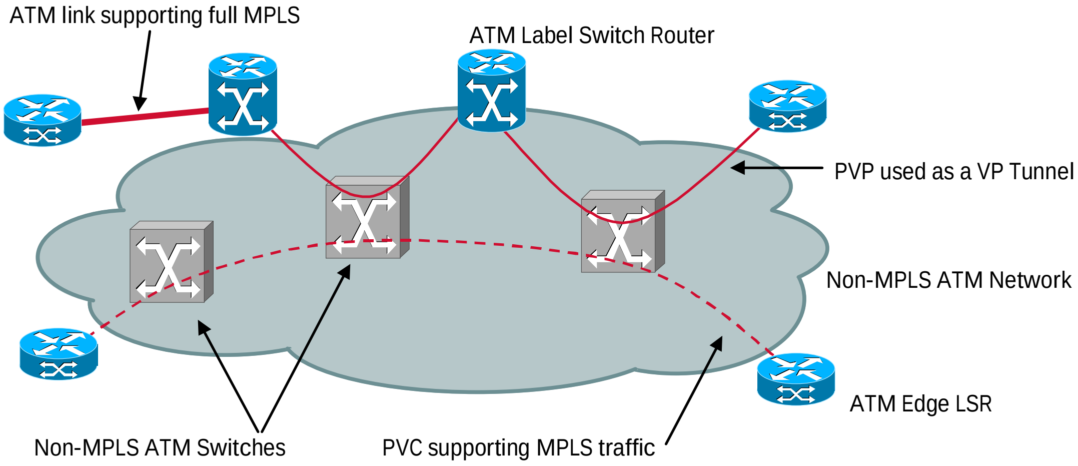


Figure 12 VP Tunnels and MPLS-over-PVCs

VP Tunnels

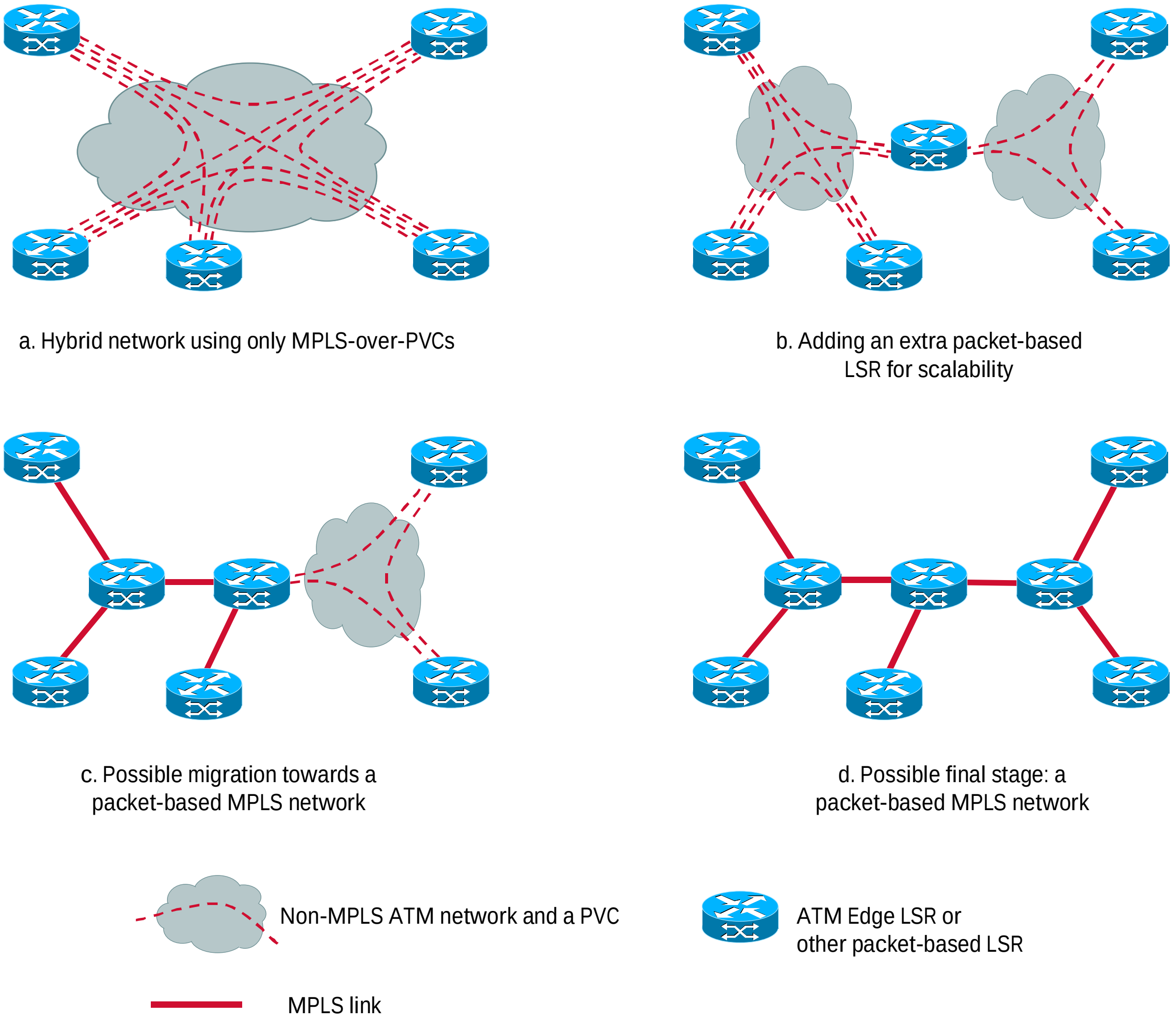
The remaining method of implementing hybrid ATM networks is the use of Virtual Path (VP) Tunnels. ATM MPLS normally involves labelling IP packets by putting them in different VCs in the same ATM trunk. Each different VC on the trunk represents a different label value. VP Tunnels are a form of 'Virtual Trunk'. ATM MPLS devices treat them almost identically to a physical trunk:

each different VC within the VP represents a different label value. The difference is that the VP Tunnel is not a physical trunk linking two adjacent ATM-LSRs. The VP Tunnel is a Permanent Virtual Path Connection (PVP) which connects ATM-LSRs by way of traditional ATM switches. VP Tunnels may also connect ATM Edge LSRs to ATM-LSRs, or connect ATM Edge LSRs to each other.

Configurations for hybrid networks, and consequent equipment requirements, are considered in the following sections.

4.2 Networks Using MPLS-over-PVCs

The simplest network structure using MPLS-over-PVCs is the full mesh shown in Figure 13(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—refer to the discussion in “1.5 MPLS and Other IP-over-ATM Schemes” on page 7. 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 13(b), or possibly a redundant pair of them. The extra ATM Edge LSRs reduce the size of the meshes, but often won’t have customer-facing interfaces. 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.



a. Hybrid network using only MPLS-over-PVCs

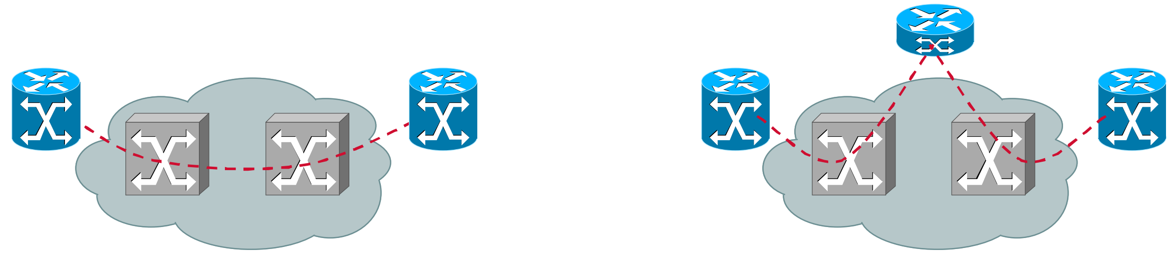
b. Adding an extra packet-based LSR for scalability

c. Possible migration towards a packet-based MPLS network

d. Possible final stage: a packet-based MPLS network

Figure 13 ATM MPLS networks using MPLS-over-PVCs

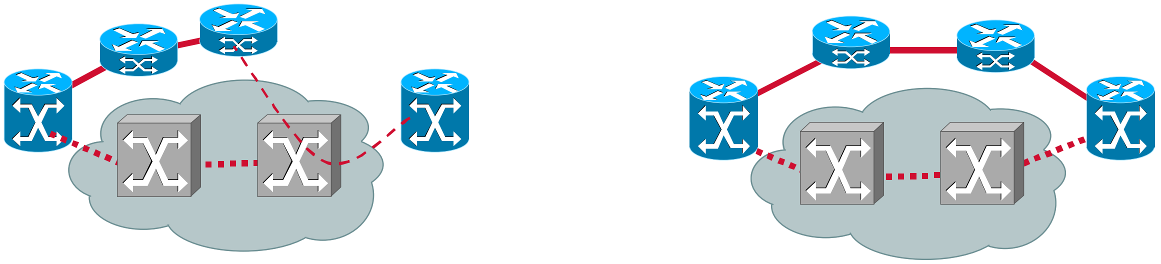
Some providers might prefer to build an infrastructure for their MPLS traffic which is separate from their traditional ATM network, in a ‘dual backbone’ configuration. This MPLS network might use ATM MPLS. Alternatively, it might use packet-based MPLS, with packet-based LSRs and links such as PPP-over-SDH. 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 Figure 13(c) and (d).



a. Hybrid network using only MPLS-over-PVCs

b. Adding an extra packet-based
LSR for scalability

a. Hybrid network using only MPLS-over-PVCs



c. Possible migration towards a
dual-backbone network

d. Possible final stage: a dual-backbone network
with no traditional ATM in the MPLS forwarding path

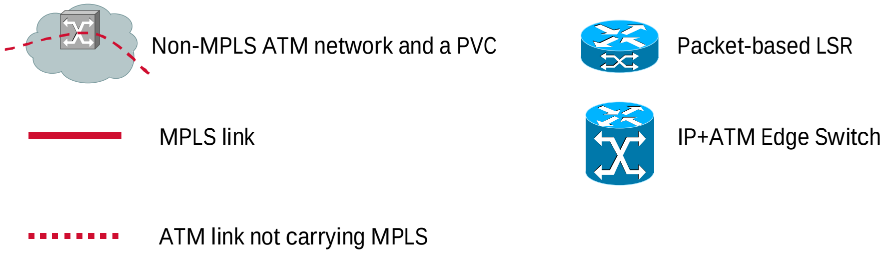


Figure 14 PVC-based migration to dual backbones: another view

Another viewpoint of migration to dual backbones is shown in Figure 14. In this example, the Edge LSRs are part of IP+ATM edge devices. The stages of migration in Figure 14 are the same as in Figure 13, but Figure 14 shows more clearly how the migration leads to the building of a new packet-based backbone alongside the existing traditional ATM backbone. If IP+ATM edge switches are used, then the same switches can connect to both the traditional ATM backbone and the MPLS backbone. Figure 14 shows an example where the new backbone uses packet-based MPLS, but it is also possible to use ATM MPLS in the second backbone.

Equipment for MPLS-over-PVCs

The core of an MPLS-over-PVCs network is a plain old ATM network which needs only to support PVCs. Virtually any ATM network can be used. Virtually any MPLS release on Cisco Edge LSRs can be used with MPLS-over-PVCs. Note that if MPLS-over-PVCs is being used, a mode of MPLS Class-of-Service using PVCs instead of LVCs, and based on Project Paris, may also be used. The “Recommended Configuration” column in Table 12 is for MPLS-over-PVCs with Class-of-Service support. Earlier releases (e.g. 11.1CT(23)) without Class-of-Service support may also be used.

Table 12 Cisco ATM Edge LSRs and MPLS-over-PVCs

Device	Minimum configuration	Recommended configuration	PVC traffic types supported with recommended configuration		
			CBR[a]	VBR	ABR
Catalyst 5500	Route Switch Module (RSM) with PA-A1 ATM Port Adaptor. IOS 11.3(5)WA4(8b).	PA-A3 ATM port adaptor, IOS 12.OT(5), or later.	✔	✔	✔
6400	Node Route Processor. IOS 12.0(6)DC	Minimum or later.	✔	✔	✔
7200 series routers	PA-A1 port adaptor, IOS 11.1CT(20) or later.	PA-A3 ATM port adaptor, IOS 12.0T(5) or 12.0T(6), or later.	✔	✔	✔
7500 series routers	PA-A1 port adaptor, IOS 11.1CT(20) or later.	PA-A3 ATM port adaptor, IOS 12.0T(5) or 12.0T(6), or later.	✔	✔	✔
MGX 8850 IP+ATM switch	Route Processor Module (RPM). IOS 12.OT(5.0.2).	Minimum or later.	✔	✔	✔
12000 series routers	IOS 12.OS(5)	Minimum or later	✔	✘	✘

[a] This means that the shaper will shape to a peak rate. The resulting shaped cell stream is guaranteed to fit within a CBR PVC, and does comply with the standard policing parameters for CBR. Note, however, that the actual rate of the stream varies up to the peak rate. In some Cisco product literature, this has been referred to incorrectly as ‘VBR’.

Using Traditional ATM Switches for Backhaul

Figure 15 shows another type of network which may occur in practice. Figure 15 shows a network which is not strictly a hybrid network, as the MPLS traffic is carried only across MPLS devices, and not on traditional ATM switches. In this network, the traditional ATM switches are used to connect the CPE into the edge LSRs, by way of ‘backhaul’ PVCs. These backhaul PVCs carry traditional IP-over-ATM, and do not carry MPLS traffic. This was briefly discussed “2.1 Structures for MPLS Networks” on page 9, and shown in Figure 2(f). Other variations are possible. For example, a network similar to Figure 15 could be used, but with VP tunnels or MPLS-over-PVCs in the network core, as well as the backhaul PVCs.

4.3 Networks Using VP Tunnels

A simple way of using VP Tunnels is use them to connect ATM Edge LSRs without using any ATM-LSRs at all in the network, as illustrated Figure 16(a). This means that all MPLS packets are carried over VP Tunnels, 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 MPLS, and some won’t. VP tunnels 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 16(b).

Migration to Full MPLS

Figure 16(a), (b) and (c) show a possible migration process for introducing MPLS to a traditional ATM network:

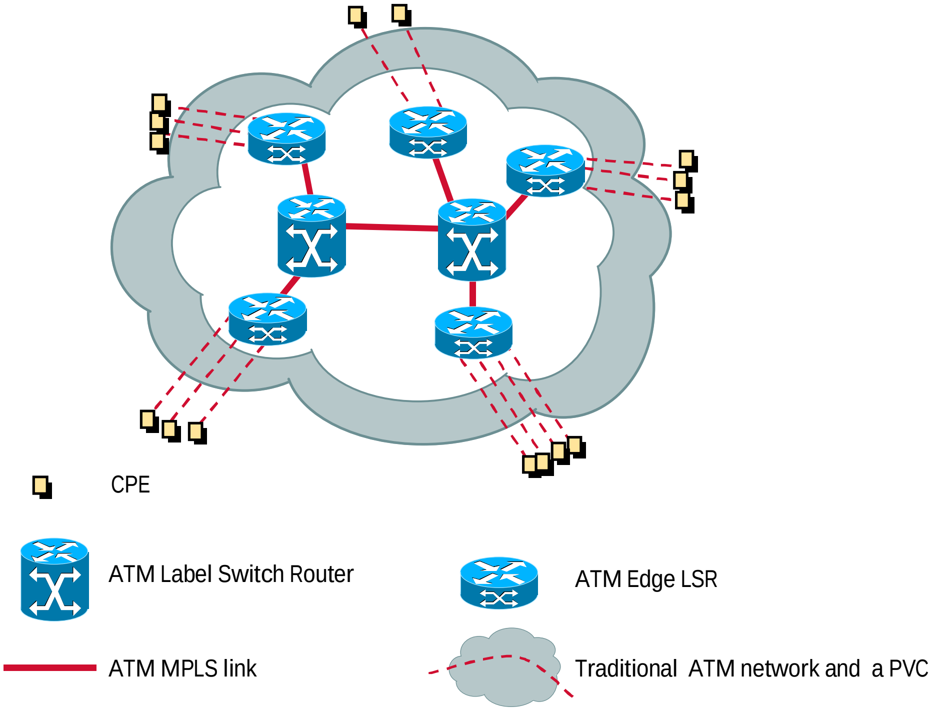


Figure 15 Network using backhaul PVCs

- MPLS edge routers ('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 VP Tunnels required, and starts to reduce some of the scalability problems of hybrid networks.
- More ATM-LSRs are added, which further reduces the number of VP tunnels, and starts to introduce native ATM MPLS links this is shown in Figure 16(c). This step naturally leads to the final one.
- Ultimately, all ATM switches are ATM-LSRs, and no VP Tunnels are used at all. The full network runs ATM MPLS, and has none of the disadvantages of hybrid networks. This is illustrated in Figure 16(d).

Other Variations of Hybrid Networks

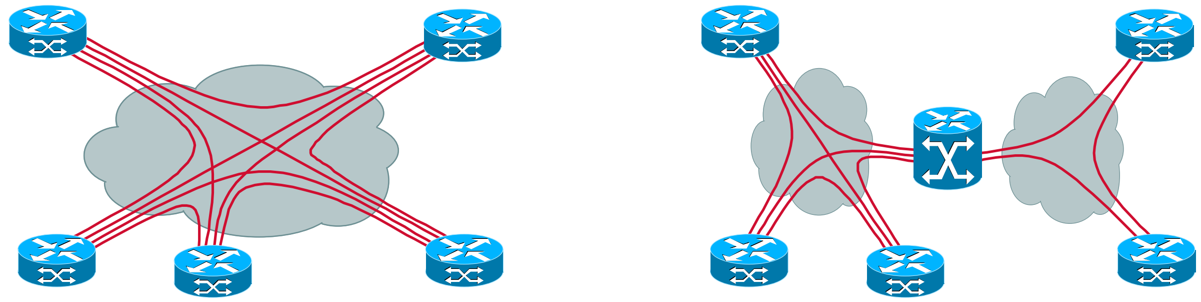
MPLS devices and traditional ATM switches can be combined in many different ways. Figure 16 shows a few examples. Figure 17 shows some other hybrid network structures which may occur. Many other hybrid network structures are possible. An ATM MPLS network must include edge LSRs, but may use any nearly any combination of zero or more ATM-LSRs and zero or more traditional ATM switches with VP tunnels.

4.4 Requirements to Support VP Tunnels

VP Tunnels are implemented using PVPs, or alternatively PNNI Soft Permanent Virtual Paths (SPVPs). Use of VP Tunnels will generally require the PVPs or SPVPs to be CBR, as shaping a VP to other than a peak rate is difficult.

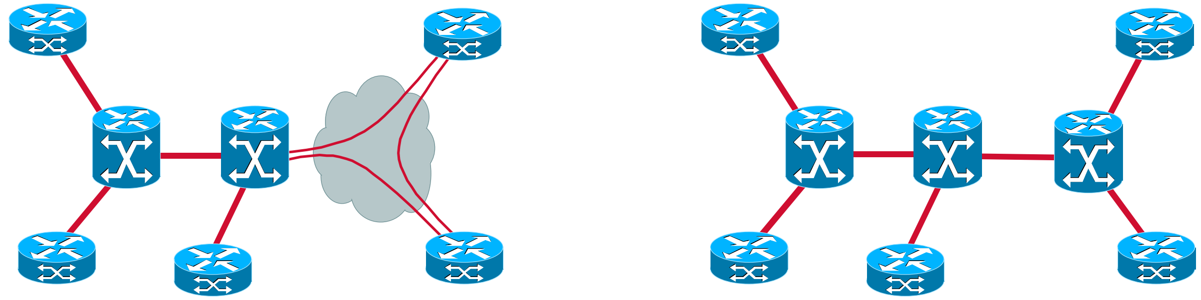
Edge LSRs

ATM Edge LSRs must meet the following requirements in order to support VP Tunnels:



a. Hybrid network using only VP Tunnels

b. Adding an ATM-LSR



c. Further simplification by adding more ATM-LSRs

d. Full ATM MPLS network

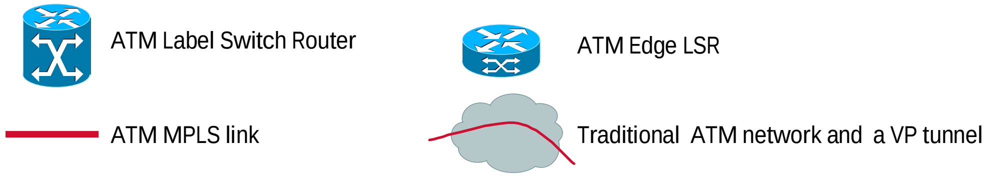
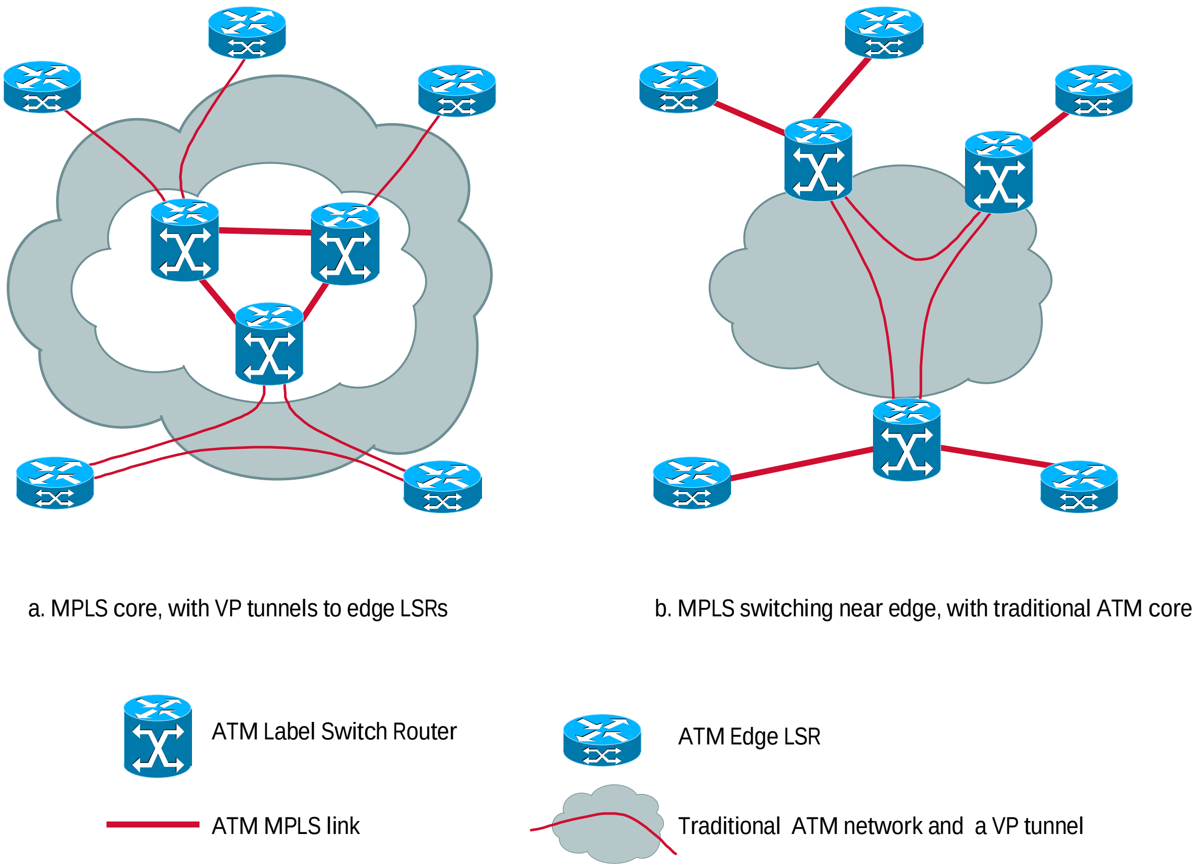


Figure 16 ATM MPLS networks using VP Tunnels

- They must be able to support distinct logical MPLS links in different sub-interfaces of a physical ATM link. If this requirement is not met, each Edge LSR in would require a separate physical ATM interface for each VP Tunnel. This would create scalability problems. For example, each Edge LSR in Figure 16(a) would require four ATM links, whereas each Edge LSR in Figure 16(d) would require only one (or two if it was dual-homed). If a particular VP tunnel uses a VPI *x* at the Edge LSR, then the LDP signalling VC for the VP Tunnel must be within *x*. It may have VPI=*x*, VCI=32, instead of the normal default VPI=0, VCI=32 for LDP signalling.
- They must support shaping of packet traffic entering the VP tunnel to the traffic parameters (e.g. a CBR peak rate) of the VP tunnel. It is highly preferable that this be done using Early Packet Discard (EPD), or at least Partial Packet Discard (PPD). The alternative to EPD or PPD is Cell Loss Multiplication, which results in a horrible congestion collapse characteristic for the VP Tunnel.


a. MPLS core, with VP tunnels to edge LSRs

b. MPLS switching near edge, with traditional ATM core

Figure 17 Other hybrid network examples

Cell loss multiplication occurs when a whole packet must be discarded at reassembly because only one cell was lost in transit. For example, assume that during a certain timeframe it is necessary to discard 100 cells at ingress to a VP tunnel due to overload. Assume that the mean packet size is 5 cells. Without EPD or PPD, the 100 cells may well be dropped from 100 different packets. This would lead to the 100 packets being discarded at reassembly. This would cause an effective loss of 500 cells when only 100 were discarded. This is cell loss multiplication. EPD ensures that if any cell in a packet is dropped, all cells are dropped. This will concentrate the cell loss in around 20 packets[1], avoiding cell loss multiplication entirely. PPD ensures that once one cell in a packet is dropped, all remaining cells in the packet will be dropped, although some will still have been accepted. This will concentrate the cell loss in about 40 packets[2], which is a reduction but not elimination of cell loss multiplication.

Capabilities of Cisco Edge LSRs are shown in Table 13.

---

[1] This is an approximation, based on the assumption that the queueing behavior at ingress to the VP tunnel is such that the queue is never completely empty. It is also assumed that the number of cells to be dropped (100) is invariant with the discard policy.

[2] In addition to the previous assumptions, this approximation assumes that packet loss first occurs at a random cell in each packet. The exact number depends on the exact traffic characteristics, the exact distribution of packet sizes, and the exact queueing behavior during the overload.

Table 13 Cisco ATM Edge LSRs and VP Tunnels

Device	Minimum software & hardware for edge LSR support for VP Tunnels	VP Tunnel traffic types supported			Smart discard for VPs		Capabilities & limitations
		CBR[a]	VBR	ABR	EPD	PPD	
Catalyst 5500	IOS 12.OT(5), Route Switch Module (RSM) with PA-A3 ATM port adaptor.	✔	✘	✘	✔		256 VP Tunnels are supported on a PA-A3 card. The total number of LVCs is limited to 4096, irrespective of the number of VP Tunnels used.
6400	Any software supporting MPLS	✔	✘	✘	✘	✘	The 6400 will also act as an ATM switch and ATM-LSR. The Edge LSR function in the 6400 can support VP tunnels only in combination with the 6400's ATM-LSR function, as the edge LSR(s) rely on the ATM-LSR to perform VP shaping. See "Edge LSRs Integrated into Edge Switches" on page 50
7200 series routers	IOS 12.0T(4), PA-A3 ATM port adaptor	✔	✘	✘	✔		256 VP Tunnels are supported on a PA-A3 card. The total number of LVCs is limited to 4096, irrespective of the number of VP Tunnels used.
7500 series routers	IOS 12.0T(4), PA-A3 ATM port adaptor	✔	✘	✘	✔		256 VP Tunnels are supported on a PA-A3 card. The total number of LVCs is limited to 4096, irrespective of the number of VP Tunnels used.
MGX 8850 IP+ATM switch	IOS 12.OT(5.0.2)	✔	✘	✘	✔		256 VP Tunnels are supported on an RPM card. The total number of LVCs is limited to 4096, irrespective of the number of VP Tunnels used. Note that the MGX 8850 can also act as an ATM switch and ATM-LSR.
12000 series routers	VP Tunnels are not currently supported.	✘	✘	✘	✘	✘	VP Tunnels are not currently supported due to lack of VP shaping capability. This will be fixed in later port card hardware.

[a] This means that the shaper will shape to a peak rate. The resulting shaped cell stream is guaranteed to fit within a a CBR PVP.

Traditional ATM Switches

The switches in the ATM clouds must support PVP or SPVP connections with ATM Forum or ITU traffic management types which match those used on the Edge LSRs.Capabilities of Cisco switches are shown in the following table. *Note that this table applies only when the switches are being used as traditional ATM switches.* Most of the switches shown below can also acts as ATM-LSRs. If non-Cisco ATM switches are used, refer to the manufacturer's documentation for those switches.

Table 14 Using Cisco ATM switches as traditional ATM switches.

Device	PVP Traffic Types Supported			Capabilities & Limitations
	CBR	VBR	ABR	
LS1010 & Catalyst 5500	✔	✔	✔	Supports up to 256 PVPs per trunk, or 8k per switch.
6400	✔	✔	✔	Supports up to 256 PVPs per trunk, or 8k per switch
IGX 8400 Series	✔	✔	✔	Supports up to 4094 PVPs per trunk, with a current maximum of 7000 PVPs per switch.
8540 MSR	✔	✔	✔	Supports up to 256 PVPs per trunk, or 24k per switch.
BPX 8600 Series	✔	✔	✔	Supports up to 4094 PVPs per trunk, with a current maximum of 12,000 PVPs per switch.
MGX 8850 IP+ATM switch	✔	✔	✔	Supports up to 4096 PVPs per trunk, with a current maximum of 12,000 PVPs per switch.
Proposed MGX 8800 with PXM-45	✔	✔	✔	Supports up to 4096 PVPs per trunk, with a maximum in excess of 12,000 PVPs per switch.

ATM-LSRs

In order to support VP Tunnels, ATM-LSRs must meet the same requirements as ATM Edge LSRs:

- They must be able to support distinct logical MPLS links in different sub-interfaces of a physical ATM link. If this requirement is not met, then the ATM-LSR would require a separate physical ATM interface for each VP Tunnel. This would create severe scalability issues. For example, the ATM-LSR in Figure 16(b) would require five ATM links, whereas the ATM-LSRs in Figure 16(d) would require three. If a particular VP tunnel uses a VPI x at the ATM-LSR, then the LDP signalling VC for the VP Tunnel must be within x. It may have VPI=x, VCI=32, instead of the normal default VPI=0, VCI=32.
- They must support shaping of packet traffic entering the VP tunnel to the traffic parameters (e.g. a CBR peak rate) of the VP tunnel. It is highly preferable that this be done using Early Packet Discard (EPD), or at least Partial Packet Discard (PPD).

Table 15 Cisco ATM-LSRs

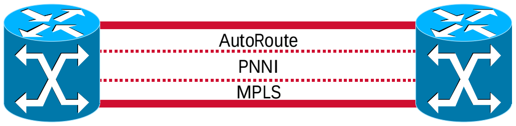
Device	Minimum Software & Hardware for VP Tunnels	VP Tunnel Traffic Types Supported (With Shaping)			Smart Discard for VPs		Capabilities & Limitations
		CBR[a]	VBR	ABR	EPD	PPD	
LS1010	Per-Flow Queueing Feature Card Required (FC-PFQ)). IOS 11.3(5)WA4(8b) or later.	✔	✘	✘	✘	✘	A maximum of 64k LVCs are supported per switch, irrespective of the VP Tunnels it supports. 64 VP Tunnels are supported.
6400	Any supported hardware, IOS I12.0(6)Db	✔	✘	✘	✘	✘	A maximum of 64k LVCs are supported per switch, irrespective of the VP Tunnels it supports. 128 VP Tunnels are supported.
8540 MSR	Any supported software release.	✔	✘	✘	✘	✘	A maximum of 256k LVCs are supported per switch, irrespective of the VP Tunnels it supports. 128VP Tunnels are supported.
BPX 8600 Series	BXM cards with firmware MEB, Switch Software 9.2.50	✔	✘	✘	✔		A maximum of 32k LVCs are supported per BXM card, irrespective of the number of ports or VP Tunnels it supports. Each BXM card supports at least 20 VP Tunnel endpoint s (12-port T3) and may support up to 31. If a BXM card is terminating MPLS VP tunnels on a particular BXM trunk, it may support no other type of connection on that BXM trunk except virtual trunks.
MGX 8850 IP+ATM switch	VP tunnels are supported on PXM-1B cards, when an LSC is used[b]. There is also a dependency of an LSC interface to the PXM-1 cards, which may not occur until H2 CY 2000, as LSC support for the PXM-45 is a higher priority.	✔	✘	✘	✔		A maximum of 16k LVCs are supported per MGX 8850 in this configuration, irrespective of the number of ports or VP Tunnels it supports. Each PXM card supports at least 28 VP Tunnel endpoint s and may support up to 31. If an MGX8850 is terminating MPLS VP tunnels on a particular PXM trunk, it may support no other type of connection on that PXM trunk.
Proposed MGX 8800 with PXM-45	AXSM cards, software and firmware to be determined. Target support timeframe is Q2 CY 2000	✔	✘	✘	✔		A maximum of 128k LVCs are supported per AXSM card, irrespective of the number of ports or VP Tunnels it supports. Each AXSM card supports at least 49 VP Tunnel endpoint s (16-port E3/T3 or OC3) and may support up to 64. If an AXSM card is terminating MPLS VP tunnels on a particular trunk, it may support no other type of connection on that trunk except virtual trunks.

[a] This means that the shaper will shape to a peak rate. The resulting shaped cell stream is guaranteed to fit within a a CBR PVP, but it is actually a VBR stream with indeterminate Sustained Cell Rate (SCR).

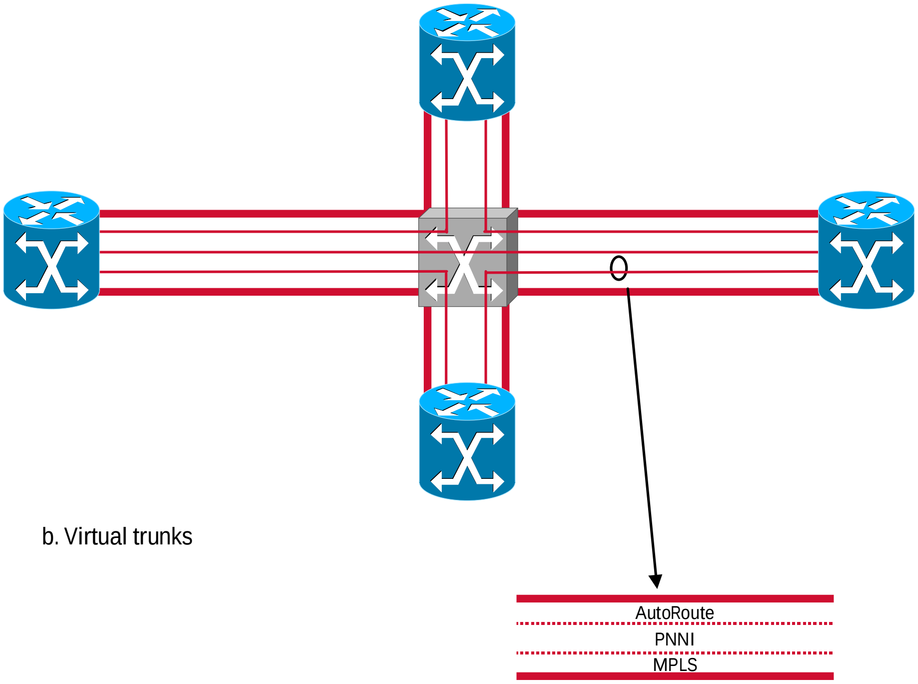
[b] If an LSC is not used, then VP tunnels can still terminate on the RPM cards. If an LSC is not used, then the MGX 8850 is acting as separate edge LSRs (Table 13) in combination with a traditional ATM switch (Table 14).

VP Tunnels and the SCI

The Virtual switch Interface (SCI) was described in “2.3 IP+ATM” on page 14. Recall that part of the configuration of SCI involves creating partitions of the bandwidth and VPI/VCI space on trunks. This is shown in Figure 18(a). MPLS VP Tunnels are a type of Virtual Trunk. Virtual trunks are PVP or SPVP connections which carry VCs through switches that don’t support that type of VC. As well as MPLS, virtual trunks can tunnel PNNI SVCs through switches that do not support them. In addition they can be used to connect a customer’s virtual circuits of any type, through a carrier’s switches.



a. Physical trunk between two IP+ATM switches



b. Virtual trunks

c. Virtual trunks between adjacent nodes

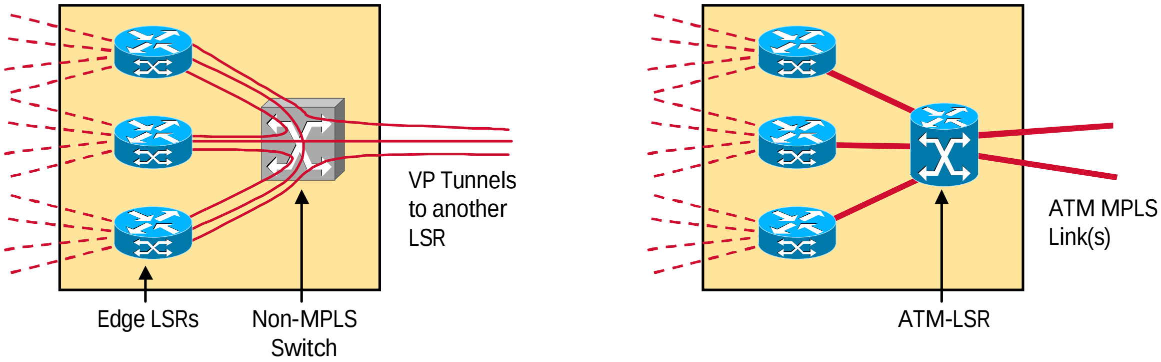
Figure 18 SCI partitions and virtual trunks

In the BPX 8650 and MGX 8850 implementation of virtual trunks, each virtual trunk will be able to have several SCI partitions as if it were a physical trunk—see Figure 18(b). This may be supported from BPX software release 9.3 onwards. In practice, the number of virtual trunks supported per card in the BPX 8650 and MGX 8850 is limited by the number of Virtual Interfaces supported by the card. A Virtual Interface is a physical device which shapes traffic and provides per-VC queueing. There are a limited number of Virtual Interfaces per card. Each BXM, PXM or UXM card has 31 Virtual Interfaces, and an AXSM has 64. They are used as follows:

- Each physical interface which does not support virtual trunks, uses one Virtual Interface. For example, the switch ports in Figure 18(a) use only one Virtual Interface each, despite having three partitions each.

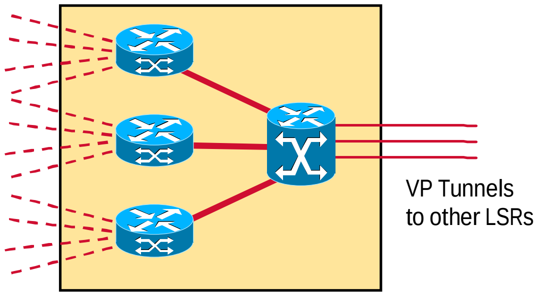
- Each endpoint of a virtual trunk uses a Virtual Interface. For example, the interfaces shown in Figure 18(c) use three Virtual Interfaces each.

Note that there are generally several SCI partitions on each Virtual Interface. Virtual Interfaces are generally not used to shape individual VSI partitions. This is because a card can support very many more partitions than Virtual Interfaces. For example, a 16-port OC3 card on an MGX 8800 platform will eventually support 16 partitions on each of its 16 interfaces, for a total of 1024 partitions. However it supports only 64 Virtual Interfaces. Note that from the limits described above, this card will support 49 virtual trunk endpoints if all endpoints are on one physical interface, and up to 64 endpoints if they are distributed over several interfaces.



a. Edge LSR/switch without MPLS switching capability (supported on Catalyst 5000 or early-release MGX 8850, but not 6400))

b. Edge LSR/switch with ATM-LSR capability (soon on 6400 and MGX 8850, but not Catalyst 5500)



c. Terminating VP Tunnels on the ATM-LSR (soon on 6400 and MGX 8850, but not Catalyst 5500)

Figure 19 Devices combining Edge Label Switch Routers and an ATM switch

Edge LSRs Integrated into Edge Switches

Edge LSR function may be integrated into an ATM access switch, as shown in Figure 19. The Cisco MGX 8850, Catalyst 5500, and Cisco 6400 are examples of such devices. An Edge LSR/switch may include one or more logically separate edge LSRs, distinct from the switch component. In the Cisco MGX 8850, the edge LSRs are Route Processor Module (RPM) cards, in the Catalyst 5500 they are Route Switch Module (RSM) cards, and in the 6400 they are Node Route Processor (NRP) cards. There are several possible combinations of MPLS function and VP Tunnels in such a device.

- Edge LSRs Only:

  If the switch component cannot perform MPLS switching, then VP Tunnels (or MPLS-over-PVCs) must be used to connect the Edge LSRs to other Edge LSRs, as shown in Figure 19(a). VP Tunnels will also be used to inter-connect the Edge LSRs in the same device. The three Edge LSR components shown are equivalent to three of the Edge LSRs in Figure 16(a), except that they are housed in a single chassis. Note that each Edge LSR component would require a separate VP Tunnel to connect to another ATM MPLS device. The VPs shown in Figure 19(a) are sufficient to connect to only one other ATM MPLS device. More would be used as necessary. The 6400 can use this combination of functions only if MPLS-over-PVCs is used, as the NRPs cannot directly support VP tunnels.

- Full LSR plus Edge LSR:
  If the switch component has MPLS control, then the Edge LSR/switch can act as a full ATM-LSR as well as Edge LSR(s). This is shown in Figure 19(b). In the MGX 8850, ATM-LSR capability will be enabled by using one of the RPMs as a Label Switch Controller (LSC) for the switch; in the 6400, the MPLS control function resides on the Node Switch Processor. An RPM in the MGX 8850 acting as an LSC may concurrently act as an Edge LSR. The ATM-LSR component interconnects the Edge LSRs using full MPLS, so VP Tunnels are no longer required within the device. Figure 19(b) shows the case where the Edge LSR/switch connects to the rest of the network via ATM MPLS links. An LSR/Edge LSR performs the functions of several of the devices in Figure 16(d) or the left side of Figure 16(c). Note that the Catalyst 5500 does not support this mode, as it cannot act as a full ATM-LSR.
- Full LSR/Edge LSR connected with VP Tunnels:
  If necessary, the full LSR/Edge LSR device can be connected by VP Tunnels to other ATM MPLS devices. This is shown in Figure 19(c). Even though VP Tunnels are still used, having internal MPLS switching simplifies the VP Tunnel mesh. For example, the three external VP Tunnels shown in Figure 19(a) are sufficient to connect to only one other ATM LSR or Edge LSR, but the VP Tunnels shown in Figure 19(c) are sufficient to connect to three other devices. This applies only to the 6400 and MGX 8850—the Catalyst 5500 does not support this.

The requirements on the components within the LSR/Edge LSR are similar to the requirements for stand-alone devices discussed earlier. In particular:

- If the Edge LSR components (RPMs or RSMs) terminate VP Tunnels as in Figure 19(a), then they must meet the requirements discussed in “Edge LSRs” on page 44.
- If the switch component does not have MPLS capability, i.e. it acts as shown in Figure 19(a), then it must meet the requirements discussed in “Traditional ATM Switches” on page 47.
- If the switch component has MPLS capability but supports VP Tunnels, i.e. it acts as shown in Figure 19(c), then it must meet the requirements discussed in “ATM-LSRs” on page 48.

VP Tunnel Restrictions and Work-Arounds on the BPX 8650

The BPX 8650 has certain restrictions on VP Tunnel use, which may be worked-around. These restrictions do not apply to the MGX 8850 or any other Cisco ATM switch. The restrictions apply to the use of an ATM interface in a BPX 8650. An ATM interface can be configured in several distinct modes:

- In either “port” or “trunk” mode, the interface can support any combination of AutoRoute-controlled PVCs and PVPs; SVCs, SPVCs and SPVPs controlled by the Service Expansion Shelf[3]; and MPLS LVCs. The interface cannot shape VP Tunnels, but it can carry the mid-points of VP Tunnels (because the middle of a VP Tunnel is just a PVP connection).
- In “AutoRoute Virtual Trunk” mode, the interface can terminate and shape Virtual Trunks which carry PVC connections only. The interface or the virtual trunks may not carry MPLS or PNNI connections. The interface may not carry PVC connections unless they are within the Virtual Trunks. The interface may not carry PVPs unless the PVPs are the Virtual Trunks which terminate on this interface.
- In “VSI Virtual Trunk” mode, the interface can terminate and shape Virtual Trunks and VP Tunnels which carry MPLS LVCs and PNNI SVCs & SPVCs. The interface may not carry PVCs at all, and may not carry any connections except those inside the Virtual Trunks. The interface may not carry PVPs unless the PVPs are the Virtual Trunks which terminate on this interface.

---

[3] From BPX software release 9.2.50 onwards

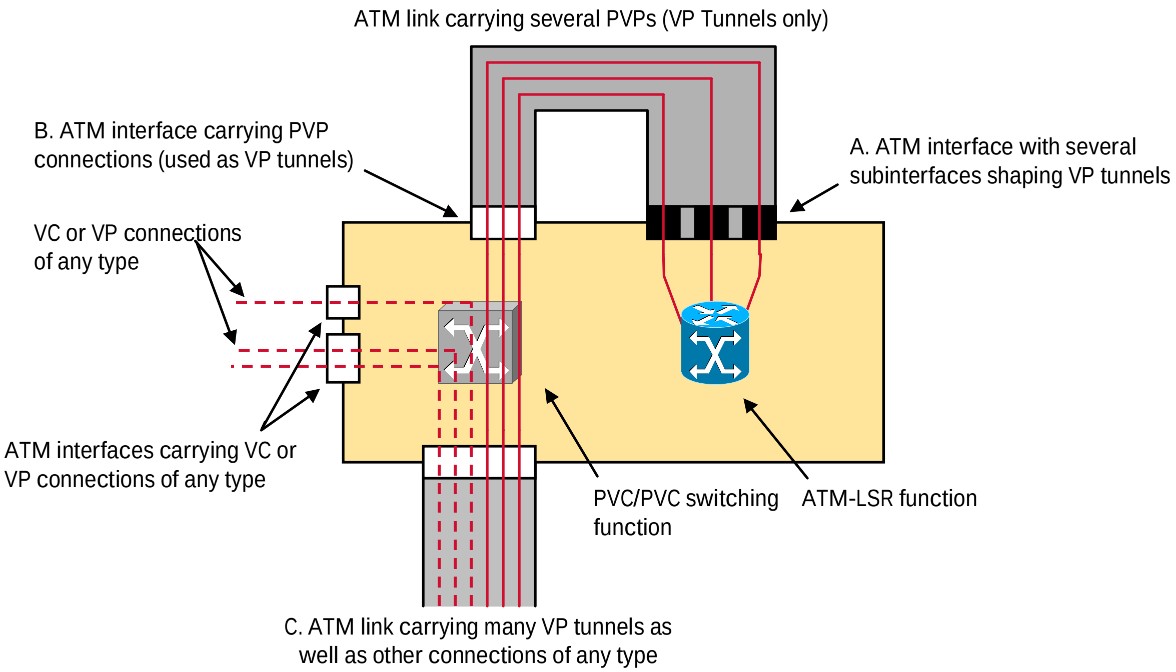


Figure 20 VP tunnels alongside ordinary PVC & PVP in the BPX 8650

These restrictions could prevent the BPX 8650 from carrying Virtual Trunks and VP Tunnels on the same links as ordinary AutoRoute, PNNI and MPLS (S)PVCs, (S)PVPs, SVCs and LVCs. However there is a work-around, as shown in Figure 20.

In the work-around, interface A is configured as a VSI Virtual Trunk interface, and it terminates several VP Tunnels. This is the normal way of terminating VP Tunnels on a BPX 8650. However the link connected to A is looped back to another interface B. B is configured as a normal Port, and it carries several PVP connections. The PVP connections are the VP Tunnels. B has no special configuration for the VP Tunnels—it treats them as ordinary PVPs. The VP Tunnels are switched, as ordinary PVPs, through to interface C. Interface C is an ordinary port or trunk, and it can carry any combination of AutoRoute, PNNI and MPLS connections, both inside and outside the VP Tunnels.

Many variations of this work-around are possible. For example, the VP Tunnels can be distributed onto several different interfaces, and not just a single interface C. Interface A might be on a different switch to interfaces B and C.

Use of Inverse Multiplexing over ATM

Inverse Multiplexing over ATM (IMA) trunks may be used in the traditional ATM parts of the network. For example, any of the links in the ‘clouds in Figure 13 or Figure 16 may be IMA links. This allows IMA to be used in an MPLS network, even if the ATM LSRs do not directly support IMA.

4.5 Examples of Hybrid ATM Network Equipment

MGX 8850 Edge LSRs connected over a big cloud

It is possible to run an MGX 8850 MPLS network without using ATM-LSRs at all. There are two ways to do this: MPLS-over-PVCs and VP Tunnels:

- MPLS-over-PVCs: Refer to Figure 13(a). For the edge LSR, RPMs in MGX 8850 IP+ATM switches will be used. These are simply connected by PVCs, and ATM-LSRs (and LSCs) are not used in the network.

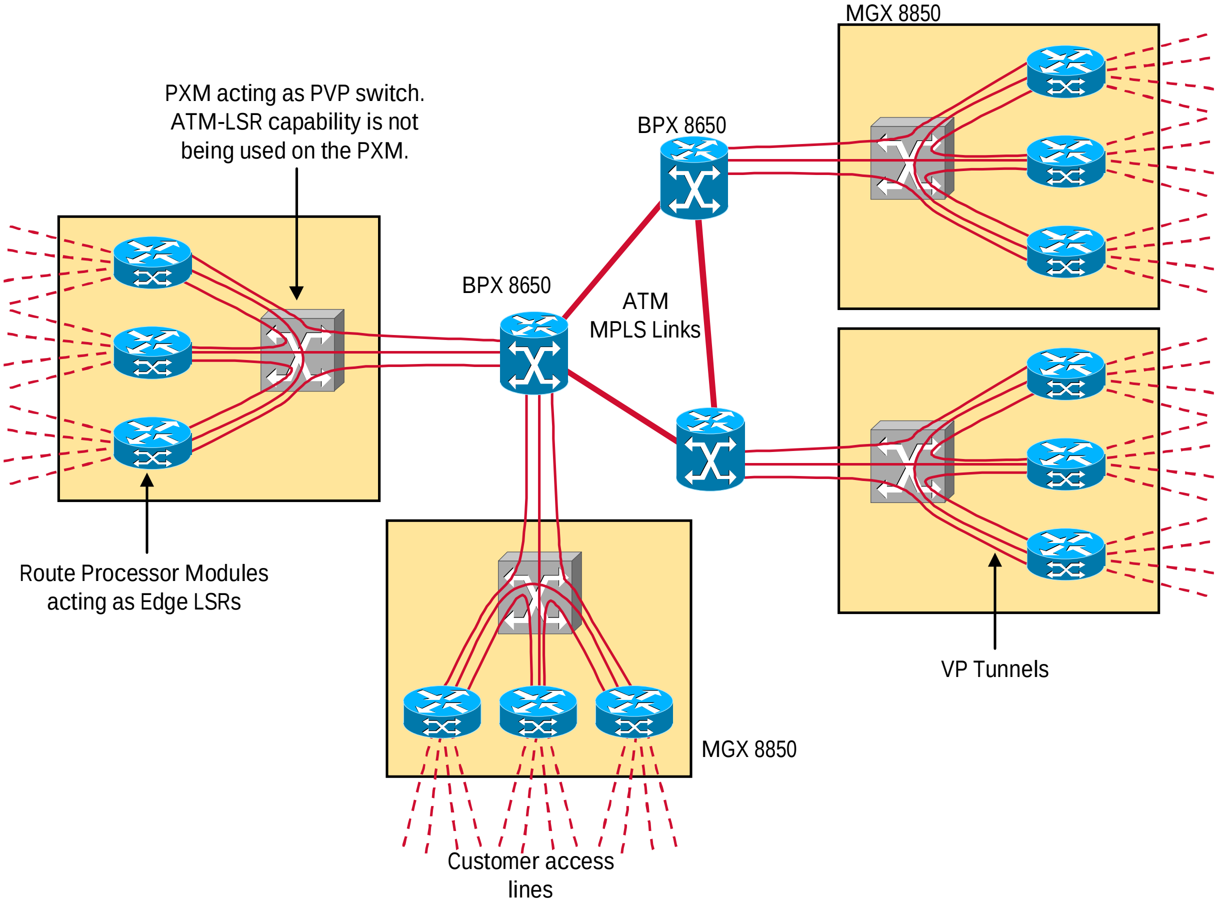


Figure 21 Deployment of a network with MGX 8850s and core ATM-LSRs, but without LSCs on the MGX 8850

- VP Tunnels: Refer to Figure 16(a) and assume that the edge LSRs are RPMs in MGX 8850 IP+ATM switches. According to Table 13, this will be supported with up to 256 VP Tunnels per RPM. The ATM cloud must support CBR PVPs or SPVPs.

MGX 8850 Deployments Without LSCs in the MGX 8850s

A more scalable deployment mechanism for MGX 8850s without LSCs is shown in Figure 21. Here, ATM-LSRs, e.g. BPX 8650s, are chosen for the core of the network. VP tunnels are used at the edge of the network to connect RPMs to each other and to the BPX 8650s. Because the VP tunnels are kept short, and kept at the edge of the network, the network is quite scalable. The BPX 8650 is a quite suitable ATM-LSR for this application, because of the excellent VP Tunnel support as shown in Table 15.

4.6 VP Tunnels and LVC Usage

Dimensioning of the number of MPLS Label VCs required per link is described in “3.4 Dimensioning MPLS Label VC space” on page 30. This dimensioning requires some modification to deal with VP Tunnels. When VP Tunnels are terminate on an interface, the LVCs on all VP Tunnels must be taken into account. For example, if 4 VP Tunnels terminate on a logical interface which supports 4000 LVCs, then an average of only 1000 LVCs will be available per VP Tunnel. This means that Equations (1) to (7) need to be modified to deal with VP Tunnels.

Equation (1), which applies to interfaces on and routers, becomes

$$l < cdt, \quad \textbf{(8)}$$

where l is the number of LVCs used per link, d is the number of destination-prefixes known in an area, and c is the number of classes-of-service used, and t is the number of VP Tunnels terminating on the interface in question. For example, if 500 destinations are known in an area, 4 classes of service are supported in an area, and 4 VP tunnels terminate on a physical interface, less than 8000 LVCs will be required on that physical interface.

Similarly, equations (2) and (3) also apply to interfaces on routers under different circumstances. They also get a multiplier of t, and become

$$l \le t(c(d - d_e) + cnd_e)\,. \tag{9}$$

and

$$l < 2tcn \tag{10}$$

Equation (4) which applies to switches with VC merge, is the same as equation (1), and is modified in the same way:

$$l < cdt\,, \tag{11}$$

Equation (5), which also applies to switches with VC merge, stays the same:

$$m < cd(k-1)\,, \tag{5}$$

where m is the number of LVCs which must be merged together in the switch, and k the number of links into the switch. However k must include VP Tunnels as well as physical links.

Equations (6) and (7). which apply to switches without VC merge, also get multipliers according to the number of virtual trunks on the interfaces, becoming

$$l < cd(n-1)t\,, \tag{12}$$

and

$$l \le c(n^2/2)t\,. \tag{13}$$

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems Europe s.a.r.l.
Parc Evolic, Batiment L1/L2
16 Avenue du Quebec
Villebon, BP 706
91961 Courtaboeuf Cedex
France
http://www-europe.cisco.com
Tel: 33 1 69 18 61 00
Fax: 33 1 69 28 83 26

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Headquarters
Nihon Cisco Systems K.K.
Fuji Building, 9th Floor
3-2-3 Marunouchi
Chiyoda-ku, Tokyo 100
Japan
http://www.cisco.com
Tel: 81 3 5219 6250
Fax: 81 3 5219 6001

Cisco Systems has more than 200 offices in the following countries. Addresses, phone numbers, and fax numbers are listed on the Cisco Connection Online Web site at http://www.cisco.com/offices.

Argentina • Australia • Austria • Belgium • Brazil • Canada • Chile • China • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE
Finland • France • Germany • Greece • Hong Kong • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia
Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Singapore
Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela