NORTEL NETWORKS

Passport 7400, 15000, 20000

Understanding IP

241-5701-805

Passport 7400, 15000, 20000

Understanding IP

Publication: 241-5701-805
Document status: Standard
Document version: 5.2S2
Document date: December 2003

Printed in Canada

Publication history

December 2003

5.2S2 Standard
General availability. Contains information on Passport 7400, Passport 15000, and Passport 20000 for the PCR 5.2 GA release.

Contents

Chapter 3
IP over ATM 61

Chapter 4
IP over frame relay using frame relay DTE 71

Chapter 5
IP over frame relay using IP-optimized DLCIs 85

Chapter 13
Border gateway protocol 4 (BGP-4) 139

Chapter 14
Static routes 157

Chapter 15
IP multicast 159

Chapter 16 Virtual router redundancy protocol 167

Chapter 17 IP class of service (CoS) 175

Chapter 18 Passport IP differentiated services 195

List of figures

List of tables

About this document

This user guide describes virtual routers, the Internet protocol (IP), and other protocols and services related to IP on the Passport system.

The following topics are discussed in this section:

Who should read this document and why

This guide is for anyone who performs the following tasks for IP on the Passport system:

- planning
- installing and provisioning
- operating and maintaining

What you need to know

This guide assumes that you are familiar with the concepts of internetworking, particularly with the IP suite and its common uses.

How this document is organized

241-5701-805 *Passport 7400, 15000, 20000 Understanding IP* contains the following sections:

- “IP on Passport fundamentals” (page 33)
- “Planning the IP on Passport configuration” (page 57)
- “IP over ATM” (page 61)
- “IP over frame relay using frame relay DTE” (page 71)
- “IP over frame relay using IP-optimized DLCIs” (page 85)
- “IP over gigabit Ethernet” (page 89)
- “IP over point-to-point protocol (PPP)” (page 93)
- “Point-to-point protocol (PPP)/ATM interworking for Passport 7400” (page 99)
- “IP routing management” (page 105)
- “Routing information protocol (RIP)” (page 117)
- “Open shortest path first (OSPF) protocol” (page 125)
- “Border gateway protocol 4 (BGP-4)” (page 139)
- “Static routes” (page 157)
- “IP multicast” (page 159)
- “Virtual router redundancy protocol” (page 167)
- “IP class of service (CoS)” (page 175)
- “Passport IP differentiated services” (page 195)
- “IP CoS to IP DiffServ Migration” (page 211)
- “IP flow filters” (page 215)
- “IP tunnels” (page 221)
- “IP accounting” (page 229)

What’s new in this document

The following features were added to this document:

- “Hitless OSPF for CP/VpnXc switchover” (page 24)
- “IP multicast” (page 24)
- “Passport flat MVR support on MS3 gigabit Ethernet FP” (page 24)
- “Passport IP differentiated services for the gigabit Ethernet FP” (page 24)
- “Passport virtual router redundancy protocol” (page 24)
- “Intermediate System to Intermediate System (ISIS) Protocol” (page 133)

Other changes made to this document include the following:

- For CR Q00707925, updated the sections “Overview of IP tunnels” (page 221) and “Point-to-point tunnels” (page 225) to remove information about provisioning multiple tunnels with the same source address or destination address on the same VR.
- Updated “Application and feature names for IP on Passport” (page 34) to include more information on IP DiffServ.
- In “IP class of service (CoS)” (page 175) “Passport IP differentiated services” (page 195), and “IP CoS and IP DiffServ drop precedence selection attributes” (page 213), “discard priority” was changed to “drop precedence”.
- In “IP class of service (CoS)” (page 175), “Scheduling class” (page 184) was added to replace the section on emission priority.
- Added information on drop precedence to “Frame relay DTE class-based forwarding” (page 184), “IP-optimized DLCI class-based forwarding” (page 188), and “ATM MPE class-based forwarding” (page 189).
- “Traffic class” (page 197) was added.
- In “Passport IP differentiated services” (page 195), the “Scheduling class” (page 197), “Connection class” (page 198), and “IP CoS and IP DiffServ local packet classification and marking attributes” (page 213), attribute names were updated.

- Added UMTS information to table “Default PHB configuration by domain type” (page 203)
- Added software activation information to “IP CoS and IP DiffServ software activation attributes” (page 212)
- Updated source and destination address information in “IP tunnels” (page 221).

Hitless OSPF for CP/VpnXc switchover

The following section was updated:

- “Hitless OSPF for CP/VpnXc switchover” (page 130)

IP multicast

The following section was added:

- “IP multicast” (page 159)

Passport flat MVR support on MS3 gigabit Ethernet FP

The following section was updated:

- “Overview of IP over gigabit Ethernet on Passport” (page 89)

Passport IP differentiated services for the gigabit Ethernet FP

The following sections were updated:

- “Overview of IP over gigabit Ethernet on Passport” (page 89)
- “IP CoS support on access media” (page 177)
- “Gigabit Ethernet class-based forwarding” (page 190)
- “Scheduling class” (page 197)
- “Domain boundary interface” (page 209)

Passport virtual router redundancy protocol

The following section was added:

- “Virtual router redundancy protocol” (page 167)

Text conventions

This document uses the following text conventions:

- **`nonproportional spaced bold type`**

  Nonproportional spaced bold type represents words that you should type or that you should select on the screen.

- *italics*

  Words that appear in italics indicate a software component or attribute name.

- `[optional_parameter]`

  Words in square brackets represent optional parameters. The command can be entered with or without the words in the square brackets.

- `<general_term>`

  Words in angle brackets represent variables which are to be replaced with specific values.

Procedure conventions

This document uses the following procedure conventions:

- You can enter commands using full component and attribute names, or you can abbreviate them. The commands used in the procedures contain the full component and attribute names in the first instance. In the second instance, the component and attribute names are abbreviated. For more information on abbreviating component and attribute names, see 241-5701-060 *Passport 7400, 15000, 20000 Components*. All component and attribute names are formatted in italics.
- The introduction of every procedure states whether you must perform the procedure in operational mode or provisioning mode. For more information on these modes, see “Operational mode” (page 26) or “Provisioning mode” (page 26).

- When you complete a procedure, you can verify your changes and then activate them as the new node configuration. For more information on completing configuration changes and exiting provisioning mode, see “Activating configuration changes” (page 27).

Operational mode

Procedures contained within this document can either be performed in operational mode or provisioning mode. When you initially log into a Passport node, you are in operational mode. Passport uses the following command prompt when you are in operational mode:

```
#>
```

where:
# is the current command number

In operational mode, you work with operational components and attributes. In operational mode, you can

- list operational components and display operational attributes to determine the current operating parameters for the node
- control the state of parts of the node by locking and unlocking components
- set certain operational attributes and enter commands to perform diagnostic tests

Provisioning mode

To change from operational mode to provisioning mode, type the following command at the operator prompt:

```
start Prov
```

Only one user can be in provisioning mode at a time. Passport uses the following command prompt whenever you are in provisioning mode:

```
PROV #>
```

where:

# is the current command number

In provisioning mode, you work with the provisionable components and attributes that contain the current and future configurations of the node. You can add and delete components, and display and set provisionable attributes. For information on completing the configuration changes, exiting provisioning mode, and returning to operational mode see “Activating configuration changes” (page 27).

For information on operational and provisionable attributes, see 241-5701-060 *Passport 7400, 15000, 20000 Components*.

Activating configuration changes

Several procedures in this document ask that you complete the configuration changes. When you complete the configuration changes, you are activating the configuration changes, confirming that you want to activate them, and saving the changes. You are instructed to complete the configuration changes only at the end of procedures that you perform in provisioning mode.

> **CAUTION**
> **Activating a provisioning view can affect service**
> Activating a provisioning view can result in a CP reload or restart, causing all services on the Passport node to fail. See 241-5701-050 *Passport 7400, 15000, 20000 Commands*, for more information.

1 Verify that the provisioning changes you have made are acceptable:

```
check Prov
```

Correct any errors and then verify the provisioning changes again.

2 If you want to store the provisioning changes in a file, save the provisioning view:

```
save Prov
```

3 If you want these changes as well as other changes made in the edit view to take effect immediately, activate, confirm, and commit the provisioning changes:

```
activate Prov
```

```
confirm Prov
```

```
commit Prov
```

4 End the provisioning session:

```
end Prov
```

Related documents

For the complete list of documents in the Passport documentation library, see 241-5701-001 *Passport 7400, 15000, 20000 Documentation Guide*.

The following sections contain documents related to the information in this guide:

- “Passport documents” (page 28)
- “Request for comments (RFCs)” (page 29)

Passport documents

The following documents containing information related to IP and the Passport system are available from Nortel Networks:

- 241-1001-506 *DPN-100 Alarm Console Indications*
- 241-5701-030 *Passport 7400, 15000, 20000 Overview*
- 241-5701-001 *Passport 7400, 15000, 20000 Documentation Guide*
- 241-5701-005 *Passport 7400, 15000, 20000 List of Terms*
- 241-5701-050 *Passport 7400, 15000, 20000 Commands*
- 241-5701-060 *Passport 7400, 15000, 20000 Components*
- 241-5701-500 *Passport 6400, 7400, 15000, 20000 Alarms*
- 241-5701-600 *Passport 7400, 15000, 20000 Configuration Guide*
- 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*
- 241-5701-650 *Passport 7400, 15000, 20000 Accounting Fundamentals*
- 241-5701-700 *Passport 7400, 15000, 20000 ATM Overview*
- 241-5701-702 *Passport 7400, 15000, 20000 ATM Routing and Signaling Fundamentals*
- 241-5701-705 *Passport 7400, 15000, 20000 ATM Traffic Management Fundamentals*
- 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*

- 241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals*
- 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management*
- 241-5701-030 *Passport 7400, 15000, 20000 Overview*
- 241-7401-200 *Passport 7400 Hardware Description*
- 241-7401-240 *Passport 7400 Hardware Installation, Maintenance and Upgrade*

Request for comments (RFCs)

The following Requests for Comments (RFCs) containing information related to IP are available from numerous sources including Internet Network Information Center (NIC) servers:

- RFC761, *DoD standard Transmission Control Protocol*
- RFC768, *User Datagram Protocol*
- RFC0791, *Internet Protocol*
- RFC792, *Internet Control Message Protocol*
- RFC793, *Transmission Control Protocol*
- RFC815, *IP Datagram Reassembly Algorithms*
- RFC821, *Simple Mail Transfer Protocol*
- RFC826, *An Ethernet Address Resolution Protocol*
- RFC854, *Telnet Protocol Specifications*
- RFC904, *Exterior Gateway Protocol Formal Specification*
- RFC950, *Internet Standard Subnetting Procedure*
- RFC951, *Bootstrap Protocol (BootP)*
- RFC959, *File Transfer Protocol*
- RFC1009, *Requirements for Internet Gateways*
- RFC1038, *Draft Revised IP Security Option*
- RFC1042, *Standard for Transmission of IP Datagrams over IEEE 802 Networks*

- RFC1122, *Requirements for Internet Hosts - Communication Layers*
- RFC1157, *Management Information Base for Network Management of TCP/IP-based Internets*
- RFC1213, *Management Information Base for Network Management of TCP/IP-based internets: MIB-II*
- RFC1253, *OSPF Version 2 Management Information Base*
- RFC1354, *IP Forwarding Table MIB*
- RFC1517, *Applicability Statement For the Implementation of Classless Inter-Domain Routing (CIDR)*
- RFC1518, *An Architecture for IP Address Allocation with CIDR*
- RFC1519, *Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy*
- RFC1541, *Dynamic Host Configuration Protocol*
- RFC1577, *Classical IP and ARP over ATM*
- RFC1583, *OSPF Version 2*
- RFC1657, *Border Gateway Protocol version 4 (BGP-4) MIB*
- RFC1701, *Generic Routing Encapsulation*
- RFC1702, *Generic Routing Encapsulation over IPv4 networks*
- RFC1723, *RIP Version 2 Carrying Additional Information*
- RFC1724, *RIP Version 2 MIB Extension*
- RFC1745, *BGP4/IDRP for IP-OSFP Interaction*
- RFC1771, *Border Gateway Protocol 4 (BGP-4*)
- RFC1772, *Application of the Border Gateway Protocol in the Internet*
- RFC2003, *IP Encapsulation within IP*
- *RFC 2236, Internet Group Management Protocol (IGMP), version 2* router functionality
- RFC2334, *Server Cache Synchronization Protocol*
- RFC2338, *Virtual Router Redundancy Protocol*

- *RFC 2362, Protocol Independent Multicast - Sparse Mode (PIM-SM)*
- RFC2474, *DiffServ Field Definition*
- RFC2597, *Assured Forwarding PHB Group*
- RFC3246, *An Expedited Forwarding PHB*

How to get more help

For information on training, problem reporting, and technical support, see the “Nortel Networks support services” section in the product overview document.

Chapter 1
IP on Passport fundamentals

Passport Internet Protocol (IP) enables Passport nodes to provide IP virtual private network (VPN) capabilities across Passport networks. Passport uses virtual routers (VRs) to provide IP connectivity between nodes. Each Passport node can support numerous VRs.

Passport IP can simultaneously manage different software applications and types of traffic. When services are running on a common network facility, Passport enables you to consolidate bandwidth usage. Passport also provides a feature-rich IP interconnect service that provides high reliability and advanced packet security.

The following sections provide an overview of how IP is implemented on Passport:

- "Related information for IP on Passport" (page 53)

Application and feature names for IP on Passport

The table "Application and feature names for IP on Passport" (page 34) lists the functionality provided by IP on Passport, the associated software application name, and the associated feature name. Use this information when you need to know the software application to download and feature name to link to a logical processor type (LPT). For information about downloading application software to a Passport node, see 241-5701-270 *Passport 7400, 15000, 20000 Software Installation Guide*.

Table 1
Application and feature names for IP on Passport

Functionality	Software application name	Feature name
IP	ip	ip
ATM MPE	wanDte	atmMpe
ATM MPE soft PVCs	wanDte	atmMpe, atmMpeSpvc
Frame relay DTE	wanDte	FrameRelayDte
IP-optimized DLCI	frameRelay	frUniIpOptimized
Gigabit Ethernet, Ethernet	ip	ip
Point-to-point protocol (PPP)	wanDte	PPP
IP class of service (CoS)	ip	ipCos
IP differentiated services (DiffServ) [1]	ip	ipDiffServ
IP flow filters	ip	ipFilter
Border gateway protocol	ip	BGP

[1] Feature ipDiffServ is required for the DiffServ profile for the interface (*Vr Ip DiffServ*) but not for the DiffServ domain for the router (*Vr Dsd*).

IP processor cards

Function processors (FPs) provide interface ports that connect network communications facilities to Passport switches. FPs support and execute real-time processes that are essential to service delivery.

IP services running on the Passport 15000 or 20000 require a CP3-based CP and a PQC 2.0-based, PQC12-based, or 4pGe FP. The CP2-based CP, CQC-based FP, and SBIC-based FPs for IP are supported on the Passport 7400 platform only. For more information on the FPs over which the IP service operates, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

The VPN extender card (VpnXc) is a special server card that you can use to increase the scalability of IP VPN services. This card has its own dedicated processor and memory and acts as the IP VPN control plane, hosting all IP VPN virtual routers (VCGs and customer VRs). For more information on the VPN extender card, see 241-7401-200 *Passport 7400 Hardware Description* and 241-1501-200 *Passport 15000, 20000 Hardware Description*.

IP protocol suite

Transmission control protocol/Internet protocol (TCP/IP) is a group of protocols that defines a common set of rules and standards that enable networks and hosts to communicate. IP is the routed, or network layer, protocol of TCP/IP and is one of the most popular internetworking protocols. Most internetworks support TCP/IP, whether or not TCP/IP end systems are present.

When you add an *Ip* component the system automatically adds supporting transmission control protocol/internet protocol (TCP/IP) processes such as address resolution protocol (ARP), internet control message protocol (ICMP), relay broadcast (RelayBC), user datagram protocol (UDP), and transmission control protocol (TCP).

Passport supports the following router management software applications, all of which are part of the TCP/IP architecture:

- “Internet control message protocol (ICMP)” (page 36)
- “Transmission control protocol (TCP)” (page 36)

- “User datagram protocol (UDP)” (page 36)
- “File transfer protocol (FTP)” (page 36)
- “Telnet” (page 37)

Internet control message protocol (ICMP)

Internet control message protocol (ICMP) provides feedback from an IP router or gateway to a source host. ICMP messages are sent in several situations: for example, to report resource or routing problems or to report a shorter available route to a destination. The Passport system uses ICMP echoes and echo replies to verify the reachability of routers or end systems. See RFC 792 for more information.

Transmission control protocol (TCP)

Transmission control protocol (TCP) is a connection-oriented transport-layer protocol. TCP provides reliable, robust, and adaptable data transfer between end-system upper layer protocols. TCP assumes that simple, potentially unreliable, data transmission services are available from lower-level protocols. See RFC 793 for more information.

User datagram protocol (UDP)

User datagram protocol (UDP) defines the use of unacknowledged datagrams. UDP packets are useful for very low-priority data or for very high-reliability networks. UDP is also useful when an application already provides an integrity function and does not need to duplicate that function by using TCP. See RFC 768 for more information.

File transfer protocol (FTP)

The file transfer protocol (FTP) provides a robust file transfer mechanism for data transfer between IP hosts. FTP transfers files between the file system of the Passport node and a UNIX server. Once a connection is established, the Passport node requests the appropriate account information (including security information) before establishing a session. See RFC 959 for more information.

Telnet

Telnet allows a valid user access to a terminal or command process on a remote system such as the operator process on a Passport system. The Passport system supports both Telnet client and server connections. See RFC 854 for more information.

IP addressing protocols

A virtual router uses IP addressing protocols to map an IP address to the correct physical address when it needs to send data across a physical network. Passport supports the following IP addressing protocols:

- "Address resolution protocol (ARP)" (page 37)
- "Reverse ARP (RARP)" (page 38)
- "Proxy ARP" (page 38)
- "Inverse ARP (InARP)" (page 38)
- "Bootstrap protocol (BOOTP)" (page 38)

Address resolution protocol (ARP)

The address resolution protocol (ARP) is a mechanism for mapping 32-bit IP addresses to 48-bit Ethernet hardware addresses. The hardware address is a concatenation, or joining, of two numbers: a vendor ID number, centrally assigned by the IEEE, and a unique serial number assigned by the vendor for each hardware unit. This hardware address, termed the media access control (MAC) address usually has significance only on the Ethernet wire.

The Passport implementation of ARP supports the following features:

- removal of out-of-date ARP cache data
- configurable cache data timeout
- translation of encapsulation information between Ethernet and IEEE 802.3 networks

Ethernet and frame relay media support ARP.

For more information about ARP, see RFC 826.

Reverse ARP (RARP)

Reverse address resolution protocol (RARP) determines or assigns a particular station’s IP address when only the station’s MAC address is known. There are many reasons why an end system does not already have an IP address. The end system can be a diskless workstation homed off a server. Or, the end system can be a portable computer belonging to an itinerant employee sharing a pool of IP addresses with other itinerant employees. The Passport system cannot currently act as a RARP server. RFC 903 defines RARP.

Proxy ARP

The proxy ARP is used to help an IP device locate a destination device, when the destination device is on a remote IP network or wire. When a source station broadcasts an ARP request on the local wire, and there is no station matching the destination IP address on the wire, the source does not receive an ARP response from the actual destination. Instead, the router derives the destination’s IP wire address and searches for a match in its IP routing table. If the destination IP wire address is present in the routing table, the router responds with its own MAC address, in effect telling the source that the router’s MAC address is the destination station’s MAC address. The source IP station has no idea that the destination is on another wire. The Passport system fully supports proxy ARP. RFC 1027 defines proxy ARP.

Inverse ARP (InARP)

The inverse address resolution protocol (InARP) is used to determine a remote router’s IP address on a particular ATM or frame relay connection. This is the local ATM or frame relay address of a permanent virtual circuit (PVC) to a remote router. The Passport system fully supports InARP. RFC 1293 defines InARP.

For more information, see “Inverse ARP scalability” (page 39).

Bootstrap protocol (BOOTP)

The bootstrap protocol (BOOTP) is a UDP/IP-based protocol which allows a booting host to configure itself dynamically and without user supervision. BOOTP provides a means to notify a host of

- its assigned IP address
- the IP address of a boot server host

- the name of a file to be loaded into memory and executed
- the local subnet mask
- the local time offset
- the addresses of default routers
- the addresses of various Internet servers

The Passport system supports the BOOTP relay agent functionality described in RFC951 and RFC1542.

Inverse ARP scalability

You can reduce the number of inverse ARP requests generated by a Passport node by linking an individual IP logical interface to a particular frame relay or ATM connection. When used on a customer VR in an IP VPN, this configuration improves IP VPN scalability.

Background

When a frame relay or ATM connection comes up, the virtual router linked to that connection is notified. The virtual router then sends an inverse ARP request to the other end.

When there is no explicit association between an IP logical connection and the connection, Passport does not know whether an individual connection is associated with only one IP subnet. Therefore, when there is more than one IP logical interface configured on a protocol port, an inverse ARP request is sent across the connection for each IP logical interface on the protocol port.

Inverse ARP scalability description

You can reduce the number of inverse ARP requests by giving each CPE its own subnet where each connection is known as a subconnection. When a frame relay or ATM subconnection comes up, the virtual router linked to that subconnection is notified and it sends only one inverse ARP request to the other end.

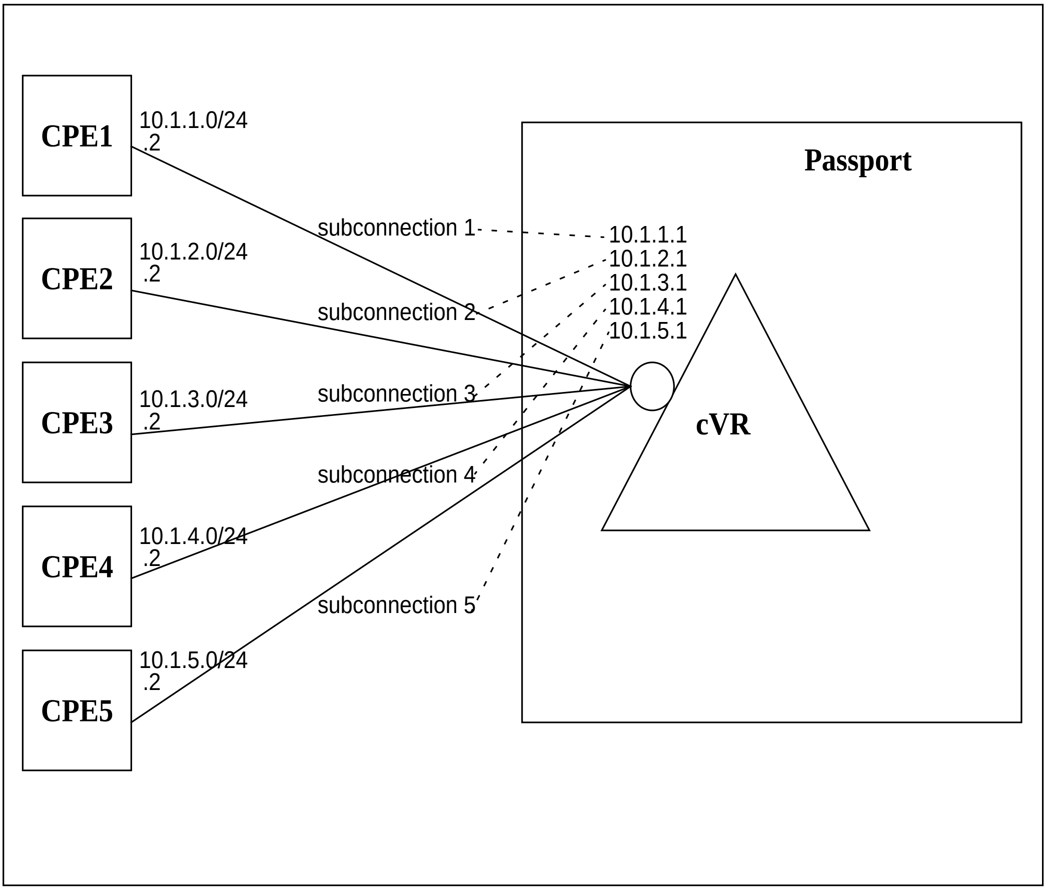
This configuration is shown in the figure “Example of inverse ARP scalability” (page 41) and includes the following steps:

1 Configure multiple frame relay or ATM connections on the same protocol port where each connection leads to a different CPE device.

2 Configure an individual IP logical interface against the port for each CPE device (i.e., for each subconnection).

3 Using attribute *Vr ProtocolPort IpPort LogicalIf linkToMediaConnection*, link each IP logical interface to the corresponding subconnection to the CPE device.

Figure 1
Example of inverse ARP scalability



When you configure OSPF on IP logical interfaces with links to individual subconnections, be aware of the following:

- It is recommended that OSPF be configured in non-broadcast mode. Broadcast/multicast mode causes hello packets to each IP logical interface to be sent over all connections. This causes unnecessary network traffic and may raise unnecessary traps or alarms.
- If an individual connection goes down, it is not detected immediately by OSPF unless all connections attached to the protocol port are down. This occurs because OSPF detects interface state changes at the protocol port

level, not at the individual connection level. Instead, when an individual connection goes down, the failure is detected after the router dead interval (attribute *Vr Pp IpPort LogicalIf OspfIf rtrDeadInt*) corresponding to the neighbor node on the far end of the connection has expired.

Virtual routers on Passport

Passport virtual routers (VRs) provide a software emulation of physical routers. A VR has two main functions:

- constructing routing tables describing the paths to networks or subnetworks
- forwarding or switching packets to the final destination network or subnetwork

Each VR is an instance of a routing protocol used over a unique set of IP ports, point-to-point protocol (PPP) sessions, frame relay data link connection identifiers (DLCI), and ATM virtual circuits (VCs). A VR coexists with other Passport facilities on the same node.

Virtual routers on a Passport node can perform the functions of independent physical routers, forwarding packets to the correct destination while isolating each customer’s traffic. Virtual routers provide a cost-effective alternative to using many separate hardware routers to provide multiple customer routing over a shared network. Carriers can therefore share backbone networks more effectively. See the figures “Traditional router configuration” (page 43) and “Passport virtual router configuration” (page 44) to see how Passport VRs eliminate the need to use separate physical routers.

VRs have independent IP routing tables and are isolated from each other. These separate routing capabilities provide each enterprise customer with the appearance of a dedicated router that guarantees isolation from other customers while running on shared switching and transmission resources. This means that a customer’s IP addressing space can overlap with another customer’s address space. The IP addresses need only be unique within a customer’s domain.

A Passport node can support multiple virtual routers. Using multiple VRs on a Passport node enables carriers to support multiple isolated networks on the same platform by assigning each network to its own virtual router. See the figure “Passport virtual router configuration” (page 44) for more information.

Figure 2
Traditional router configuration

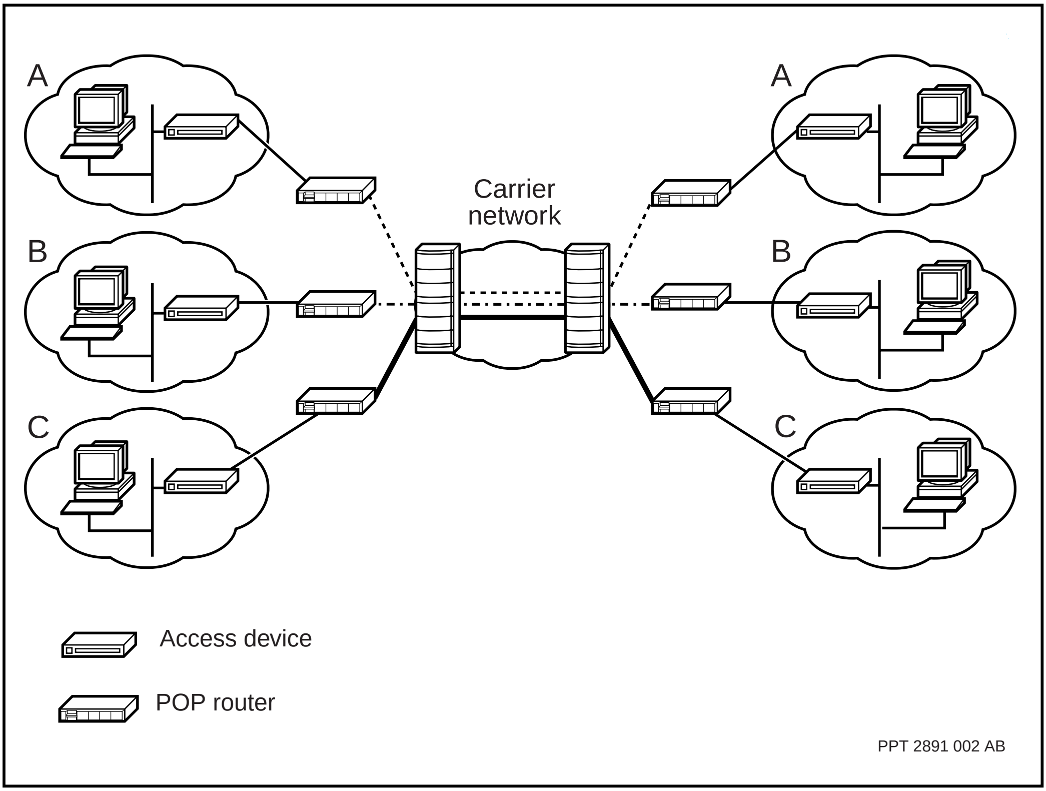
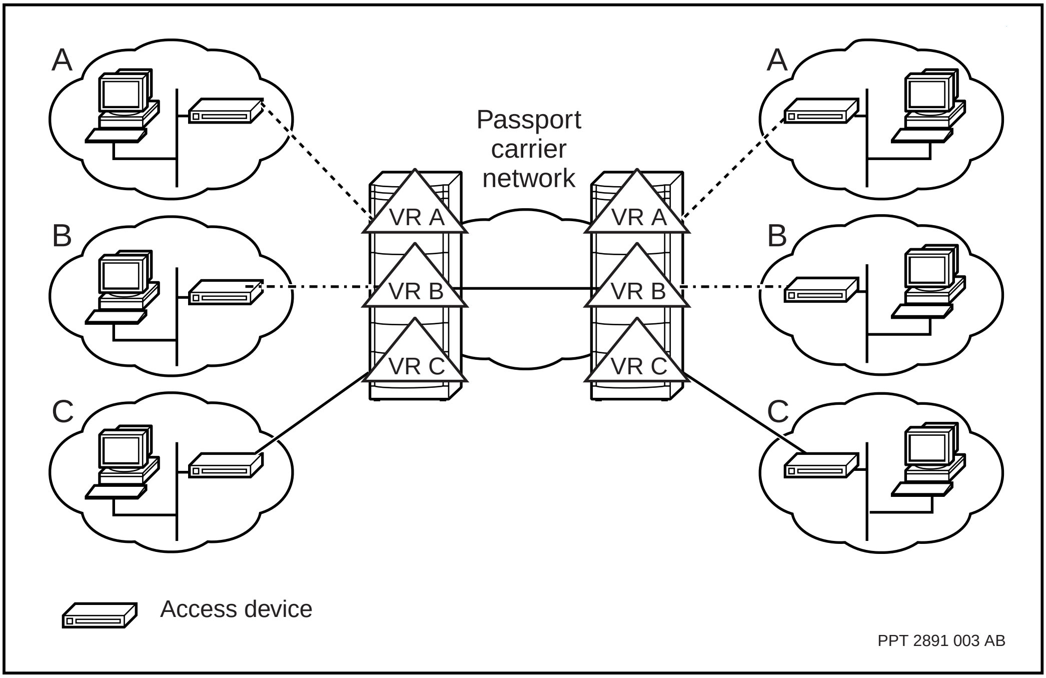


Figure 3
Passport virtual router configuration



For more information on virtual routers and their functionality, see the following sections:

- “Management virtual router” (page 45)
- “Customer virtual router” (page 45)
- “Virtual connection gateway” (page 46)
- “Virtual router memory management” (page 47)
- “Source routing option” (page 47)
- “Cache table size” (page 48)

Management virtual router

The management VR is a Passport virtual router that provides a single point of external entry into the Passport node. You can also use the management VR to manage all customer VRs that reside on the Passport node. The figure “Management access for customer VRs” (page 46) illustrates the use of a management VR.

The first VR you create on a Passport node becomes, by default, the management VR. This means that even on a Passport running a single VR, that VR has all the features associated with the management VR. Once you activate your provisioning view, you cannot designate any other VR as the management VR.

A single TCP agent running under the management VR allows external access to the Passport node from a workstation running network management system software through telnet, using TCP or FTP. You can also manage all VRs on the Passport node through a single SNMP agent running under the management VR.

For information about configuring a management virtual router, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Customer virtual router

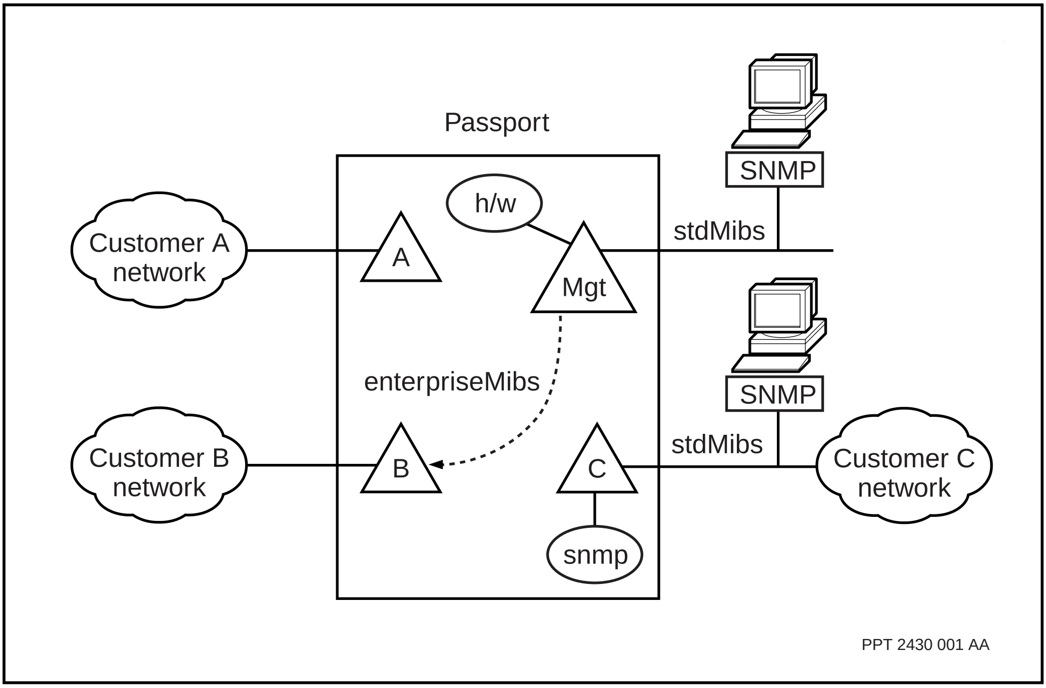
Configuring the customer VR is the same as the management VR with the exception that the customer VRs are restricted as to the protocols and interfaces they support. Management access for customer VRs is disabled meaning that users cannot set up telnet sessions to any of the interfaces on the customer VR.

You can manage these VRs through SNMP or Preside Multiservice Data Manager on each VR (stdMibs), or enterpriseMibs on the management VR. The figure “Management access for customer VRs” (page 46) illustrates the management of customer VRs.

> ***Note:*** For security purposes, you should restrict SNMP access to customer VRs to designated personnel only.

For information about configuring a customer virtual router, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Figure 4
Management access for customer VRs



Virtual connection gateway

In a typical Passport IP VPN implementation, CPE routers connect to a customer VR assigned to that enterprise. Each customer VR on the Passport node connects to a common VR for the switch, called the Virtual Connection Gateway (VCG).

The VCG aggregates traffic from the customer VRs and provides a single outbound connection into the wide area network (WAN) for all individual customer traffic on the node. The VCGs link all Passport nodes that provide IP VPN functionality, and provide connectivity between customer VRs in the same IP VPN through point-to-multipoint (PTMP) IP tunnels.

For information about configuring a virtual connection gateway, see 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management*.

Virtual router memory management

The number of routes in a VR’s routing database affects its memory consumption. A memory limit is assigned to a VR by setting the value of attribute *Vr Mm vrMaxHeapSpace* as a percentage of total available CP heap memory. As memory usage for the VR increases and the pre-defined thresholds are crossed, alarms are generated. See the table “VR memory thresholds” (page 47).

If memory usage reaches 101% of *Vr Mm vrMaxHeapSpace*, the VR is automatically locked. When this occurs, either reduce the number of routes propagated to the VR (for example, through route summarization) or reconfigure *Vr Mm vrMaxHeapSpace* to a larger value. Then, manually unlock the VR so that it can provide service again. You need to take these steps before unlocking the VR or else it will continue to exhaust its allocated memory and repeat the locking behavior.

It is strongly recommended that *vrMaxHeapSpace* be left to its default value (100%) for VRs configured as virtual connection gateways (VCGs).

Table 2
VR memory thresholds

Percentage used of *Vr Mm vrMaxHeapSpace*	**Network action**
101	Set critical alarm and lock VR
99	Clear critical alarm and replace with major alarm ***Note:*** VR must be unlocked manually
90	Set major alarm
85	Clear major alarm and replace with minor alarm
80	Set minor alarm
75	Clear minor alarm

Source routing option

Source routing is an option specified in the IP header that allows the originator of a packet to specify a particular route to its destination.

You can enable or disable the processing of input datagrams that have a source IP option on a VR basis using the *sourceRoute* provisionable attribute. See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP* for information about how to configure this option.

Cache table size

The cache management system (CMS) allows you to configure the IP local cache table size. This enables you to fine tune network performance by provisioning cache table size based on resource demand.

When planning your CMS

- determine the optimum memory requirements for all LPs, and adjust your cache table sizes accordingly. Carefully consider the type and amount of traffic being run on LPs.
- ensure an increase in cache table size does not adversely impact IP traffic (as long as the cache table sizes are optimized as discussed above).
- be aware a decrease in cache table size can impact IP traffic in the case where the number of cache entries is larger than the newly provisioned cache table size permits. In this situation, the related protocol traffic is blocked during the adjustment period. However, if cache table sizes are optimized, there is no adverse impact on IP traffic.
- consult the *Passport IP VPN Engineering Guidelines* for information on cache table entry allocation behavior and recommendations.

The Passport system creates the local IP cache table on a logical processor (LP) as soon as the first inbound protocol port is enabled on that LP. The system creates cache tables using default values. For provisioning information, see the section on configuring IP on a virtual router in 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

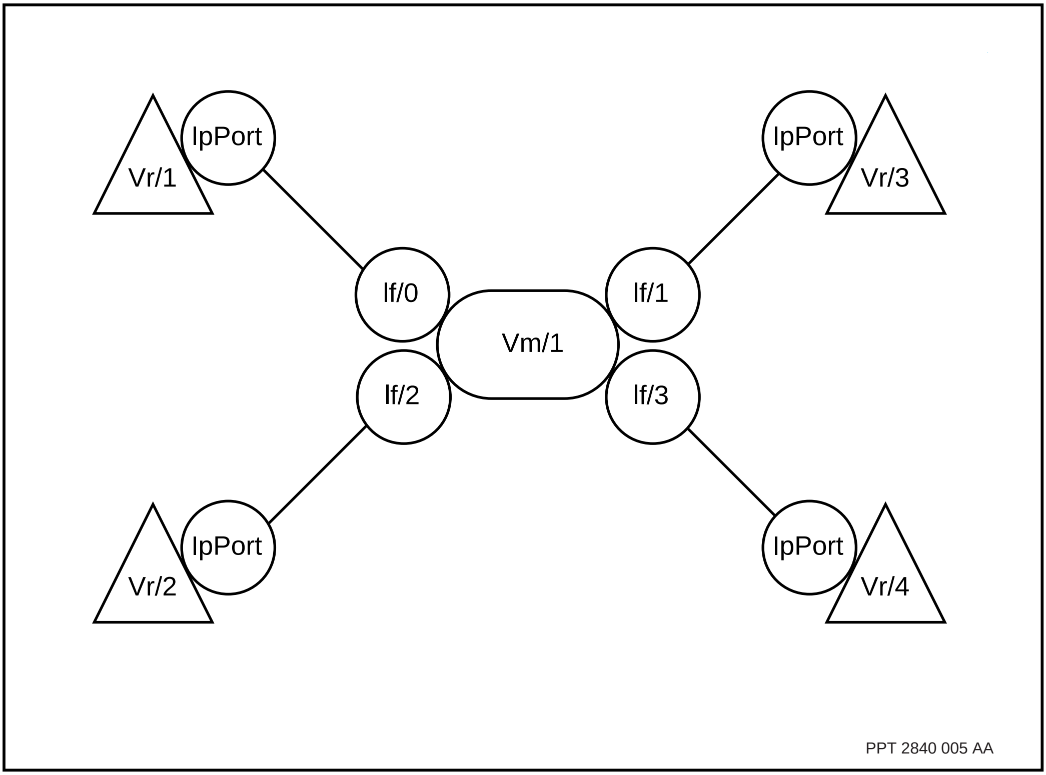
The CMS also offers local cache monitoring and control capabilities through the component administration system (CAS) standard interface. The *Cache* component is a dynamic subcomponent of the *Ip* component. It represents the IP cache table on an LP and contains the operational attributes that allow for cache table monitoring. For information on monitoring IP cache tables, see the section on monitoring and testing in 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Virtual media on Passport

You can configure connectivity between some of your VRs on a Passport node. By default, VRs on a Passport shelf are completely isolated from one another, for security purposes. You can add connectivity through the use of a physical hairpin connecting the physical ports through which the different VRs link. This method, however, is wasteful of physical resources.

Instead, you can emulate a physical connection between VRs by configuring IP-only connectivity in the software using the *Interface* (*If*) subcomponent of the *VirtualMedia* (*Vm*) component. The *Vm If* component provides virtual (as opposed to physical) next-hop functionality between VRs. You can enable connectivity between different VRs by linking them through an IP port to different instances of the *If* subcomponent under the same *Vm* component. “VR connectivity through a software link” (page 50) illustrates the relationship between the *Vm* component and the VRs.

Figure 5
VR connectivity through a software link



You can add a *Vm* component if you want to provision an always-up IP interface for applications such as open shortest path first (OSPF) protocol, routing information protocol (RIP), and border gateway protocol (BGP). A virtual media application is not associated with a physical port. Since logical IP interfaces under the virtual media application are defined independently of any physical media, they remain up even though individual links to the Passport might lose connectivity. An IP address associated with the virtual media protocol is always reachable as long as the Passport node itself remains connected to the network.

Virtual media on Passport support the following routing and forwarding functions:

- IP ARP and IP ARP Reply
- IP datagram forwarding through ARP and static route definitions
- OSPF
- RIP
- BGP-4

For information about configuring virtual media, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

IP virtual private networks (VPNs)

An IP VPN is a managed IP service offered by a carrier to an enterprise customer. The IP VPN service provides secure and reliable connectivity, management, and addressing (equivalent to that available on a private network) over a shared public network infrastructure.

See 241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals* and 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management* for more information on the Passport IP VPN service.

Provisioning MTU size

The maximum transmission size (MTU) is the largest unit of data that a network's physical medium can transmit. It can be set at the media link or the protocol port. The default MTU size is set as follows:

- 9188 for AtmMpe
- 9180 for IP tunnels
- 1604 for FrDte
- 1500 for Ethernet
- MRU (maximum receive unit) of 18000 for PPP
- must be manually provisioned on protocol port

You can provision lower MTU values for certain media. Smaller MTU sizes can help control jitter in a real-time voice stream because small voice packets are no longer delayed by large data packets.

The minimum MTU value depends on the type of processor card type. See the table “Minimum supported MTU sizes” (page 52).

Set MTU size at

- *AtmMpe mtu*
- *Vr Ip Tunnel Msep mtu*
- *FrDte Rg mtuSize*
- *Ppp Link configInitialMru*
- *Vr ProtocolPort IpPort mtu*

If set at the protocol port, the MTU must be within the valid range of the *IpPort* media type. If both the media and the protocol port MTU are set, the lowest of the two values becomes the MTU.

> ***Note:*** PQC-based FPs used to have the same minimum supported MTU sizes as CQC and SBIC-based FPs. Setting the MTU size on PQC-based FPs to less than its previous minimum (576 for IP tunnels and the protocol port and 262 for FrDte) can have a real time impact on FPs and throughput, so when using these small sizes contact your Nortel Networks representative for detailed engineering assistance.

Table 3
Minimum supported MTU sizes

Media	CQC and SBIC	PQC
AtmMpe	256	256
IP tunnels	576	100
FrDte	262	80
PPP	68	68
Protocol port	576	80

Related information for IP on Passport

This section describes where to find information related to the following topics:

- “IP media” (page 53)
- “IP routing protocols” (page 54)
- “IP features” (page 55)

IP media

Passport can provide customer access to the carrier network using the media listed in the table “Passport-supported access media” (page 53).

Table 4
Passport-supported access media

Passport 7400	ATM, frame relay using FRDTE, frame relay using IP-optimized DLCIs, PPP, 10BaseT Ethernet, 100BaseT Ethernet
Passport 15000 and 20000	ATM, frame relay using FRDTE, frame relay using IP-optimized DLCIs, PPP, gigabit Ethernet

Passport-supported core media are ATM, frame relay using FRDTE, and MPLS.

Protocol ports represent physical instances of data link or media protocols. When you configure protocol ports, you must link them to the corresponding media. You can configure protocol port designations that follow a descriptive numbering convention to allow easy recognition of protocol port-attached media.

The table “Where to find IP media information” (page 54) tells you where to find more information about specific IP media.

Table 5
Where to find IP media information

Media	Fundamentals	Configuration	RFC
IP over ATM	“IP over ATM” (page 61)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 1483
IP over frame relay	“IP over frame relay using frame relay DTE” (page 71) “IP over frame relay using IP-optimized DLCIs” (page 85)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 2427 RFC 1490
IP over gigabit Ethernet	“IP over gigabit Ethernet” (page 89)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 894
IP over point-to-point protocol (PPP)	“IP over point-to-point protocol (PPP)” (page 93)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 1661
IP over multiprotocol label switching (MPLS)	241-5701-445 *Passport 7400, 15000, 20000 Multiprotocol Label Switching Guide*	241-5701-445 *Passport 7400, 15000, 20000 Multiprotocol Label Switching Guide*	RFC 2702

IP routing protocols

Passport supports static routes as well as interior and exterior dynamic routing protocols. Interior routing protocols determine best paths within an autonomous system or enterprise. Exterior routing protocols determine best paths between autonomous systems. You can configure multiple dynamic routing protocols on one virtual router.

The table “Where to find IP routing information” (page 55) lists the IP traffic routing methods that Passport supports. It also tells you where to find fundamental and configuration information about specific IP traffic routing methods.

Table 6
Where to find IP routing information

Routing method	Fundamentals	Configuration	RFC
Routing information protocol (RIP)	"Routing information protocol (RIP)" (page 117)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 1723 RFC 1724
Open shortest path first protocol (OSPF)	"Open shortest path first (OSPF) protocol" (page 125)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 2178
Border gateway protocol 4 (BGP-4)	"Border gateway protocol 4 (BGP-4)" (page 139)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 1771 RFC 1772 RFC 1745
Static routes	"Static routes" (page 157)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	n/a
Bootstrap protocol (BOOTP)	"Bootstrap protocol (BOOTP)" (page 38)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC951 RFC1542
Virtual router redundancy protocol (VRRP)	"Virtual router redundancy protocol" (page 167)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC2338

IP features

The table "Where to find IP feature information" (page 56) tells you where to find more information about specific IP features on Passport.

Table 7
Where to find IP feature information

Service	Fundamentals	Configuration	RFC
IP multicast	“IP multicast” (page 159)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC2236 RFC2262
IP class of service (CoS), including IP CoS to QoS mapping	“IP class of service (CoS)” (page 175)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 2474
IP flow control	“IP flow filters” (page 215)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 2267
IP tunnels between IP networks	“IP tunnels” (page 221)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 1701 RFC 1702 RFC 2003
IP differentiated services	“Passport IP differentiated services” (page 195)	241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*	RFC 2474 RFC 3246 RFC 2597
IP accounting	“IP accounting” (page 229)	241-5701-650 *Passport 7400, 15000, 20000 Accounting Fundamentals*	n/a

Chapter 2
Planning the IP on Passport configuration

This section describes the things you need to consider when planning your IP on Passport configuration. It covers the following topics:

- “Network considerations” (page 57)
- “Mapping the IP network” (page 57)
- “IP on Passport configuration sequence” (page 60)

Network considerations

If there are other routers, of any manufacture, included in your network plans, you need to complete some the following steps:

- Select a routing protocol. In cases where you are integrating the Passport system into an existing network, choose the routing protocol to conform or interoperate with the existing routers.
- Gather relevant information about the networks on the other side of the remote routers including server addresses and special needs.
- If your network connects to other networks that are not under the control of your organization, you must plan security firewalls to prevent unauthorized access to the network.

Mapping the IP network

It is very important to have a usable representation of your network before configuring IP. If IP is already in use in your network, you probably only need a rough diagram showing the network numbers you need and the IP addresses assigned to the ports. “Simple network diagram” (page 58) illustrates a simple network map.

The networks shown in “Simple network diagram” (page 58) are established and only need to be joined by the Passport system. The only information that the installers and administrators need to understand the network are the network addresses for the ports and the routing protocol currently in use. Each connected segment must have a unique network or subnetwork number.

However, if you are introducing IP at the same time you are installing the Passport system, you can benefit from a map showing each node and its IP address. “Detailed network diagram” (page 59) is an example of one page of such a network map.

Figure 6
Simple network diagram

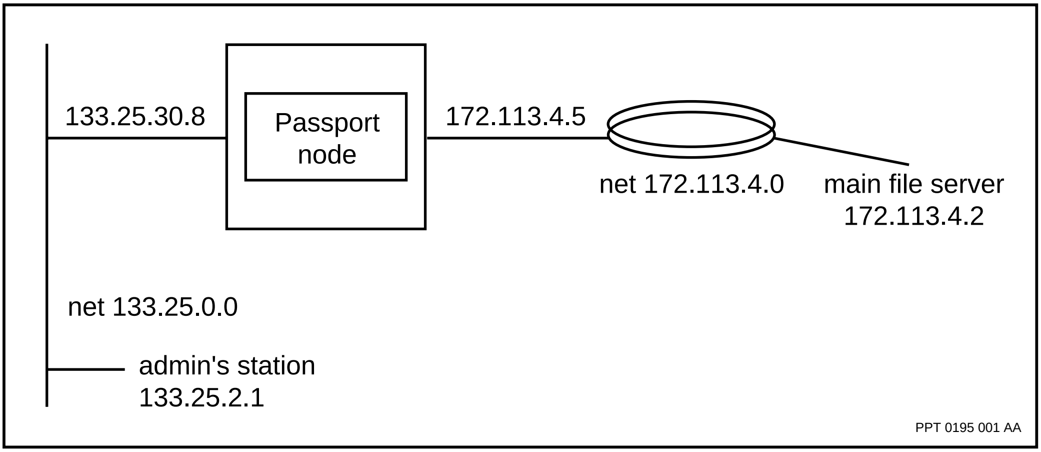
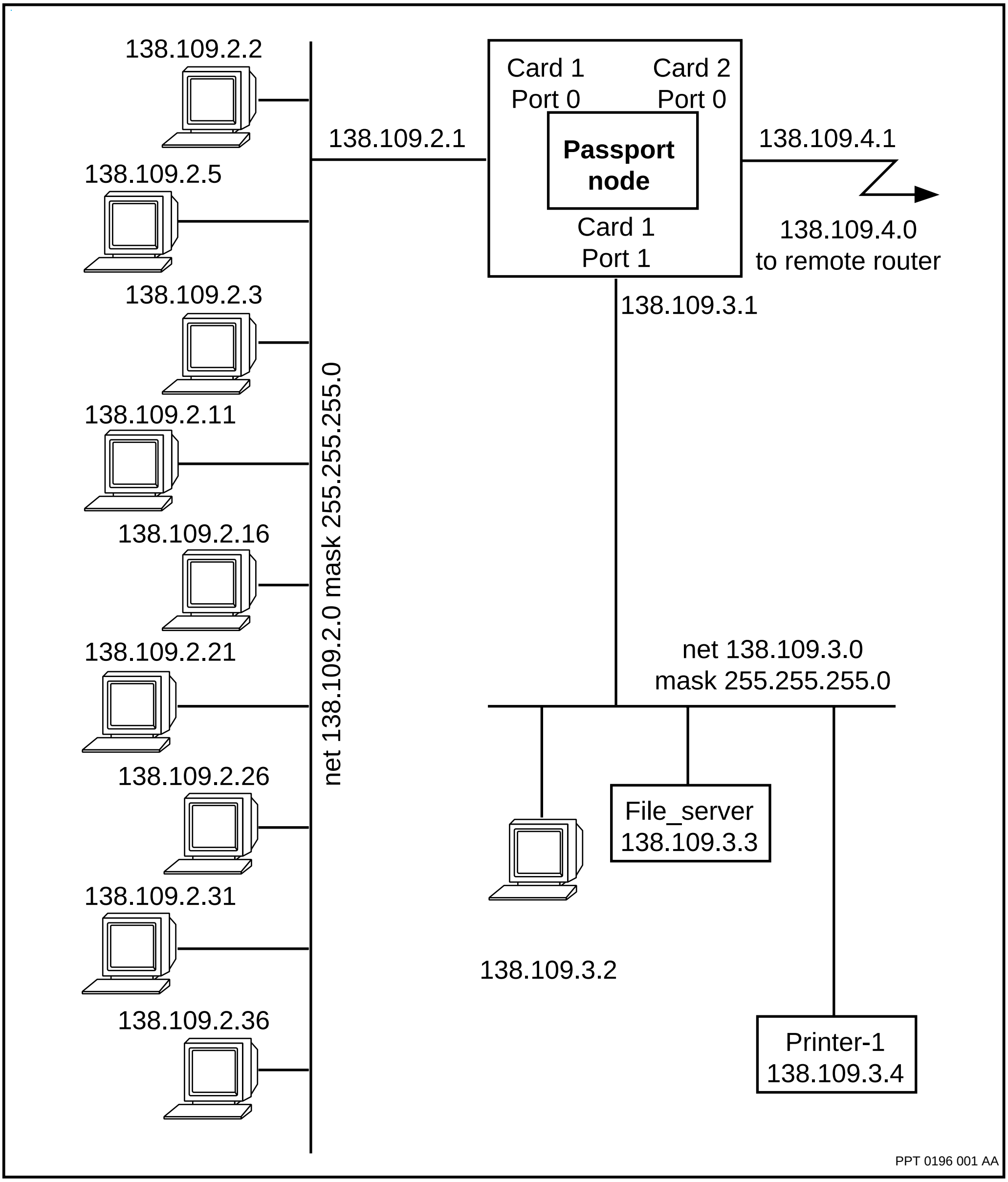


Figure 7
Detailed network diagram



IP on Passport configuration sequence

The table “IP on Passport configuration sequence” (page 60) provides a high-level view of the IP on Passport configuration process. You can use it to plan the end-to-end configuration of IP on a Passport node.

You might find it more efficient to download all the software for IP, virtual routers, and access media together, then proceed to configure all the LPs, LPTs, and FPs you need to operate IP, virtual routers, IP routing protocols, and IP services.

You must also configure all required IP media, such as ATM MPE, frame relay DTE, and point-to-point protocol (PPP), before you can configure IP. When you configure IP you will link the protocol ports of the virtual router to the IP media.

Table 8
IP on Passport configuration sequence

	Task
1	Download all required software. See 241-5701-270 *Passport 7400, 15000, 20000 Software Installation Guide*.
2	Configure all required LPs and LPTs. See 241-5701-270 *Passport 7400, 15000, 20000 Software Installation Guide*.
3	Configure all required FPs. See 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*, 241-1501-210 *Passport 15000, 20000 Hardware Installation Guide*, 241-7401-240 *Passport 7400 Hardware Installation, Maintenance and Upgrade*.
4	Configure all required IP media. See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.
5	Configure virtual routers and IP. See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.
6	Configure all required routing protocols. See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.
7	Configure IP features. See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Chapter 3
IP over ATM

This section describes the implementation of IP over ATM on Passport. It covers the following topics:

- “Overview of ATM MPE on Passport” (page 61)
- “ATM MPE media” (page 61)
- “Encapsulation methods” (page 63)
- “Inverse ARP on ATM” (page 65)
- “Frame forwarding for IP traffic” (page 66)

For information about configuring IP over ATM on Passport, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of ATM MPE on Passport

The ATM multiprotocol encapsulation (MPE) interface is an access service that allows IP encapsulation over ATM in accordance with RFC 1483. You can use the ATM MPE service to transmit IP traffic to interconnected external routers and other Passport virtual routers over an ATM network.

For information about FPs that support the ATM MPE service, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

ATM MPE media

The ATM MPE service allows IP traffic to be transmitted across the ATM network using the following two types of ATM MPE media:

- permanent virtual circuits (PVCs)

- soft PVCs

ATM MPE over PVCs

In a PVC, all the connection points through the network are defined, or nailed up. In this ATM MPE medium, standard encapsulation (using RFC 1483) allows the system to interoperate with any other RFC 1483 implementation (Nortel Networks or other vendor's equipment).

The Passport ATM MPE service follows the specifications detailed in RFC 1483, which describes two methods for carrying connectionless network traffic over ATM Adaptation Layer 5 (AAL5). The first method is Logical Link Control (LLC) encapsulation, where multiple upper layer protocols (ULPs) are carried over a single ATM Virtual Channel Connection (VCC). The second method is Virtual Circuit (VC) encapsulation, where only the IP protocol is permitted for each ATM VCC.

When you run the ATM MPE service on a CQC-based ATM FP, you must run the ATM MPE service in conjunction with an ILS Forwarder FP. IP forwarding over ATM on a CQC-based ATM FP alone is restricted to network management connectivity only.

ATM MPE over soft PVCs

The ATM MPE medium allows you to transmit IP traffic over soft PVCs in an ATM private network-to-network interface (PNNI) network. In a soft PVC, only the endpoints of the PVC are defined. PNNI routing provides route selection through the network between the endpoints. This ATM MPE medium interoperates only within Passport networks.

After the soft PVC is established, the dynamic component *AtmConnection* (*AtmCon*) is created by the system under the *Ac* component at both ends of the connection. The *AtmCon* component links to the ATM VCC through the *AtmIf Vcc Ep* component.

Carriers typically use soft PVCs to connect virtual connection gateways (VCG) across the backbone. However, any customer VR in an ATM PNNI network can be connected over soft PVCs. ATM soft PVCs support only the LLC encapsulation method. For more information on IP over soft PVCs, see

241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals* and 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management*.

Encapsulation methods

There are two methods for carrying connectionless network interconnect traffic over ATM Adaptation Layer 5 (AAL5):

- "LLC encapsulation" (page 63)
- "VC encapsulation" (page 65)

LLC encapsulation is supported on ATM MPE over PVCs and soft PVCs. VC encapsulation is supported on ATM MPE over PVCs only.

LLC encapsulation

Logical link control (LLC) encapsulation allows the ATM virtual circuits (VCs) associated with the ATM MPE interface to carry multiple protocols. For more information on LLC encapsulation, see RFC 2684.

For more information on LLC encapsulation on Passport, see

- "LLC encapsulation for routed protocols" (page 63)
- "LLC encapsulation for bridged protocols" (page 64)

LLC encapsulation for routed protocols

The protocol of the protocol data unit (PDU) is identified by prefixing the PDU with an IEEE 802.2 LLC header. The first octet is the Destination Service Access Point (DSAP), the second octet is the Source Service Access Point (SSAP), and the third octet is the Control (Ctrl) field. These fields indicate the type of PDU that follows.

An LLC header is followed by a SNAP header. A SNAP header exists when the LLC header has a value of 0xAA-AA-03. The first three octets of the SNAP header represent the Organizationally Unique Identifier (OUI) field; the next two octets represent the Protocol Identifier (PID) field.

The meaning of the PID field value depends on the OUI field value. For example, if the OUI field has the value 0x000000, the PID specifies an EtherType. The EtherType for IP, for example, is 0x0800.

LLC encapsulation for bridged protocols

LLC encapsulation for bridged protocols allows Passport to internetwork with a bridge over an ATM link. Currently the only type of supported bridged media over ATM MPE is Ethernet. While Passport does not provide bridging functionality, it can perform the following:

- receive bridged Ethernet LLC encapsulation packets and provide IP forwarding on them
- transmit bridged Ethernet LLC encapsulation packets into the bridge network

As in LLC encapsulation for routed protocols, the LLC header must be equal to 0xAA-AA-03. The LLC header is followed by a SNAP header. The Organizationally Unique Identifier (OUI) field in the SNAP header must be the 802.1 organization code 0x00-80-C2. The type of the bridged media must be specified by the Protocol Identifier (PID) field. The PID must also identify whether the original Frame Check Sequence (FCS) is preserved within the bridged protocol data unit (PDU). See the table “LLC encapsulation for bridged Ethernet PDUs” (page 65).

Do the following to enable LLC encapsulation for bridged protocols:

- Set the *AtmMpe encaptype* attribute to llcBridgeEncap. Also, it is recommended that you set the *AtmMpe maxTransmissionUnit* attribute to 1524.
- Ensure that the other end of the connection is a bridged port running over an ATM interface.

When *AtmMpe encaptype* is set to llcBridgeEncap, a MAC address is automatically assigned to the protocol port of the VR to which the *AtmMpe* component is linked. Address Resolution Protocol (ARP) is then used on demand to discover the layer 2 addresses of the Ethernet hosts that internetwork with the Passport through the remote bridge.

Bridged termination is supported only on PQC2-based OC-3 and OC-12 FPs.

Table 9
LLC encapsulation for bridged Ethernet PDUs

LLC header	OUI	PID	Padding	Supported media
0xAA-AA-03	0x00-80-C2	0x00-01 with preserved FCS	0x00-00	Ethernet
0xAA-AA-03	0x00-80-C2	0x00-07 without preserved FCS	0x00-00	Ethernet

VC encapsulation

Virtual circuit (VC) encapsulation allows the ATM virtual circuits (VCs) associated with the ATM MPE interface to carry one (and only one) protocol. Therefore, no protocol identifier is required since the Virtual Channel Connection (VCC) distinguishes between different protocols.

You configure the protocol type that is carried over a VCC, and must ensure that the protocol type is configured to the same value at both ends of the connection.

If the encapsulation type is *IpVcEncap*, Address Resolution Protocol (ARP) is not supported on that ATM MPE service. Since ARP is a protocol distinct from IP, no ARP packets can be transported on the ATM MPE service. If you use VC encapsulation, you must configure static ARP entries to ensure IP connectivity across the ATM network.

Inverse ARP on ATM

Note: A full implementation of RFC 1577 is not used, just the use of inverse ARP.

Inverse ARP provides a method for dynamically discovering the IP address at the remote end of a VCC. When inverse ARP is absent (for example, when the remote end does not support inverse ARP or VC encapsulation is used), the IP address of the remote end must be provisioned.

For more information related to inverse ARP on ATM, see “Inverse ARP scalability” (page 39).

Frame forwarding for IP traffic

Passport 7400 ATM MPE supports VCCs that terminate on CQC-based ATM FPs and on ATM IP FPs. If you are using CQC-based ATM FPs, you must configure the ATM MPE service in conjunction with an ILS Forwarder FP. The Passport 15000 and 20000 support IP forwarding on ATM IP FPs only.

For more information, see the following sections:

- "Frame forwarding on CQC-based ATM FPs for Preside Multiservice Data Manager connectivity" (page 66)
- "Frame forwarding using the ILS Forwarder FP" (page 67)
- "Frame forwarding on ATM IP FPs" (page 69)

Frame forwarding on CQC-based ATM FPs for Preside Multiservice Data Manager connectivity

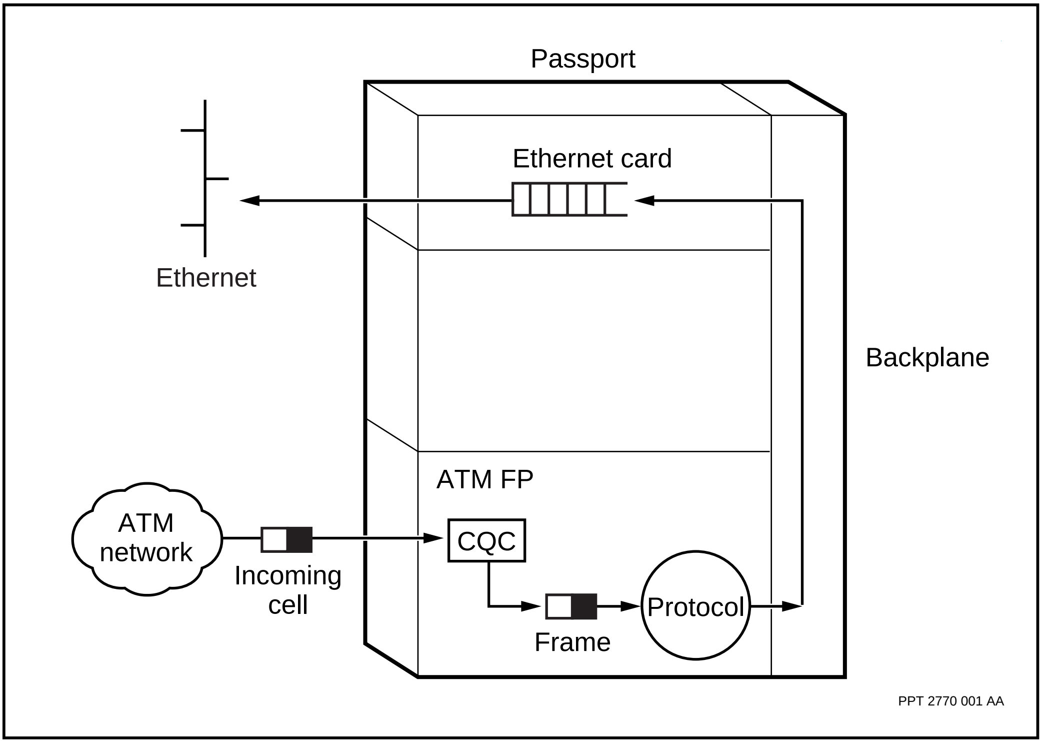
For Preside Multiservice Data Manager connectivity only, you can configure the ATM MPE service on a CQC-based ATM FP alone, for IP forwarding over ATM.

By default, the Passport 7400 frame forwarding decisions for IP traffic over ATM are made on the ATM FP's cell queue controller (CQC). The applicable protocol stack resides on the ATM FP. See the figure "Frame forwarding on an ATM FP" (page 67) for an illustration of the frame forwarding process when you use a Passport 7400 CQC-based ATM FP.

There are no special configuration procedures to enable this capability. If you do not configure the *ilsForwarder* attribute under the *AtmMpe* component and link it to an *Lp IlsForwarder* component, the ATM FP performs all the tasks necessary to handle the encapsulated IP frames. If you delete an existing *IlsForwarder* component from the current software view, all references to the *AtmMpe* component are also removed; this enables the ATM FP's frame forwarding functions.

For information about using a Passport 7400 CQC-based ATM FP in conjunction with an ILS Forwarder FP, see "Frame forwarding using the ILS Forwarder FP" (page 67).

Figure 8
Frame forwarding on an ATM FP



Frame forwarding using the ILS Forwarder FP

The ILS Forwarder FP is designed specifically for handling frames, and enhances the frame handling capability of frames coming in on a Passport 7400 CQC-based ATM FP. You can use an ILS Forwarder FP in conjunction with a CQC-based ATM FP to provide higher frame forwarding performance. In addition, other services running on the CQC-based ATM FP do not have to share resources (such as CPU and memory) with the ATM FP's frame forwarding service.

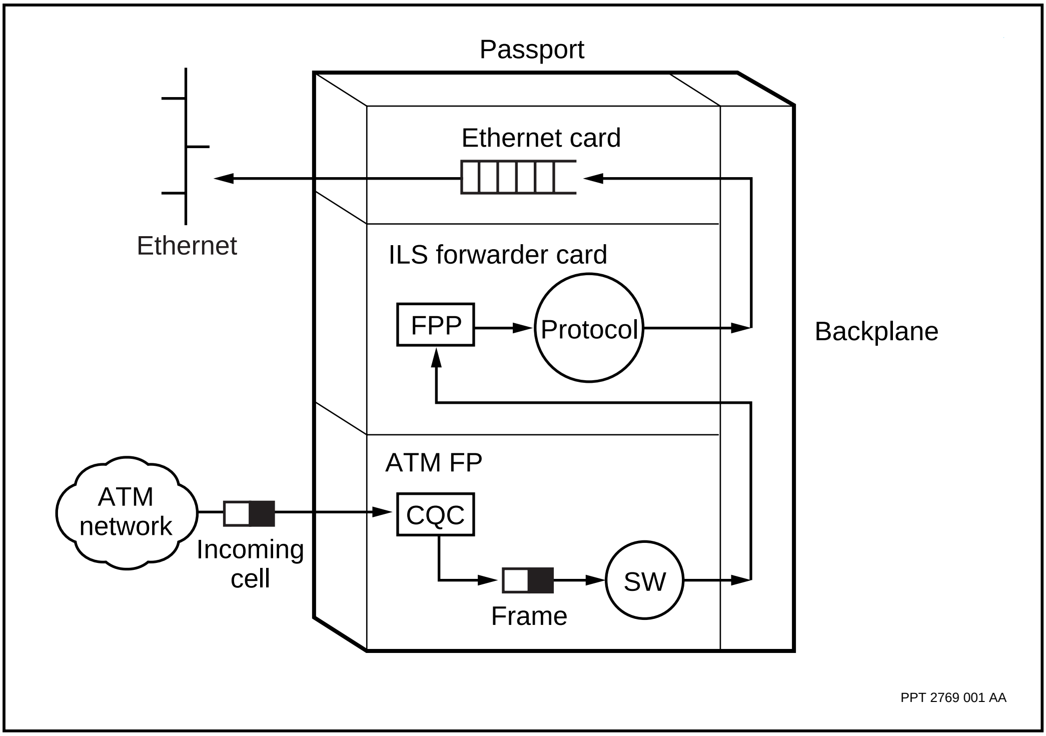
An ILS Forwarder FP makes the forwarding decisions, where

- one ATM FP links with multiple ILS Forwarder FPs
- multiple ATM FPs link to one ILS Forwarder FP

When you configure the *ilsForwarder* attribute under the *AtmMpe* component and link it to an *Lp IlsForwarder* component, all ATM MPE traffic that arrives on the CQC-based ATM FP is forwarded directly to the ILS Forwarder FP. The applicable IP protocol stack resides on the ILS Forwarder FP, and forwarding decisions are made with the assistance of the ILS Forwarder FP's fast packet processor (FPP) hardware.

The figure "Frame forwarding on an ILS Forwarder card" (page 68) illustrates the frame forwarding process when you use an ILS Forwarder FP.

Figure 9
Frame forwarding on an ILS Forwarder card



ILS Forwarder FP restrictions for ATM MPE

The ILS Forwarder FP has the following restrictions on the *AtmMpe* component:

- The total throughput of all the ATM connections forwarding to an ILS Forwarder FP should not exceed the maximum throughput of the ILS Forwarder FP.
- The maximum size (4475 bytes) of the frame that an *AtmMpe* component can accept is limited to the maximum size of the frame that passes through the FPP on the ILS Forwarder FP.
- The ILS Forwarder FP is supported on the Passport 7400 platform only. It is not supported on the Passport 15000 or 20000.

Frame forwarding on ATM IP FPs

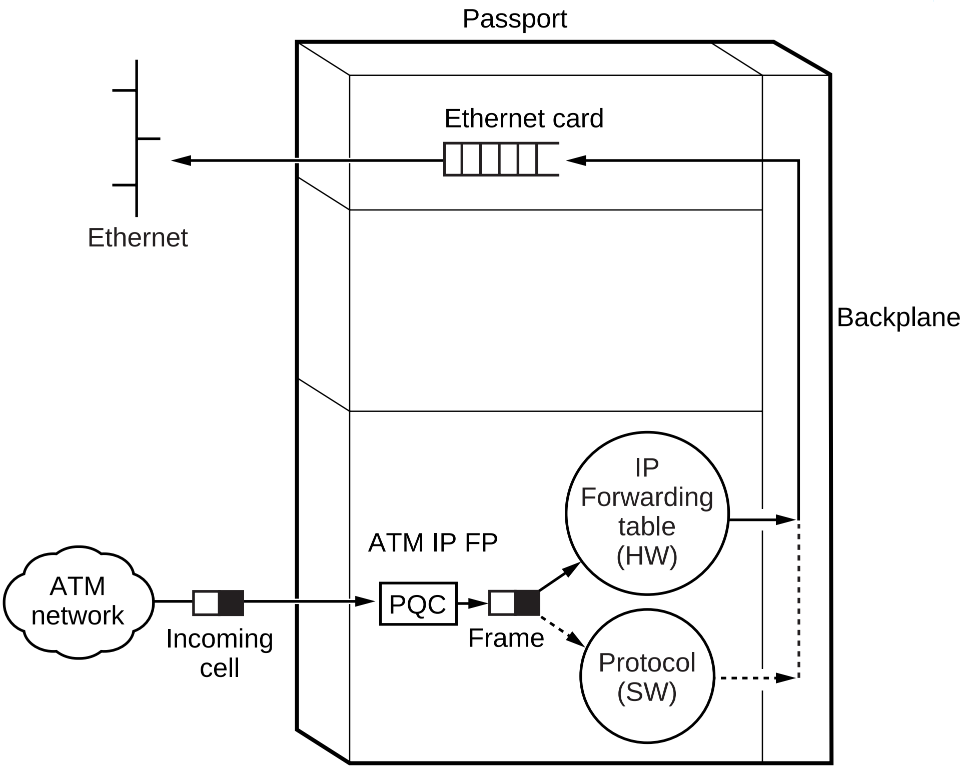
The ATM IP FP performs both hardware and software forwarding functions. Hardware forwarding is much faster than software forwarding, and operates independently of the software forwarding function. The ATM IP FP has specialized hardware to forward IP packets autonomously: it supports hardware lookups into IP forwarding tables, so it can forward almost all IP packets without the aid of the CP or the ATM IP FP's processor. See the figure “Frame forwarding on an ATM IP FP” (page 70) for an illustration of the frame forwarding process when you use an ATM IP FP.

To route IP packets properly, the hardware forwarding function must have a full IP forwarding table for each configured instance of a virtual router (VR). The IP software distributes a copy of these tables to each ATM IP FP. You can limit the number of routes stored in the ATM IP FP's hardware IP forwarding tables by configuring the *ipRoutesPoolCapacity* attribute under the *Lp Eng Fcrc Pqc Ov* component. This value applies to the entire LP, so all virtual routers on the LP should be taken into account and the capacity for IP routes set accordingly.

The ATM IP FP's software forwarding function processes traffic for routes that are not found in the hardware. Packets remain in the software datapath as long as the condition that caused them to take the software datapath persists. This can occur as a result of IP packet fragmentation, or when the maximum size of the hardware forwarding table is exceeded. Once the condition clears, traffic flow reverts back to the hardware path.

IP software running on an ATM IP FP does not require the use of an ILS Forwarder card. The *ilsForwarder* attribute under the *AtmMpe* component has no bearing on traffic received on the ATM IP FP. You can, however, use an ILS Forwarder card with the Passport 7400 CQC-based ATM FPs in the same node as ATM IP FPs. If you set the *ilsForwarder* attribute under the *AtmMpe* component, IP packets that arrive on CQC-based ATM FPs are sent to the ILS Forwarder for processing, and packets that arrive on ATM IP FPs are processed locally.

Figure 10
Frame forwarding on an ATM IP FP



PPT 2770 001 AB

Chapter 4
IP over frame relay using frame relay DTE

This section describes the implementation on Passport of IP over frame relay using frame relay DTE, which is an alternative to "IP over frame relay using IP-optimized DLCIs" (page 85) as an access media. It covers the following topics:

- "Overview of frame relay DTE (FR DTE) on Passport" (page 71)
- "Data link connection identifiers (DLCIs)" (page 73)
- "FrDte to FrUni connectivity" (page 75)
- "Congestion control" (page 82)
- "Committed information rate (CIR)" (page 83)

For information about configuring IP over frame relay on Passport, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

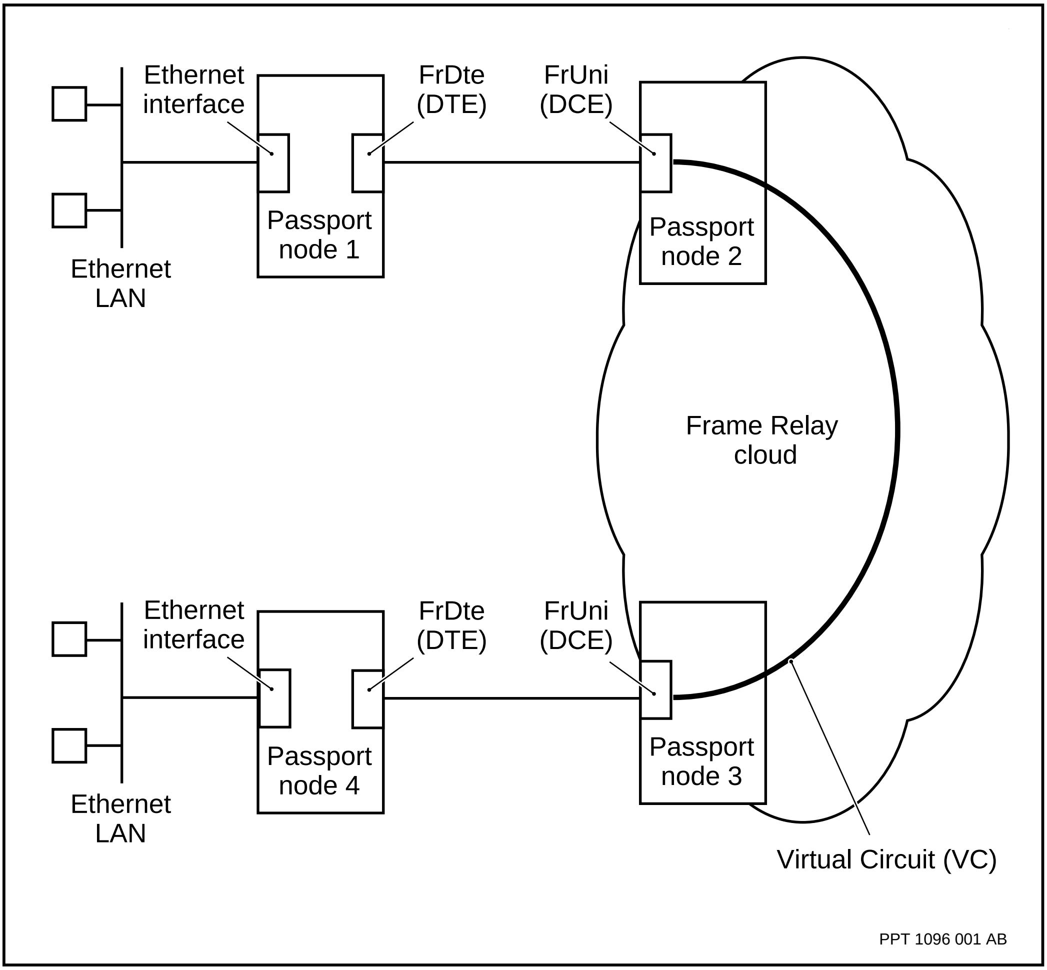
Overview of frame relay DTE (FR DTE) on Passport

The frame relay connection, called a virtual circuit (VC), is provided through a standard interface between the user device and the network. The interface is called the user-to-network interface (UNI). The connection that attaches UNI to the VC is provided by data circuit terminating equipment (DCE). The connection that attaches UNI to a device is provided by data terminating equipment (DTE). The figure "Frame relay overview" (page 72) illustrates how frame relay works. This section describes the access software associated with the DTE endpoint of frame relay UNI. This access software is referred to as frame relay DTE and the provisionable component of the software is *FrameRelayDte* (*FrDte*).

Encapsulation as defined in RFC 2427 and RFC 1490 is supported. This specifies recognition of the control field and a one byte padding field. It specifies use of the NLPID for IP protocol and SNAP encapsulation for other protocols. Passport supports only IP protocol.

Inverse ARP for IP protocol is supported as specified in RFC 1293.

Figure 11
Frame relay overview



Data link connection identifiers (DLCIs)

The FR DTE software is capable of automatically setting up Data Link Connection Identifier (DLCI) dynamic subcomponents for dynamically learned circuits. This feature is enabled or disabled through the *acceptUndefinedDlci* attribute located under the *FrDte* component. When enabled, a *DynamicDlci* subcomponent will be created when the frame relay network notifies the FR DTE of a new permanent virtual circuit (PVC) through the LMI protocol and a corresponding *StaticDlci* subcomponent does not exist. A *DynamicDlci* subcomponent will also be created when a frame is received on the FR DTE interface over a PVC which does not yet exist.

DynamicDlci subcomponents are always linked to the *RemoteGroup/1* instance (a mandatory component) and inherit the attributes provisioned under the *DynamicDlciDefaults* (*DynDlciDefs*) subcomponent. A *DynamicDlci* subcomponent can be removed using the Clear verb or by replacing it with a *StaticDlci* subcomponent. Although a *DynamicDlci* can have a Committed Information Rate (CIR) enforced on egress frames sent out on it, it cannot have a *HibernationQueue* (Hq) subcomponent to buffer frames in violation of the rate enforcement policy. Only *StaticDlci* components, on a Passport 7400 with SBIC-based FPs, can have *Hq* subcomponents.

You can provision dynamic DLCIs using commands if you are in operational mode. Static DLCIs are provisioned using commands in provisioning mode. See “Operational mode” (page 26) and “Provisioning mode” (page 26).

See also...

- “Local management interface (LMI)” (page 73)
- “Remote groups” (page 75)

Local management interface (LMI)

LMI is used between a frame relay end station and the local switch that is directly attached. (The Passport system acts as a frame relay DTE end station or as the local switch.) It allows each end of the frame relay UNI service to verify that the other end is operational, and also allows the end station to learn from the switch which PVCs are active. LMI is provisioned through the *LinkManagementInterface* (*Lmi*) component. Several standards apply:

- Vendor Forum LMI support:

- — Frame Relay Specification with Extensions, Doc. No. 001-208966.
- — “Section 4: Physical Interfaces” is dependent on the attached link.
- — “Section 5: Data Link Interface” in entirety. The default value of dN1=1604.
- — “Section 6: LMI” in entirety with the frame relay interface defined as “DTE”.
- — “Section 7: Optional Extensions” supports the PVC status of the update status message. The D, R, and PVC bandwidth fields are ignored.

- ITU-T, Annex A is supported, with the following exceptions:
  - — “Section A.3.3: PVC Status” is supported except that only two-byte DLCIs are recognized.
  - — “Section A.6: Optional Network Procedures” is not supported.
  - — “Section A.7: System Parameters” is supported as follows: full compliance with default parameters (timer T392 pertains to the network and is not applicable to frame relay DTE); full Status Polling Counter N391 is set to 6 polling cycles; Error Threshold N392 is set to 3 errors; Monitored Events Count N392 is set to 4 events; Link Integrity Verification Polling Timer T391 is set to 10 seconds.
- ANSI T1.617, Annex D is supported, with the following exceptions:
  - — “Section D.3.3: PVC Status” is supported except that only two-byte DLCIs are recognized.
  - — “Section D.6: Optional Network Procedures” is not supported.
  - — “Section D.7: System Parameters” is supported as follows: full compliance with default parameters (timer T392 pertains to the network and is not applicable to frame relay DTE); full Status Polling Counter N391 is set to 6 polling cycles: Error Threshold N392 is set to 3 errors; Monitored Events Count N392 is set to 4 events; link Integrity Verification Polling Timer T391 is set to 10 seconds.

Remote groups

IP over frame relay on Passport supports remote groups as follows:

- Multicast frames are transmitted across each PVC in the associated remote group.
- Each frame relay DTE remote group is modeled as a fully connected mesh network by IP. If a network is not fully connected, it can be divided into smaller subnetworks until it is fully connected under each remote group.

FrDte to FrUni connectivity

The frame relay user-to-network interface (FrUni) is the standard interface between the user device and the network. The FrDte is the frame relay interface into an IP network. There are three methods of implementing FrDte to FrUni connectivity: physical (hairpin), logical, and direct. The table “FrDte to FrUni connectivity on Passport FPs” (page 75) summarizes which Passport FPs support each method.

Table 10
FrDte to FrUni connectivity on Passport FPs

	Passport 7400 SBIC	**Passport 7400 MSA32**	**Passport 15000 and 20000**
Physical (hairpin)	supported	not supported	not supported
Logical	supported	supported	supported
Direct	not supported	supported	supported
Note: Where supported, a direct connection is the recommended method.			

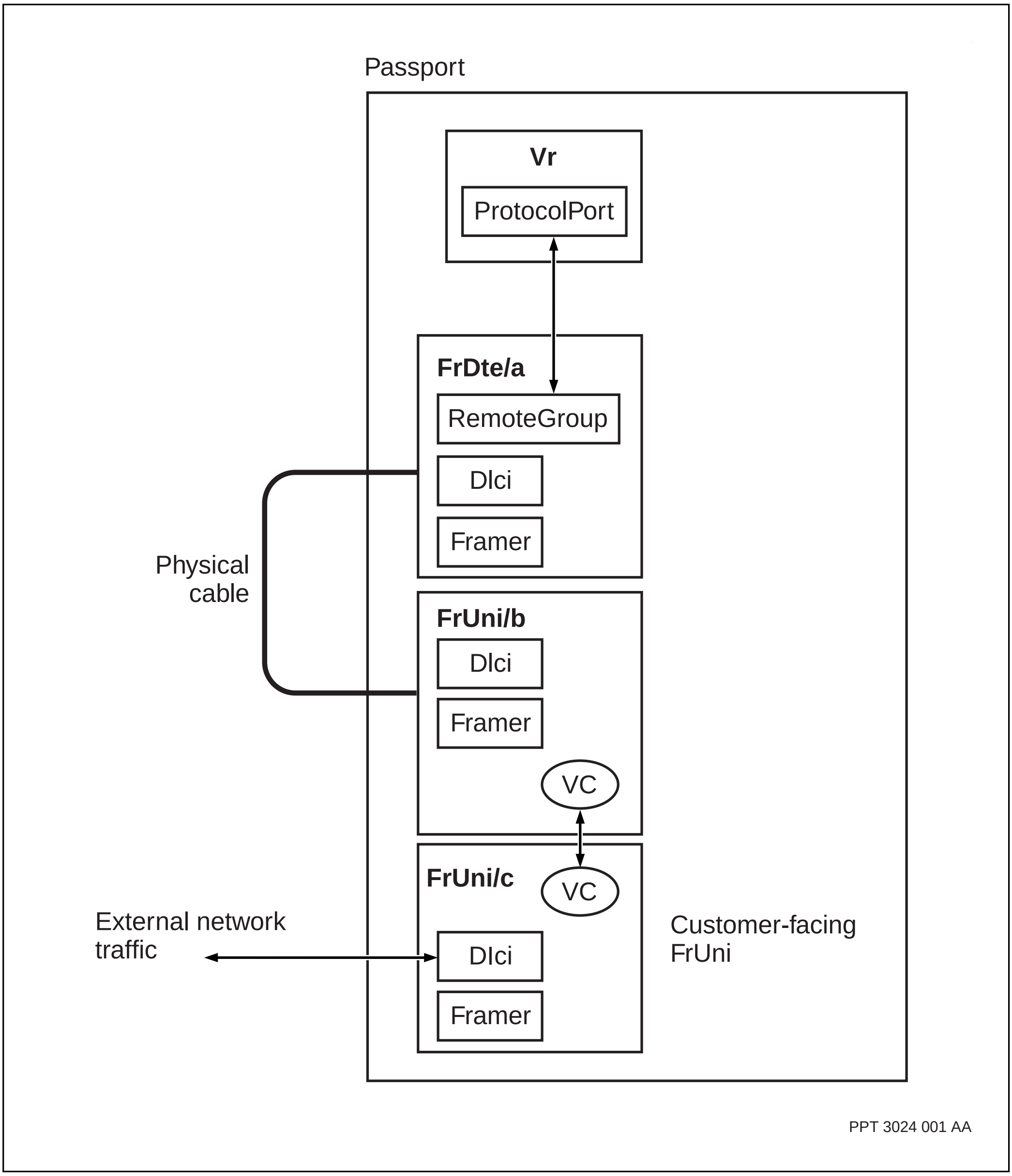
See the following sections for more information:

- “Physical (hairpin) connection” (page 76)
- “Logical connection” (page 78)
- “Direct connection” (page 80)

Physical (hairpin) connection

Use this configuration if you are using a hardware connection to link the FrDte and FrUni. In this configuration, each interface must be linked to a physical port through their respective *Framer* components. In addition, the two ports must be physically linked through a cable. Configure a PVC to the customer-facing FrUni on the Passport.

Figure 12
FrDte to FrUni connection with a physical link



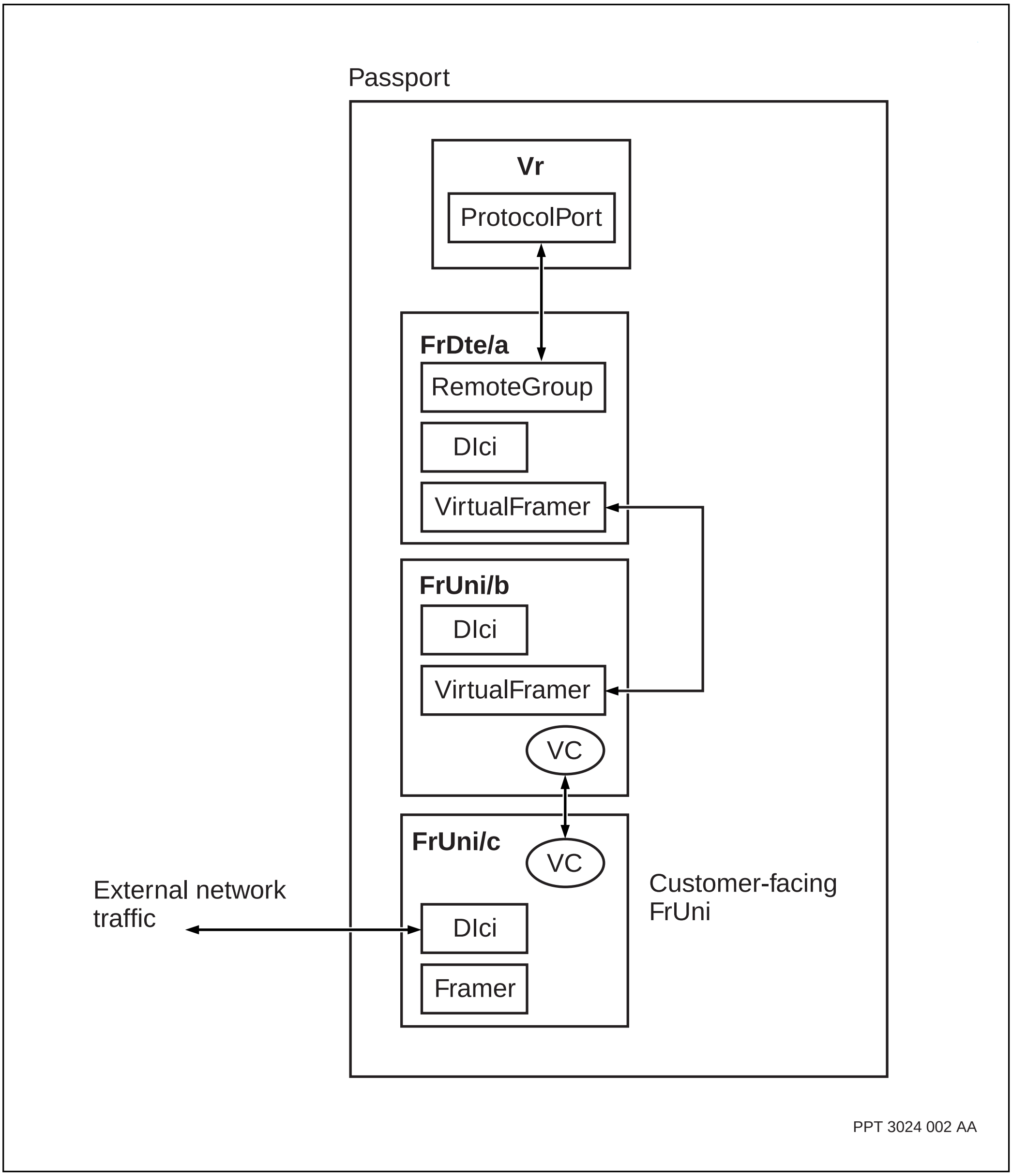
Logical connection

You can conserve physical ports by using an internal software connection to link the FrDte and FrUni interfaces through their respective *VirtualFramer* (*VFramer*) components. Configure a PVC to the customer-facing FrUni on the Passport.

If you use a logical connection, you need to delete the *Framer* components, which are automatically created on installation, and add *VFramer* components in their place. The components linked by the *VFramer* components must reside on the same card.

CoS to QoS mappings over a single DLCI are not supported on the logically connected FrDte and FrUni ports.

Figure 13
FrDte to FrUni connection with a logical link



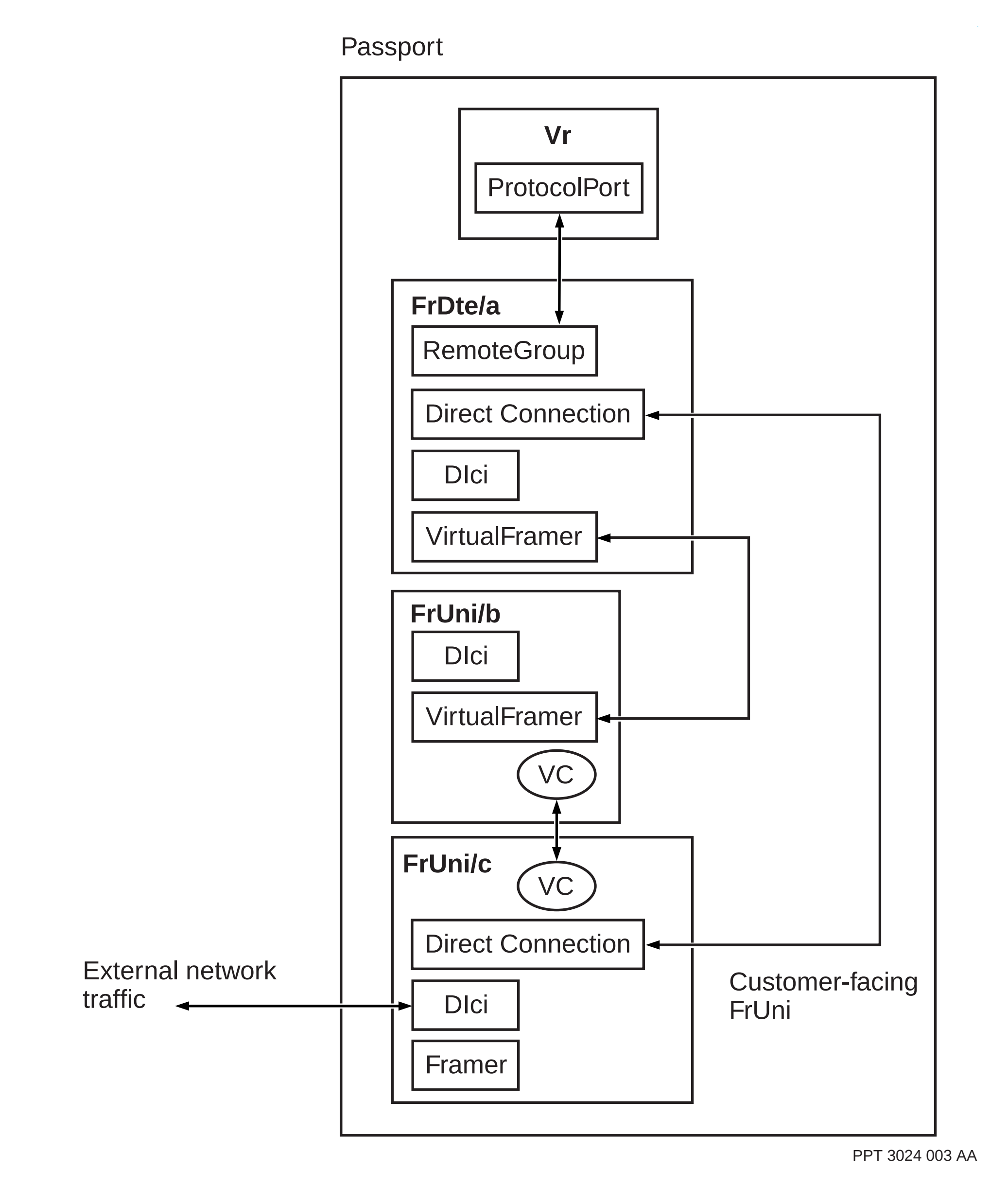
Direct connection

This configuration enables you to create an alternative data path (direct connection) between the FrDte and the customer-facing FrUni using a single DLCI. Using a direct connection causes a significant improvement in performance.

The direct connection uses the same frame relay and IP configuration as a logical connection, except that you need to add DirectConnection (Dconn) components to the FrDte and customer-facing FrUni and link them.

> ***Note 1:*** When you use a direct connection, the FrDte and FrUni must reside on the same card.
>
> ***Note 2:*** When you use a direct connection, you must use a single DLCI for CoS to QoS mapping. See “CoS to QoS mapping over a single DLCI” (page 186) for more information.

Figure 14
FrDte to FrUni connection with a direct connection link



Congestion control

The primary objective of congestion control is to maintain the quality of service by minimizing frame discards, warning end users of congestion, and minimizing the possibility of one end user consuming network resources at the expense of other users. It is recommended that all traffic management, including congestion control, be performed at the entry point of the network, on the port of the customer-facing FrUni. See the figure “FrDte to FrUni connection with a physical link” (page 77) for an example of a customer-facing FrUni.

The following are the congestion control mechanisms found in this software:

- The Committed Information Rate (CIR) is the rate that the network agrees to transfer information over a a virtual circuit under typical conditions. CIR enforcement prevents one user from utilizing an inequitable share of the network resource. In FrDte, CIR enforcement is a provisionable feature on a per DLCI basis. *committedburst* (*bc*), *committedInformationRate* (*cir*), Excess Information Rate (EIR), and Time Measurement Interval (Tc) are used to define this feature.

  In this feature, CIR enforcement is applied on the egress data path (i.e. frames sent out the interface to the WAN cloud) on a per DLCI basis as provisioned by the user. The ingress data path (i.e. frames received on the interface from the WAN cloud) is not monitored by the FrDte for rate enforcement.

- The Explicit Congestion Notification (ECN) bits are two bits in the Q.922 address part in the frame header that may be set by the frame relay network to notify the user that frames are encountering congestion. The objective of ECN is to reduce demands on resources so the network can return to normal operation mode. There are two types of congestion notification: Forward Explicit Congestion Notification (FECN) and Backward Explicit Congestion Notification (BECN). The user may also set the FECN or BECN bit.

  Unlike DE bit, FECN and BECN bits are not set by FrDte during processing of the frame. All arriving frames at FrDte, the FECN and BECN bits are noticed and counted in appropriate statistics. However, no action is taken to throttle traffic on the basis of these indications.

- The Discard Eligibility (DE) bit in Q.922 address part of the frame header is also used for congestion management. The DE bit determines whether a frame should be discarded in preference to other frames. It may be set by either the user or the service. The discard of discard eligible frames during congestion ensures that priority is given to frames within the CIR limit.
- *LinkEmissionQueue* (*Leq*), which is an optional component under *FrDte*, provides services that help to control congestion by putting limits on the size of the queue, the maximum number of multicast packets allowed to transmit, and a time to live limit for each packet in the queue. One use of this feature is rate shaping. The *Leq* is supported only on the Passport 7400 with SBIC-based FPs.

Committed information rate (CIR)

The CIR is the rate (in bit/s) that the network agrees to transfer information over a PVC under typical conditions. On the FR DTE interface, CIR enforcement is applied on the egress data path (i.e. frames sent out the interface to the frame relay cloud) on a per DLCI basis as provisioned by the user. The ingress data path (i.e. frames received on the interface from the frame relay cloud) is not monitored by the FR DTE for rate enforcement.

There are five attributes (provisionable on each *Dlci* subcomponent) that control the rate shaping performed on each PVC:

- *rateEnforcement* (*re*) - this attribute enables or disables rate enforcement on the PVC i.e. if it is disabled, the attributes *committedInformationRate* (*cir*), *committedBurst* (*Bc*), *excessBurst* (*be*), and *excessBurstAction* (*beAction*) cannot be used to define rate enforcement.
- *committedInformationRate* (*cir*) - this attribute is the CIR (in bit/s) to be enforced on egress frames sent over the PVC.
- *committedBurst* (*bc*) - this attribute is the committed burst size (in bits) for the PVC. This value represents the maximum amount of data that can be sent as a burst of traffic over the PVC at a rate above the CIR.

- *excessBurst* (*be*) - this attribute is the excess burst size (in bits) for the PVC. This value represents an additional amount of data which may be sent in excess of the BC amount at a rate above the CIR. Traffic operating in this range may have the Discard Eligible flag set in the frame relay header if provisioned to do so by the excessBurstAction attribute.
- *excessBurstAction* (*beAction*) - this attribute controls whether the frame relay header Discard Eligible flag is set in frames which are exceeding the committed burst level (i.e. traffic operating in the excess burst range). Frames marked as Discard Eligible are discarded at the discretion of the network.

The Leaky Bucket algorithm implements the rate enforcement mechanism. This algorithm enforces conformance to a single rate.

Traffic that violates the rate shaping policy is typically discarded. However on a Passport 7400 with SBIC-based FPs, a *HibernationQueue* (*Hq*) subcomponent can be provisioned under a *StaticDlci* component as a temporary holding area for violating frames. While at least one frame remains on the *Hq*, all frames being forwarded out that PVC are automatically placed on the *Hq*. This is done to preserve ordering. Frames are then forwarded from the *Hq* in the order they were placed on the queue (higher priority frames first) as the rate shaping policy allows.

The *Hq* provides the same services as the *LinkEmissionQueue* (*Leq*) with the following noted exceptions:

- an *Hq* holds frames destined to be sent out a particular PVC. Therefore, the Indexed and Balanced service classes are not applicable
- the Hardware Forced flag is ignored

All data frames which make it through the CIR enforcement are placed on the high priority link queue for transmission. Frames that do not undergo CIR enforcement (that is, *rateEnforcement* is disabled for the PVC) are placed on the normal priority link queue for transmission and are also subject to be placed on the Link Emission Queue if one is provisioned.

Chapter 5
IP over frame relay using IP-optimized DLCIs

This section describes the implementation on Passport of IP over frame relay using IP-optimized DLCIs, which is an alternative to “IP over frame relay using frame relay DTE” (page 71) as an access media. It covers the following topics:

- “Overview of IP-optimized DLCIs” (page 85)
- “Frame relay congestion notification” (page 86)
- “LMI and A-bit status” (page 86)

For information about configuring IP over frame relay on Passport, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP-optimized DLCIs

An IP-optimized data link connection identifier (DLCI) can directly bind to a virtual router protocol port. This type of DLCI is linked to the Passport frame relay user-to-network interface (FRUNI), which eliminates the need for a frame relay DTE as described in “IP over frame relay using frame relay DTE” (page 71) and simplifies provisioning.

IP-optimized DLCIs can be combined with existing DLCI types (PVC, SVC, or SPVC across DPRS, or BnxIwf DLCI) so that a single FRUNI can have several DLCIs that are PVCs across DPRS and several IP-optimized DLCIs that are linked to a protocol port. The ability to have several types of DLCIs on a single interface facilitates the migration to IP-optimized DLCIs.

Due to its flexibility and performance, using IP-optimized DLCIs is the preferred method for frame relay access to an IP VPN. For IP VPN information, see 241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals* and 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management*.

Frame relay congestion notification

Existing frame relay congestion notification is used to signal local congestion to the CPE device. In the egress direction, the CPE is notified with BECN when the receive queue is congested, and with FECN when the transmit queue is congested.

In the ingress direction, FECN and BECN are counted by the DLCI but do not trigger any special behavior. When local congestion occurs, frames tagged with discard eligibility (DE) bits are discarded before frames whose DE bits are not set.

For more information on FECN and BECN, see 241-5701-901 *Passport 7400, 15000, 20000 Frame Relay Fundamentals*.

LMI and A-bit status

The Passport frame relay local management interface (LMI) can run network side procedure (NSP), user side procedure (USP), or both procedures at the same time. All existing protocols (Vendor Forum, ITU, and ANSI) are supported and can be used with IP-optimized DLCIs with no restrictions.

See also...

- “Network side procedure” (page 86)
- “User side procedure” (page 87)

Network side procedure

When the LMI is running the network side procedure, it must periodically report the state of its DLCI. Attribute *FrUni Dlci aBitReasonToIf* is set as shown in the table “A-bit status and reason signaled to the CPE device” (page 87).

An IP-optimized DLCI is reported active when it is unlocked and the protocol port to which it is attached is active.

Table 11
A-bit status and reason signaled to the CPE device

aBitReasonToIf	**Cause**
notApplicable	Used when aBitStatusToIf is active.
localLmiError	Used when the local FRUNI is down.
localLinkDown	Used when the local link is down.
pvcSpvcDown	Used when the DLCI is locked or its protocol port is disabled.
remoteUserSignaled	The remaining values of *aBitReasonToIf* are not used by IP-optimized DLCI.
remoteLinkDown	
remoteLmiError	
userNotAuthorized	
resourceNotAvailable	
dlciCollisionAtNni	

User side procedure

When IP-optimized DLCI is deployed at the edge of the network, the LMI should be set to perform the network side procedure. However you can set the LMI to run the user side procedure by setting attribute *FrUni Lmi side* to user or both.

With the user side procedure, the LMI periodically polls the local CPE for the status of each DLCI. The status reported by the LMI (network side procedure) is combined with the local status (user side procedure) to create the PVC state reported to the virtual router.

Attribute *FrUni Dlci aBitReasonFromIf* is set as shown in the table “A-bit status and reason signaled from the CPE device” (page 88).

Table 12
A-bit status and reason signaled from the CPE device

aBitReasonFromIf	Cause
notApplicable	Used when the status of the DLCI is active. An active status is reported to the virtual router.
remoteUserSignaled	Used when the remote user is down. An inactive status is reported to the virtual router.
localLmiError	Used when the LMI is down. An inactive status is reported to the virtual router.
localLinkDown	Used when the link is down. An inactive status is reported to the virtual router.
missingFromLmiReport	Used when the LMI is up, but the adjacent LMI was not reported by the adjacent LMI. An inactive status is reported to the virtual router.

Chapter 6
IP over gigabit Ethernet

This section describes the implementation of IP over gigabit Ethernet on Passport 15000 and Passport 20000. It covers the following topics:

- “Overview of IP over gigabit Ethernet on Passport” (page 89)
- “IP datapath interworking” (page 90)
- “IP packet sizes” (page 91)

For information about configuring IP over gigabit Ethernet on Passport, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP over gigabit Ethernet on Passport

Gigabit Ethernet is an IP forwarding and routing service that enables the use of gigabit Ethernet media on a VR protocol port, in combination with legacy media.

IP over gigabit Ethernet is available on Passport 15000 and Passport 20000. It supports IP encapsulation over Ethernet at a rate of up to 1 gigabit per second per IP port for IP routing and forwarding. The following functionality is supported for IP over gigabit Ethernet on Passport:

- full compliance with RFC 894
- TCP and UDP transport layer protocols
- ICMP and ARP control protocols
- static routes
- RIPv2, OSPF, and BGP-4 routing protocols

- IP class of service (CoS)
- IP differentiated services

IP datapath interworking

IP datapath interworking between gigabit Ethernet ports and the following ports is supported over the backplane:

- ATM MPE media ports on PQC-based ATM FPs. See the table "Supported ATM FP types for IP datapath interworking" (page 90) for details.
- Voice Services Processors 2 and 3 (VSP2 and VSP3) ports for Packet Voice Gateway (PVG) VoIP applications. See 241-5701-780 *Passport 7400, 15000, 20000 Packet Voice Gateway Technology Fundamentals* for more information.
- CP3 OAM Ethernet 100BaseT ports for network management traffic

For more information on FPs, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

For more information on CPs, see 241-1501-200 *Passport 15000, 20000 Hardware Description*.

Table 13
Supported ATM FP types for IP datapath interworking

ATM FP type	PEC code
12-port E3 ATM	NTHR25DA
16-port OC3/STM-1 ATM	NTHW21, NTHW31
1-port OC-12/STM-4 ATM	NTHR29DA
4-port OC-12/STM-4 ATM (PQC12)	NTHW86xx
4-port OC-3 ATM (multimode) (PQC2)	NTHR17DA
4-port OC-3 ATM (multimode) (PQC1)	NTHR17CA
4-port OC-3 ATM (multimode) (PQC12)	NTHW05AA

IP packet sizes

The table “IP packet sizes” (page 91) shows the minimum and maximum supported sizes for gigabit Ethernet packets. Maximum transmission unit (MTU) refers to the maximum IP packet size carried by the Ethernet frames.

The minimum packet size that is transported by the Ethernet link is 64 bytes. If a packet that is less than 64 bytes is received over the plane from another FP, the FP pads the packet to 64 bytes with little or no impact on performance.

Table 14
IP packet sizes

Packet type	Minimum (octets)	Maximum (octets)	MTU (octets)
Ethernet V2.0	64	1600	1500
Ethernet 802.3 LLC-SNAP	64	1600	1492

Chapter 7 IP over point-to-point protocol (PPP)

This section describes the implementation of IP over point-to-point protocol (PPP) on Passport. It covers the following topics:

For information about configuring IP over PPP on Passport, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP over PPP

Point-to-point protocol (PPP) is a link layer protocol (LLP). It provides a simple, reliable method of transporting IP datagrams by encapsulating them into a high-level data link (HDLC) frame.

PPP is designed for simple links that transport packets between two peers. These links provide full-duplex (simultaneous and bi-directional) operation and are assumed to deliver packets in order. PPP can provide a common solution for easy connection of a wide variety of hosts, bridges, and routers.

Although PPP operates at the link layer, which is layer 2 of the OSI model, it performs several functions for the network layer, which is layer 3. Its unit of information consists of 8-bit bytes with no parity. The WAN link required for

PPP does not require modem status signalling. Any type of leased line including copper, fiber optic, microwave or satellite can be used to transmit PPP messages.

PPP consists of three main functional components:

- a method for encapsulating multi-protocol datagrams
- a link control protocol (LCP) for establishing, configuring, and testing the data-link connection
- a family of network control protocols (NCPs) for establishing and configuring different network protocols

IP over PPP implementation on Passport

IP over PPP is supported on Passport 7400 (both SBIC-based and MSA32 FPs), Passport 15000 and 20000. For more information on which FPs support IP over PPP, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

> ***Note:*** On MSA32 FPs and Passport 15000 and 20000 FPs, IP over PPP is supported on PQC 2.0 or later FPs.

The following features are supported on the Passport implementation of IP over PPP:

- On all supported FPs, PPP’s maximum receive unit (MRU) is 18,000 octets.
- Cold standby equipment protection is provided on the 4-port DS3CH FR FP.
- A maximum of 200 PPP services are supported per MSA32 FP and Passport 15000 or 20000 FP.
- Weighted fair queuing is supported on Passport 7400 MSA32FPs and on Passport 15000 and 20000.

Link transmission and monitoring features

On SBIC-based FPs only, the *Ppp Leq* component provides a link emission queuing (LEQ) feature, which allows the service to assign transmit priorities to ensure that selected high priority packets are transmitted before lower priority packets. In addition, LEQ controls the maximum number of packets and bytes that can be present in the queue at any time, and the amount of time that each packet can remain in the queue. These parameters can be tailored to a particular PPP implementation through provisioning. The LEQ is typically used for low data rate connections with delay sensitive traffic.

Using the *Ppp Link continuityMonitor* attribute, the link continuity monitoring (LCM) feature can continually confirm the link connection with the peer PPP application. LCM uses LCP echo packets to continually communicate with the peer. If more than five echo packets in succession are not received from the peer, the link is considered unusable, marked disabled, and attempts renegotiation of LCP.

The *Ppp Link configMagicNumber* attribute provides a method for detecting looped back links.

For Passport 15000 and 20000 and Passport 7400 MSA32 FPs, direct hardware forwarding allows, for example, an IP packet to be routed directly to the link queue of the card where PPP resides without software intervention on the FP. This dramatically improves full duplex packet switching performance over PPP. In order to use direct hardware forwarding, the link quality monitor (LQM) must be disabled. Do this with the *Ppp Lqm configStatus* attribute. On Passport 7400 MSA32 FPs and on Passport 15000 and 20000 FPs, these statistics are not enabled and should not be used for monitoring the link.

Weighted fair queuing (WFQ) provides a mechanism to queue traffic into separate traffic flows according to traffic class definitions, and ensures that low priority traffic has some opportunity to transmit. For Passport 15000 and 20000 (4pDS3Ch and 1pSTM1Ch) and Passport 7400 MSA32 FPs, the *Ppp Ewfq* component provides configurable weighted fair queuing using a priority guaranteed WFQ algorithm based on a frame rate for IP traffic over PPP. This component provides the ability to reconfigure the WFQ table to redistribute the transmission opportunities among the four egress traffic queues. For each of the four egress queues the transmission opportunity values range from 0% to 100%. The default values for the four queues are 91%, 3%, 3%. and 3% for queues 0 to 3 respectively.

For more information on provisioning the features in this section, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

PPP Framer statistics

There are twelve statistics available under the *PPP Framer* component and on SBIC-based FPs, these statistics are all valid. But on MSA32 and Passport 15000 and 20000 FPs, these statistics should never be used as they are not valid for FPs that are not SBIC-based.

Two of these statistics, *frmToIf* and *frmFromIf*, have equivalent counts available under the *Vr IfTableEntry* component. For the remaining ten statistics, there are no equivalent statistics on MSA32 and Passport 15000 and 20000 FPs. See the table “PPP Framer statistics” (page 96) for details.

Table 15
PPP Framer statistics

PPP-related statistics available on SBIC-based FPs under the *PPP Framer* component	**Equivalent statistics on Passport 7400 MSA32 and Passport 15000 and 20000 FPs**
frmToIf	Vr IfTableEntry ifInUcastPkts
frmFromIf	Vr IfTableEntry ifOutUcastPkts
aborts	None
crcErrors	None
lrcErrors	None
nonOctetErrors	None
overruns	None
underruns	None
largeFrmErrors	None
frmModeErrors	None
outOfSequenceFrm	None
repeatedFrm	None

PPP outages

On all supported FPs, PPP experiences a brief outage during some provisioning changes. Specifically, for any protocol port that is linked to a *Ppp* component, a PPP outage occurs if any component or attribute under the *Vr Pp IpPort* component is changed. The outage is a result of PPP renegotiating the Network Control Protocol (NCP) layer. Effectively, any provisioning change to a PPP-linked *IpPort* component is treated as a critical change.

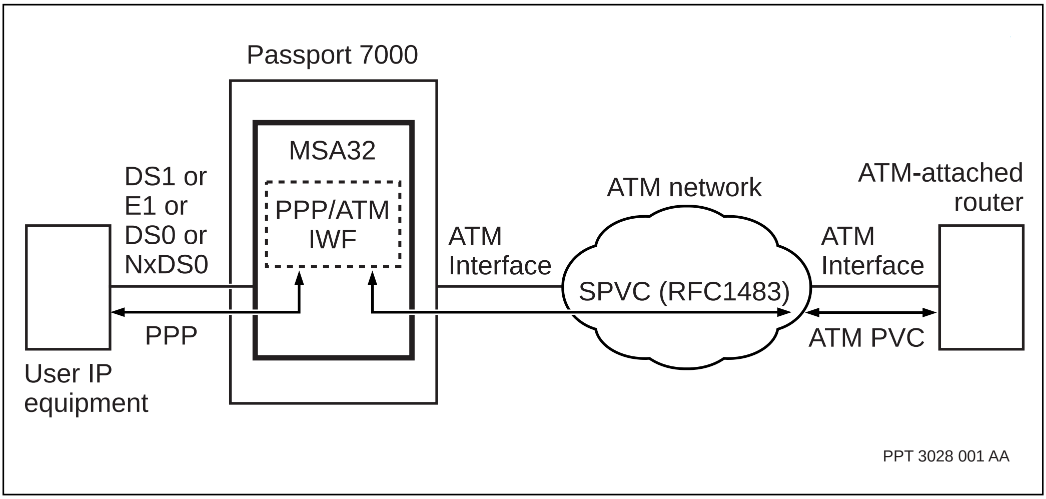
The NCP layer goes down for a short period, typically less than one second under non-stress conditions, and then automatically re-establishes itself. A potential effect is that the PPP link’s routing protocol, for example, OSPF, may experience a brief outage.

Chapter 8
Point-to-point protocol (PPP)/ATM interworking for Passport 7400

The PPP/ATM interworking feature on 32-port MSA function processors, enables IP transport between PPP-attached user devices and ATM attached routers as shown in figure “PPP/ATM interworking on MSA32” (page 100).

In the ingress direction, the IP packets received on DS1 or E1 ports (or DS0 or NxDS0 channels) are extracted from the PPP protocol, re-encapsulated into ATM (RFC1483), and forwarded onto ATM SVPCs toward ATM-attached routers. In the egress direction, ATM-encapsulated IP packets are extracted, encapsulated into PPP, and forwarded on DS1 or E1 ports (or DS0 or NxDS0 channels). Routing or switching does not take place, and there is a 1-to-1 relationship between the port or channel carrying the PPP protocol and a corresponding ATM SPVC.

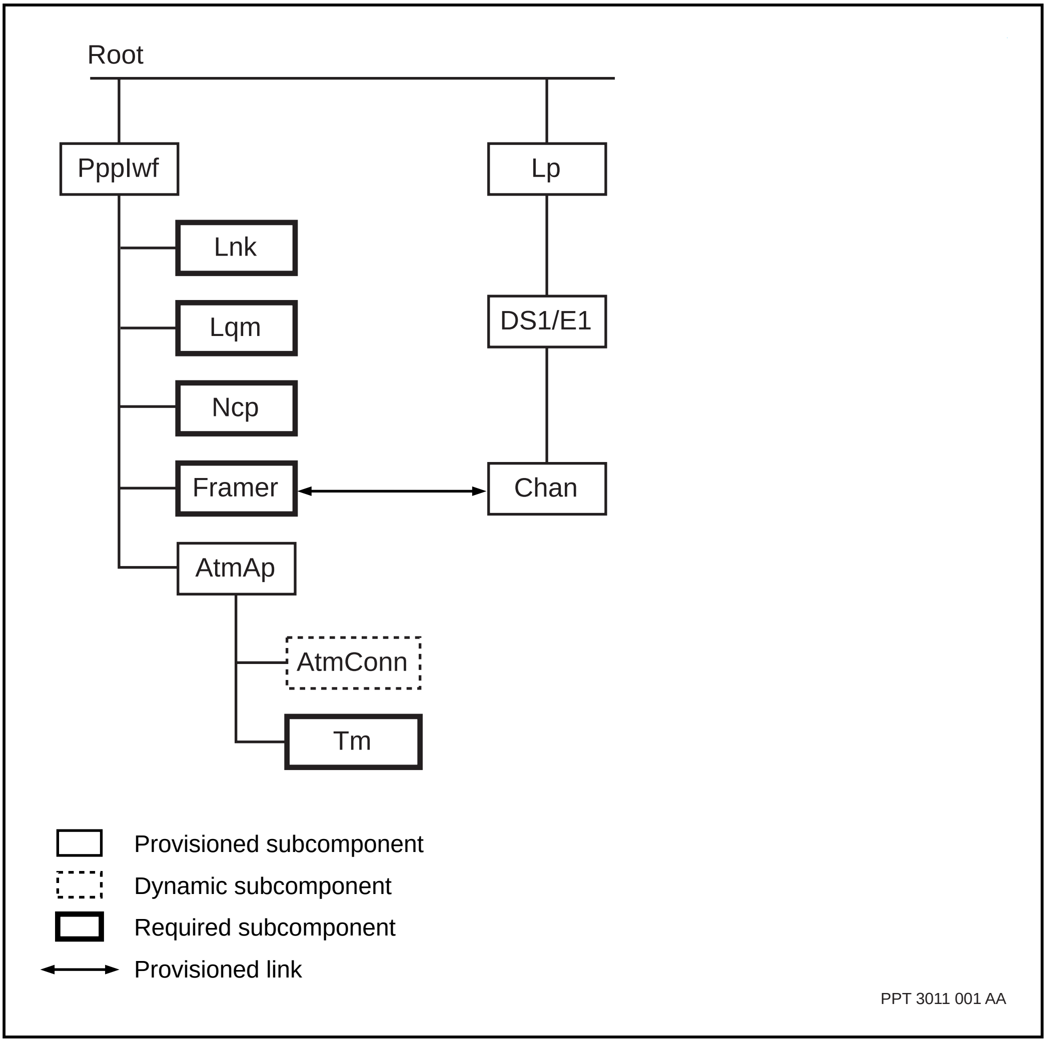
Figure 15
PPP/ATM interworking on MSA32



Software architecture of PPP/ATM interworking

The figure “Component hierarchy for PPP/ATM interworking” (page 101)shows the component hierarchy for the PPP/ATM interworking function.

Figure 16
Component hierarchy for PPP/ATM interworking



The components supporting the PPP/ATM interworking function are as follows:

- *PppIwf* defines an instance of the Ppp interworking function which provides the means for PPP to interwork with different forms of layer two protocols. For example, AtmMpe. The component instance number is a unique number.
- *Lnk* contains all the attributes related to the Ppp link.
- *Lqm* contains all the operational attributes for the Ppp link quality monitor.
- *Ncp* contains all the operational attributes of the Ppp network control protocols (NCP).
- *Fr*amer controls link layer framing for application components that send and receive data on a link interface. It is also through this component that an application component is associated with a specific hardware link interface.
- *AtmAp* represents the Soft PVC between *PppIwf* and *AtmIf*.
- *AtmConn* displays where the data traffic for this connection is directed.
- *Tm* contains ATM QoS information used by the SPVC connection to request a particular QoS from ATM.

Components and attributes of PPP/ATM interworking

This section describes the components and attributes of the PPP/ATM interworking function:

- “PppIwf/n component” (page 102)
- “PppIwf/n AtmAdaptationPoint component” (page 103)
- “PppIwf/n AtmAp TrafficManagement component” (page 103)

PppIwf/n component

This component defines an instance of the PPP interworking function.

Table 16
Configurable attributes for PppIwf/n

Attributes	Description
customerIdentifier (cid)	Identification of the customer who owns the CES.
ifAdminStatus	The desired state of the interface.
ifIndex	This is the index for the IfEntry.

PppIwf/n AtmAdaptationPoint component

This component is used to set up Soft PVCs to transmit and receive encapsulated IP data from the ATM interface. It also performs encapsulation of the IP data on the Ppp port.

Table 17
Configurable attributes for AtmAdaptationPoint

Attributes	Description
maxTransmissionUnit	The size of the largest datagram that can be sent on the Soft PVC.
encapType	The RFC1483 encapsulation type to be used.
localAddress	The local NSAP address.
calledVpiVci	The identity of the PVC at the remote ATM node on which the soft PVC connection terminates.
addressToCall	The remote NSAP address called a PNNI address.
firstRetryInterval	The time to wait in seconds, before attempting to establish the connection after the first failed attempt.
retryLimit	The maximum number of consecutive unsuccessful connection setup attempts that may be made before further attempts are abandoned.

PppIwf/n AtmAp TrafficManagement component

This component is used to request a particular QoS from ATM.

Table 18
Configurable attributes for PppIwf/n AtmAp TrafficManagement

Attributes	Description
atmServiceCategory	The desired ATM service category for an SPVC connection.
peakCellRate	The desired peak cell rate for an SPVC connection.

Chapter 9
IP routing management

To determine the optimum route to a destination, routers in a network must exchange route information. This function is accomplished by dynamic routing protocols and static routes. After you have configured IP and virtual routers on your Passport system, you must configure static routes or one or more dynamic routing protocols to enable the exchange of route information. The exchange of route information between two different routing protocols is called route redistribution.

This section is intended to help you plan the redistribution of route information among dynamic routing protocols operating on Passport VRs. It contains the following sections:

- “Overview of IP routing management” (page 106)
- “Route preferences” (page 111)
- “Example routing topologies” (page 114)
- “IP differentiated services for routing packets” (page 116)

For background information about IP routing, see the following sections:

- “Static routes” (page 157)
- “Routing information protocol (RIP)” (page 117)
- “Open shortest path first (OSPF) protocol” (page 125)
- “Border gateway protocol 4 (BGP-4)” (page 139)

For information about configuring IP routing, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP routing management

You can configure multiple dynamic routing protocols on a single VR to connect networks that use different dynamic routing protocols. For example, you can run RIP on one subnetted network, and OSPF on another subnetted network, and exchange routing information between them in a controlled fashion. Import and export policies control the exchange of information between protocols and between routers.

This section covers the following topics:

- “Routing policies” (page 107)
- “Flow of routing information” (page 108)

A Passport VR supports simultaneous operation of one instance of

- routing information protocol (RIP). See “Routing information protocol (RIP)” (page 117).
- open shortest path first (OSPF) protocol. See “Open shortest path first (OSPF) protocol” (page 125).
- border gateway protocol 4 (BGP-4). See “Border gateway protocol 4 (BGP-4)” (page 139).

Each of these dynamic routing protocols collects different types of information and reacts to topology changes in its own way. For example, RIP uses a hop-count metric to compute the shortest paths, while OSPF uses an interface-cost metric. RIP and OSPF are interior routing protocols used by routers to exchange information within a single autonomous system. BGP-4 is an exterior routing protocol (EGP) used by routers to exchange information among different autonomous systems, although it can also function as an interior routing protocol. BGP-4 uses an AS_PATH length metric to compute the shortest paths. Passport also uses static routes and local routes from a directly-connected interface to exchange information between virtual routers.

In the case where routing information is being exchanged between different networks that use different dynamic routing protocols, there are many configuration options that enable the exchange of routing information. For more information on some of these configuration options, see “Example routing topologies” (page 114).

Routing policies

Routing policy is a set of rules used to control the exchange of information between protocols and between routers. A protocol uses routing policy to determine which routes it learns from a peer to install in the routing database (RDB) and which information from the RDB to announce to other routers in the network.

All of the routing protocols have a default policy. In addition, each of the routing protocols can be configured with many different import and export policies on top of its default policy. If no other policies are configured, the routing protocol uses the actions contained in the default policy to filter route information.

A routing protocol has two sets of routing policies:

- import policy determines if a route learned from a peer can be installed in the RDB (used) or cannot be installed in the RDB (ignored)
- export policy determines if a route in the RDB can be sent to a peer (send) or cannot be sent to a peer (blocked)

Routing protocols on Passport use import and export polices in the following ways:

- RIP (versions 1 and 2) uses both import and export policies
- BGP-4 uses both import and export policies. By default, BGP-4 on Passport accepts all routes from any internal peer and rejects all routes from any external peer. BGP-4 also blocks any routes to its exterior BGP (EBGP) peers and advertises all routes learned from EBGP to its interior BGP (IBGP) peers.
- OSPF uses only export policy. It accepts all advertisements and therefore does not need to filter routes with import policy. However, it uses export policy to determine which non-OSPF routes can be sent to its neighbors.

Import and export statements use attributes (also called selection keys) to define information found in the protocol update packets or RDB. This information helps to discriminate between packets or table entries. Two examples of selection keys are the IP address and the protocol. See “Route preferences” (page 111) for more information on selection keys.

You can use multiple policy statements when configuring a router. The IP forwarding table uses a best match policy, where policy decisions are made from the longest (that is, the most specific) match between the policies and the routes stored in the RDB. To determine which match is most specific between policies, the selection keys available to each protocol’s import or export statements get a rating integer to order their importance. If two policies with different selection keys both match in selection criteria, then the rating integers of the selection keys that are set on each policy are added up. The one with the highest sum of rating integers is the best match policy. The selection keys available and their respective rating integer are listed with the import and export statements of each protocol.

Flow of routing information

Each dynamic IP routing protocol installs its routes in the routing database (RDB). The RDB stores all sources of routing information. The routing table manager (RTM) manages the RDB.

As part of the filtering process, the routing protocol applies local import policy to the information learned from a peer and places the best route to each destination in the RDB. Thus, the RDB contains the best routes computed by each protocol.

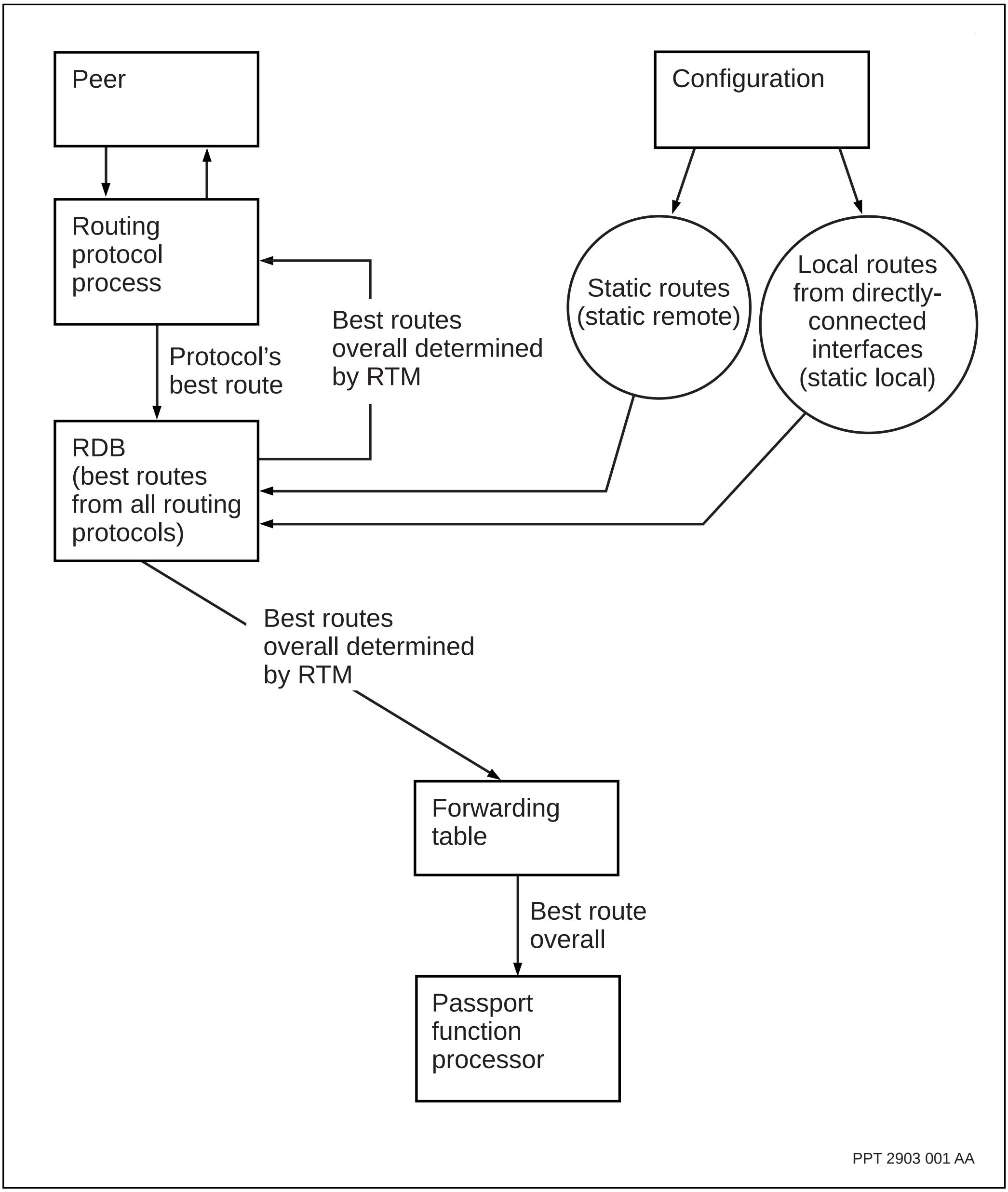
The RTM examines the RDB and selects the best overall route to each destination. If there is only one route to a given destination in the RDB, then it is the best overall route. If there are multiple routes to a given destination in the RDB, for example a route installed by OSPF and a route installed by BGP, then the RTM selects the best overall route according to a specific order. For more information on preferences in route selection, see “Route preferences” (page 111).

The RTM informs each routing protocol of the best overall routes. Each routing protocol applies its export policy to determine if it will send these routes to peers. See “Routing policies” (page 107) for more information.

The RTM also places the best overall routes in the IP forwarding table of the Passport node. The IP forwarding table stores the routes used by the forwarding process. The routes in the forwarding table are also sent to the Passport node’s function processors.

The “Flow of routing information through the Passport system” (page 110) illustrates how information flows from peers through the routing protocol, to the RDB, then to each routing protocol as well as to the IP forwarding table and the Passport’s FPs.

Figure 17
Flow of routing information through the Passport system


Peer
Configuration
Routing protocol process
Static routes (static remote)
Local routes from directly-connected interfaces (static local)
Best routes overall determined by RTM
Protocol's best route
RDB (best routes from all routing protocols)
Best routes overall determined by RTM
Forwarding table
Best route overall
Passport function processor
PPT 2903 001 AA

Route preferences

See the following sections for information on route preferences and their purpose:

- “Overview of route preferences” (page 111)
- “Forwarding classes and route preferences” (page 112)
- “Override recommendations” (page 113)

For information on the provisioning procedures used to modify the route preference for each routing protocol, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP* and 241-5701-445 *Passport 7400, 15000, 20000 Multiprotocol Label Switching Guide*.

Overview of route preferences

You can have more than one valid route to the same destination. When this happens, the routing manager uses the route preference to distinguish the different routes. The route preference is a number that is assigned to the route entry and it represents the dependability of the routing information. The lower the number, the more dependable the routing information. See the table “Default route preference values” (page 112).

You can assign different values to a route preference, allowing you to engineer how IP traffic is routed through your network. It is highly recommended that you do not change the default route preference unless you have done thorough engineering and modelling of the network and you fully understand the potential impacts.

When you change a route preference using a provisioning procedure, it is recommended that the change is applied to all virtual routers in the network/ autonomous system during low traffic periods in order to avoid potential routing loops. You do not need to stop and restart the protocol for the change to take effect.

If you use the same value for more than one route preference, the routing protocols’ metrics are used to determine which routing protocol is used. The protocol that has the best metric is used.

Table 19
Default route preference values

Protocol	Default route preference
Local, static discard	0 *
MPLS	10
Reserved for RIP migration	20 *
Reserved for RIP migration	21 *
OSPF internal	30
BGP external	70
Static remote	72
OSPF external type 1	80
RIP	82
OSPF external type 2	120
BGP internal	122
BGP aggregate	124 *
Reserved for internal use	254 *
UN_PREF	255 **

* reserved default values for internal protocols that are non-configurable

** used by the software and represents routes that are never put in the forwarding table

Forwarding classes and route preferences

Each route preference belongs to one of three forwarding classes. The lower the value of the forwarding class, the more important the routing information. The table “Forwarding classes” (page 113) shows how a route preference is assigned a forwarding class. It is recommended that whenever you change a route preference, to keep the new value within its existing forwarding class.

The forwarding class has an effect on routing. Specifically, a new more specific route can only be inserted into the forwarding table as long as less specific routes do not exist in a more important forwarding class.

Example
Assume an OSPF internal route (forwarding class 0) exists for address 10.1.0.0/16 and BGP internal (forwarding class 1) finds a route for address 10.1.1.0/24.

The BGP route, which is more specific, is not allowed in the forwarding table because its forwarding class is less important than the forwarding class of the OSPF route.

Table 20
Forwarding classes

Route preference range	Forwarding class
0-63	0
64-127	1
128-255	2

Override recommendations

In addition to changing the route preference, some protocols have an override facility that allows you to prefer one protocol over another. When you use these overrides, it is recommended that you use the following values. It is assumed that all other protocols are using their original route preference values.

- To prefer BGP internal routes over OSPF internal routes, the recommended setting for attribute *ibgpRtePref* is 6.
- To prefer BGP external routes over OSPF internal routes, the recommended setting for attribute *ebgpRtePref* is 6.
- To prefer static remote routes over OSPF internal routes, the recommended setting for attribute *staticRemoteRtePreference* is 5.

For information on the provisioning commands to use these overrides, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Example routing topologies

There are many configuration options that enable the redistribution of route information between different dynamic routing protocols. Some common redistribution scenarios include:

- “Route redistribution between two interior routing protocols within a single autonomous system (AS)” (page 114)
- “Route redistribution from an interior routing protocol to EBGP” (page 115)

Route redistribution between two interior routing protocols within a single autonomous system (AS)

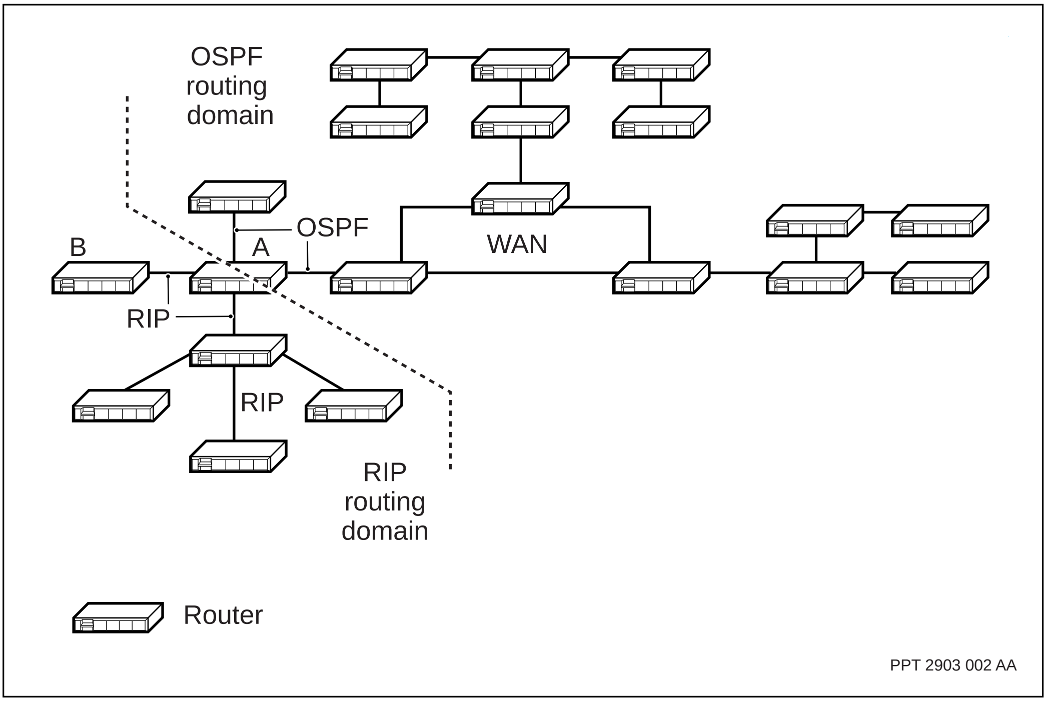
An autonomous system (AS) is a series of gateways or routers that fall under a single administrative entity and cooperate using the same Interior Gateway Protocol (IGP). The figure “Route redistribution between two interior routing protocols within a single AS” (page 115) illustrates the redistribution of route information from a RIP routing domain to an OSPF routing domain.

In this example, router B sends RIP routes to router A. Router A’s RIP process applies import policy to the routes it receives from router B. It might reject some routes. The RIP process selects the best routes from among the routes received from all RIP peers and installs these routes in the RDB.

The RTM selects the best routes overall from the RDB and installs them in the forwarding table. Some of the best overall routes might be RIP routes.The routes installed in the forwarding table are sent to the Passport FP and to the other routing protocols operating on router A.

The OSPF process on router A receives the routes sent to it by the RTM. The OSPF process applies export policy to these routes to determine if it will send them to OSPF neighbors. For more information on how OSPF manages the routes it receives, see “Route redistribution from an interior routing protocol to EBGP” (page 116).

Figure 18
Route redistribution between two interior routing protocols within a single AS



Route redistribution from an interior routing protocol to EBGP

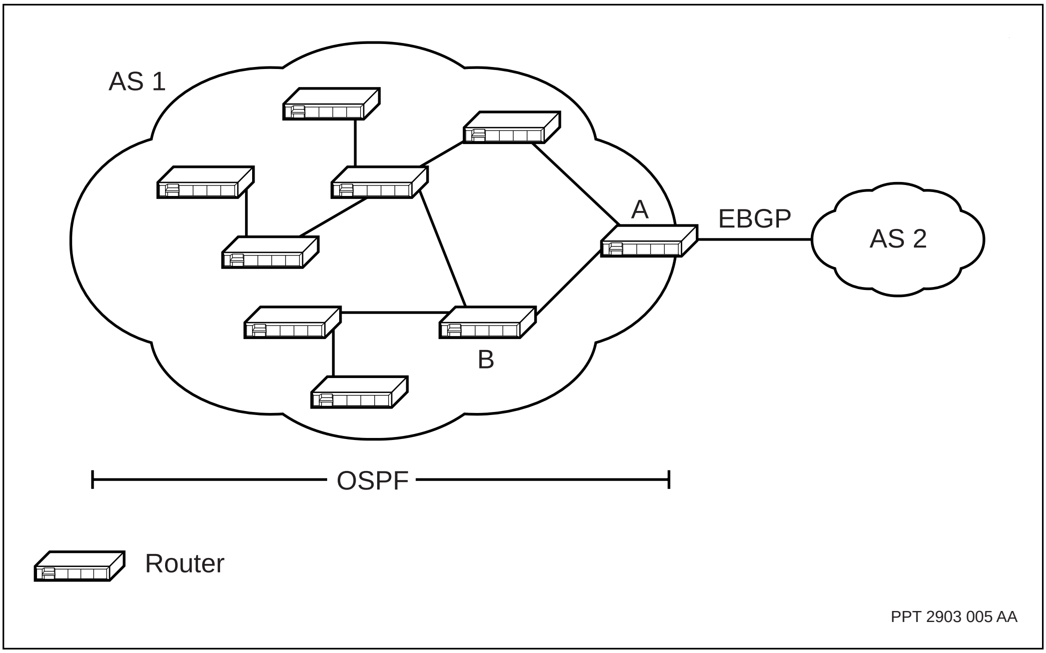
The figure “Route redistribution from an interior routing protocol to EBGP” (page 116) illustrates the redistribution of route information from an OSPF routing domain to a BGP-4 routing domain.

In this example, the OSPF process running on router A is exchanging link state advertisements (LSAs) with its neighbors in AS 1. Since OSPF has no import policies, it cannot filter these LSAs. OSPF computes the shortest paths to all destinations described in the LSAs and installs these routes in the RDB.

The RTM selects the best routes overall from the RDB and installs them in the forwarding table. Some of these overall best routes may be OSPF routes. The routes installed in the forwarding table are sent to the Passport FP and to the other routing protocols operating on router A.

The BGP process operating on router A receives the routes sent to it by the RTM. The BGP process applies export policy to these routes to determine if it will send them to the BGP peers in AS 2.

Figure 19
Route redistribution from an interior routing protocol to EBGP



IP differentiated services for routing packets

The “Differentiated services code point” (page 196) field of routing packets is controlled by the *VirtualRouter Ip dscpRoutingSource* attribute. See the *Vr Ip* section of the NTP 241-5701-060 *Passport 7400, 15000, 20000 Components* for more information on the *dscpRoutingSource* attribute.

Chapter 10
Routing information protocol (RIP)

This section describes the implementation of the routing information protocol (RIP) on Passport. It covers the following topics:

- “Overview of RIP on Passport” (page 117)
- “RIP policies” (page 118)
- “Migrating from RIPv1 to RIPv2” (page 118)
- “Migrating from RIP to OSPF” (page 122)

For information about configuring RIP, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of RIP on Passport

Passport provides seamless support for both routing information protocol version 1 (RIPv1) and RIP version 2 (RIPv2).

RIPv1 requires that entries in the routing database age and expire periodically. Either the old route is re-entered in the database or a new route takes its place. In this way, the database contains the latest possible data, and topology changes have a limited impact on packet delivery. RIPv1 uses a fixed subnet mask rule associated with different classes of IP addresses.

RIPv2 supports the same functionality as RIPv1, but adds support for variable length subnet masks, multicast, and next hop, as defined in RFC 1723.

RIP policies

RIP import policies define which learned set of routing information is given to the RIP routing process, as well as which metrics to use. You can assign non-default metrics so that traffic must use a particular route unless that route becomes unusable. You can also use import policies to define which routing processes (for example, OSPF) provide routing information to the RIP routing process.

RIP export policies define how to advertise routing information on specific interfaces. You can assign non-default metrics so that traffic must use a particular route unless that route becomes unusable. RIP export policies also define which routing processes, learned from a specific interface, can be exported. RIP export policies are optional.

> ***Note:*** If you provision a RIP export policy make sure you are aware of the export policies of the “send all staticLocal routes” and “send all RIP learned routes” as well as their possible impact on other configured policies.

For more information on import and export policy, see “Routing policies” (page 107). To configure import and export policies for RIP, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Migrating from RIPv1 to RIPv2

You can optimize the RIP configuration by migrating from RIPv1 to RIPv2. To migrate the Passport nodes in a network from RIPv1 to RIPv2, all Passport nodes in the network must be running a release of software that supports RIPv2 (R5.1 and later). Migrate all Passport nodes on a link-by-link basis, until all the nodes in the network are set to RIPv2.

The following example describes the sequence of steps involved in migrating from RIPv1 to RIPv2 using two Passport nodes. The steps correspond to the figure “Example migration from RIPv1 to RIPv2 using two Passport nodes” (page 120).

1 Add RIPv2 to Passport 2, but provision the RIP interface on Passport 2 to be backwards compatible with RIPv1 on Passport 1.

For example, set the ifConfSend attribute on the RIP interface of Passport 2 to v2b.

See 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP* for details on configuring the *ifConfSend* and *ifConfReceive* attributes under the *RipIf* component. See the table “Example migration: RIP behavior on two Passport nodes with different RIP configuration” (page 121) for the meaning of the values of the *ifConfSend* and *ifConfReceive* attributes.

The table “Example migration: RIP behavior on two Passport nodes with different RIP configuration” (page 121) also illustrates the behavior of the RIP interface on these two nodes for different combinations of attribute values provisioned for the *ifConfSend* and *ifConfReceive* attributes. This table can be useful in helping you planning your migration. The attribute values appear in the table in italics.

2 Configure the RIP interface on Passport 1 to support RIPv2 only.

For example, set the ifConfSend attribute on the RIP interface of Passport 1 to v2, and the ifConfReceive attribute to v2.

3 Change the RIP interface on Passport 2 to support RIPv2 only.

For example, set the ifConfSend attribute on the RIP interface of Passport 2 to v2, and the ifConfReceive attribute to v2.

4 Remove all RIPv1 components from Passport 2.

Figure 20
Example migration from RIPv1 to RIPv2 using two Passport nodes

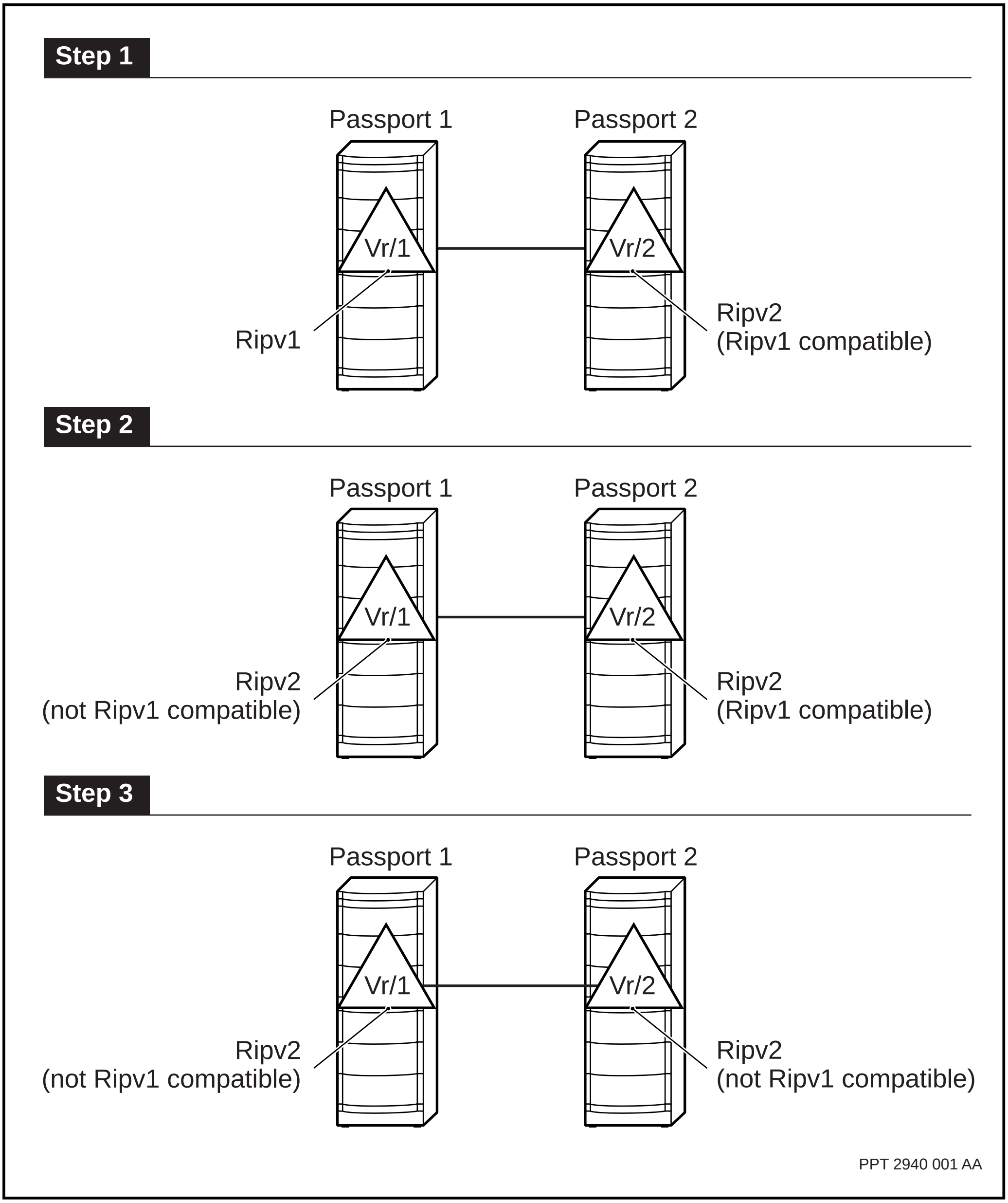


Table 21
Example migration: RIP behavior on two Passport nodes with different RIP configuration

ifConfSend attribute value on Passport 1 (Vr/1) (transmitting)	**ifConfReceive attribute value on Passport 2 (Vr/2) (receiving)**			
	v1 **(RIP 1)**	***v2*** **(RIP 2)**	***both*** **(RIP 1 or 2)**	***reject*** **(do not accept)**
silent (do not send)	No transmission	No transmission	No transmission	No transmission/ updates are rejected.
v1 (RIP 1)	RIP 1 updates broadcast by Vr/1. RIP 1 updates accepted by Vr/2.	RIP 1 updates broadcast by Vr/1. RIP 1 updates rejected by Vr/2.	RIP 1 updates broadcast by Vr/1. RIP 1 updates accepted by Vr/2. The Vr/2 RIP interface processes the updates as RIP 1 updates.	RIP 1 updates broadcast by Vr/1. RIP 1 updates are rejected by Vr/2.
v2b (RIP 1 compatible)	RIP 2 updates broadcast by Vr/1. RIP 2 updates accepted by Vr/2. The Vr/2 RIP interface processes the RIP 2 updates as RIP 1 updates. (The Vr/2 RIP interface ignores the subnet mask and next hop fields in the RIP 2 update.)	RIP 2 updates broadcast by Vr/1. RIP 2 updates accepted by Vr/2. Vr/2 RIP interface will process the subnet mask and next hop fields in the RIP 2 updates.	RIP 2 updates broadcast by Vr/1. RIP 2 updates accepted by Vr/2. Vr/2 RIP interface will process the subnet mask and next hop fields in the RIP 2 updates.	RIP 2 updates broadcast by Vr/1. RIP 2 updates are rejected by Vr/2.
(Sheet 1 of 2)				

Table 21 (continued)
Example migration: RIP behavior on two Passport nodes with different RIP configuration

ifConfSend attribute value on Passport 1 (Vr/1) (transmitting)	**ifConfReceive attribute value on Passport 2 (Vr/2) (receiving)**			
	***v1* (RIP 1)**	***v2* (RIP 2)**	***both* (RIP 1 or 2)**	***reject* (do not accept)**
v2 (RIP 2)	RIP 2 updates are multicast by Vr/1. Because the Vr/2 RIP interface is set for RIP 1 only, Vr/1 will not send RIP 2 updates to Vr/2.	RIP 2 updates are multicast by Vr/1. RIP 2 updates are accepted by Vr/2. Vr/2 RIP interface will process the subnet mask and next hop fields in the RIP 2 updates.	RIP 2 updates are multicast by Vr/1. RIP 2 updates are accepted by Vr/2. Vr/2 RIP interface will process the subnet mask and next hop fields in the RIP 2 updates.	Updates are rejected

(Sheet 2 of 2)

Migrating from RIP to OSPF

If your routing requirements change, you can migrate from RIP to OSPF. Use one of the following methods to migrate your network from RIP to OSPF. Using a route preference is preferable to the *migrateRip* attribute because enabling and disabling *migrateRip* restarts the protocol.

> ***Note:*** After enabling *migrateRip*, any subsequent route preference changes are not enabled until *migrateRip* is disabled.

Using a route preference

Change the route preference of either RIP or OSPF internal so that RIP routes are preferred over OSPF internal routes. The route preference change must be done on all RIP and OSPF virtual routers in the network/autonomous system. See “Route preferences” (page 111) for more information.

Using *migrateRip*

When making a transition from RIP to OSPF, the *migrateRip* attribute of the *Ospf* and *Rip* components can be enabled during the transition phase. This process allows both RIP and OSPF routes to be learned, but gives preference to RIP-learned routes. When OSPF has proven its stability, disable the *migrateRip* attribute to allow OSPF to take over route selection duties.

All RIP and OSPF virtual routers in the network/autonomous system must have the *migrateRip* attribute in the same state and the changeover must be made quickly to prevent routing loops.

For more information about migrating from RIP to OSPF using *migrateRip*, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Chapter 11
Open shortest path first (OSPF) protocol

This section describes the implementation of the open shortest path first (OSPF) protocol on Passport. It covers the following topics:

- "Overview of OSPF on Passport" (page 125)
- "OSPF export policy" (page 129)
- "OSPF equal-cost multipath routing" (page 129)
- "OSPF optimization" (page 129)
- "Migrating from RIP to OSPF" (page 132)

For information about configuring OSPF, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of OSPF on Passport

OSPF is a link-state route management protocol. Neighboring routers send routing updates, called link state advertisements (LSAs), that include the cost of internal and external routes. From information provided by neighboring routers, the OSPF process creates a tree-like map of the network with itself at the root. From this tree, it creates a routing table using the LSAs to calculate the shortest paths.

OSPF is defined in RFC1583.

This section covers the following topics:

- "OSPF areas" (page 126)
- "OSPF routing types" (page 127)

- “OSPF router types” (page 128)
- “OSPF virtual links” (page 128)

OSPF areas

An autonomous system (AS) is a series of gateways or routers that fall under a single administrative entity and cooperate using the same Interior Gateway Protocol (IGP). An OSPF autonomous system (AS) can be subdivided into areas. An area is a contiguous collection of networks and hosts. The topology of an area is invisible to routers outside the area, and the topologies of other areas are unknown from within an area. As a result, the topological database of each router in the AS can be different from other routers within the AS.

There are five OSPF area types: transit, stub, summary stub, not-so-stubby-areas (NSSA), and backbone:

- Transit areas can carry data traffic that neither originates nor terminates in the area itself. If an active virtual link transits an area, then the area must be a transit area.
- Stub areas are used when there is a single exit point from the area or when the choice of an exit point does not have to be made on a per-external-destination basis. Stub areas cannot include an autonomous system boundary router (ASBR) connecting the AS to an external AS.
- Summary stub areas are similar to stub areas except that they accept summary LSAs. Summary stub areas cannot include an ASBR.
- Not-so-stubby-areas (NSSA) are another kind of stub area that have the capability of importing external routes in a limited fashion. NSSAs can include an AS boundary router. External information learned from the ASBR propagates throughout the OSPF AS.
- Backbone areas consist of those networks not contained in any area, their attached routers, and those routers that belong to multiple areas (area border routers). The backbone must be contiguous, or have virtual links provisioned to make it contiguous. The backbone is a specialized transit area with a special area ID of 0.0.0.0.

OSPF supports two main configurations of areas:

- Hierarchical OSPF divides the internetwork into contiguous logical areas, with each area usually corresponding to a community of interest. Each area, including the area 0 backbone, runs a separate copy of the OSPF algorithm. Each router in a particular area knows how to reach every subnet in that area as well as how to reach the backbone. Once the datagram is given to the backbone, it is the backbone's job to route the datagram to the destination area. The destination area's interior routers know how to complete the routing to the destination end system. Hierarchical OSPF is characterized by smaller routing tables because of OSPF partitioning, but usually requires more router hardware.
- Collapsed backbone refers to the practice of having only one routing area defined, with all routers and networks in the AS belonging to the same area. Collapsed backbone configurations are ideal for small (under 20 routers) IP networks.

OSPF routing types

The routes that are learned from the OSPF protocol can be divided into two types, internal and external. Internal routes are internal to the OSPF routing domain and external routes are learned from other routing protocols and are for destinations outside the OSPF routing domain. By default, internal routes are preferred over external routes.

External routes are imported into the OSPF routing domain by autonomous system boundary routers (ASBRs). There are two types of external routes:

- OSPF external type 1, where the metric assigned to the external part of the route has the same order as the metric assigned to the internal part (source to ASBR) of the route. Add the internal and external costs to calculate the total cost of the route. For example, using hop-count as the metric in the OSPF routing domain, you can import RIP routes as external type 1 routes.
- OSPF external type 2, where the metric assigned to the external part of the route is more significant than the metric assigned to the internal part of the route. The total cost of the route is equal to the external cost.

By default, type 1 routes are preferred over type 2 routes. For more information on route preferences, see "Route preferences" (page 111).

OSPF router types

Passport can function as any of four types of OSPF routers. To configure any of these router types, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

- Internal routers are routers whose directly-connected networks belong to the same area. There can be internal routers in the backbone area if all of their interfaces are in the backbone. It is sufficient to configure a single instance of the *AreaTable* subcomponent for internal or routers, since they are part of only one area.
- Area border routers (ABR) connect to one or more areas and the backbone. Area border routers can condense or summarize the topological data of their attached areas for distribution on the backbone. The backbone in turn distributes the information to other areas. A single instance of the AreaTable subcomponent is sufficient for internal or backbone routers, since they are part of only one area. You must configure instance for each area of the *AreaTable* subcomponent for area border routers.
- Backbone routers are routers that have an interface to the backbone (including ABRs). A backbone router that has connections only to other backbone routers is also considered an internal router. A single instance of the *AreaTable* subcomponent is sufficient for backbone routers, since they are part of only one area.
- Autonomous system boundary routers (ASBR) exchange routing information with routers belonging to other autonomous systems. The path to each ASBR is known by every router in the autonomous system (AS), except in the case of stub areas. See “OSPF areas” (page 126) for more information. This classification is completely independent of the previous classifications. ASBRs can be internal routers or area border routers, and may or may not participate in the backbone.

OSPF virtual links

OSPF does not allow a non-contiguous backbone (area 0.0.0.0), nor does it allow a router to connect two or more areas, becoming an area border router, unless the router is part of the backbone. One solution is to provision a virtual link reaching from the isolated area border router, through one of the attached areas, to the backbone. For information on configuring virtual links, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

OSPF export policy

Export policies define how routing information is shared between routing processes (for example, between OSPF and RIP). Export policies are optional and are not present in cases where you don't want to share routing information. For information on configuring OSPF export policy, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Since OSPF normally uses IP routing information from all sources, OSPF import policies are not used. OSPF allows only export policies.

For more information on import and export policies, see "Routing policies" (page 107).

OSPF equal-cost multipath routing

OSPF supports an equal-cost multipath routing function where equal-cost paths are used in load-sharing mode. Equal-cost multipath routing supports up to three next hop addresses. For more information on equal-cost multipath routing, see "Static routes" (page 157).

OSPF optimization

See the following sections for more information on various ways to optimize OSPF:

- "Optimizing OSPF memory allocation" (page 129)
- "Hitless OSPF for CP/VpnXc switchover" (page 130)

There is additional information about OSPF optimization in "Inverse ARP scalability" (page 39).

Optimizing OSPF memory allocation

You can optimize OSPF memory allocation by configuring estimates for the number of routes in the network. Passport uses these numbers to calculate queue sizes for OSPF. In this way, you can make more effective use of memory by allocating only what is necessary.

You can set estimated values for the number of

- internal OSPF routes (routes learned from OSPF neighbors)

- external OSPF routes (routes learned from BGP-4 and RIP)
- OSPF areas in the AS
- OSPF interfaces in each area
- OSPF neighbors for each OSPF interface

For information on configuring route estimates, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*. If you do not set values for the parameters described, Passport uses default values to calculate OSPF queue sizes. These default values result in queue sizes similar to those allocated to the IP service prior to the introduction of the user-configurable parameters.

Hitless OSPF for CP/VpnXc switchover

Hitless OSPF refers to the ability to maintain a synchronized OSPF instance on the standby card. Its purpose is to further secure the core of an IP VPN. On a VCG-based IP VPN network configuration, hitless OSPF can be enabled on both VCGs and/or customer VRs. On a direct VR-to-VR configuration, it can be enabled on any virtual router.

Hitless OSPF can protect OSPF against failure in the following spaces: from the CPE to the customer VR, and from VCG to VCG.

See 241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals* and 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management* for more information on IP VPNs.

Hitless OSPF is available on the following:

- On a Passport that has either a control processor (CP) or both a CP and a VPN extender card (VpnXc) along with their respective standby modules
- On IP traffic over ATM MPE protocol ports on PQC2 and PQC12-based FPs
- On IP traffic over IP-optimized DLCI protocol ports on PQC2 and PQC12-based FPs
- On access connections with supported media (ATM MPE and IP-optimized DLCI) that use OSPF

Background

All routers running OSPF in the network maintain identical databases of link state advertisements (LSAs). OSPF accomplishes this by having routers synchronize their link state databases with neighboring routers by exchanging LSAs. Whenever the LSAs change in the database, the IP routes are recalculated. These new routes and other IP routing protocol routes make up the routing table.

Since an OSPF instance establishes relationships with its neighbors, an unplanned restart of OSPF affects not only the restarting router but also the general behavior of IP traffic within an OSPF network. IP traffic is interrupted by an unplanned OSPF restart due to two main reasons:

- In the restarting router’s routing table, OSPF-discovered IP routes are identified and removed, which limits its forwarding ability.
- The OSPF neighbors detect the restart and inform the network that the OSPF instance has restarted. This causes traffic to be redirected around the restarting router.

Hitless OSPF description

Enabling the hitless OSPF feature allows information such as the neighbor relationships and the link state database to be preserved during a planned or unplanned switchover. In particular, remote OSPF routers within the network, especially neighbors, do not detect any change in the topology. Generated IP routes prior to the switchover are not affected, allowing the standby card to seamlessly assume normal neighboring interactions.

The benefits of hitless OSPF are apparent when the switchover is executed on the card (CP or VpnXc) where OSPF resides. If OSPF resides on the VpnXc and a CP switchover is executed, OSPF is unaffected whether or not hitless OSPF is enabled. However under the following conditions, OSPF remains unaffected only when hitless OSPF is implemented:

- OSPF resides on the VpnXc and a VpnXc switchover is executed
- OSPF resides on the CP and a CP switchover is executed

Enable hitless OSPF by setting attribute *Vr Ip Ospf spareInstance* to enable. In order for hitless OSPF to function, you must set attribute *Shelf cpEquipmentProtection* to hot.

Migrating from RIP to OSPF

If your routing requirements change, you can migrate from RIP to OSPF. Use one of the following methods to migrate your network from RIP to OSPF. Using a route preference is preferable to the *migrateRip* attribute because enabling and disabling *migrateRip* restarts the protocol.

> ***Note:*** After enabling *migrateRip*, any subsequent route preference changes are not enabled until *migrateRip* is disabled.

Using a route preference

Change the route preference of either RIP or OSPF internal so that RIP routes are preferred over OSPF internal routes. The route preference change must be done on all RIP and OSPF virtual routers in the network/autonomous system. See “Route preferences” (page 111) for more information.

Using *migrateRip*

When making a transition from RIP to OSPF, the *migrateRip* attribute of the *Ospf* and *Rip* components can be enabled during the transition phase. This process allows both RIP and OSPF routes to be learned, but gives preference to RIP-learned routes. When OSPF has proven its stability, disable the *migrateRip* attribute to allow OSPF to take over route selection duties.

All RIP and OSPF virtual routers in the network/autonomous system must have the *migrateRip* attribute in the same state and the changeover must be made quickly to prevent routing loops.

For more information about migrating from RIP to OSPF using *migrateRip*, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Chapter 12

Intermediate System to Intermediate System (ISIS) Protocol

ISIS is a link-state routing protocol suitable for use as an Interior Gateway Protocol (IGP) within an Autonomous System (AS).

ISIS was originally specified in ISO 10589 as a protocol for exchanging CLNP routing information. The protocol was adapted, allowing it to be used for IP routing. The changes to the protocol are specified in RFC 1195. On Passport, the ISIS protocol is used for IP routing only.

For additional information about ISIS, see the following sections:

- “ISIS terminology” (page 133)
- “ISO based node identification” (page 135)
- “Default route” (page 136)
- “Media types” (page 137)

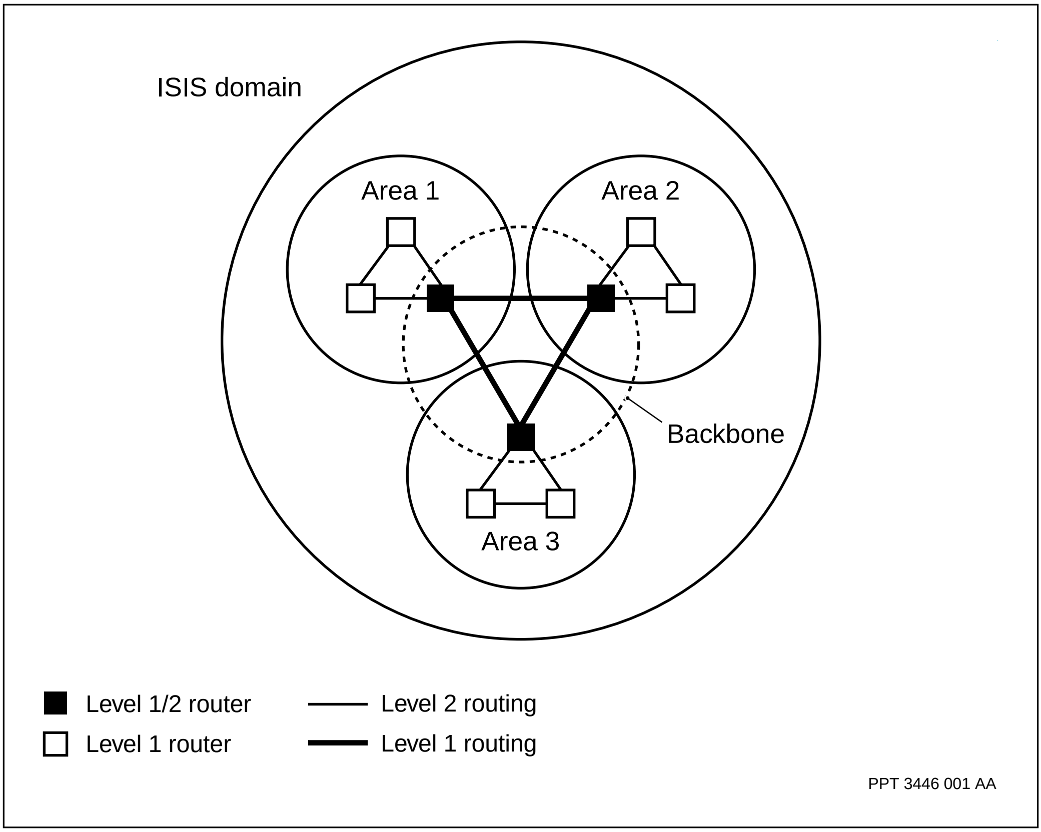
ISIS terminology

An ISIS routing domain is a network in which all the routers run ISIS to support intra-domain exchange of routing information. Routers within such a domain are called Intermediate Systems (ISs). An ISIS domain can be partitioned into smaller segments known as areas. Routers belonging to a common area engage in Level 1 routing, which involves the exchange of intra-area IP prefix information. Border routers in different areas may exchange inter-area routing information, this process is known as Level 2 routing. For information, see Figure 21, “Level 1/Level 2 routing,” (page 135).

A router engaged in Level 1 routing generates a Level 1 link state packet (LSP). A Level 1 LSP contains intra-area routing information. A router engaged in Level 2 routing generates a Level 2 LSP, which contains inter-area routing information.

For Passport, only Level 1 routing is supported. Passport nodes will not form an adjacency with a router that does not belong to the same area. A Passport node will not generate or accept Level 2 LSPs.

Figure 21
Level 1/Level 2 routing



ISO based node identification

Even when used for IP routing only, ISIS is based on ISO concepts. For example, ISIS uses an ISO addressing scheme for identifying an ISIS node.

ISO network layer addresses are called Network Service Access Points (NSAPs). An NSAP address consists of an Area Address, System ID, and Network Selector (NSEL). The Area Address uniquely identifies an area within an ISIS domain; the System ID uniquely identifies a node within an area. The NSEL identifies a network layer service on the node. On an ISIS node used exclusively for IP routing, there is only one network layer service, the ISIS routing engine itself. When the routing engine is specified as the network layer service, the NSEL is set to zero and the NSAP is called a Network Entity Title (NET).

Multiple NETs are permitted per node. These NETs must have the same System ID and are differentiated only by the Area Address. This does not mean that the router is connected to multiple separate areas, rather the router belongs to one area, which is known by multiple synonymous Area Addresses. Normally, a router would be configured with only a single Area Address and would therefore have only a single NET. However, the ability to configure multiple Area Addresses is useful for migration purposes. For example, renumbering, merging, or splitting areas. This allows operators to perform a migration of their ISIS area topologies without suffering service interruptions during the reconfiguration period. The Passport ISIS implementation allows for provisioning of up to 3 NETs per ISIS instance to support this functionality. These NETs must have identical System ID components and only differ in Area Address.

The Passport ISIS implementation uses a fixed-size, non-configurable System ID length of 6 bytes.

As an example, the following address illustrates the ISO format:

```
49.000001.12ca.0065.90ab.00
```

The first portion (49.0001) is the Area Address. The area Address can be from 1 to 13 bytes in length. The first byte of the Area Address (49) is referred to as the Authority and Format Identifier (AFI). The next 6 bytes (12ca.0065.90ab) are the System ID. The System ID can be any 6 bytes that allow the ISIS node to be uniquely identified within the domain. Common methods for choosing a System ID include using one of the MAC Addresses on the node or some derivation of an IP address belonging to the node (e.g. the IP address 47.202.187.168 could be transformed into the System ID 0472.0218.7168). The last byte (00) is the NSEL.

Default route

ISIS Level 1 Routers exchange only intra-area prefix information and, therefore, do not know about any routes outside their areas. Backbone routers exchange Level 2 LSPs with routers in other areas, but also exchange Level 1

LSPs with routers in their own area. A backbone router will set the attached bit in its Level 1 LSP, to indicate that it is attached to the backbone. A Level 1 router will install a default route to the nearest Level 2 router that has set the attached bit. All traffic for destinations outside the local area will be sent to that backbone router.

The Passport ISIS implementation functions as a Level 1 router only. If a Level 2 router with its attached bit set exists in the ISIS area, a default route to that Level 2 router will be installed into the routing table by ISIS.

Media types

The ISIS protocol distinguishes 2 main types of link layer media, general topology (point-to-point) and broadcast.

On Passport, ISIS supports the following types of media:

- Broadcast: Ethernet (4 port Gig E card only)
- General Topology: ATM (ATM PQC cards only)

Chapter 13
Border gateway protocol 4 (BGP-4)

This section describes the implementation of the border gateway protocol 4 (BGP-4) on Passport. It covers the following topics:

- “Overview of BGP-4 on Passport” (page 139)
- “BGP-4 routing policies” (page 143)
- “BGP-4 route selection” (page 148)
- “BGP-4 optimization” (page 150)

For information about configuring BGP-4, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of BGP-4 on Passport

The border gateway protocol 4 (BGP-4) is a routing protocol for exchanging network reachability information between autonomous systems (ASs). BGP-4 is replacing the external gateway protocol (EGP) as the routing protocol of the Internet backbone.

BGP-speaking routers establish peer relationships with other BGP-speaking routers over a TCP connection and exchange routing information. Using this reachability information, the BGP-speakers construct a map of AS connectivity that allows them to eliminate routing loops and enforce policy decisions at the AS level. BGP-4 also supports classless inter-domain routing as described in RFC 1519, and route aggregation.

BGP-4 is defined in RFC 1771 and RFC 1772.

This section covers the following topics:

- "BGP-4 peers" (page 140)
- "BGP-4 updates" (page 142)
- "BGP-4 path attributes" (page 142)

BGP-4 peers

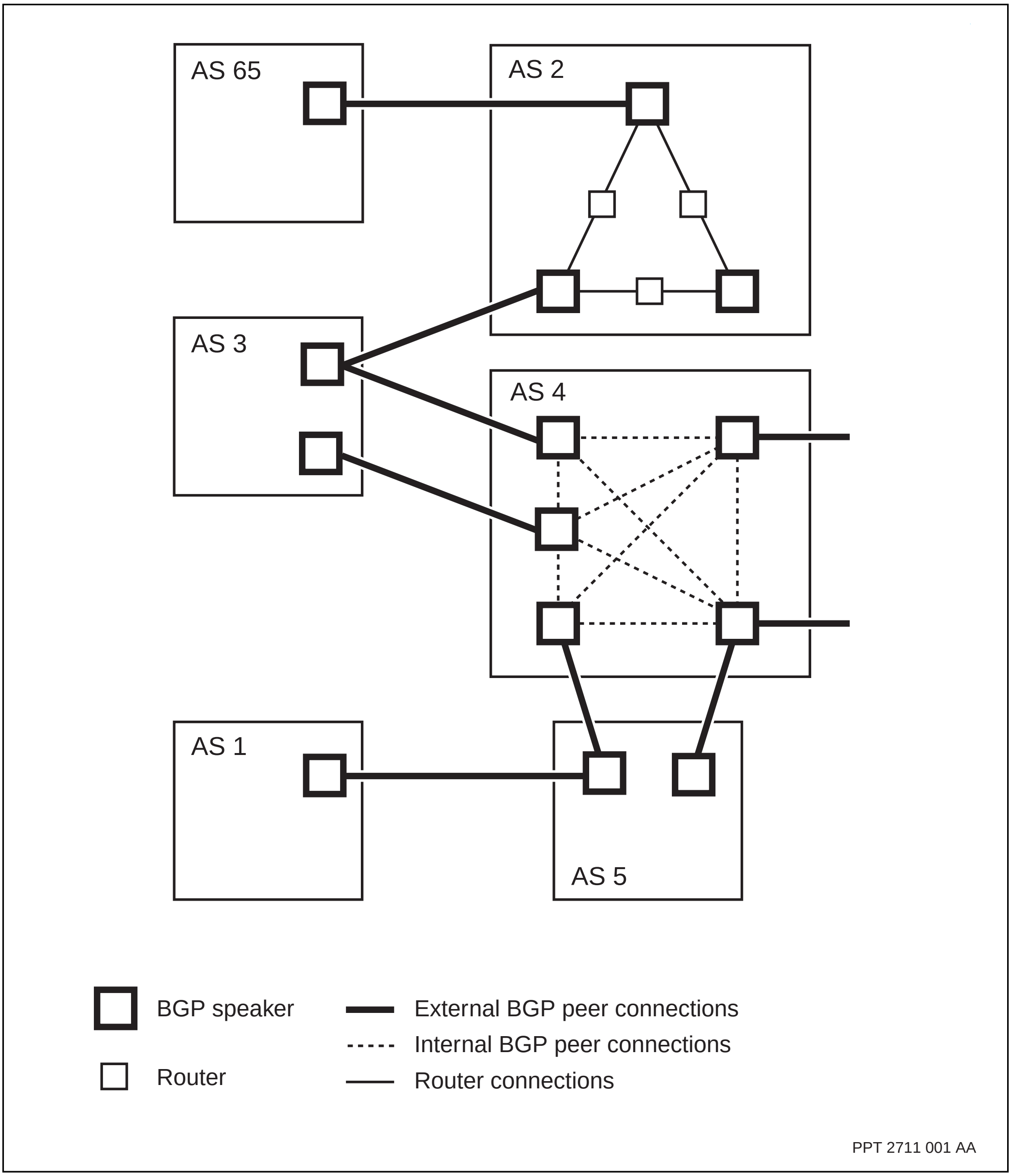
BGP-4 peers are a pair of BGP-4 speakers that exchange routing information about reachable destinations in different ASs. If the two BGP-4 speakers are in the same AS, they are internal BGP (IBGP) peers. If they are in different ASs, they are external BGP (EBGP) peers. The figure "BGP-4 peering" (page 141) shows an example of BGP peer relationships.

BGP-4 sets up peer relationships between BGP-speaking routers over a TCP connection. TCP exchanges the peer connection requests, responses, and route updates. When the peer relationship is first established, the BGP-4 speakers exchange the entire BGP routing table. After the initial exchange of the complete routing table, peers exchange only route changes whenever network topology changes. In the meantime, peers exchange periodic keep alive messages to confirm connectivity.

Any given AS can have several different types of peerings with another AS:

- stub AS singly-homed, for example, AS 65 and AS 1 in the figure "BGP-4 peering" (page 141).
- multi-homed non-transit A, for example, AS 5 in the figure "BGP-4 peering" (page 141).
- multi-homed transit AS, for example, AS 2, AS3, and AS 4 in the figure "BGP-4 peering" (page 141).

Figure 22
BGP-4 peering



BGP-4 updates

BGP-4 peers exchange routing information about reachable destinations through network layer reachability information (NLRI) update messages. If route information changes, BGP-4 informs its peers by withdrawing the invalid routes and advertising new route information.

To conserve bandwidth and processing resources, BGP-4 uses incremental updates. BGP-4 also uses route aggregation to represent information about a group of networks as a single entity. In addition to learning routes from peers, a BGP-4 router can originate routing updates to advertise networks that belong to its own AS.

BGP-4 path attributes

Path attributes associated with the NLRI updates provide detailed information about an advertised route. Routers can use path attribute information when making policy decisions.

To avoid routing loops, BGP-4 prepends the AS number of the BGP speaker to the AS path attribute of a route when that route is advertised to an EBGP peer.

Under the AS path attribute, you can configure

- path attributes
- AS weights. For more information see“AS weights” (page 149).
- private AS number removal. For more information, see “Private AS number removal” (page 152).

The table “BGP-4 path attributes” (page 143) lists the path attributes supported on Passport.

Table 22
BGP-4 path attributes

Path attribute	Description
Origin	Specifies whether an NLRI originated from an interior gateway protocol (IGP), an exterior gateway protocol (EGP), or an unknown protocol, usually a static route (incomplete).
AS path	Contains a list of all the ASs traversed by the NLRI.
Next hop	Indicates the IP address of the AS border router to use to reach the route advertised in the NLRI.
Multi exit discriminator (MED)	Specifies a preferred entry point for traffic coming back to the advertising AS. A lower MED value indicates a higher preference.
Local preference	Specifies a preferred route in an NLRI update between internal BGP peers. A higher value indicates a higher preference.
Atomic aggregate	Indicates that BGP-4 has performed some aggregation on the NLRI and removed previous path information.
Aggregator	Indicates the AS and BGP-4 peer that performed the last aggregation on a route. A BGP-4 peer that performs route aggregation adds the aggregator path attribute to the NLRI.
Communities	Associates one or more properties with the route.
Extended communities	Used in IP VPN auto discovery to carry the VPN ID and VPN peering topology value.
MP reach NRLI	Used in IP VPN auto discovery to carry public/private address mapping information.
Originator ID	Indicates the router ID of the BGP speaker that originated the route in the AS. This attribute is inserted by the route reflector.
Cluster list	Contains a list of the clusters traversed by the route. This attribute is inserted by the route reflector.

BGP-4 routing policies

A BGP-4 policy is a set of rules that determine which routes BGP-4 uses or advertises to other peers. BGP-4 uses import policies to filter NLRI updates received from other BGP-4 peers. BGP-4 export policies determine which NLRI updates to advertise to other BGP-4 peers.

Policies consist of keys and actions. A routing policy specifies an action for routing information received from BGP peers defined by a key.

For more information, see the following sections:

- “BGP-4 import policy” (page 144)
- “BGP-4 export policy” (page 146)

BGP-4 import policy

By applying import policies, BGP-4 allows certain routing information into the IP routing database or blocks it. Through import policies, you can configure BGP-4 to

- use or ignore an update by setting the *usageFlag* attribute
- set a route preference for IBGP and EBGP routes using the *ibgpRtePreference* and *ebgpRtePreference* attributes
- add a community number to the routing information using the *appendCommunity* attribute
- set a preference for updates to be used by the local BGP instance using the *localPreference* attribute

As routing updates arrive, BGP-4 applies the policy with the most specific key matches. When using path matching expressions, the *expressPreference* attribute acts as the tie-breaker between equally valid expressions.

By default, BGP-4 on Passport rejects all routes from external BGP (EBGP) peers and accepts all routes from internal BGP (IBGP) peers. See the table “BGP-4 import keys” (page 145) for a summary of BGP-4 import policy criteria.

Table 23
BGP-4 import keys

Import key	Associated attribute	Description
Network prefix and length	*network*	Applies the policy action to NLRI updates that match, or are more specific than, the specified network.
Peer AS number	*peerAS*	Applies the policy action to updates from peers in a specified AS.
Peer IP address	*peerIpAddress*	Applies the policy action to updates from a peer specified by an IP address.
Originating AS number	*originAs*	Applies the policy action to NLRI originating from the specified AS.
Originating protocol	*originProtocol*	Applies the policy action to routes with the specified origin protocol.
AS path	*asPathExpression*	Applies the policy action to routes with an AS path that matches the specified expression.
Community path	*communityExpression*	Applies the policy action to routes with a communities attribute that matches the specified expression.

For example, a Passport configured with two BGP-4 import policies, x and y, receives a routing update originating from an EBGP peer along the AS path 1 2 3 4. The two policy keys are defined as follows in the table “Example BGP-4 import policy key definition” (page 145):

Table 24
Example BGP-4 import policy key definition

Keys for import policy x	Keys for import policy y
Protocol = all	Protocol = all
AS path = 1 2 3	AS path = 2
expressPreference = 90	expressPreference = 100

The path matches both AS path expression keys. However, the expression for policy y has a greater preference value. BGP-4 applies the actions of policy y.

BGP-4 export policy

By applying export policies, BGP-4 distributes certain IP routing information to other BGP-4 peers. Through export policies, you can configure BGP-4 to

- send or block routing updates by setting the *advertiseStatus* attribute
- include the preferred entry point to the AS in updates to external peers by setting the *sendMultiExitDiscToEbgp* attribute. The MED value sent is the value specified in the *muliExitDisc* attribute.
- send your community number to BGP peers through the *sendCommunity* attribute
- alter the outgoing route's AS path by setting the *insertDummyAs* attribute
- set a preference for routes sent out to internal peers through the *localPreference* attribute

As BGP-4 sends out routing updates, it applies the policy with the most specific key matches. When using path-matching expressions, the *expressPreference* attribute acts as the tie breaker between equally valid expressions.

By using the export policies of other protocols, BGP-4 can distribute BGP-4-learned routes into other routing protocols such as OSPF, RIP, and EGP.

Table 25
BGP-4 export keys

Export key	Associated attribute	Description
Peer AS number (see Note)	*peerAS*	Applies the policy action to routing updates of a specific protocol.
Peer IP address	*peerIpAddress*	Applies the policy action to updates destined for a peer specified by its IP address.
Protocol	*protocol*	Applies the policy action to routes learned from peers defined by specific protocols.
(Sheet 1 of 2)		

Table 25 (continued)
BGP-4 export keys

Export key	Associated attribute	Description
Local RIP interface	*ripInterface*	Applies the policy action to routes learned from a local RIP interface (defined by an IP address).
Protocol-specific peers	*egpAs, bgpAs, ripNeighbor, ospfTag*	Applies the policy action to routes learned from peers defined by specific protocols.
Network prefix and length	*Network*	Applies the policy action to NLRI updates that match, or are more specific than, the specified network.
AS path	*matchAsPath*	Applies the policy action to routes with an AS path that matches the specified expression.
Community path	*matchCommunity*	Applies the policy action to routes with a communities attribute that matches the specified expression.
Note: To configure an export policy for an internal BGP peer, specify its local AS number in the *peerAs* attribute; otherwise, the internal default export policy is preferred.		
(Sheet 2 of 2)		

For example, a Passport configured with two export policies, a and b, has a route with an AS path of 123, learned from an external peer. The two policy keys are defined as follows in the table "Example BGP-4 export policy key definition" (page 147):

Table 26
Example BGP-4 export policy key definition

Keys for export policy a	Keys for export policy b
Protocol = all	Protocol = EBGP
AS path = 1 2 3	

BGP-4 would apply the actions of policy b. In this example, the protocol key matched in policy b is more specific than the AS path key of policy a, and therefore has greater weight.

BGP-4 route selection

This section describes the mechanisms that BGP-4 uses to handle routing information. It covers the following topics:

- “BGP-4 routing information bases (RIBs)” (page 148)
- “Tie-breaking rules” (page 149)
- “AS weights” (page 149)

BGP-4 routing information bases (RIBs)

BGP-4 keeps track of all BGP-4 updates in a BGP-4 route database. When there are multiple routes to the same destination, BGP-4 selects the best route and places it in the IP forwarding table of the Passport node.

Each BGP-4 speaker maintains three sets of routing information bases (RIBs):

- The Indb RIB contains all NLRI updates received from IBGP and EBGP peers.
- The Localdb RIB contains routes that have been selected by the BGP-4 decision process and propagated from the Indb.
- The Outdb RIB contains all NLRI updates advertised by the BGP-4 instance to its IBGP and EBGP peers.

BGP-4 stores all incoming routes learned from BGP-4 peers in the Indb. The BGP-4 instance applies local import policies to all routes stored in the Indb, and filters out some of the routes. Only the routes that remain are potential candidates for the next stage of the BGP-4 process. For more information on import policy, see “BGP-4 import policy” (page 144).

BGP-4 then selects the most preferred route in the Indb for propagation to the Localdb. To select the best BGP-4 route, BGP-4 uses the tie breaking rules For more information, see “Tie-breaking rules” (page 149).

BGP-4 then applies its local export policies to the routes in the Localdb, and filters out some of the routes. BGP-4 stores the remaining routes in the Outdb, and advertises them to internal and external BGP-4 peers. For more information on export policy, see “BGP-4 export policy” (page 146).

Tie-breaking rules

If a route specifies an unreachable next hop, BGP-4 does not use it. However, if more than one route to a destination is available, BGP-4 uses the criteria in the table “Tie-breaking rules” (page 149) in the order shown to choose the best BGP-4 route.

Table 27
Tie-breaking rules

Order	Rule
1	BGP-4 uses the route with the highest calculated local preference.
2	If the local preferences are the same for each route, BGP-4 uses the route with the lowest weight in the AS path or the shortest AS path.For more information, see “AS weights” (page 149).
3	If the weights and the AS paths are the same for each route, BGP-4 uses the route with the lowest multiExitDiscriminator attribute.
4	If the multiExitDiscriminator attributes are the same for each route, BGP-4 uses the lowest cost route to the next hop.
5	If the interior cost is the same for each route, BGP-4 uses the route learned from an external peer over the route from an internal peer.
6	If the criteria in rules 1 to 5 are true, BGP-4 uses the route with the lowest BGP-4 router ID. You set the BGP-4 router ID in the *bgpIdentifier* attribute of the *Bgp* component. The BGP-4 router ID must be unique in the BGP-4 network.

AS weights

You can set a preference for an AS and discriminate against other ASs by using AS weights. The AS weight attribute is local to the BGP-4 speaker, and is not propagated in route advertisements.

You can assign a weight to each AS in an AS path attribute. The weight is an integer value between 0 and 255 inclusive. BGP-4 prefers the path with the lowest weight. BGP-4 considers the value 255 as infinity and does not use that path. If you do not assign a weight to an AS, BGP-4 uses 128 as the default weight.

You can configure a BGP-4 instance to add a dummy AS to a path to decrease its preference by setting the *insertDummyAs* attribute under the *Bgp Export* component.

BGP-4 optimization

You can use optional subcomponents to optimize the BGP-4 configuration. For more information on these subcomponents and their functions, see the following sections:

- "Route aggregation" (page 150)
- "Route reflection" (page 150)
- "BGP-4 communities" (page 152)
- "Private AS number removal" (page 152)
- "Dynamic default aggregation (DDA) mode" (page 153)

Route aggregation

Through the aggregate policy, BGP-4 combines the characteristics of different routes and advertises this combination as a single route. Aggregation reduces the data a BGP-4 speaker stores and exchanges with another BGP-4 speaker. See RFC 1771 for a complete list of rules for route aggregation.

Route reflection

BGP requires that all the IBGP speakers be fully meshed. Route reflection is an IBGP peering mechanism that reduces the IBGP mesh. This implementation of route reflection is compliant with RFC 1966.

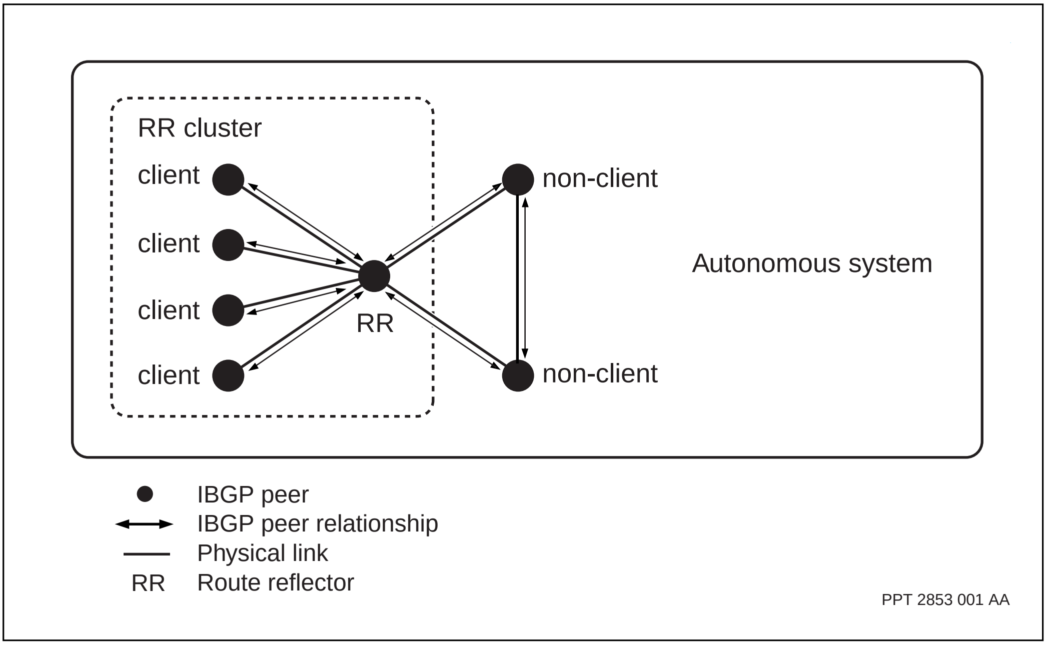
When you configure an IBGP speaker as a route reflector, you make it the focal point for internal BGP sessions. Multiple BGP routers in an AS have a peer relationship with a route reflector as either clients or non-clients. See the figure "BGP-4 route reflection" (page 151). The route reflector and client peers form a cluster. Unlike client peers, non-client peers must be fully meshed with each other.

A route reflector advertises the best IBGP route based on the following rules:

- If the route is received from a non-client peer, advertise to client peers only

- If the route is received from a client peer, advertise to all non-client and client peers, except for the originator of the route

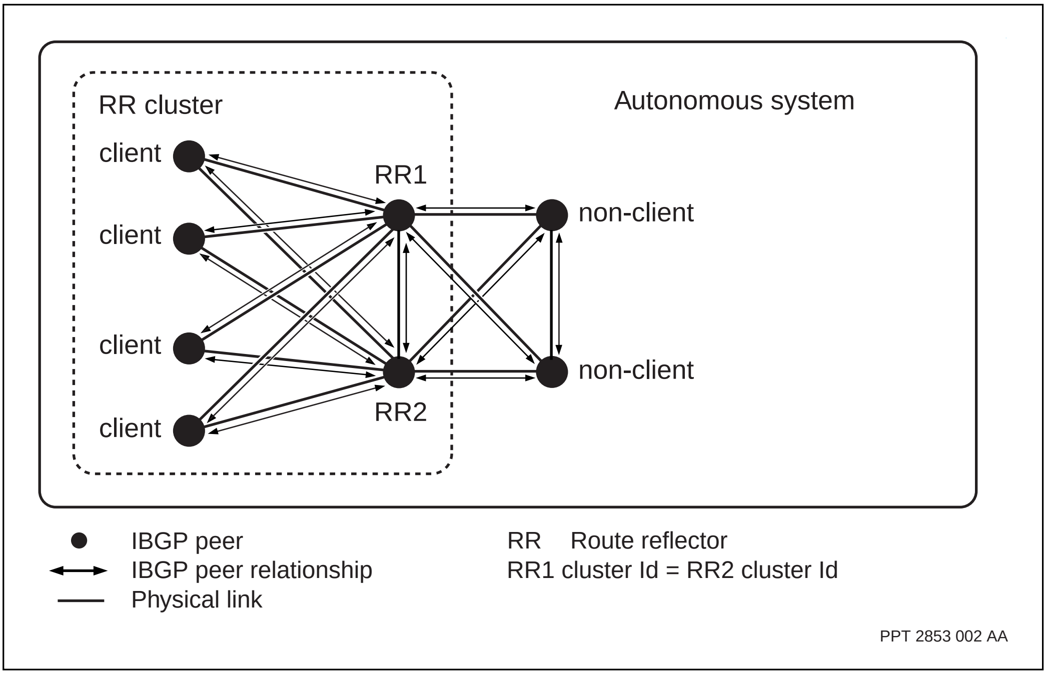
Figure 23
BGP-4 route reflection



BGP-4 allows redundant route reflectors in each cluster, and multiple clusters in an AS. See the figure “BGP-4 route reflection with redundant reflector” (page 152). The cluster identifier of both route reflectors is set to the same IP address using the *routeReflectorCluster* attribute.

It is recommended that you configure at most two route reflectors per cluster.

Figure 24
BGP-4 route reflection with redundant reflector



BGP-4 communities

A BGP-4 community is a group of destinations sharing some common property. RFC 1997 defines the BGP communities attribute (note that in this instance, attribute refers to a configured decimal identifier of a community). Each system administrator may define which community a destination belongs to. The communities attribute adds another route filtering mechanism, giving users greater flexibility in defining import and export policies.

Private AS number removal

Every BGP-4 speaker belongs to an AS. Under normal circumstances, the AS number must be unique if routes from that AS are advertised to the Internet. However, as described in RFC 1930, an enterprise customer that is homed to a single carrier need not be allocated a globally unique AS number, but may use a private AS number. When the carrier advertises its customer’s routes to the Internet, the private AS number must be removed from the AS path attribute in the routing updates.

You can configure BGP-4 on the Passport node to remove any AS numbers from the AS path attribute that are within the private AS range of 64 512 to 65 535. This applies to routes advertised through EBGP peering only. It does not apply to internal BGP peers.

Dynamic default aggregation (DDA) mode

The dynamic default aggregation (DDA) feature allows the Passport node to aggregate routes learned from an external BGP peer to a single default route. The Passport switch then propagates the DDA route to other nodes through IBGP peering.

For more information, see the following sections:

- “DDA routes” (page 153)
- “DDA mode for Internet access” (page 154)

DDA routes

You can activate an aggregate policy for incoming routes learned from a specific EBGP peer.

When the first valid route is learned from the EBGP peer, the policy generates a single default route. This DDA route is the only route considered valid among all those learned from the EBGP peer. If the import policy permits, BGP-4 propagates the DDA route to the routing database and then to the IP forwarding table. The route is tagged as a BGP external route whose next hop is the remote IP interface of the EGBP peer.

When BGP-4 generates the single DDA route, it sets the AS path attribute to the AS number of the EBGP peer and the next hop attribute to the IP address of the EBGP peer. In addition, it sets the multi-exit discriminator (MED) attribute to the value specified in the *defaultInAggMed* attribute (under the *Vr Ip Bgp Peer* component).

If you configure an OSPF export policy that specifies BGP external routes (by setting the *protocol* attribute under the *Vr Ip Ospf Export* component), BGP-4 advertises the DDA route into the OSPF domain. Therefore, there is a DDA route exported into OSPF from each node that runs DDA mode for a configured EBGP peer.

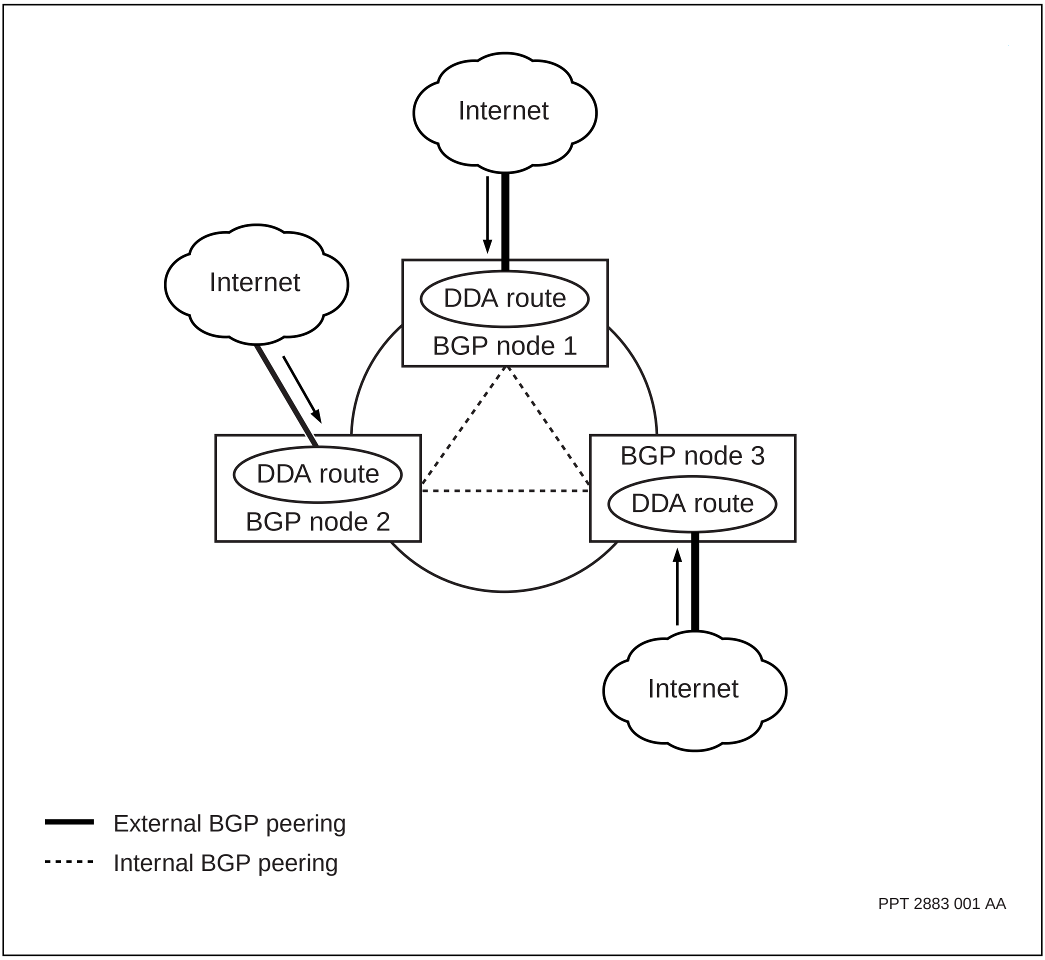
DDA mode for Internet access

The DDA feature allows for full connectivity to the Internet without the necessity of maintaining large forwarding tables. A Passport node that connects to the Internet can use DDA mode to aggregate all routes learned from its Internet EBGP peer. The Passport node can then propagate the DDA route to other BGP nodes through IBGP peering.

For example, in the figure “DDA mode on EBGP peers” (page 155), BGP nodes 1, 2 and 3 connect to the Internet through EBGP peering, and to each other through IBGP peering. Through DDA, each node aggregates all incoming routes from the Internet into a single default route.

By default, this DDA route is advertised over each IBGP peer to other BGP nodes. Thus, each node learns one DDA route from its EBGP peering with the Internet and two DDAs from its IBGP peering with the other BGP nodes. The DDA route that is learned through EBGP peering is preferred, but the propagation of DDA routes from other IBGP peers ensures redundant access to the Internet without the risks associated with routing loops. The DDA route is withdrawn if the node loses its connectivity to its external peer.

Figure 25
DDA mode on EBGP peers



Chapter 14
Static routes

This section describes the implementation of static routes on Passport. It covers the following topics:

- “Overview of static routes on Passport” (page 157)
- “Equal-cost multipath routing” (page 157)
- “Static route definition” (page 158)
- “Route entry discard” (page 158)

For information about configuring static routes, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of static routes on Passport

Static route definition allows the Passport system to specifically identify routes to remote IP networks or hosts. The definition includes a destination address, address mask, and one or more next hop addresses (gateways).

Equal-cost multipath routing

Static routes (and OSPF) support an equal-cost multipath routing function where equal-cost paths are used in a load-sharing mode. Equal-cost multipath routing has the following limitations:

- It allows a maximum of three next hop addresses.
- The entire route is discarded if, in a multiple-hop address, any next-hop IP address cannot be resolved through the address resolution protocol (ARP) process.

Static route definition

The *Static* subcomponent of the *Ip* component allows you to add, delete, and modify static route information. You can configure one static route instance on each VR.

The *RouteEntry* subcomponent of the *Static* subcomponent is used to define static routes by specifying one host, a subnetwork, or a network. There is one *RouteEntry* subcomponent for each static route. The *RouteEntry* subcomponent has one subcomponent: called *NextHop*. You must provision at least one *NextHop* component for each *RouteEntry* component.

Route entry discard

The *DiscardRouteEntry* subcomponent of the *Static* component is an optional subcomponent used to identify destination networks and nodes that do not receive packets through IP. The system discards packets addressed to these destinations immediately. No notification is sent to the sending host that the Passport system has discarded the packets. There is one *DiscardRouteEntry* subcomponent for each route that you wish to restrict.

Chapter 15
IP multicast

This section describes the implementation of IP multicast on Passport. It covers the following topics:

- “Overview of IP multicast” (page 160)
- “Supported media” (page 160)
- “Dense and sparse mode protocols” (page 160)
- “Source specific and shared trees” (page 161)
- “IGMP” (page 163)
- “PIM-SM” (page 163)
- “Multicast domains” (page 166)

Passport 7400, 15000, 20000 supports IP multicast as defined in the following RFCs:

- *RFC 2362, Protocol Independent Multicast - Sparse Mode (PIM-SM)*

  Note: PIM Border Router is currently not supported.

- *RFC 2236, Internet Group Management Protocol (IGMP), version 2* router functionality

Passport multicast

- supports IGMP version 2 router functionality.
- interworks with IGMP version 1 or version 2 hosts.

> ***Note:*** Passport multicast does not support the IGMP version 1 router functionality, nor can it interwork with an IGMP version 1 router (that is, it cannot share a LAN segment).

For information about configuring IP multicast, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP multicast

IP multicast is receiver initiated, in that receivers may join a group at any time. Receivers use the Internet group management protocol (IGMP) to inform local routers about groups they are interested in receiving. Multicast routers use a multicast routing protocol to build trees to group members. Multicast packets are forwarded along these trees. Trees may be either rooted at a source or at a core node in the network. The IP multicast model is receiver oriented in that receivers control the 'building of tree' by joining groups. In order to send to a group, a host need not perform any routing operation. A host simply sources multicast packets on its interfaces, which are then forwarded by multicast routers to all group members.

Supported media

Passport supports multicast forwarding over the following media:

- Ethernet (100BaseT)
- ATM MPE
- FR DTE

Dense and sparse mode protocols

IP multicast routing protocols generally fall into two categories, depending on the assumption of the distribution of the receiving hosts. A group is considered dense if there are many receivers within a region. A group is considered sparse if receivers are sparsely distributed across a larger area, or where a small group of receivers is concentrated in one portion of the network. DVMRP, PIM-DM, MOSPF are all dense mode protocols. BGMP, CBT and PIM-SM are sparse mode protocols.

Dense mode protocols assume dense membership within a region, and either flood group membership information (MOSPF), or multicast packets (DVMRP, PIM-DM) using the flood and prune model. Sparse mode protocols assume sparse membership over a larger region, and do not flood group information or data packets. Sparse mode protocols use an explicit joining mechanism.

Note: Passport only supports PIM-SM.

Source specific and shared trees

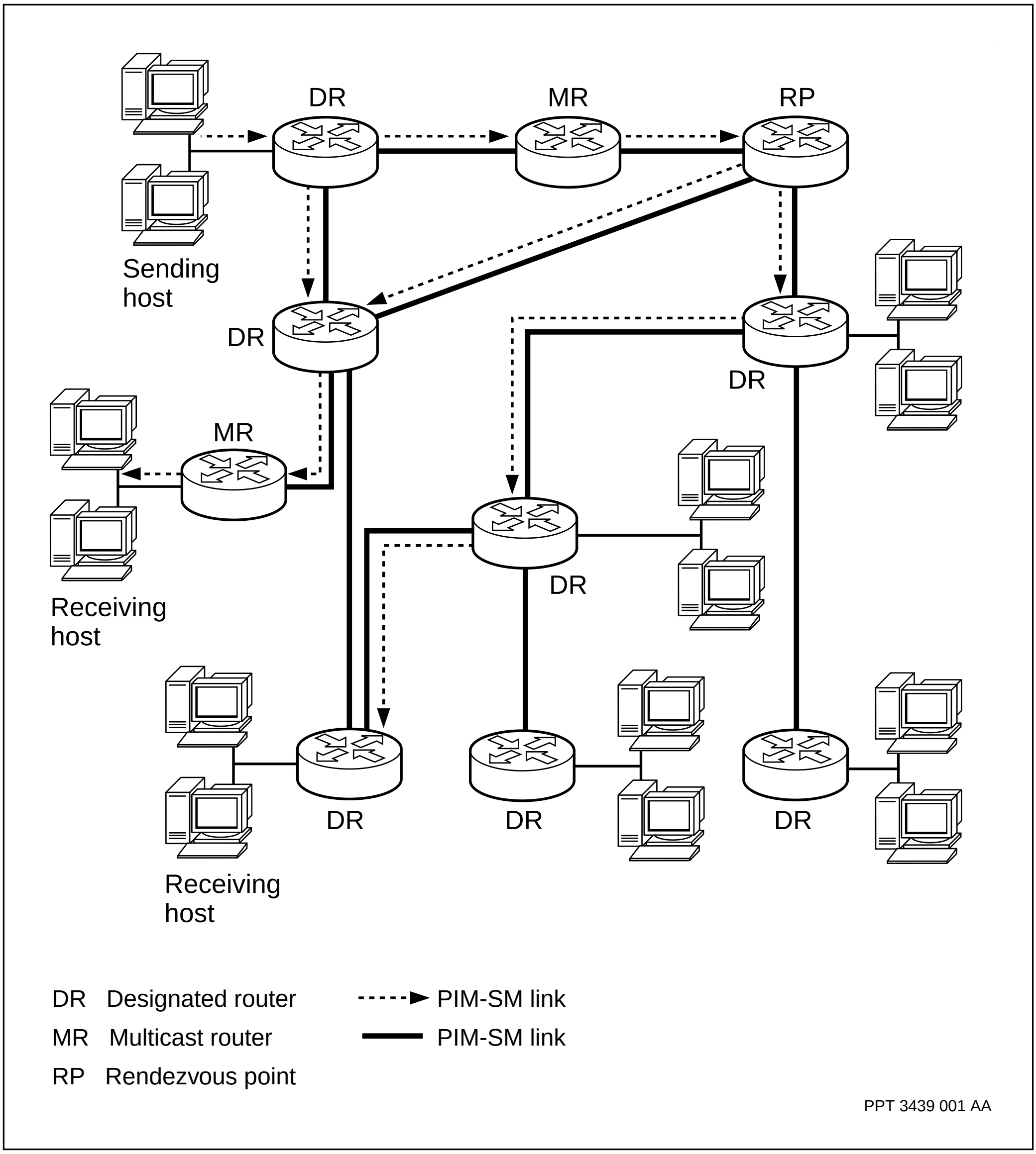
Multicast routing protocols build spanning trees to all group members for forwarding of multicast data packets. These spanning trees may be either source-based trees (DVMRP, PIM-DM, MOSPF) or shared trees (CBT, BGMP, PIM-SM).

PIM-SM can build both source-specific and shared distribution trees. Shared trees are used by default, but when a particular data threshold is reached, PIM-SM may switch to source-specific trees.

Source-based trees are rooted at a source and span to all group members. If N-N communication is desired (for example: conferencing) when a source-specific tree routing protocol is in use, then multiple source-specific trees comprise a given group. These trees have the ability to offer lowest delay, but with the cost of extra state. Source-specific trees require keeping (S,G) state, or state that is per source per group and thus scales as the number of sources per group. Source-specific trees may also be based on source prefixes, (Sp,G), and represent all the sources summarized by the prefix. Source-specific trees are almost always used in dense mode routing protocols and are inherently uni-directional.

Shared trees are used in sparse mode routing protocols. They require the definition of a root node, which acts as a sink for all source data and group joins. The root node, also known as the core node or rendezvous point (RP), is used as a meeting point between sources and receivers. When a source transmits, its data is sent towards the root node; when a host joins, its join is propagated towards the root node. Shared trees do not guarantee the lowest delay, but can offer state savings within routers.

Figure 26
IP multicast - RP and shared tree for a multicast group



IGMP

The Internet group management protocol (IGMP) [DEERING] is the protocol used by hosts to communicate their desired group memberships to their local multicast router. IGMP exists as part of a host's IP implementation, as well as part of a router's multicast implementation. IGMP uses a query / response mechanism to solicit membership status from hosts. Periodically, routers send a query message, to which hosts respond with a report per group desired. When a router receives a report from a host, the action it takes depends upon the multicast routing protocol in use.

IGMPv1 [DEERING] was the first implemented version of IGMP, and exists in many host (and router) implementations. IGMPv2 [FENNER1] added support for a fast leave message so that hosts can inform the network immediately when they wish to leave a group. This message improves the multicast "leave latency" (with IGMPv1, routers simply time out the state for a group). IGMPv3 [CAIN1] enhancements support source-specific join and leave messages. This allows hosts to individually join or leave sources or sets of sources.

PIM-SM

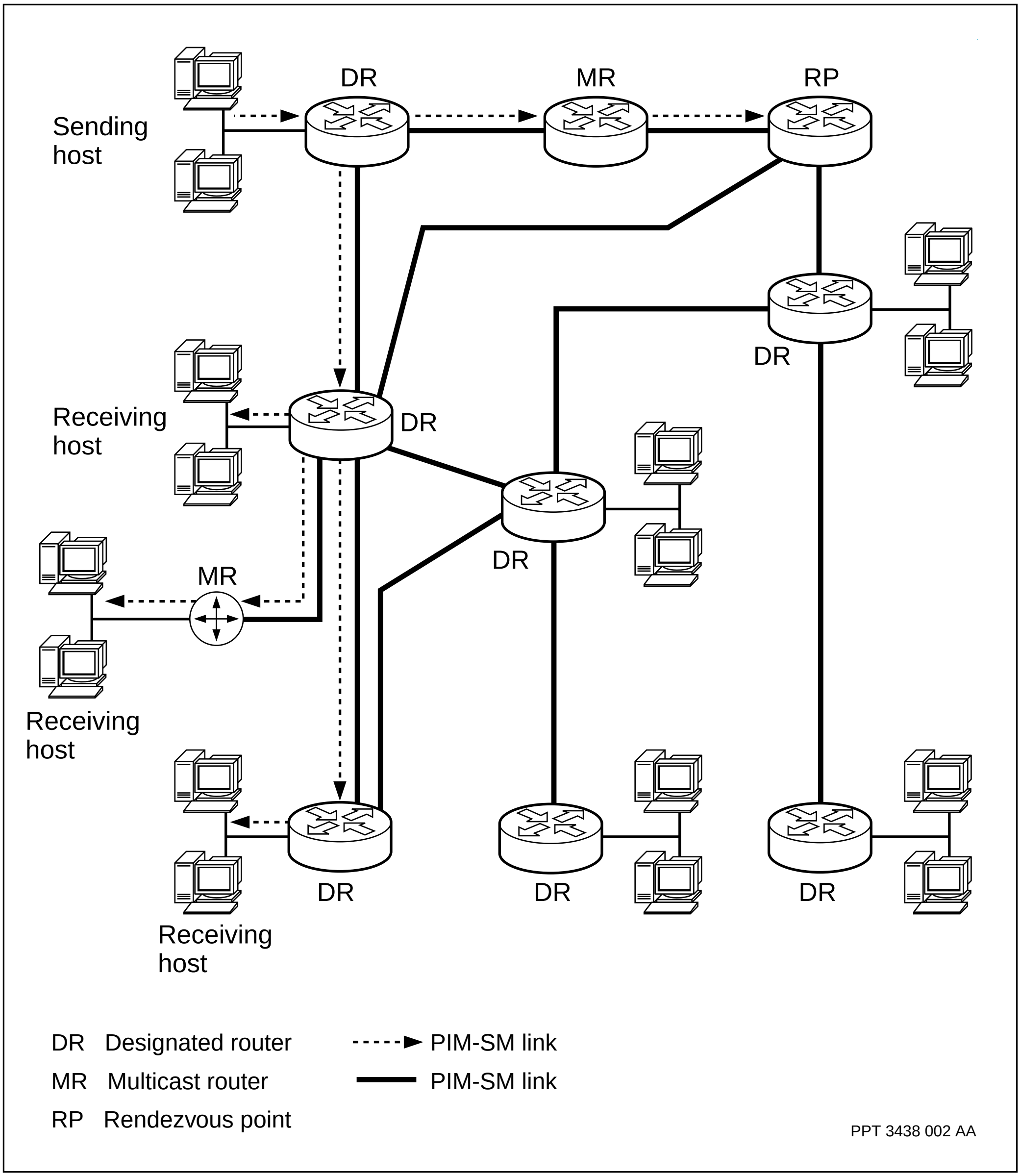
Protocol independent multicast sparse mode (PIM-SM) is a multicast routing protocol designed for use in networks where receivers make up a small portion of the overall network topology. Under PIM-SM, multicast traffic is distributed to only those routers with group members downstream. PIM-SM builds soft-state, multicast-group based, shared trees for forwarding multicast traffic. PIM-SM employs the use of a rendezvous point (RP), where receivers and sources meet. The RP is the node toward which the join/prune and register messages are sent for the shared tree. Each group has one specified RP. Routers with downstream receivers wishing to join a group must explicitly inform the RP of the join. This in turn requires RPs to maintain state information pertaining to group membership.

Routers running PIM-SM respond to changing group membership by issuing join or prune messages towards an RP. Spanning trees created under the PIM-SM model are defined by the actions of joining and pruning.

PIM-SM has the flexibility to operate using a shared tree or a shortest-path tree (SPT). Use of the SPT can only be initiated by an RP, or by a PIM router with locally attached hosts. A Passport 7400, 15000, 20000 acting as an RP will initiate the SPT by default, on receipt of the first packet from a source.

A Passport 7400, 15000, 20000 acting as a last-hop DR can be configured to join the SPT on receipt of a first packet from a source. Passport does not support the use of traffic thresholds to trigger SPT joins.

Figure 27
IP multicast - RP and shortest path tree for a multicast group



Multicast domains

Passport supports the use of up to four independent multicast domains per virtual router. Domains are used to partition a network into sections where multicast operates independently under the control of the same or different multicast routing protocols. Forwarding of multicast traffic between domains requires the use of a multicast border router which supports the multicast routing protocols used in the neighboring domains. Passport does not currently support the multicast border router functionality.

Chapter 16
Virtual router redundancy protocol

This section describes the implementation of the virtual router redundancy protocol (VRRP) on Passport 7400. It covers the following topics:

For more information on configuring VRRP, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of VRRP

Passport 7400 uses the virtual router redundancy protocol (VRRP) version 2, to provide router redundancy and availability to IP routing. Router redundancy is achieved with VRRP virtual routers, and router availability by monitoring critical IP interfaces. RFC2338 describes VRRP in detail. Passport 7400’s implementation of VRRP supports

- IP over the 2-port 100BaseT Ethernet FP
- multiple instances of Vrrp virtual routers on each Passport VR

Note: The term ‘virtual router’ is used by Passport and the VRRP protocol to describe different entities. To minimize the confusion, the terms Passport VR and VRRP virtual router are used to differentiate between the two.

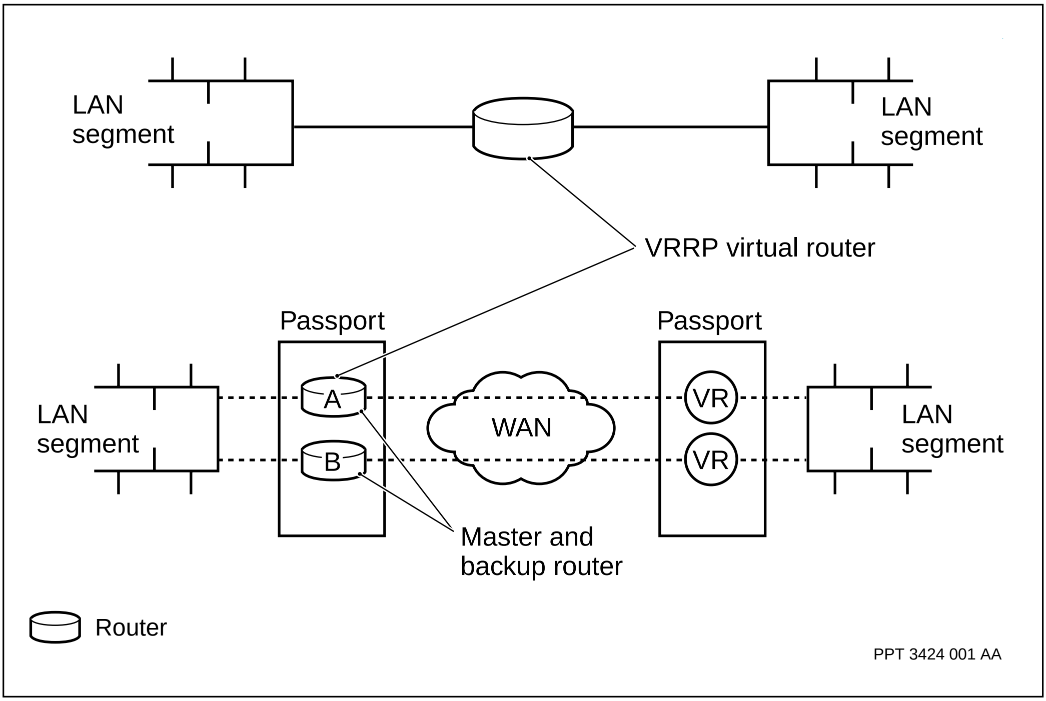
VRRP virtual routers

Implementing VRRP involves creating a VRRP virtual router made up of two or more routers sharing IP addresses and a virtual MAC address. Within the VRRP virtual router, one router will act as the master and the others as backups. To an end-host, this VRRP virtual router appears as a single router. “VRRP virtual router” (page 169) depicts this arrangement. The VRRP routers communicate with each other using IP multicasts through the local Ethernet interfaces (Passport VR LAN protocol port).

The VRRP protocol communicates entirely over the local LAN media without taking into consideration the Passport VR configuration. A VRRP virtual router can consist of Passport VRs configured as either virtual IP routers (VIPR with individual WAN connectivity per VR) or as RFC2764-based IP-VPN routers (customer VRs aggregated through a virtual connection gateway). Alternatively, a VRRP virtual router can consist of Passport VRs and non-Passport routers compliant to RFC2338

Each VRRP router has a priority value that determines whether it will act as a master or backup. The VRRP master router typically owns the IP addresses of the VRRP virtual router and has a priority of 255.

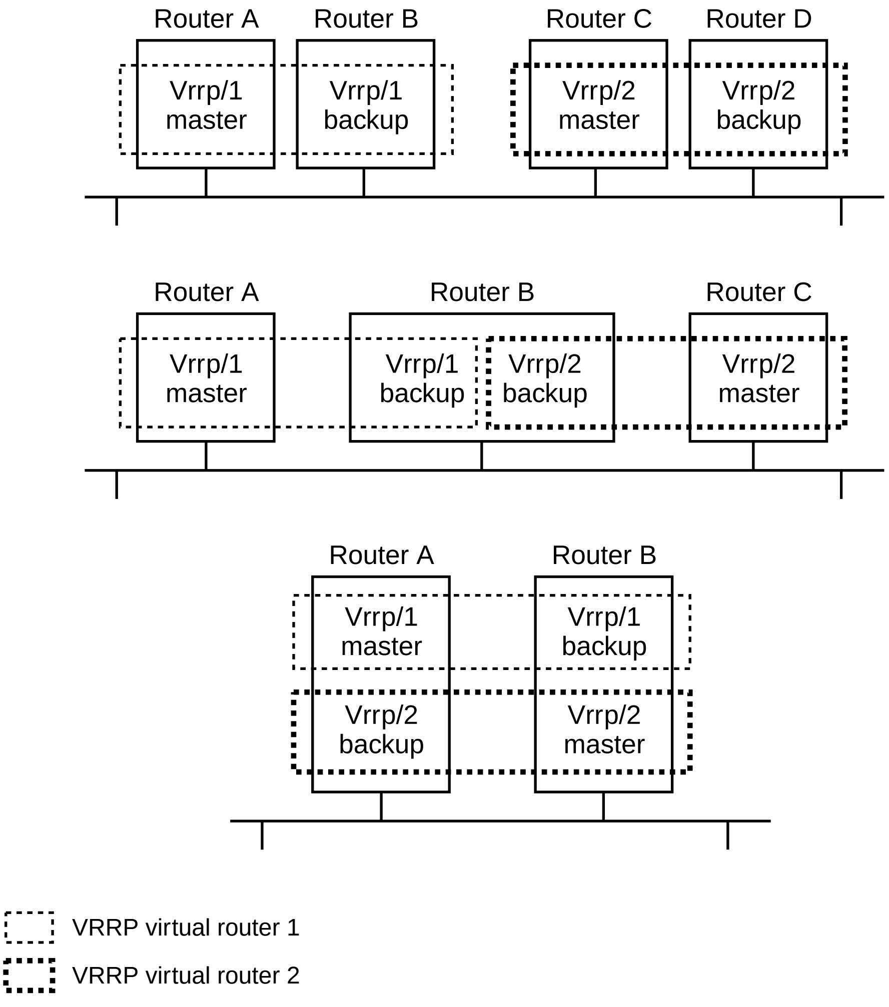
Figure 28
VRRP virtual router



Router redundancy

You can configure a single Passport VR with multiple instances of VRRP. This allows one Passport VR to participate in more than one VRRP virtual router. How you configure router redundancy will depend on the unique characteristics of your network. “VRRP virtual router” (page 169) depicts three possible scenarios where a LAN segment uses multiple routers for load balancing and static routes to end hosts.

Figure 29
Example router redundancy topologies



Note: Routers A, B, C, and D are either Passport VRs or a combination of Passport VRs and external non-Passport VRRP-compliant routers.

PPT 3424 003 AA

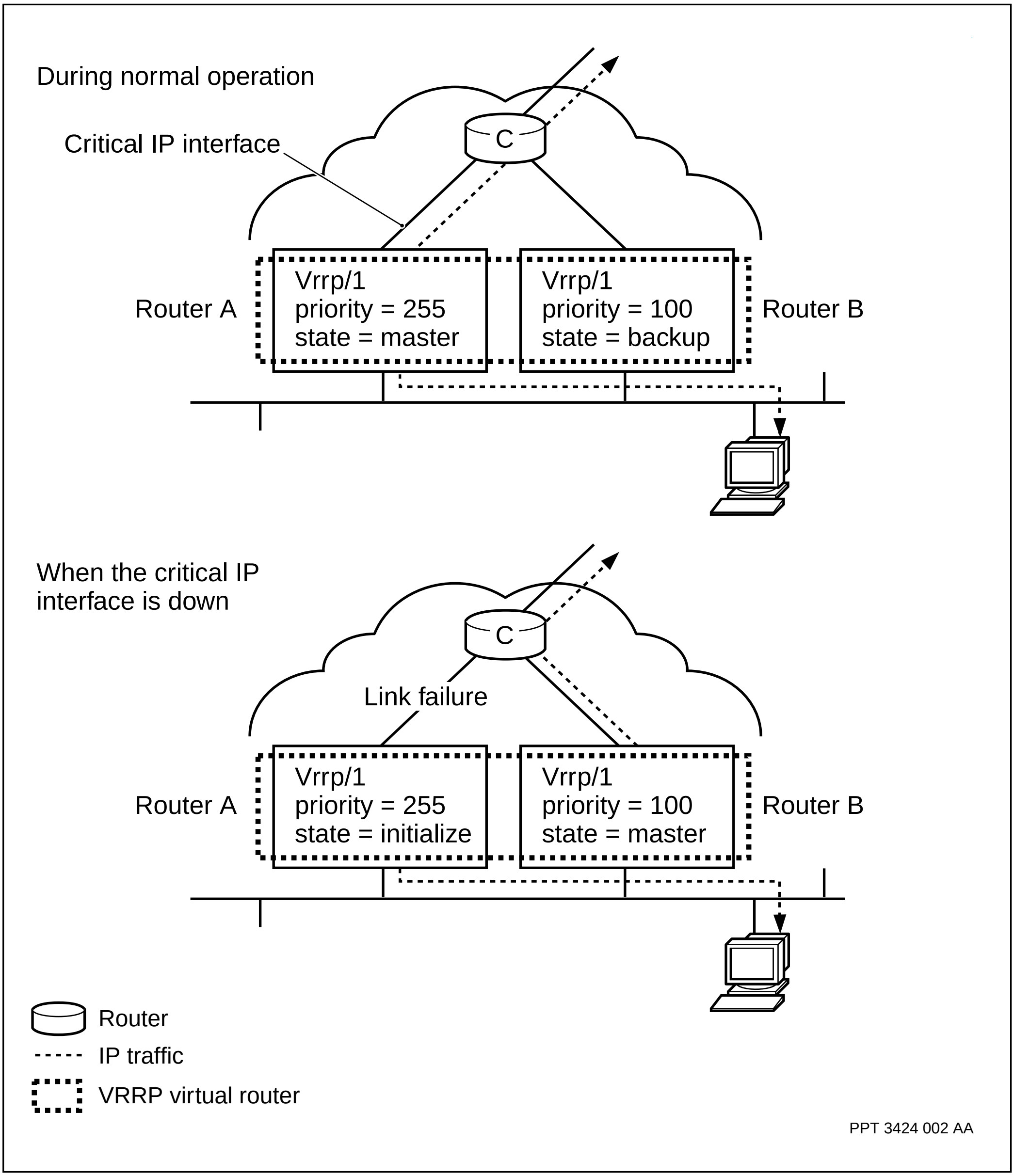
Router availability

Specific IP interfaces can be monitored by a VRRP router. These interfaces are called critical IP interfaces. A VRRP router can be operational but have a key interface that’s down, resulting in the loss of an IP traffic flow. By configuring VRRP with critical IP interfaces, you can better guarantee router availability for the IP traffic.

When a critical interface goes down or becomes locked, the current master goes into an initialize state. One of the VRRP backup routers, routing via a different interface, becomes the new master until the interface comes back up or is unlocked.

Critical IP interfaces can be passing IP traffic over media other than ethernet. “Critical IP interfaces” (page 172) depicts a scenario where VRRP and a critical IP interface is used to provide router availability.

Figure 30
Critical IP interfaces



The VRRP process

When operational, VRRP routers are in one of three states: master, backup, or initialize. Routers in the master state perform the routing duties for addresses associated with the VRRP virtual router. Routers in the backup state monitor the availability of the master router. The priority parameter of a VRRP router determines if it acts as a master or backup. A router is in the initialize state when its *Vrrp* component is locked, or when its IP interface or linked critical IP interface is down. Table “Summary of the VRRP router states in relation to network conditions” (page 173) summarizes the states of the VRRP routers under different conditions.

Table 28
Summary of the VRRP router states in relation to network conditions

	VRRP router state and activities	
network condition	**VRRP router A** **priority = 255**	**VRRP router B** **priority = 100**
start up	master	backup
normal	master As master, router A multicast messages at determined intervals, advertising to the backup router that it is operational.	backup As backup, router B listens for the multicast advertisements. When it receives multicast, an advertisement wait timer resets
router A failure	na	master Router B transitions to master when the advertisement wait timer expires.
critical IP interface goes down	initialize Advertises to backup to take over as master router.	master
critical IP interface comes back up	master Advertises to backup that it is resuming the role as master router.	backup Transitions to backup state upon receiving startup multicast from router A

Chapter 17
IP class of service (CoS)

This section describes the implementation of IP CoS on Passport. It covers the following topics:

For information about configuring IP CoS, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP CoS

IP CoS is one of two ways differentiated services can be deployed on Passport for IP. "Passport IP differentiated services" (page 195) offers several advantages over IP CoS and should be considered as the best choice for differentiated services on a Passport 15000 or 20000.

Passport supports differentiation of IP traffic for different levels of service. It can examine layer 2, layer 3, and layer 4 parameters to classify IP packets. Once classified, the node has the option of marking the DiffServ codepoint (DSCP) and forwarding the packets using different qualities of service (QoS) over media that support multiple QoS.

For information on which Passport function processors support IP CoS, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

The table “IP CoS support on access media” (page 177) summarizes support for IP CoS over different access media. Layer 3 and layer 4 classification on PQC-based FPs is only on PQC2-based FPs.

Table 29
IP CoS support on access media

	Frame relay		ATM MPE	PPP	Ethernet		
	FR DTE	IP-optimized DLCI			MS3 (GigE)	PQC	SBIC
Packet classification at ingress							
Layer 2 classification (VC, protocol port)	yes	yes	yes	yes	yes	yes	yes
Layer 3 classification [1] (DSCP-based)	yes	yes	yes	yes	no	yes	yes
Layer 3/4 classification [1] (flow-based)	yes	yes	yes	yes	no	yes	yes
Packet marking at egress							
DSCP marking [1] (modify DSCP field)	yes	yes	yes	yes	no	yes	yes
Class-based packet forwarding at egress							
CoS to VC mapping (select virtual connection)	yes	yes	yes	no	no	no	no
CoS to DP mapping [2] (apply drop precedence)	yes	yes	yes	no [3]	yes	yes	no
CoS to EP mapping [2] (apply emission priority)	yes	yes	no [4]	yes	yes	yes	no

[1] Feature ipCos is required in the feature list of the ingress FP. Ip CoS capability is based on the ingress media, not the egress media.

[2] Feature ipCos is required in the feature list of the ingress FP if CoS to DP mapping or CoS to EP mapping is required and is to be functional.

[3] CoS to DP mapping on PPP is supported on SBIC FPs.

[4] EP is determined by CoS to VC mapping.

You can also configure IP CoS functionality on both IP tunneling protocol ports and virtual media protocol ports. For more information about configuring IP CoS functionality on IP tunneling protocol ports and virtual media protocol ports, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Packet classification at ingress

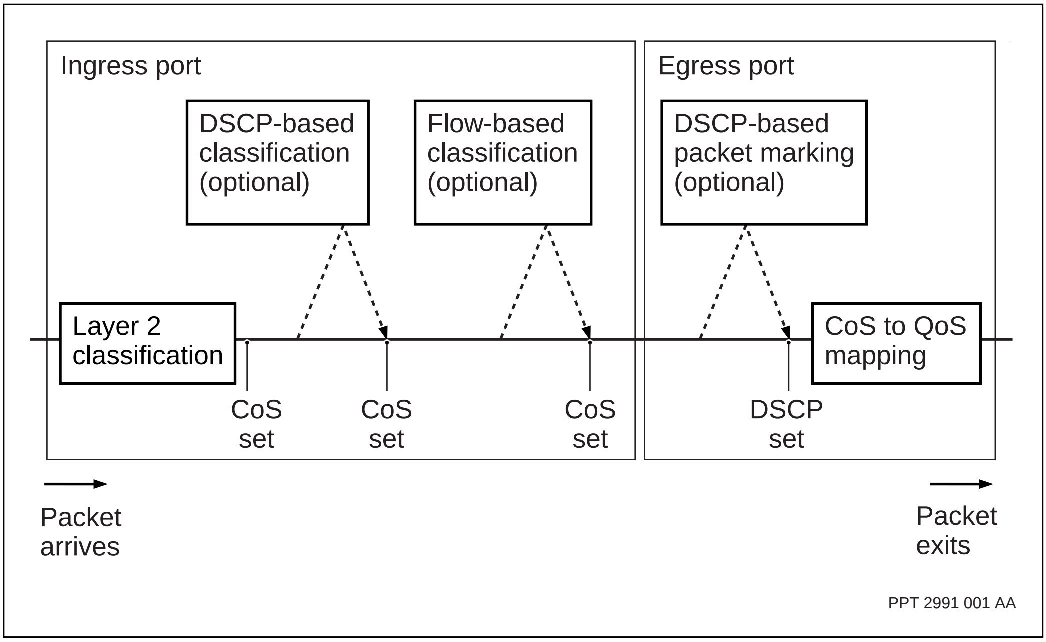
IP CoS classifies IP packets at different points in the transmission path, based on one or more of the following criteria:

- incoming media link
- DiffServ codepoint (DSCP)
- IP address (source or destination), layer 4 (transport) protocol, TCP or UDP port numbers

See the figure “IP CoS assignment on transmission path” (page 179). Each classification supersedes any preceding classification. For more information on how IP CoS manages packet classification, see the following sections:

- “Layer 2 classification” (page 179)
- “DSCP-based classification” (page 180)
- “Flow-based classification” (page 180)
- “IP CoS policies” (page 181)

Figure 31
IP CoS assignment on transmission path



Layer 2 classification

IP CoS performs layer 2 classification on each packet as it enters the Passport switch. The packet's first assigned CoS value is based on the layer 2 link on which the packet arrives. Classification can be VC-based or port-based.

For access media that support VCs (frame relay and ATM MPE), the CoS value assigned to packets corresponds to the CoS value associated with the incoming connection through the appropriate *ipcos* attribute. See "Frame relay DTE class-based forwarding" (page 184) and "ATM MPE class-based forwarding" (page 189) for more information.

For other access media capable of bearing IP traffic (PPP and Ethernet), the CoS value assigned to packets arriving on a particular port corresponds to the value configured under the *Vr Pp IpPort ipCoS* attribute.

For layer 2 classification on ATM MPE, CQC-based ATM FPs require the assistance of an ILS Forwarder card. For more information on Passport FPs, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

Layer 2 classification is supported for any egress access and core media even if the ipcos feature is not on the feature list of a logical processor type (LPT).

DSCP-based classification

The IP packet header contains an 8-bit type of service (ToS) byte. The six most significant bits of the ToS field represent the DiffServ codepoint (DSCP). The DSCP field is used to specify a quality of service (QoS) for the packet that can affect the packet delay, throughput, and reliability.

You can enable a virtual router to classify IP packets by configuring a list of ToS byte values in attribute *Vr Ip Pg Policy TosMap tos*. Attribute *Vr Ip Pg EgressCosTreatment tosMask* determines which bits of the ToS byte are examined. If the ToS value of an incoming packet matches any of the values in the configured list, the packet is assigned the CoS value specified for that IP CoS policy.

It is recommended you keep attribute *Vr Ip Pg EgressCosTreatment tosMask* set to its default value of 0xFC, which represents the DSCP bits.

DSCP-based classification on CQC-based ATM FPs requires the assistance of an ILS Forwarder card. For more information on Passport FPs, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

Flow-based classification

Flow analysis identifies the flow of IP packets between a source and destination. If enabled, IP CoS can use a flow identification policy to determine the CoS value assigned to packets in the traffic flow.

You can configure IP CoS to distinguish IP traffic flow based on any combination of the following parameters:

- IP address (source or destination) or range of IP addresses
- layer 4 protocol (TCP, UDP or ICMP)
- TCP/UDP port or range of ports

You can specify IP addresses and port numbers as a single value, as a range of values, or as no value (matches all values). IP CoS tries to match the destination IP address and port number first. If there is no match, IP CoS tries to match the source IP address and port number. If the IP address and port number for an incoming packet matches any one of the values in the configured list, IP CoS assigns the CoS value specified for that IP CoS policy.

IP CoS does not classify ICMP packets using IP addresses. If you configure IP CoS to match ICMP packets, IP CoS assigns the same CoS to all ICMP packets, regardless of IP source or destination address. Flow identification also excludes fragmented IP packets or IP packets with options.

For flow-based classification on ATM MPE, CQC-based ATM FPs require the assistance of an ILS Forwarder card. For more information on Passport FPs, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*

Transport protocols using static port assignment

Static port assignments simplify IP flow identification. Most protocols fall into the static port assignment protocol category (examples include HTTP, SMTP, and Telnet). The client-server flow is easily identified based on information supplied by the user, typically the client’s IP address and port number.

Port numbers in the transport (TCP/UDP) protocol port identify the ends of the logical connections. Some of these ports refer to specific protocols and are often referred to as well-known ports. Protocols running on well-known ports include FTP, Telnet, and HTTP.

IP CoS policies

IP CoS policies determine how packets are classified and the treatment they receive according to their classification. Policy groups contain sets of policies that you can assign to a virtual router or protocol port.

Each policy has an assigned CoS value that references a packet treatment profile under the parent policy group. When there is a policy match, Passport classifies the packet with the referenced CoS value.

Each policy has a set of criteria for applying DSCP mapping (specified under the *Vr Ip Pg Policy TosMap* subcomponent) and flow identification (specified under the *Vr Ip Pg Policy IpAddrLayer4Flow* subcomponent). The actions for a match with policy criteria are defined for the policy group through the *Vr Ip Pg ingressCosTreatment* and *Vr Ip Pg egressCosTreatment* components.

You can assign the policy group to the virtual router by setting attribute *Vr Ip cosPolicyAssignment*. If you assign a policy group to a virtual router, all of the policies defined for that policy group are applied to each protocol port under the virtual router.

Alternatively, you can assign policy groups on a port by port basis by setting attribute *Vr Pp IpPort cosPolicyAssignment*. If you assign a policy group to a protocol port, its policies are applied on that particular protocol port only and override any policy groups assigned to the parent virtual router.

Packet treatment at egress

See the following sections for information on how IP CoS can process a packet after it has been classified:

- “Packet marking” (page 182)
- “Class-based packet forwarding” (page 183)

Packet marking

Once a packet has been classified, you can configure IP CoS to mark the packet so that the system forwards information about the assigned CoS value to subsequent nodes without the nodes having to repeat the flow classification process.

Enable packet marking by setting attribute *Vr Ip Pg EgressCosTreatment setTosByte* to yes. When enabled, the Passport marks the packet by setting the ToS byte in the IP header, which is based on the CoS value assigned to the packet, the initial value of the packet’s ToS byte, and attribute *Vr Ip Pg EgressCosTreatment tos* (new ToS). Attribute *Vr Ip Pg EgressCosTreatment tosMask* determines which bits of the packet’s ToS byte are marked. It is recommended you keep attribute *Vr Ip Pg EgressCosTreatment tosMask* set to its default value of 0xFC, which represents the DSCP bits.

PQC-based FPs use the default ToS mask of 0xFC regardless of the value of attribute *Vr Ip Pg EgressCosTreatment tosMask*.

The ToS byte is set as follows:

```
[(initial ToS)&(~tosMask)] | [(new ToS)&(tosMask)]
```

Example
Using the default value of 0xFC for *tosMask*, the ToS byte is set as follows:

```
(initial ToS & 00000011) | (new ToS & 11111100)
```

Class-based packet forwarding

IP CoS allows IP traffic differentiation into four separate classes of service. You can map each CoS value to a specific set of QoS parameters for a given media type. In this way, you can control the service that packets receive when they are forwarded out of the Passport node, based on their CoS classification.

For more information, see the following sections:

- “Drop precedence” (page 183)
- “Scheduling class” (page 184)

Drop precedence

IP CoS allows you to configure the drop precedence of a packet. The drop precedence determines a packet’s importance when being forwarded. It is used to determine whether or not an IP packet should be dropped to reduce traffic load during traffic congestion.

You can associate a drop precedence with a CoS value by setting attribute *Vr Ip Pg IngressCosTreatment discardPriority*. There are four drop precedence settings: unchanged, low, medium, and high. A value of low indicates a low drop precedence, meaning that the packet is less likely to be dropped than packets with a different drop precedence. A value of high indicates a high drop precedence, meaning that the packet is more likely to be dropped than packets with a different drop precedence. If you assign a *discardPriority* of unchanged, the drop precedence of IP packets is not modified by IP CoS but is instead determined by ingress layer 2 media.

Drop precedence is supported on all WAN media (ATM MPE, frame relay, and PPP) on all applicable SBIC-based and most PQC2-based FPs. It is not supported on PPP with PQC2-based FPs and Ethernet on SBIC-based FPs.

Scheduling class

IP CoS lets you configure the scheduling class of a packet. The scheduling class is used to determine when a packet is scheduled to be transmitted at its egress interface relative to other packets.

Scheduling class values are derived from the *Vr Ip Pg EgressCosTreatment emissionPriority* attribute used by the IP port where the packet is transmitted. The emission priority attribute of the *EgressCosTreatment* whose instance value matches the CoS index assigned to the packet is used.

> ***Note:*** Generally, a numerically lower value of the emission priority attribute indicates the packet is less likely to be delayed. A numerically higher value indicates the packet is more likely to be delayed. The actual scheduling behavior of packets depends on the scheduling mechanism and number of queues at the egress interface.

Frame relay DTE class-based forwarding

Passport frame relay DTE (FrDte) supports CoS to QoS mapping over multiple DLCIs or a single DLCI.

For more information, see the following sections:

- “CoS to QoS mapping over multiple DLCIs” (page 184)
- “CoS to QoS mapping over a single DLCI” (page 186)

CoS to QoS mapping over multiple DLCIs

For CoS to QoS mapping over multiple DLCIs, IP CoS lets you create up to four DLCIs to each next hop router, with each DLCI having separate QoS parameters and a different CoS index.

All dynamic DLCIs under the same *FrDte* component share the same assigned CoS index, which is set in the *FrDte DynamicDlciDefaults ipcos* attribute, unless configuration on a static DLCI overrides it. For full service differentiation, use static DLCIs and assign a unique CoS value to each DLCI using its *FrDte StaticDlci ipcos* attribute.

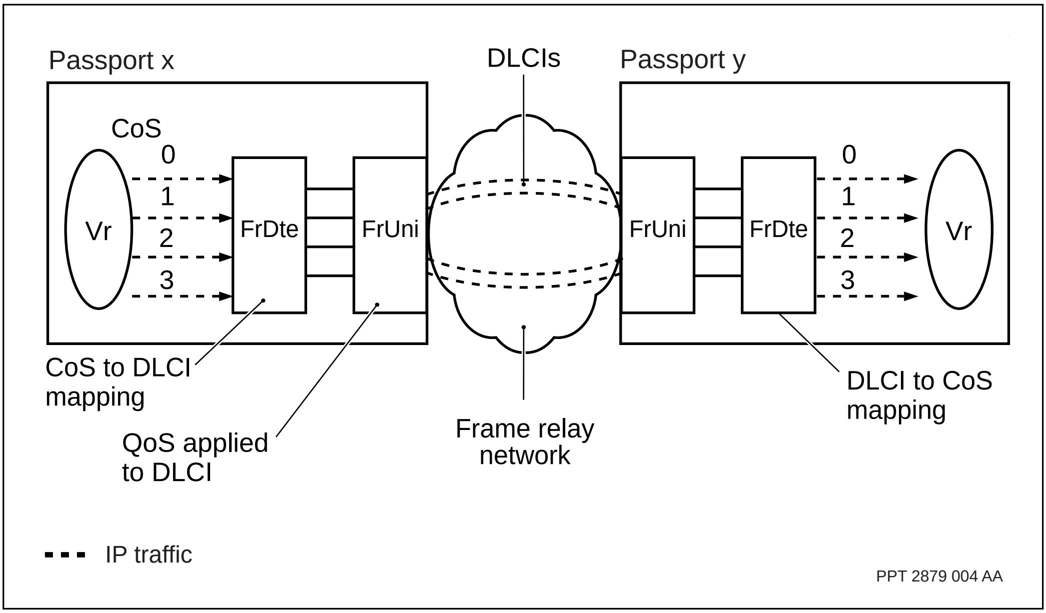
The drop precedence assigned to an IP packet arriving on a frame relay connection is determined by the discard eligibility (DE) bit to drop precedence mapping configured on the FRUNI. This can be overridden by the CoS to drop precedence mapping configured on the CoS policy group used by the ingress protocol port.

When a packet goes out a frame relay DTE interface where there are multiple DLCIs available for the same IP next hop, Passport forwards the packet over the DLCI whose associated *ipcos* attribute matches the packet's CoS value. See the figure "CoS to QoS mapping on multiple DLCIs" (page 186).

If the CoS assigned to the packet does not match an operational DLCI (it exists and is operationally active) with the same CoS value, the packet is transmitted through an operational DLCI with the next lowest CoS value. If no operational DLCIs with a lower CoS value exist, the packet is transmitted through an operational DLCI with the next highest CoS value.

The remote end of each DLCI controls the QoS characteristics for that DLCI. To obtain a particular set of characteristics within a Passport node, route the DLCI through a virtual framer to a FR UNI, where you can set the QoS parameters. For more information, see 241-5701-901 *Passport 7400, 15000, 20000 Frame Relay Fundamentals*.

Figure 32
CoS to QoS mapping on multiple DLCIs



CoS to QoS mapping over a single DLCI

Passport IP CoS allows flows of IP packets that are transmitted over a single frame relay DTE DLCI to be differentiated by QoS characteristics. The differentiation is done by setting the emission priority and drop precedence of packets in certain IP flows based on their assigned CoS.

The drop precedence assigned to an IP packet arriving on a frame relay connection is determined by the discard eligibility (DE) bit to drop precedence mapping configured on the FRUNI. This can be overridden by the CoS to drop precedence mapping configured on the CoS policy group used by the ingress protocol port.

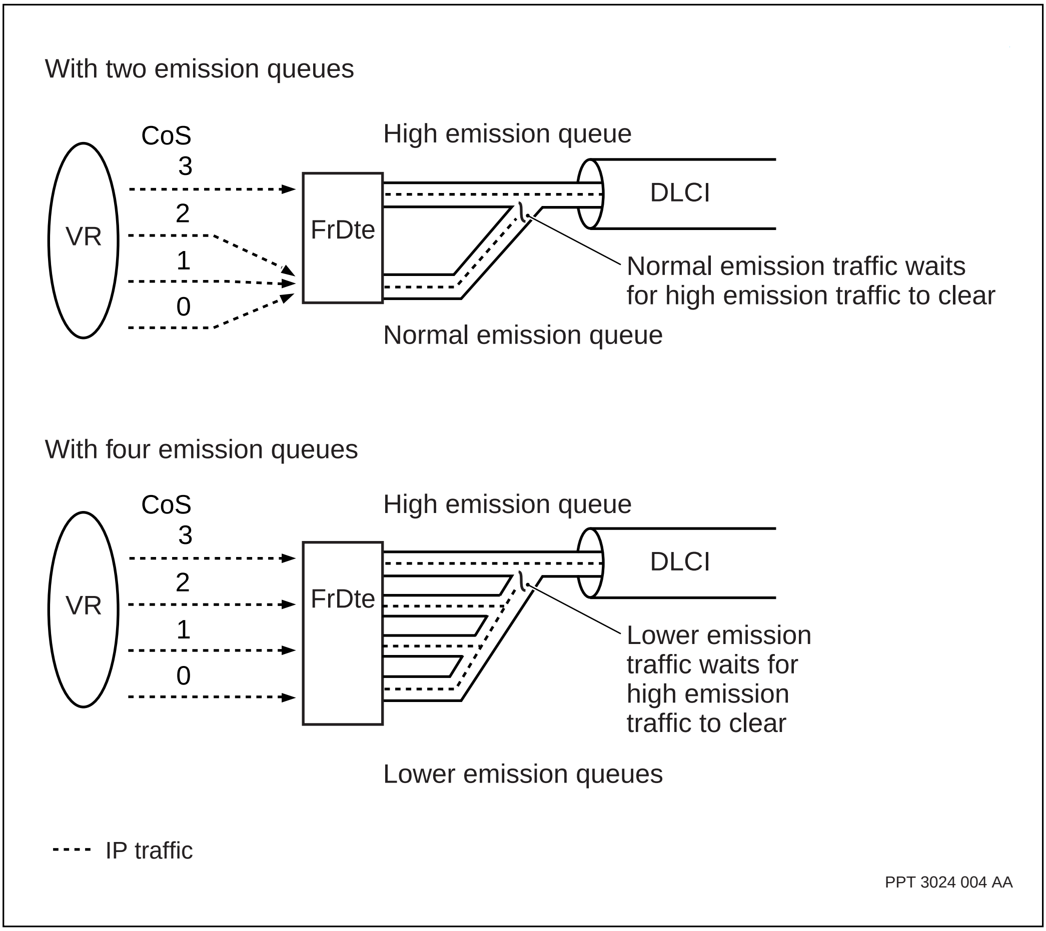
The Passport node uses the policy group configured on the egress protocol port for emission priority mapping. When a packet goes out on a frame relay DTE interface on a non-channelized FP, Passport assigns an emission priority to the packet based on the packet’s CoS value. For CoS to emission priority mapping over a single DLCI, you can specify up to four different emission priorities, one for each CoS value.

CoS to emission priority mapping is not supported over *VirtualFramer* components (a logical connection). Instead, the mapping must be over either *Framer* (a hairpin connection) or *Dconn* (a direct connection) components. See “FrDte to FrUni connectivity” (page 75) for more information on these connection types.

Emission priority mapping based on CoS over a single frame relay DTE DLCI is not supported on channelized SBIC-based WAN FPs.

There are four emission queues on Passport 15000 and 20000 FPs and MSA32 FPs. There are two emission queues on non-channelized SBIC-based WAN FPs. See the figure “Emission priority mapping on a single DLCI” (page 188).

Figure 33
Emission priority mapping on a single DLCI



IP-optimized DLCI class-based forwarding

IP-optimized DLCIs support CoS to QoS mapping over a single DLCI. Passport IP CoS allows flows of IP packets that are transmitted over a single IP-optimized DLCI to be differentiated by QoS characteristics. This differentiation is done by setting the emission priority and drop precedence of packets in certain IP flows based on their assigned CoS.

The drop precedence assigned to an IP packet arriving on a frame relay connection is determined by the discard eligibility (DE) bit to drop precedence mapping configured on the FRUNI. This can be overridden by the CoS to drop precedence mapping configured on the CoS policy group used by the ingress protocol port.

The Passport node uses the policy group configured on the egress protocol port for emission priority mapping and assigns an emission priority to the packet based on the packet's CoS value. IP-optimized DLCIs support four emission priorities.

ATM MPE class-based forwarding

ATM MPE supports CoS to QoS mapping over multiple virtual channel connections (VCCs) or over a single VCC.

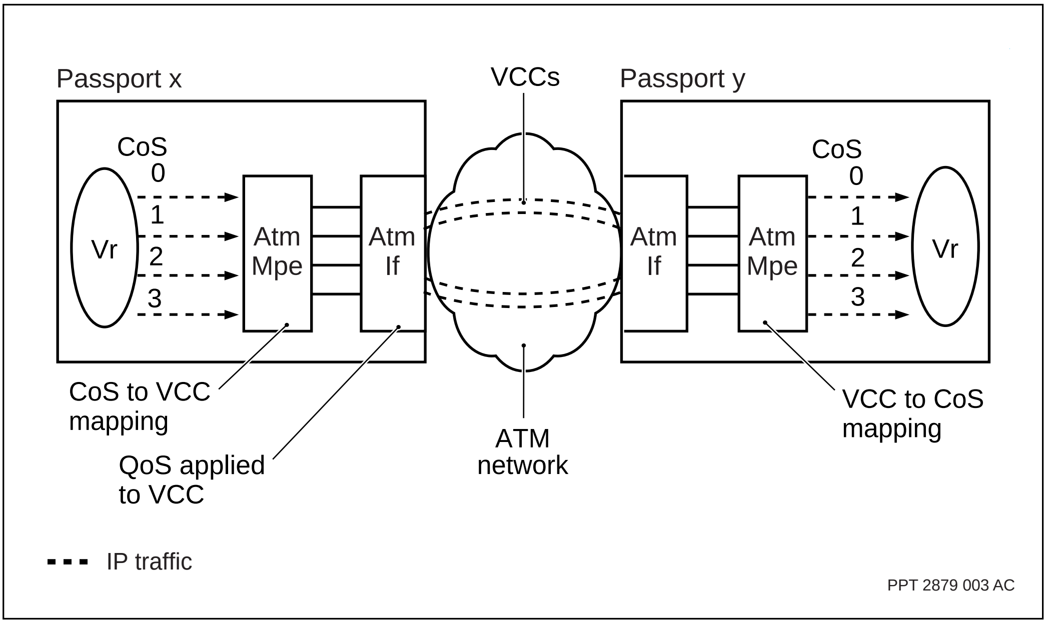
For CoS to QoS mapping over multiple VCCs, IP CoS lets you create up to four VCCs to each next hop router, with each VCC having separate QoS parameters and a different CoS index. Set each associated *AtmMpe AtmConnection ipcos* attribute.

When a packet goes out on an ATM MPE interface, Passport forwards the packet over the VCC whose associated *ipcos* attribute matches the packet's CoS value. See the figure "CoS to QoS mapping on multiple VCCs" (page 190).

The drop precedence assigned to an IP packet arriving on an ATM connection is determined by the cell loss priority (CLP) bit to drop precedence mapping configured on the ATM connection. This can be overridden by the CoS to drop precedence mapping configured on the CoS policy group used by the ingress protocol port.

You can configure the service class of each VCC under the *AtmIf* component as required, in order to provide different levels of service. Service classes include CBR, rtVBR, nrtVBR, UBR, and ABR. In addition, you can use the virtual path connection (VPC) functionality to apply traffic management functions to multiple VCCs. For more information, see 241-5701-705 *Passport 7400, 15000, 20000 ATM Traffic Management Fundamentals*.

Figure 34
CoS to QoS mapping on multiple VCCs



Gigabit Ethernet class-based forwarding

Quality of service for gigabit Ethernet is applied using IP CoS when component *Vr Ip Dsd* is not provisioned. When *Vr Ip Dsd* is provisioned, quality of service is applied using differentiated services as described in “Passport IP differentiated services” (page 195).

Gigabit Ethernet supports only layer 2 packet classification. Use attribute *Vr Pp IpPort ipCos* to assign a CoS index. The drop precedence is always set to high (3) and the scheduling class is always set to the lowest priority (0) for IP packets arriving on gigabit Ethernet ports.

Gigabit Ethernet supports CoS interworking with PQC-based ATM media. The interworking media combinations are described in the following figures:

- “Ingress on gigabit Ethernet; egress on PQC-based ATM” (page 192)
- “Ingress on PQC-based ATM; egress on gigabit Ethernet” (page 193)

You may see the following semantic warning coming from the IP port associated with a gigabit Ethernet interface:

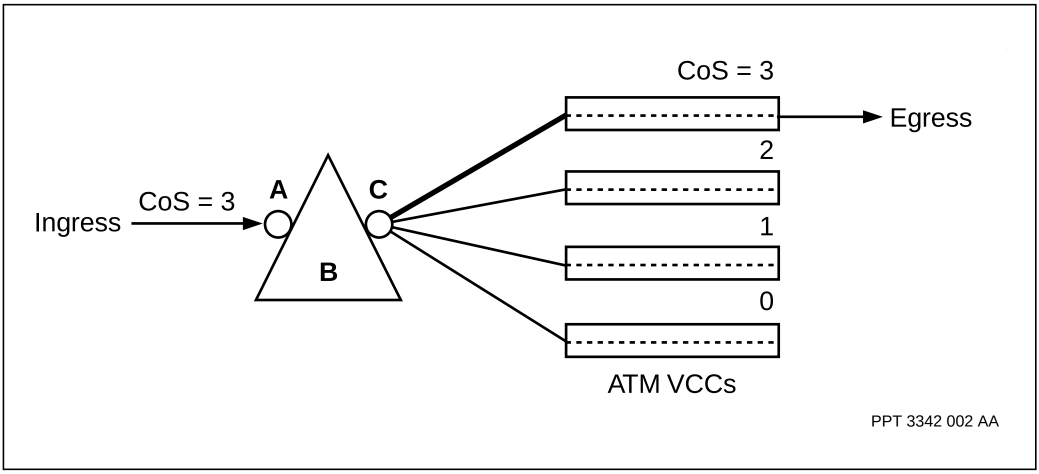
Warning: If a CosPolicyGroup is linked to the Ip or IpPort, then ipCos must be included in the featureList of the LogicalProcessorType linked to the Logical Processor that is used by the media linked to the ProtocolPort.

A Check Prov command produces this warning if

- the Ip or IpPort is linked to a CosPolicyGroup and
- the LogicalProcessorType linked to the LogicalProcessor that is used by the media linked to the ProtocolPort does not include ipCos in the featureList.

To eliminate this warning, you would normally add ipCos to the featureList of the LogicalProcessorType linked to the LogicalProcessor that is used by the media linked to the ProtocolPort, or ensure that the Ip and IpPort are not liked to a CosPolicyGroup; however, since the ipCos feature is not supported on the 4pGe FP, the warning cannot be eliminated.

Figure 35
Ingress on gigabit Ethernet; egress on PQC-based ATM



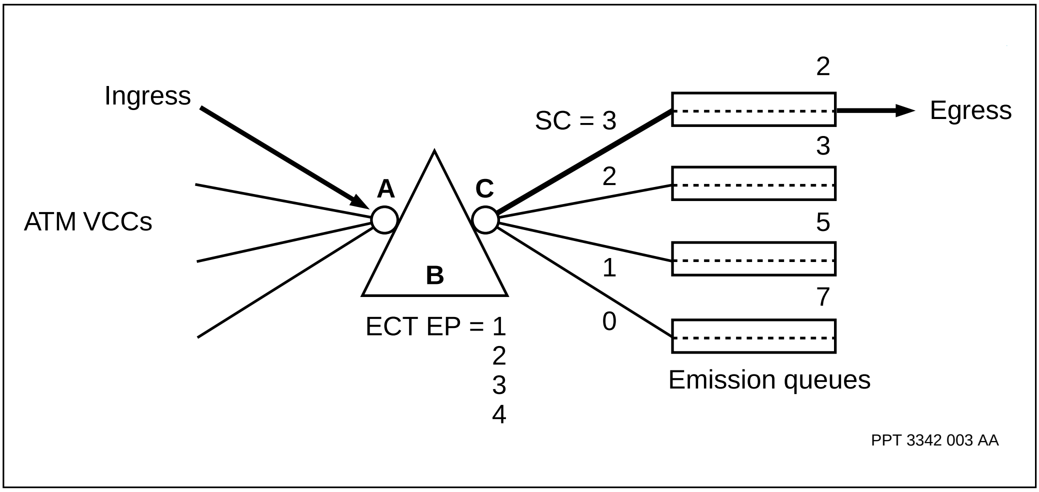
In the above figure, the following sequence occurs:

A The IP packet arrives on the ingress gigabit Ethernet protocol port. The IP packet is assigned a CoS index, using layer 2 classification, in attribute *Vr Pp IpPort ipCos*.

B The CoS index of the IP packet determined in step A is matched against the CoS indexes of the ATM connections belonging to the egress protocol port via attribute *AtmMpe Ac ipCos*. See “ATM MPE class-based forwarding” (page 189) for more information.

C The IP packet is transmitted via the egress ATM connection determined by step B.

Figure 36
Ingress on PQC-based ATM; egress on gigabit Ethernet



In the above figure, the following sequence occurs:

A The IP packet arrives on the ingress protocol port. The IP packet is assigned a CoS index, using layer 2, layer 3, or layer 4 classification.

B The CoS index is matched against a corresponding emission priority (EP), which is defined in attribute *Vr Ip Pg Ect ep*. This EP is mapped to one of eight emission queue of the egress gigabit Ethernet port. For the mapping values, see the table below. If no policy group is assigned to the egress gigabit Ethernet port, then queue 7 is used, regardless of the CoS index assigned to the that corresponds to the ingress PVC.

ECT emission priority	GigE queue number
1	2
2	3
3	5
4-8	7

C The IP packet is transmitted at the egress gigabit Ethernet protocol port on the queue selected by the EP that is now associated with the packet. The traffic management characteristics associated with each EP are defined in component *Lp Ethernet Tm Ep*.

Point-to-point protocol (PPP) class-based forwarding

Passport IP CoS allows flows of IP packets that are transmitted over a point-to-point protocol (PPP) interface to be differentiated by QoS characteristics. The differentiation is done by setting the emission priority and drop precedence of packets in certain IP flows based on their assigned CoS.

The Passport node performs layer 2 classification on every IP packet that arrives on a PPP interface based on the incoming protocol port. The Passport node uses the policy group configured on the ingress protocol port for drop precedence mapping.

The Passport node uses the policy group configured on the egress protocol port for emission priority mapping. Therefore, when a packet goes out a PPP interface, Passport assigns an emission priority to the packet based on the packet’s CoS value. For CoS to emission priority mapping over PPP, you can specify up to four different emission priorities, one for each CoS value.

Emission priority mapping based on CoS over PPP is not supported on channelized SBIC-based WAN FPs or PQC-based FPs.

IP CoS over virtual media

Virtual media protocol ports support layer 2 classification for multicast packets only. Multicast packets transmitting from one VR to another receive a CoS value at the ingress to the second VR, as configured through the *Vr ProtocolPort IpPort ipCoS* attribute of the associated ingress virtual media protocol port.

Chapter 18
Passport IP differentiated services

This section describes the implementation of IP differentiated services on Passport. It covers the following topics:

For information about configuring IP differentiated services, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

IP DiffServ advantages

IP DiffServ is the best solution for deploying differentiated services on Passport 15000 and 20000. IP CoS is limited to four treatments based on four classes associated with the IP packet. DiffServ offers the following improvements.

- more available treatments for each IP packet forwarded by the virtual router.
- greater control of DSCP values that can be marked into the DSCP field of each IP packet header.
- greater control of packets that are generated locally at the virtual router.

- Simplified provisioning model, network planning, and deployment.
- Improved standards compliance.
- Availability of future deployment and enhancements to IP differentiated services that are not feasible with IP CoS.

IP CoS uses who wish to migrate their network to IP DiffServ see the section “IP CoS to IP DiffServ Migration” (page 211).

Overview of IP DiffServ

Passport IP DiffServ is a framework for IP traffic management at each node in an IP network.

Each virtual router (VR) can be individually configured to support a set of per-hop-behaviors (PHBs). The “Per-hop-behaviors” (page 196) define how the IP packets are treated relative to each other.

Differentiated services are available on VR-to-VR IP VPN configurations. For information on which FPs support differentiated services, see 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

Differentiated services code point

The differentiated services code point (DSCP) is a six bit field in the IP packet header that determines how an IP packet is treated relative to others as it is forwarded by the virtual router. The DSCP is used to select “Per-hop-behaviors” (page 196) for the IP packets.

The possible DSCP values range from 0 to 63 but only twenty-one of these have standard definitions. The remaining values are for local or experimental use.

Per-hop-behaviors

Per-hop-behaviors (PHBs) define the treatment of an IP packet. The “Traffic class” (page 197), “Scheduling class” (page 197), “Drop precedence” (page 198), and “Connection class” (page 198) attributes of each PHB determine the importance of the IP packets relative to each other.

Each per-hop-behavior is identified by a particular DSCP. See the table “Default PHB configuration by domain type” (page 203) for the relationship between PHBs and DSCPs.

There are four general categories of per-hop-behaviors.

- “Default effort forwarding” (page 200)
- “Assured forwarding” (page 201)
- “Class selector forwarding” (page 201)
- “Expedited forwarding” (page 202)

Traffic class

The *trafficClass* attribute of a per-hop-behavior lets you aggregate per-hop-behavior component instances into groups that specify the same scheduling attribute values. The *trafficClass* attribute must be equal to an existing instance of component *Vr Dsd TrafficClass*.

For more information, see the description of attribute *Vr Dsd Phb trafficClass* in 241-5701-060 *Passport 7400, 15000, 20000 Components*.

Scheduling class

The scheduling class attribute of a per-hop-behavior determines how a packet is scheduled at the egress interface relative to other packets. Scheduling class is provisioned as a number from 0 to 7 depending on the number of queues supported at an interface.

In general, a higher value means that a packet is less likely to be delayed. A lower value means a packet is more likely to be delayed. The actual scheduling of packets depends on the scheduling mechanism at the egress interface.

The 4pGe FP is the only Passport FP that supports IP queuing over eight queues. You can use all eight queues only with IP traffic where the ingress and egress FPs are both 4pGe.

For more information, see the description of attributes *Vr Dsd Phb schedulingClass8Queues* and *schedulingClass4Queues* in 241-5701-060 *Passport 7400, 15000, 20000 Components* and the section on traffic management on the 4-port gigabit Ethernet FP in 241-5701-615 *Passport 7400, 15000, 20000 FP Configuration Reference*.

Drop precedence

The *dropPrecedence* attribute of a per-hop-behavior controls the loss sensitivity of IP packets relative to other IP packets in a queue. The drop precedence can be provisioned to *low*, *medium*, or *high*.

In general, packets with a drop precedence of *low* are less likely to be discarded when the queue is congested. Packets with a drop precedence of *high* have a lower probability of being forwarded and are most likely to be discarded. The actual loss sensitivity of packets depends on the discarding mechanism of the queue.

For more information, see the description of attribute *Vr Dsd Phb dropPrecedence* in 241-5701-060 *Passport 7400, 15000, 20000 Components*.

Connection class

Connection class is a value between 0 and 3 that is used to select egress virtual connections when packets are transmitted over the connection.

A packet is dynamically assigned a connection class value according to the per-hop-behavior that matches the DSCP of the packet header. The packet is transmitted on the link that has the same connection class value. If that link does not exist then the link with the next lower value is used. If that alternate link also does not exist, then the link having the next higher value is used.

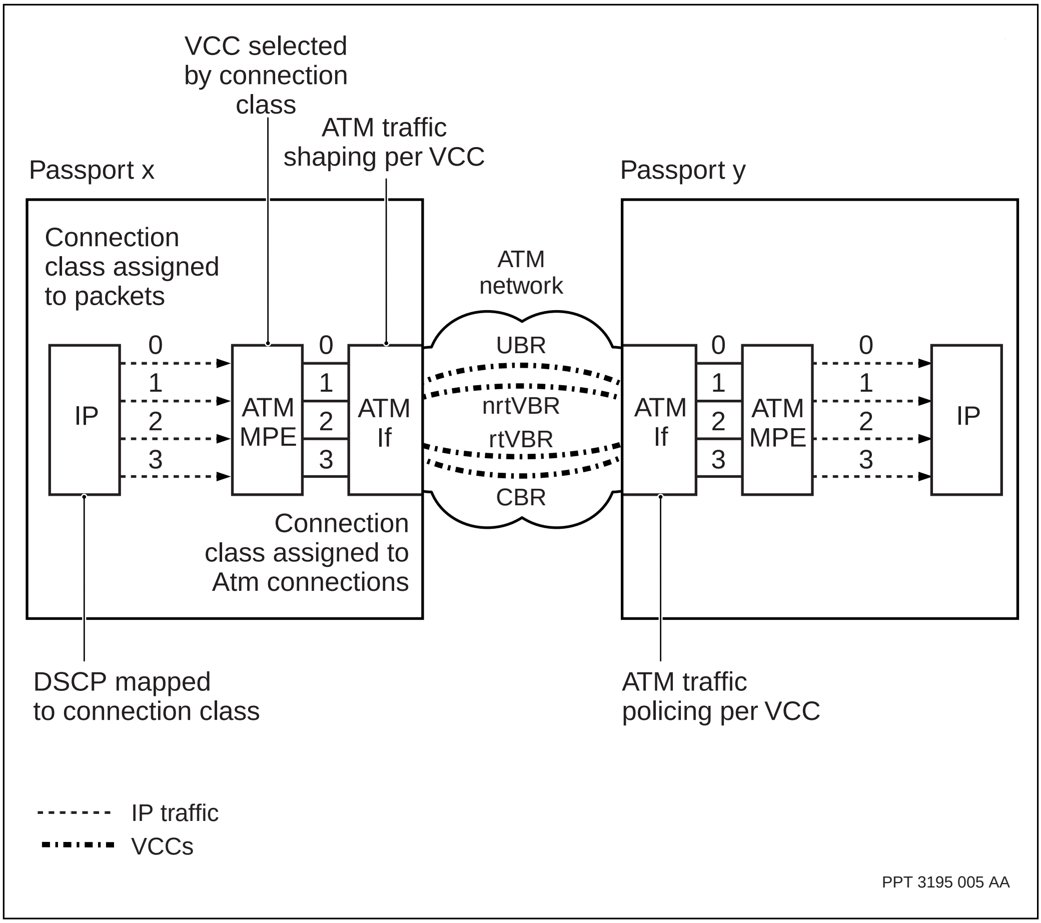
Each virtual connection is configured to provide different policing and scheduling characteristics. The forwarding behaviors of IP packets are differentiated as they move through the link layer network. In general, the scheduling characteristics of the connections should be configured to be better for numerically greater connection class values. For example, if four AtmConnection components are associated with AtmIf components configured for ATM service classes UBR, nrtVBR, rtVBR, and CBR, then the ipCos attributes of the AtmConnection should be 0, 1, 2, and 3, respectively.

> ***Note:*** In all default configurations, IP packets where the DSCP field value is CS6 (48) are sent on the connection where the IP CoS value is 2. This includes all locally generated IP routing control packets.

The connection class of a PHB is controlled by the scheduling class of the PHB according to a fixed mapping.

For more information, see the description of attribute *Vr Dsd Phb connectionClass* in 241-5701-060 *Passport 7400, 15000, 20000 Components*.

Figure 37
Example: IP service differentiation on AtmMpe using connection class


VCC selected by connection class
ATM traffic shaping per VCC
Passport x
Passport y
Connection class assigned to packets
ATM network
0
1
2
3
IP
ATM MPE
ATM If
UBR
nrtVBR
rtVBR
CBR
Connection class assigned to Atm connections
DSCP mapped to connection class
ATM traffic policing per VCC
IP traffic
VCCs
PPT 3195 005 AA

Default effort forwarding

Packets classified with default effort forwarding (DF) are given “best effort” quality-of-service treatment. Packets with this PHB are considered the least urgent and are the most likely packets to be discarded in a DiffServ enabled IP network.

Assured forwarding

Assured forwarding (AF) provides four bandwidth classes and three discard priorities. There are 12 assured forwarding PHBs that are grouped into four classes. Each class has three levels of loss sensitivity.

- af43
- af42
- af41
- af33
- af32
- af31
- af23
- af22
- af21
- af13
- af12
- af11

Class selector forwarding

There are 8 class selector (CS) PHBs. Each value is associated with a different level of urgency.

- cs7
- cs6
- cs5
- cs4
- cs3
- cs2
- cs1
- cs0

Expedited forwarding

Packets classified with expedited forwarding (EF) are considered delay sensitive and the least likely to be discarded. This PHB provides a very high quality of service treatment with special considerations for delay intolerance.

Differentiated service domains

IP networks can be organized into IP DiffServ regions and IP DiffServ domains. An IP DiffServ domain is a group of virtual routers that share the same per-hop-behavior definitions. An IP DiffServ region is a network of interconnected DiffServ domains.

Passport provides 5 different pre-configured IP DiffServ domains. Each virtual router is configured with one of the following domains:

- Passport domain (pa)
- ClassSelector domain (cs)
- AssuredForwarding domain (af)
- PacketVoice domain (pv)
- WirelessUmts domain (umts)
- Custom domain (cu)

Each differentiated service domain offers a subset of the “Per-hop-behaviors” (page 196) (PHBs) available to the IP packets being transported within the domain. Any domain may be customized by adding or deleting PHBs.

The table “Default PHB configuration by domain type” (page 203) lists the per-hop-behaviors supported for each DiffServ domain.

Each PHB in an IP DiffServ domain is given a default “Traffic class” (page 197), “Drop precedence” (page 198), and “Connection class” (page 198) value when it is created.

Any default PHB can be removed from a domain except for DSCP=0. The scheduling class, drop precedence, and connection class can also be modified for each PHB.

To view the default values for a PHB when it is created, see the *Vr Ip Dsd Phb* section of 241-5701-060 *Passport 7400, 15000, 20000 Components*,

Table 30
Default PHB configuration by domain type

DSCP		Passport (pa)	Class Selector (cs)	Assured Forwarding (af)	Packet voice (pv)	Wireless Umts (umts)	Custom (cu)
decimal format	RFC format						
56	CS7	yes	yes		yes	yes	
48	CS6	yes	yes	yes	yes	yes	yes
46	EF	yes			yes	yes	
40	CS5	yes	yes		yes		
38	AF43	yes		yes			
36	AF42	yes		yes			
34	AF41	yes		yes			
32	CS4	yes	yes				
30	AF33	yes		yes		yes	
28	AF32	yes		yes		yes	
26	AF31	yes		yes		yes	
24	CS3	yes	yes				
22	AF23	yes		yes		yes	
20	AF22	yes		yes		yes	
18	AF21	yes		yes		yes	
16	CS2	yes	yes				
14	AF13	yes		yes	yes	yes	
12	AF12	yes		yes	yes	yes	
10	AF11	yes		yes	yes	yes	
8	CS1	yes	yes		yes	yes	
(Sheet 1 of 2)							

Table 30 (continued)
Default PHB configuration by domain type

DSCP		Passport (pa)	Class Selector (cs)	Assured Forwarding (af)	Packet voice (pv)	Wireless Umts (umts)	Custom (cu)
decimal format	RFC format						
0	DF	yes	yes	yes	yes	yes	yes
0	CS0	yes	yes	yes	yes	yes	yes
(Sheet 2 of 2)							

Your Nortel Networks technical support representative can help you determine the DiffServ domain that is the most suitable for your network.

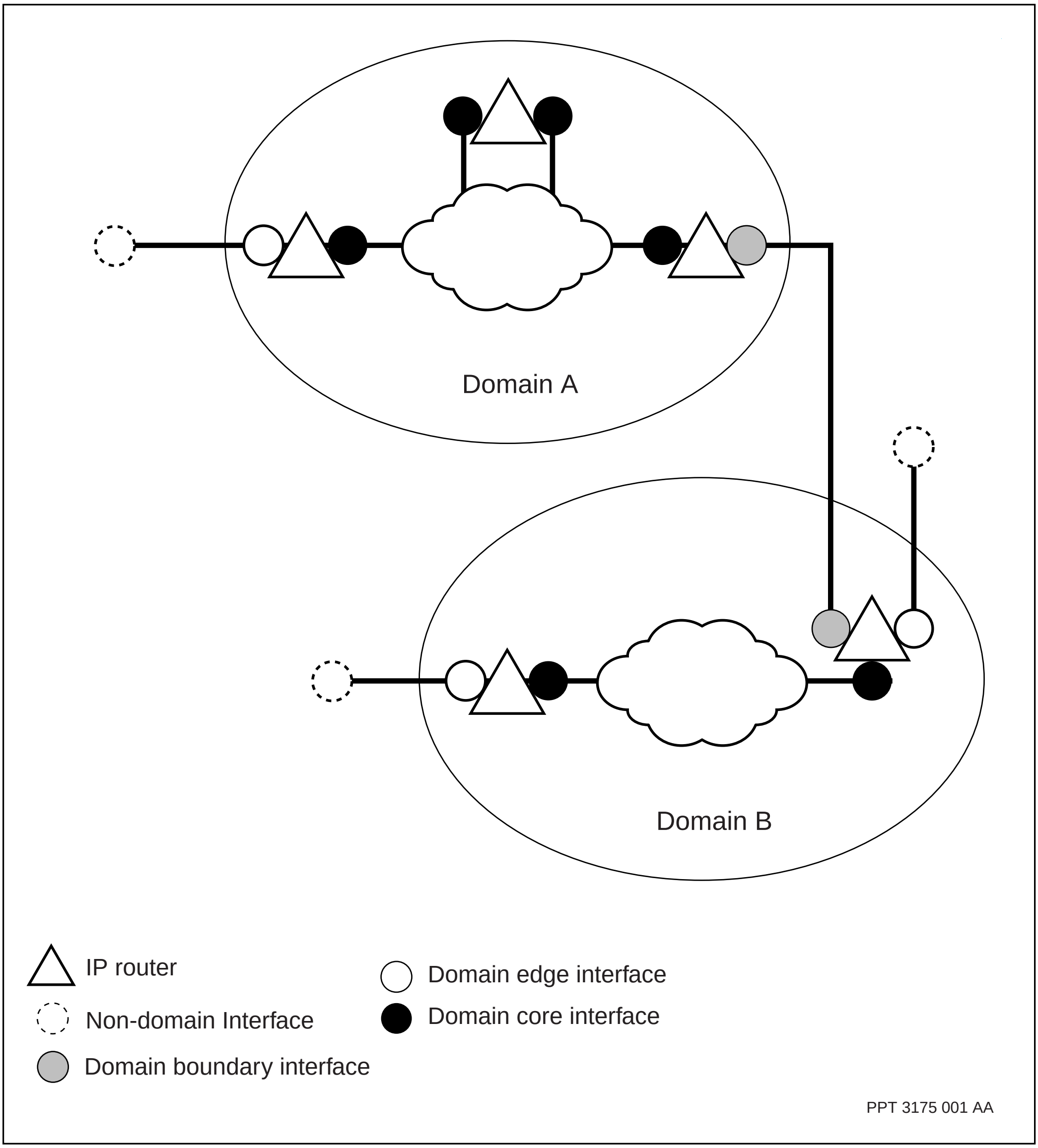
IP DiffServ domain interface types

There are three IP DiffServ domain interface types for differentiated services.

- “Domain core interface” (page 207)
- “Domain edge interface” (page 208)
- “Domain boundary interface” (page 209)

Each interface has a different method of classifying and marking the DSCP field of the IP packets that are received and transmitted by the interface. The figure “IP Diffserv domain interface relationships” (page 205) shows where each interface type appears in a network.

Figure 38
IP Diffserv domain interface relationships



On Passport, an IpPort is configured as an IP DiffServ domain edge, core, or boundary interface by configuring a DiffServ profile for the IpPort (*Vr Ip DiffServ*).

The table “IP DiffServ profile settings for DSCP marking at the IpPort” (page 206) lists the specific values that are required to configure IP DiffServ interface profiles to be either domain core, domain edge, or domain boundary.

Table 31
IP DiffServ profile settings for DSCP marking at the IpPort

Component	Core interface	Edge interface	Boundary interface
Vr Ip DiffServ Is baMode (receive)	preserve	link	preserve or translate
Vr Ip DiffServ Es baMode (transmit)	preserve	preserve	preserve or translate

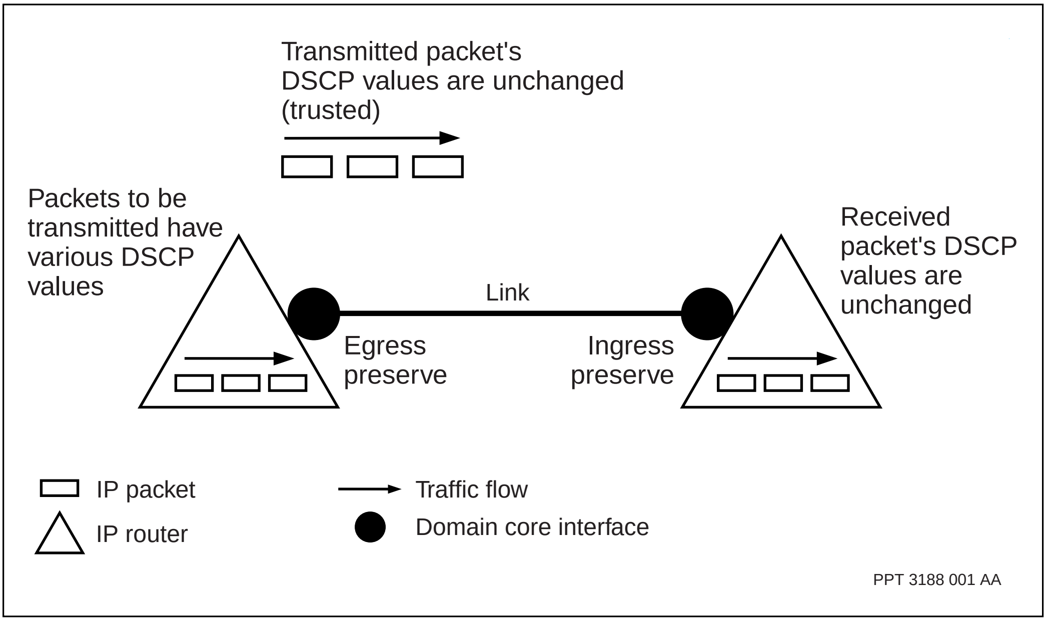
Note: If the ingress and egress behavior aggregate modes at one end of a link between domain boundary interfaces are set to *translate*, then the modes at the other end should both be set to *preserve*.

Domain core interface

A domain core interface transmits and receives packets from an IP router that is also in the same differentiated services domain. The DSCP value of a packet is preserved as it is received or transmitted through a core interface.

The figure “DSCP marking through a core interface” (page 207) show how packets are treated at a core interface.

Figure 39
DSCP marking through a core interface

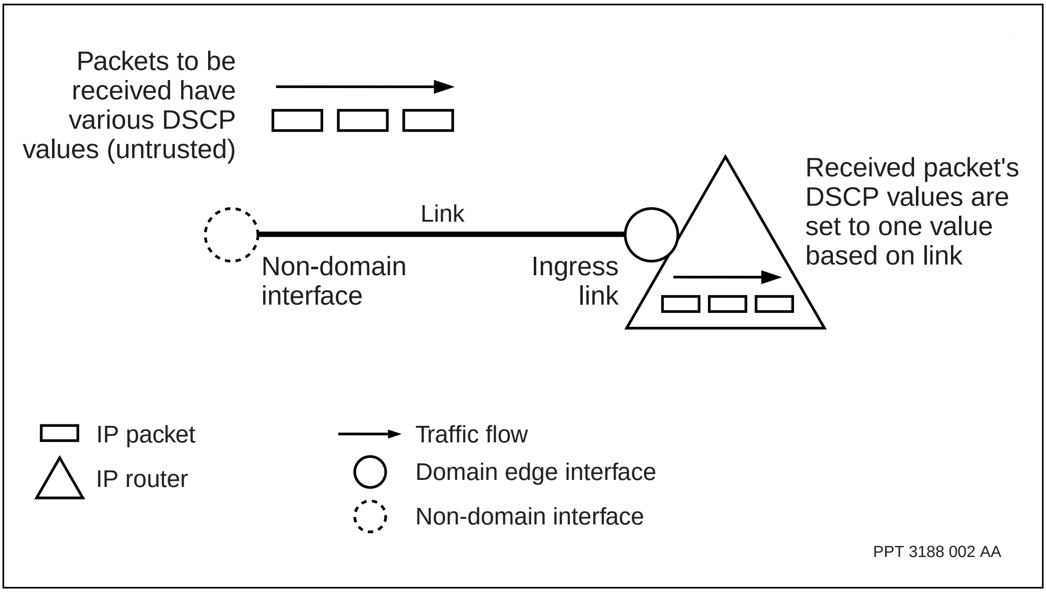


Domain edge interface

A domain edge interface transmits or receives packets from a device outside of a DiffServ domain. The DSCP value of a packet is changed as it is received by a domain edge interface according to the configuration of the link.

The figure “DSCP marking through a domain edge interface” (page 208) shows how the DSCP of the packet is changed.

Figure 40
DSCP marking through a domain edge interface


Packets to be received have various DSCP values (untrusted)
Link
Non-domain interface
Ingress link
Received packet's DSCP values are set to one value based on link
IP packet
IP router
Traffic flow
Domain edge interface
Non-domain interface
PPT 3188 002 AA

Domain boundary interface

A domain boundary interface transmits or receives packets from an IP router that is not in the same DiffServ domain but is in another DiffServ domain.

There are two options for handling the DSCP at a boundary interface.

1 The DSCP of the packet may be translated as the packet is transmitted by the boundary interface.

  This process is known as re-marking or egress IP DSCP marking. It is typically used for boundary interfaces where the next IP hop does not have the ability to translate a received DSCP value into a value with a PHB that its own DiffServ domain can understand.

  Egress IP DSCP marking is not supported when the ingress interface that receives the packet is physically located on a 4pGe FP. It is supported when the egress interface that transmits the packet is physically located on a 4pGe FP as long as the FP of the ingress interface supports egress IP DSCP marking. See the table “Egress IP DSCP marking” (page 209).

2 The DSCP of the packet may be translated as the packet is received when it enters the boundary interface of the next domain.

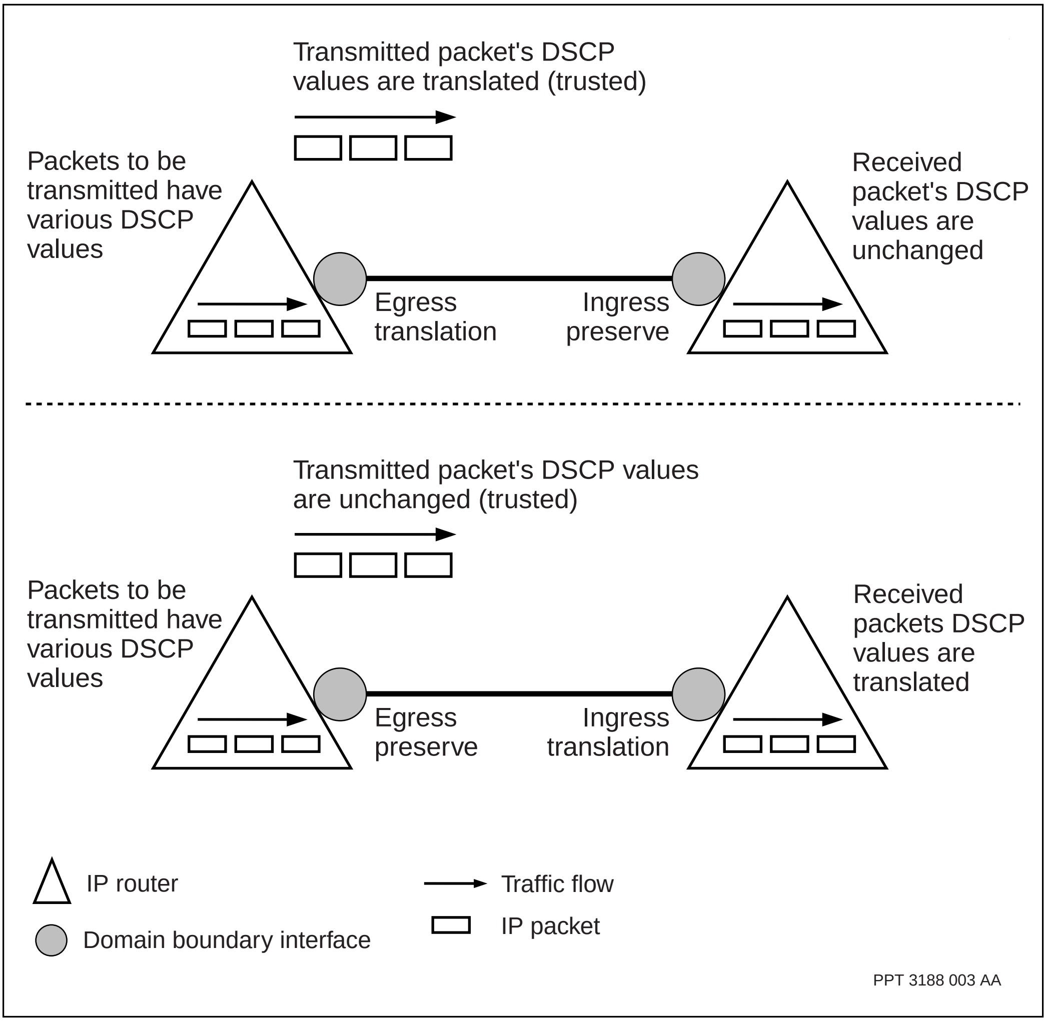
Table 32
Egress IP DSCP marking

		Egress FP	
		4pGe	**Other FPs** [1]
Ingress FP	**4pGe**	not supported	not supported
	Other FPs [1]	supported	supported

[1] other FPs that support the ipDiffServ feature and egress IP DSCP marking

The figure “DSCP marking through a domain boundary interface” (page 210) shows how the DSCP of the packets are mapped to new values.

Figure 41
DSCP marking through a domain boundary interface


Transmitted packet's DSCP values are translated (trusted)
Packets to be transmitted have various DSCP values
Received packet's DSCP values are unchanged
Egress translation
Ingress preserve
Transmitted packet's DSCP values are unchanged (trusted)
Packets to be transmitted have various DSCP values
Received packets DSCP values are translated
Egress preserve
Ingress translation
IP router
Domain boundary interface
Traffic flow
IP packet
PPT 3188 003 AA

Chapter 19
IP CoS to IP DiffServ Migration

IP CoS and IP DiffServ are the two ways IP differentiated services can be deployed on Passport. DiffServ offers several advantages over IP CoS and should be considered as the best choice for differentiated services on a Passport 15000 or 20000. “Passport IP differentiated services” (page 195) describes DiffServ and its advantages over IP CoS.

To migrate from IP CoS to IP DiffServ refer to the related procedures in 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

The section “Comparison between IP CoS and IP DiffServ provisioned attributes” (page 211) compares the components used to provision either IP CoS or DiffServ.

Comparison between IP CoS and IP DiffServ provisioned attributes

There is no direct mapping in component provisioning between IP CoS and IP DiffServ. IP CoS is provisioned using *Vr Ip CosPolicyGroup* components, while DiffServ is configured using the *Vr DifferentiatedServicesDomain* and *Vr Ip DifferentiatedServices* components.

The following tables show the configuration relationship between DiffServ and IP CoS.

- “IP CoS and IP DiffServ software activation attributes” (page 212)
- “IP CoS and IP DiffServ layer 2 configuration attributes” (page 212)
- “IP CoS and IP DiffServ layer 3 configuration attributes” (page 212)

- “IP CoS and IP DiffServ PHB definition attributes” (page 213)
- “IP CoS and IP DiffServ drop precedence selection attributes” (page 213)
- “IP CoS and IP DiffServ local packet classification and marking attributes” (page 213)
- “IP CoS and IP DiffServ behavior aggregate classification and marking attributes” (page 213)

Table 33
IP CoS and IP DiffServ software activation attributes

IP CoS configuration	**IP DiffServ configuration**
Sw Lpt fl ipCos	*Sw Lpt fl ipDiffServ*
Feature ipDiffServ is required for the DiffServ profile for the interface (*Vr Ip DiffServ*) but not for the DiffServ domain for the router (*Vr Dsd*)	

Table 34
IP CoS and IP DiffServ layer 2 configuration attributes

IP CoS configuration	**IP DiffServ configuration**
Vr Pp IpPort ipCos	*Vr Pp Ipport ipCos*
Vr Pp IpPort ipDscp	*Vr Pp Ipport ipDscp*
AtmMpe Ac ipCos	*AtmMpe Ac ipCos*
AtmMpe Ac ipDscp	*AtmMpe Ac ipDscp*

Table 35
IP CoS and IP DiffServ layer 3 configuration attributes

IP CoS configuration	**IP DiffServ configuration**
Vr Ip cosPolicyAssignment	*Vr Ip diffServAssignment*
Vr Pp IpPort cosPolicyAssignment	*Vr Pp diffServAssignment*

Table 36
IP CoS and IP DiffServ PHB definition attributes

IP CoS configuration	IP DiffServ configuration
Vr Ip CosPolicyGroup Ect ep	*Vr DiffServDomain Phb sc*
Vr Ip CosPolicyGroup Ict dp	*Vr DiffServDomain Phb dp*

Table 37
IP CoS and IP DiffServ drop precedence selection attributes

IP CoS configuration	IP DiffServ configuration
Vr Ip CosPolicyGroup Ict dp	*Vr Ip DiffServ Is dpMode*

Table 38
IP CoS and IP DiffServ local packet classification and marking attributes

IP CoS configuration	IP DiffServ configuration
Vr Ip dscpGeneralLocalSource	*Vr Ip phbGeneralLocalSource*
Vr Ip dscpRoutingLocalSource	*Vr Ip phbRoutingLocalSource*

Table 39
IP CoS and IP DiffServ behavior aggregate classification and marking attributes

IP CoS configuration	IP DiffServ configuration
n/a	*Vr Ip DiffServ Is baMode*
Vr Ip Pg Policy assignedCos	*Vr Ip DiffServ Is Ba dscpToMark*
Vr Ip Pg Policy TosMap tos	*Vr Ip DiffServ Is Ba dscpToMatch*
Vr Ip Pg Ect tosMask	(always OxFC, not provisioned)
(Sheet 1 of 2)	

Table 39 (continued)
IP CoS and IP DiffServ behavior aggregate classification and marking attributes

IP CoS configuration	IP DiffServ configuration
Vr Ip Pg Ect setTos	*Vr Ip DiffServ Es baMode*
Vr Ip Pg Ect tos	*Vr Ip DiffServ Es Ba dscpToMark*
Vr Ip Pg Policy assignedCos	*Vr Ip DiffServ Es Ba dscpToMatch*
Vr Ip Pg Ect tosMask	(always OxFC, not provisioned)
(Sheet 2 of 2)	

Chapter 20
IP flow filters

This section describes the implementation of IP flow filters on Passport. It covers the following topics:

- “Overview of IP flow filters on Passport” (page 215)
- “IP flow filter definition” (page 218)

For information about configuring IP flow filters, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP flow filters on Passport

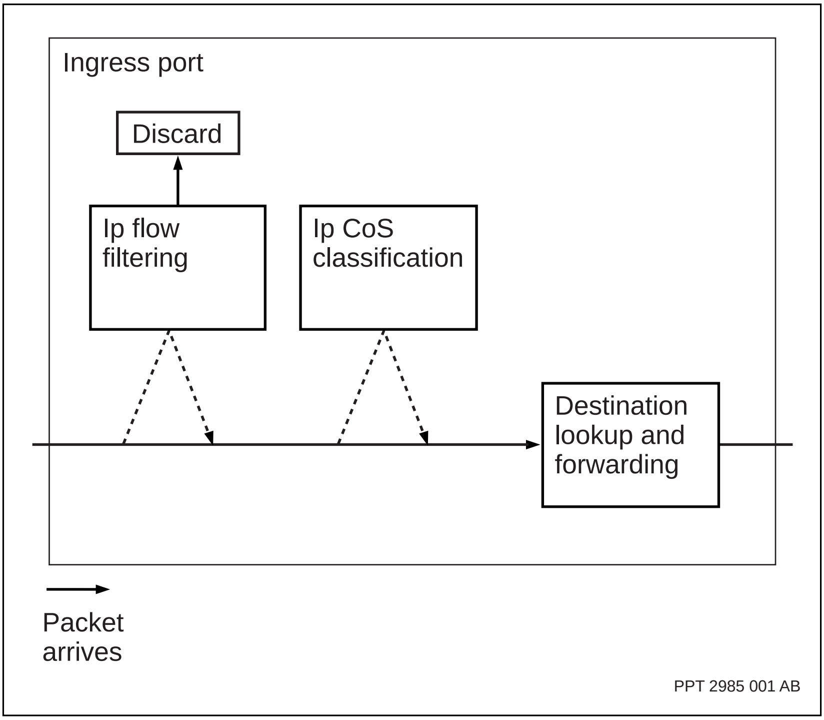
IP flow filters help you to create a secure network by allowing you to decide which IP packet flows will be permitted or denied entry to the network. Permit or deny actions are based on the most specific match to the IP packet flow’s source IP address, destination IP address, or both the source and destination IP addresses.

If permitted, packets are processed and forwarded to their destination. If denied, packets are dropped. If packets do not match any flow of the assigned filter, they are dropped.

When the source IP address, destination IP address, or both the source and destination IP addresses match more than one of the specified flows under a filter, the match with the most precise source address is given precedence. If, once that source address specification is selected, the destination address matches more than one of the specified flows under a filter, the match with the most precise destination address is given precedence.

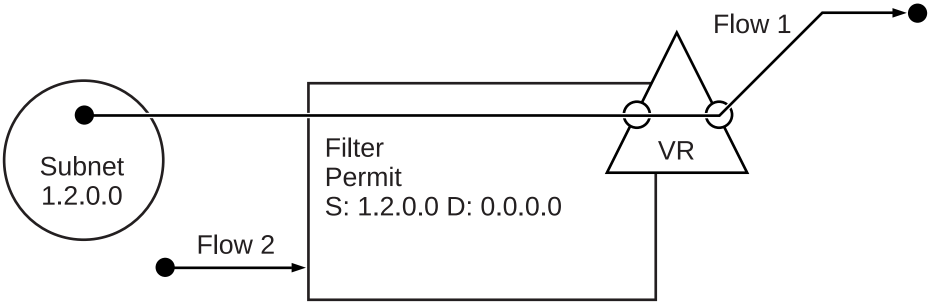
The figure “IP flow filter placement on the transmission path” (page 216) shows the placement of the IP flow filters on the transmission path.

Figure 42
IP flow filter placement on the transmission path



To see permit and deny actions for IP packet flows based on source address, see the figure “Filtering based on source address” (page 217). The figure “Filtering based on destination address” (page 217) illustrates permit and deny actions for IP packet flows based on destination address.

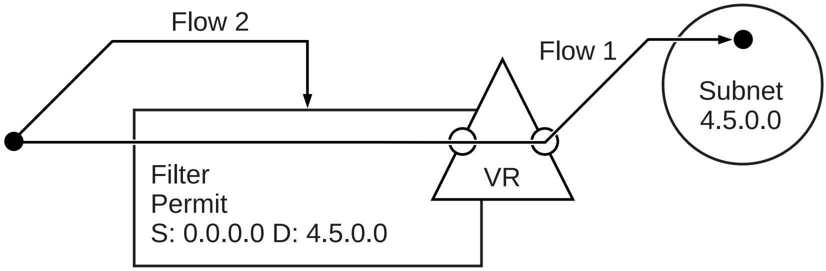
Figure 43
Filtering based on source address



Flow 1 (from inside 1.2.0.0 to anywhere) permitted
Flow 2 (from outside 1.2.0.0 to anywhere) denied

PPT 3027 002 AA

Figure 44
Filtering based on destination address



Flow 1 (from anywhere to inside 4.5.0.0) permitted
Flow 2 (from anywhere to outside 4.5.0.0) denied

PPT 3027 001 AA

You can configure IP flow filters under the *Ip* subcomponent of a VR. After the flow filter is created for a VR, you can assign the filter to any of the protocol ports on that VR. To assign a flow filter to a protocol port on a VR, link the filter to the *IpPort* component under the protocol port. To assign a flow filter to all of the protocol ports on a VR, link the filter to the *Ip* component under the VR. The filter assignment for a protocol port will override the filter assignment for a VR. A filter created under one VR cannot be assigned to another VR or to the protocol ports of another VR.

IP flow filtering actions will not be applied to any protocol port that is linked to IP tunnel media or to any protocol port assigned to a VR that is linked to IP tunnel media. Since no filtering actions are applied to protocol ports linked to IP tunnel media, all IP packets are forwarded in this situation.

IP flow filters are only supported on PQC2-based cards.

IP flow filter definition

The *Filter* subcomponent of the *Ip* component allows you to add, delete, and modify IP flow filters. A filter is assigned to a virtual router by linking the corresponding *Filter* subcomponent to the *Ip* component under the same virtual router. A filter is assigned to a protocol port by linking the corresponding *Filter* subcomponent to the *IpPort* subcomponent of the desired *Protocol Port* component, under the same virtual router. If no filter is assigned, all packets are permitted entry to the network.

A filter associated with a protocol port is applied to all ingress IP traffic, but not egress traffic, at that port. Filtering is performed prior to making the forwarding decision that determines the egress port where the packet will be transmitted. If no filter component is associated with a port (because there is neither an assignment to the *Vr Ip* component, nor an assignment to the *Vr Pp IpPort* component) all IP packet flows are admitted at that port.

The *FilterFlow* subcomponent of the *Filter* subcomponent is used to define the filter action (whether an IP flow is permitted or denied) that is required for an individual IP flow. A *Filter* subcomponent must have at least one *FilterFlow* subcomponent whose filter action is permit.

Two filter flow subcomponents should not be configured under one filter component or subcomponent that have equivalent source and destination address specifications.

IP flow filters can block IP packets that are used by routing protocols. When configuring IP flow filters, it is important to configure IP flow filters that permit or deny the IP control packets.

Chapter 21
IP tunnels

This section describes the implementation of IP tunnels on Passport. It covers the following topics:

- “Overview of IP tunnels” (page 221)
- “Encapsulation techniques” (page 223)
- “Point-to-point tunnels” (page 225)
- “Point-to-multipoint tunnels” (page 226)

For information about configuring IP tunnels, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*.

Overview of IP tunnels

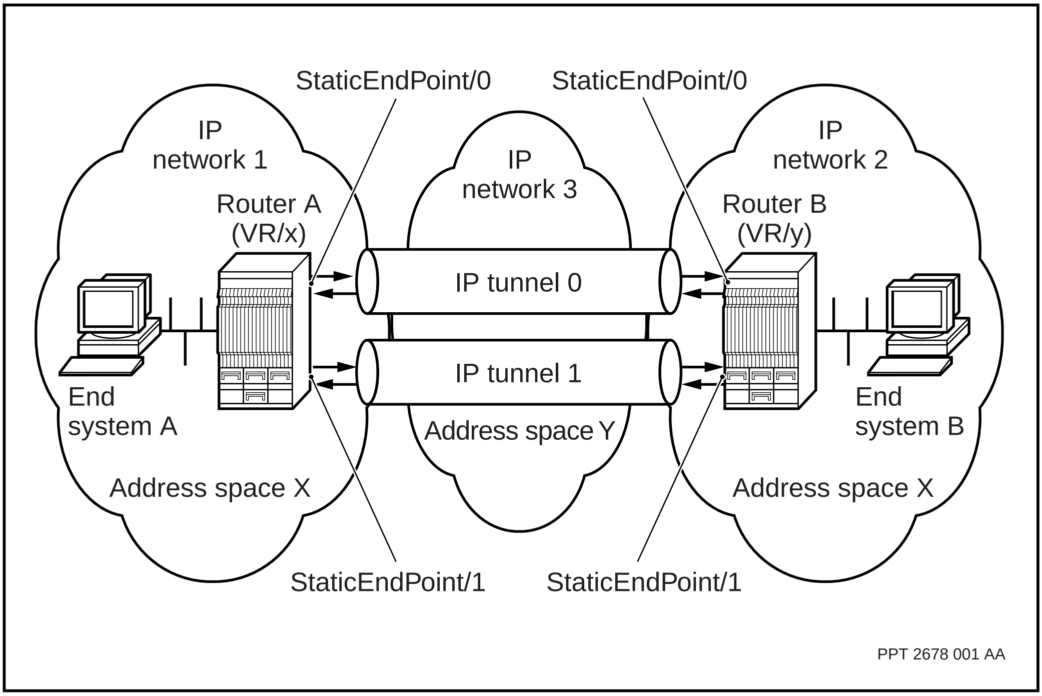
The IP tunnels feature enables you to connect two physically separate networks that share the same address space through an IP network with a different address space by encapsulating the original packet in an IP header. This outer IP header contains the routing information to traverse, or tunnel through, a network with a different address space. Passport switches implement IP tunnel functionality through the *Tunnel* component and *StaticEndPoint* subcomponent.

Passport supports point-to-point (PTP) tunnels and point-to-multipoint (PTMP) tunnels.

The IP tunneling feature on Passport supports static point-to-point tunnels. “IP tunnel example” (page 222) illustrates the following characteristics of static point-to point tunnels:

- You identify tunnel end points using a *StaticEndPoint* component.
- You identify the source address of a tunnel using the *sourceAddress* attribute under the *StaticEndPoint* component. You identify the destination address of a tunnel using the *destinationAddress* attribute under the *StaticEndPoint* component. The source and destination addresses are what define a tunnel instance.You must manually provision source and destination addresses on each Passport switch in the shared address space.
- The tunnels themselves are subnets of the address space of the network being tunneled through which, in the figure “IP tunnel example” (page 222), is IP Network 3.

Figure 45
IP tunnel example



Encapsulation techniques

IP tunneling, regardless of the protocol of the originating network, is made possible by adding an outer IP header to the original packet. This outer IP header contains the routing information to traverse, or tunnel through, a network with a different address space.

You can use two encapsulation methods for point-to-point or point-to-multipoint tunnel traffic originating in IP networks, but the method must be the same at both ends of the tunnel:

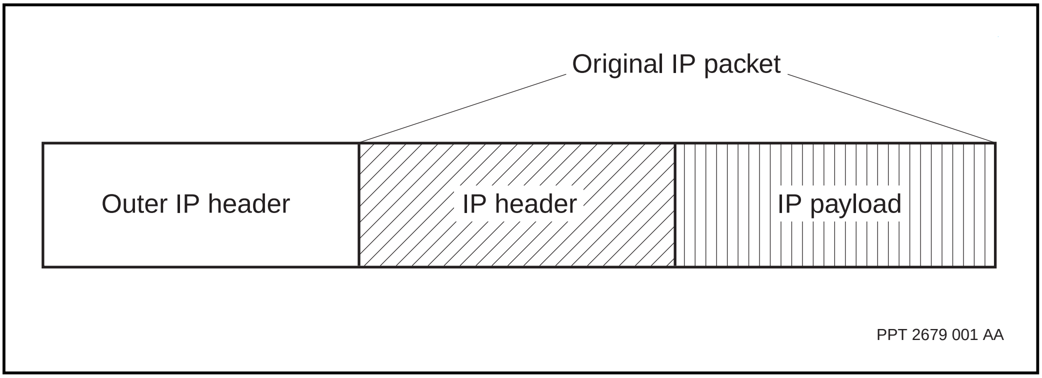
- IP in IP encapsulation as defined in RFC 2003. For more information see "IP in IP encapsulation" (page 223).
- Generic Routing Encapsulation (GRE) over IP as defined in RFC 1702. For more information see "Generic routing encapsulation (GRE)" (page 224).

IP in IP encapsulation

IP in IP is the encapsulation method for tunnel traffic originating in IP networks. In this method, the Passport in the originating IP network adds a new (outer) IP header, which contains the routing information to tunnel through the network with a different address space, to the original IP packet. The Passport *Vr* component adds this header before entering the tunnel. The Passport *Vr* component in the IP network at the far end of the tunnel strips off the outer IP header as the packet exits the tunnel. The original packet is then forwarded to its destination as usual. "IP in IP encapsulation format" (page 224) shows the format of an IP packet encapsulated in IP.

> ***Note:*** IP in IP tunnel encapsulation is not compatible with RFC 2003 when using an ATM IP FP as a backbone FP.

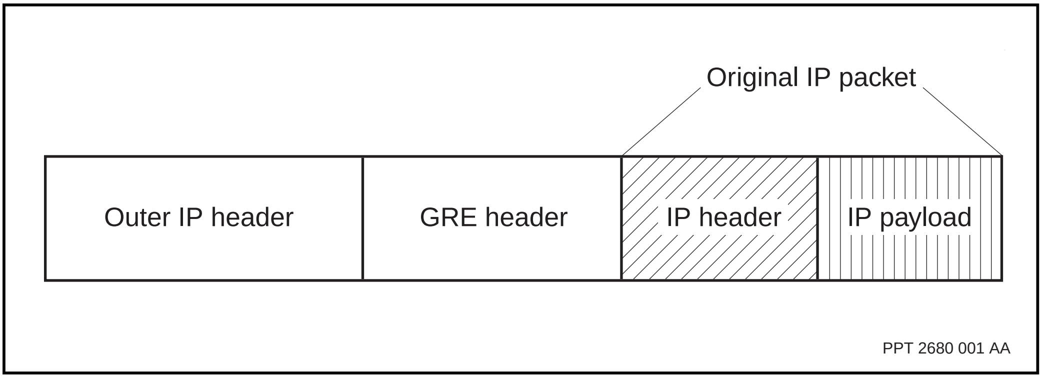
Figure 46
IP in IP encapsulation format



Generic routing encapsulation (GRE)

GRE is the other encapsulation method for tunnel traffic originating in IP networks. The Passport *Vr* component in the originating IP network adds a GRE header and then an outer IP header to the original IP packet before it enters the tunnel. The network being tunneled through treats the encapsulated packet as a regular IP packet while it is in transmission through the tunnel. The Passport *Vr* component in the IP network at the far end of the tunnel strips off the outer IP header and the GRE header as the packet exits the tunnel. The original packet is then forwarded to its destination as usual.

Figure 47
GRE over IP encapsulation format for IP

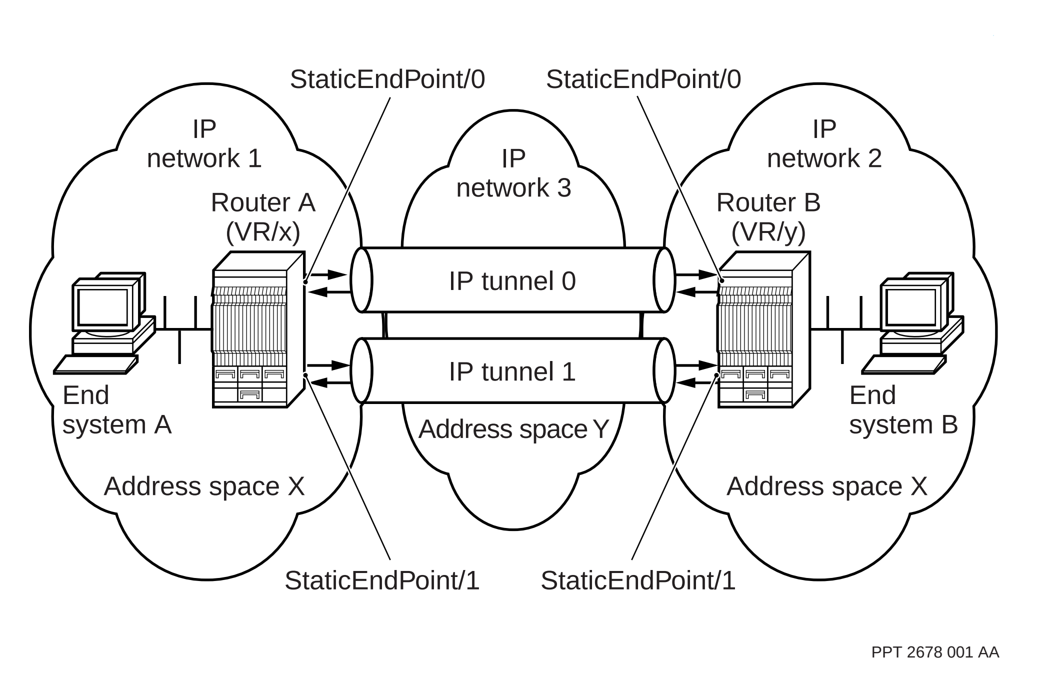


Point-to-point tunnels

The Passport IP service uses point-to-point (PTP) IP tunnels to provide connectivity between customer VRs that reside on different Passport nodes.

The figure “Point-to-point IP tunnels” (page 225) provides an example of two IP networks sharing address space X connected through two IP tunnels through a third IP network on address space Y.

Figure 48
Point-to-point IP tunnels



You define a tunnel instance by its source and destination addresses. The source address at one end of the tunnel must be the same value as the destination address at the other end of the tunnel. In addition, the source address at one end and the destination address at the other end must belong to the same address space. The tunnels themselves are subnets of the address space of the network being tunneled through which, in the figure “Point-to-point IP tunnels” (page 225), is IP Network 3.

You must manually configure source and destination addresses on each Passport node in the shared address space. For example, in the figure “Point-to-point IP tunnels” (page 225) an operator must configure the source and destination addresses on the Passport node in IP network 1, and another operator must configure the same in IP network 2.

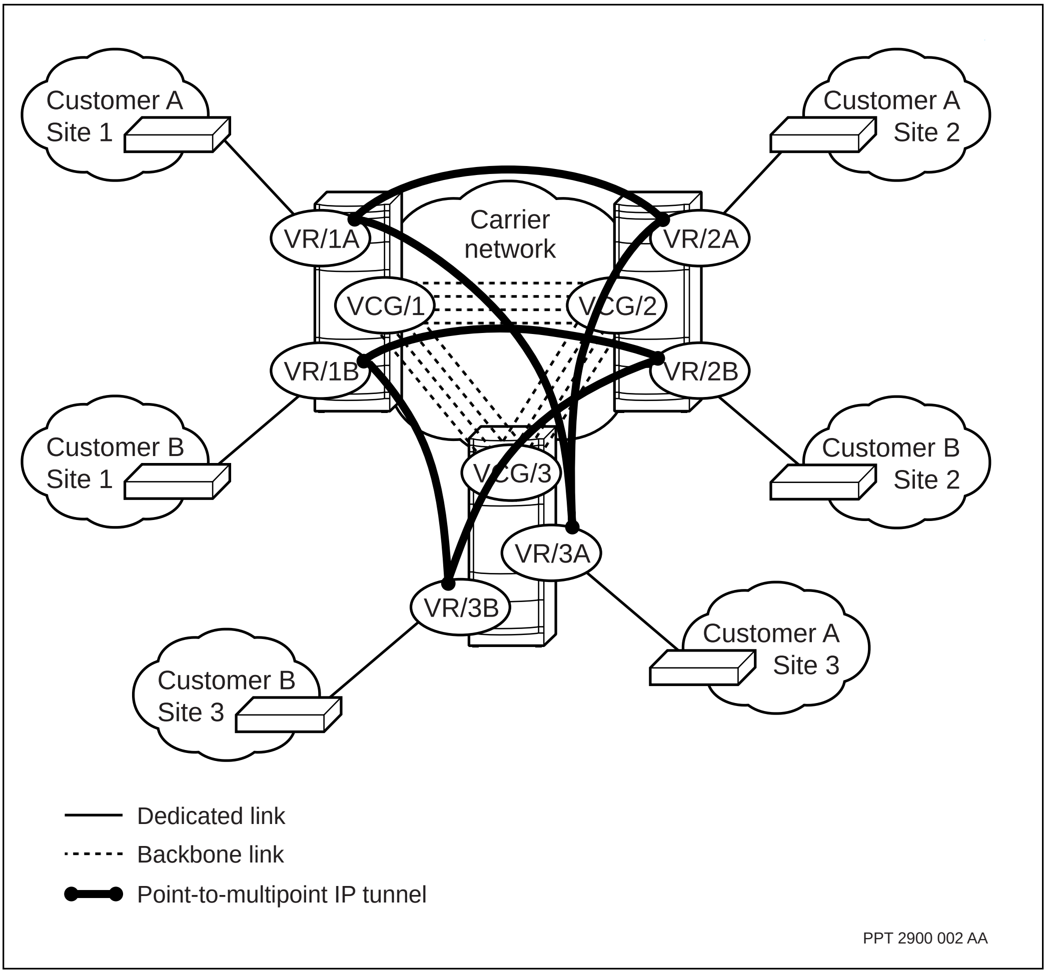
For IP VPN accounting, IP layer 3 usage measurements are collected for VRs connected by PTP tunnels. With IP accounting enabled, tunnel encapsulation and decapsulation counts are collected for PTP tunnel configurations on ATM functional processors. These statistics are categorized by CoS.

For information on configuring PTP IP tunnels, see 241-5701-810 *Passport 7400, 15000, 20000 Configuring IP*. For information on IP accounting, see “IP accounting” (page 229) and 241-5701-650 *Passport 7400, 15000, 20000 Accounting Fundamentals*.

Point-to-multipoint tunnels

The Passport IP virtual private network (VPN) service uses point-to-multipoint (PTMP) IP tunnels through a virtual connection gateway (VCG) to provide connectivity between customer VRs that reside on different Passport nodes. An IP VPN consists of multiple customer VRs, each representing a private customer VPN site. VCGs on different Passport nodes connect to each other through backbone logical connections. See the figure “VCG-based IP VPN with point-to-multipoint IP tunnels” (page 227).

Figure 49
VCG-based IP VPN with point-to-multipoint IP tunnels



For full site-to-site connectivity, the carrier must configure the source and multiple destination addresses of the PMTP tunnel on every customer VR in the IP VPN. The customer VR performs IP in IP encapsulation (as defined in RFC 2003) at the ingress. The VCG performs decapsulation at the egress. With IP VPN accounting enabled, the IP tunnel encapsulation and decapsulation counts are collected for PTMP tunnel configurations on ATM IP functional processors. These accounting statistics are further broken by CoS.

> ***Note:*** IP in IP tunnel encapsulation is not compatible with RFC 2003 when using an ATM IP FP as a backbone FP.

For more information about PTMP IP tunnels, IP VPN, and VCGs, see 241-5701-581 *Passport 7400, 15000, 20000 Basics: VPN Fundamentals* and 241-5701-582 *Passport 7400, 15000, 20000 VPN Configuration Management*.

Chapter 22
IP accounting

IP accounting allows you to collect, record and report usage measurements for each customer virtual router (VR) that is part of an IP virtual private network (VPN). The IP accounting statistics provide a breakdown of the volume of IP packets sent and received by VPN customers. IP accounting statistics are collected for three VPN configurations:

- point-to-point (PTP) tunnels
- point-to-multipoint (PTMP) tunnels
- layer 2 data connections (virtual circuits)

Accounting records are generated for each VR within these VPN configurations. VPN site-to-site information is generated for VRs within a VPN connected through IP tunnels. Local site information is available for VRs connected to the network through direct data link connections. The breakdown of the IP traffic based on tunnel usage and aggregate packet counts gives the carrier the ability to analyze the IP packets sent and received by VPN customers.

Layer 3 usage statistics are generated for VRs connected by IP tunnels. IP tunnel encapsulation and decapsulation counts are collected for PTP and PMPT tunnel configurations. If IP class of service (CoS) is applied to the customer traffic, the statistics are categorized into four possible IP traffic classifications.The CoS breakdown is generated at tunnel entry. Source and destination counts are provided at tunnel exit.

The CoS breakdown is generated at tunnel ingress and tunnel egress. There is information on tunnel source address, tunnel destination address, and packet counts per CoS.

> ***Note:*** Tunnel ingress statistics are recorded as outbound statistics for the protocol port at which the tunnel originates. Similarly, tunnel egress statistics are recorded as inbound statistics for the protocol port at which the tunnel terminates.

Network statistics are provided for layer 2 connections. These statistics are collected for AtmMpe, IP-optimized DLCI, FrDte, and PPP media. Since the VPN site address is unknown, the accounting records gathered at the outgoing traffic ports are aggregate statistics of the number of packets received and sent by the VR to the network. These statistics are broken down by CoS.

For more information about Passport accounting, see 241-5701-650 *Passport 7400, 15000, 20000 Accounting Fundamentals*.

For more information on IP accounting, see the following sections:

- “IP accounting fundamentals” (page 230)
- “Collecting records” (page 232)
- “Troubleshooting IP accounting” (page 232)

IP accounting fundamentals

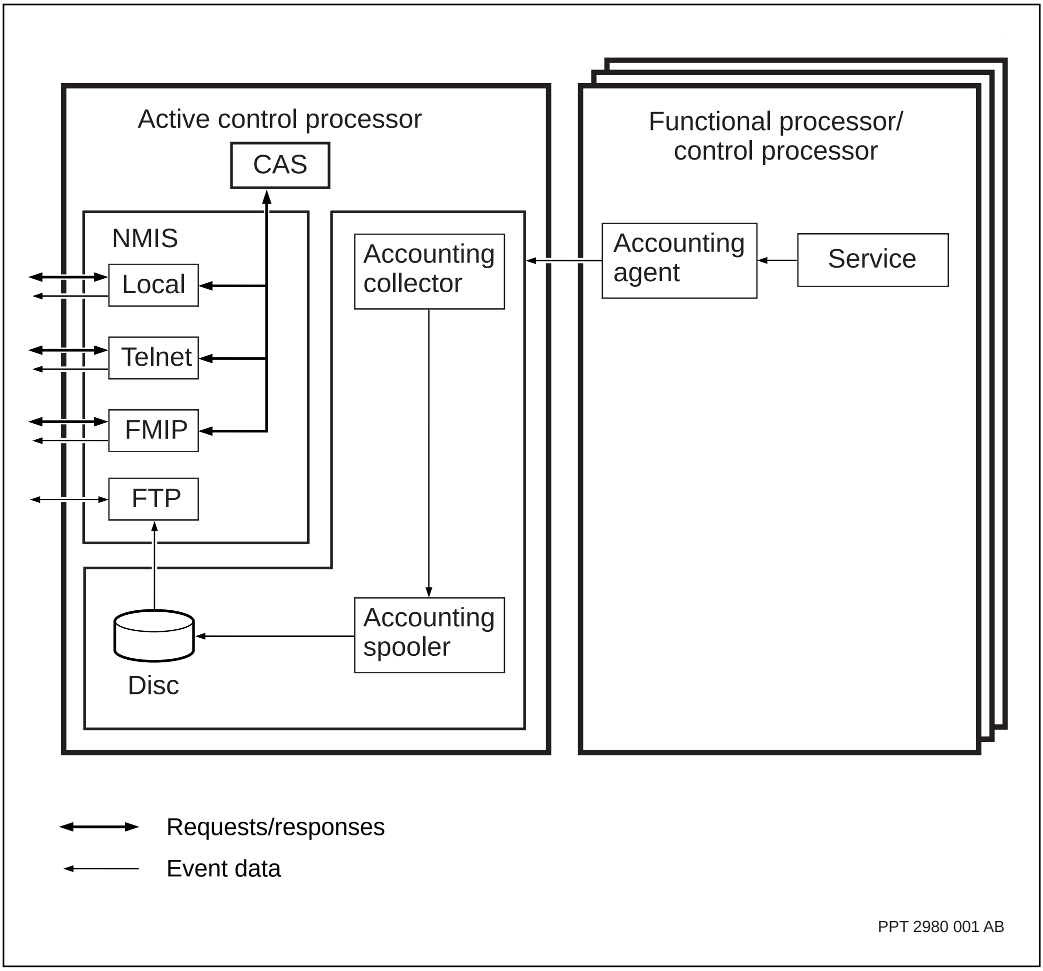
The Passport system manages the IP accounting information by controlling, recording, and reporting the usage statistics. The IP accounting system provides the statistical data to the nodal and network management system through the interactions between Preside Multiservice Data Manager and the disk on the control processor (CP).

The IP accounting information is managed by an accounting controller, located on the CP of the Passport node. The accounting controller works with Preside Multiservice Data Manager to:

- start and stop the collection of usage statistics
- identify the usage statistics collected
- create and format the accounting records

The figure “Accounting structure” (page 231) illustrates the interactions of the subsystems on the Passport node and the path followed to collect the accounting data at the protocol port to the output of the accounting record at the management data provider (MDP).

Figure 50
Accounting structure



Collecting records

The accounting record presents the usage statistics collected over a specified time. A record is generated for each protocol port, for each source address, and for each destination address for PTP and PTMP tunnels.

An IP accounting record contains the following information:

- accounting collection reasons
- static data such as VR name, source and destination address, virtual private network identitfier (VPN ID), and protocol port identity
- usage data such as collection time and duration, and packet and byte counts

For further information on accounting collection reasons see attribute *Vr Ip accountCollection* in 241-5701-060 *Passport 7400, 15000, 20000 Components*.

An accounting record is generated for each protocol port with accounting enabled. The carrier can determine the record collection time by configuring the time of day accounting (TODA) or by allowing the system to collect statistics at the default 12-hour period. For more information on TODA and accounting collection times, see 241-5701-650 *Passport 7400, 15000, 20000 Accounting Fundamentals*.

Troubleshooting IP accounting

Specific circumstances can affect the collection of the accounting records. During a CP switchover, for example, all IP VPN accounting statistics are cleared and the records for that accounting interval are lost.

The configuration of the protocol ports within the VPN determines if single-ended or double-ended accounting is possible. For layer 2 data connections, only single-ended accounting is possible. For IP tunnel configurations, if accounting is enabled at both ends of the tunnel, then two accounting records are generated for each tunnel (double-ended accounting).

Passport 7400, 15000, 20000

Understanding IP

Release 5.2

Publication: 241-5701-805
Document status: Standard
Document version: 5.2S2
Document date: December 2003
Printed in Canada