

# IP to ATM Class of Service Phase 1 Design Guide

---

## Scope and Audience

The Internet and intranet networks are migrating toward enhanced service offerings and in particular toward support for differentiated quality of service (QoS). This migration represents a key milestone in their continuing phenomenal worldwide success.

Cisco's Internetwork Operating System (Cisco IOS) software is gradually enhanced to offer more and more sophisticated QoS support at higher and higher performance levels. To extend existing IP class of service (CoS) features for operation over an ATM network, the IP to ATM CoS feature is gradually introduced in Cisco IOS software releases.

The IP to ATM CoS feature implements solutions for coarse-grained mapping of QoS between IP and ATM in Cisco router products, using ATM port adapters (PAs) that support ATM traffic shaping and have rich ATM service category support. This category of coarse-grained IP QoS is also often referred to as CoS, hence the name of the feature: IP to ATM CoS.

Phase 1 of the IP to ATM CoS feature was introduced in Cisco IOS 11.1(22)CC.

This document provides guidelines and suggestions for the deployment of the Cisco IOS IP to ATM CoS Phase 1 feature in operational networks. Another software revision is expected to cover Phase 2 of the IP to ATM CoS feature.

This document is meant for technical staff in charge of network architecture and network design. Although some background is provided on IP/Internet QoS, familiarity with IP QoS mechanisms and concepts is generally assumed; references are provided for readers wishing to obtain more information. Familiarity with ATM and ATM QoS concepts is also assumed.

## Differentiated Services in Internet/Intranets and IP to ATM CoS

### IP Differentiated Services in the Internet and Intranets

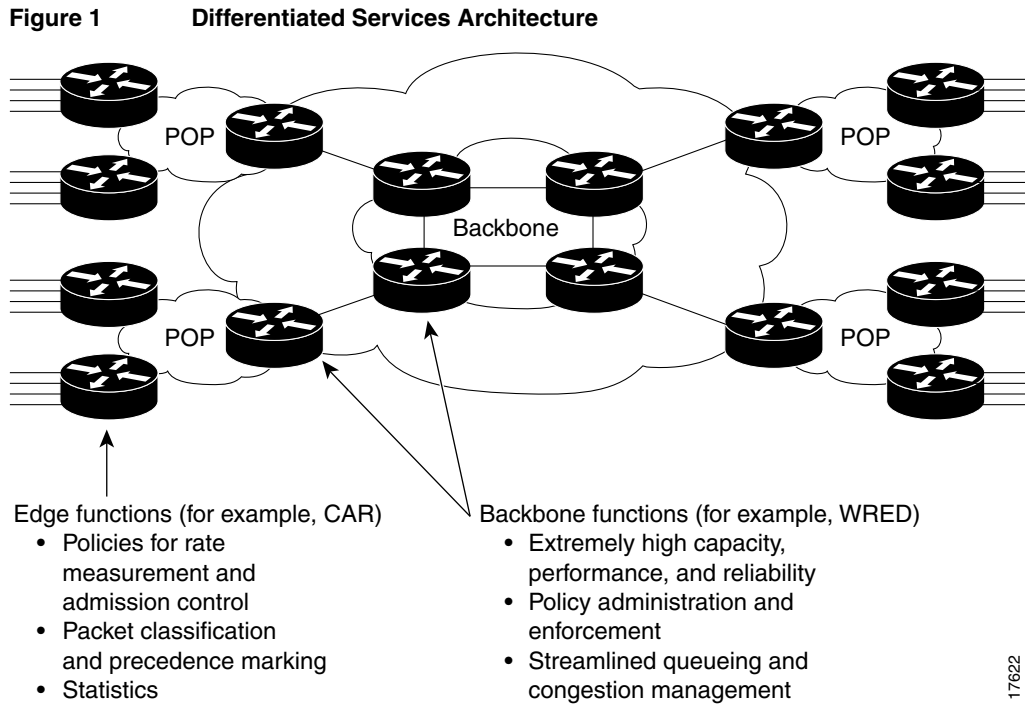
Rapid growth in Internet and intranet deployment and usage has resulted in a major shift in both corporate and consumer computing paradigms. This shift has resulted in massive increases in demand for both higher volume of communication and for predictable performance for some existing and emerging applications. However, this demand has often left Internet service providers (ISPs) or intranets with insufficient network capacities to indiscriminately offer predictable performance to all of the traffic. Thus, it became necessary to differentiate among the traffic so that the applications or users requiring (and willing to pay for) predictable performance can be identified and granted the appropriate QoS. In response to this need, Cisco introduced in Cisco IOS Release 11.1 CC new QoS extensions to Cisco IOS software network services. This set of services enables a new "Intelligent Internet" business model wherein ISPs and intranets can generate profitable

business growth by rapidly defining, deploying, and charging for differentiated services targeted at specific customer requirements including efficient handling of mission-critical, bandwidth-hungry web applications, and real-time traffic such as Voice over IP or multimedia applications.

A brief overview of the architecture for large scale deployment of differentiated services follows, but details on this architecture and on the individual IP QoS features can be found in the following documents accessible to Cisco’s customers and partners on Cisco’s Web site, Cisco Connection Online (CCO):

- *White Paper: Advanced QoS Services for the Intelligent Internet*  
[http://www.cisco.com/warp/public/732/net\\_enabled/qos\\_wp.htm](http://www.cisco.com/warp/public/732/net_enabled/qos_wp.htm)
- *White Paper: Cisco IOS(TM) Software Quality of Service Solutions*  
[http://www.cisco.com/warp/public/729/c8500/qosio\\_wp.htm](http://www.cisco.com/warp/public/729/c8500/qosio_wp.htm)
- *Cisco IOS IP QoS Features Data Sheet*  
[http://www.cisco.com/warp/customer/732/netflow/qos\\_ds.htm](http://www.cisco.com/warp/customer/732/netflow/qos_ds.htm)

Figure 1 illustrates the architecture for differentiated services.



Cisco IOS QoS features provide network operators with the means to distribute network functionality and responsibility between edge functions and backbone functions. This distribution of functionality enables simultaneous performance and services scalability for large scale support of differentiated services (also known as CoS).

At the edge of the network, ISPs gain the capability to do the following:

- Flexibly specify policies that establish traffic classes and service levels
- Flexibly specify policies that define how network resources are allocated and controlled to handle the traffic classes established
- Efficiently map packets into the traffic classes “on the fly” by marking in each IP packet the 3-bit precedence field in the ToS field of the IP header to a particular value specifying the CoS assigned to the packet

- Flexibly apply policies and “high touch” services to meet customer application and security requirements
- Flexibly collect and export detailed measurements concerning network traffic and service resource utilization

The Cisco IOS committed access rate (CAR) feature is one of the key edge features. CAR can perform packet classification, rate measurement, and enforcement as well as precedence marking, all at very high performance.

In the backbone of the network, Cisco IOS QoS and supporting technologies provide the capabilities to do the following:

- Scale the network to provide extremely high capacity, performance, and reliability
- Provide policy administration and enforcement
- Provide streamlined queueing and congestion management

Random Early Detection (RED) is a technique documented by the IETF and the Internet research community to perform congestion avoidance in an Internet backbone. RED provides network operators with the ability to flexibly specify traffic handling policies to maximize throughput under congestion conditions. RED works in conjunction with robust transport protocols (for example, Transmission Control Protocol [TCP]) to intelligently avoid network congestion by implementing algorithms that do the following:

- Distinguish between temporary traffic bursts that can be accommodated by the network and excessive offered load likely to swamp network resources. This distinction is achieved through management of a calculated average queue size, rather than a measured instantaneous queue, which provides the ability for momentary traffic bursts and fluctuations to be absorbed without necessitating a corresponding fluctuation in packet drops.
- Work cooperatively with traffic sources to avoid TCP slow start synchronization that can create periodic waves of network congestion (phenomenon referred to as “TCP Synchronization”).
- Provide bandwidth reduction through intelligent selective packet drop to reduce traffic sources in proportion to the bandwidth being utilized.

In short, RED interacts with TCP to anticipate and manage congestion during periods of heavy traffic to maximize throughput through managed packet loss.

Details on RED can be found in the following papers:

- *Random Early Detection (RED) Gateways*  
<http://www-nrg.ee.lbl.gov/floyd/red.html>
- *Random Early Detection (RED) Gateways for Congestion Avoidance*  
<http://www-nrg.ee.lbl.gov/papers/red/red.html>

Weighted Random Early Detection (WRED) is a Cisco IOS feature that combines IP Precedence and RED congestion management capabilities to provide differentiated performance characteristics for different CoS, thus providing preferential traffic handling for higher priority packets. WRED also provides preferential traffic handling under congestion conditions without exacerbating the congestion.

Thus, WRED is a key IP QoS backbone function for the deployment of IP differentiated services over the Internet or intranets.

Weighted Fair Queueing (WFQ) is another Cisco IOS feature that implements a sophisticated scheduling algorithm that can allocate a configurable share of the link bandwidth to each CoS based on the packet’s precedence. As an IP QoS backbone feature, WFQ can be used in combination with WRED, or on its own, for the deployment of IP differentiated services.

## IP to ATM CoS for IP Differentiated Services over ATM

The IP to ATM CoS feature extends support of IP differentiated services where the underlying technology for IP transport is ATM. It provides rich and flexible mapping between the IP differentiated service mechanisms into the ATM QoS mechanisms to achieve consistent end-to-end differentiated services where part (or all) of the IP network uses ATM as the underlying transport technology.

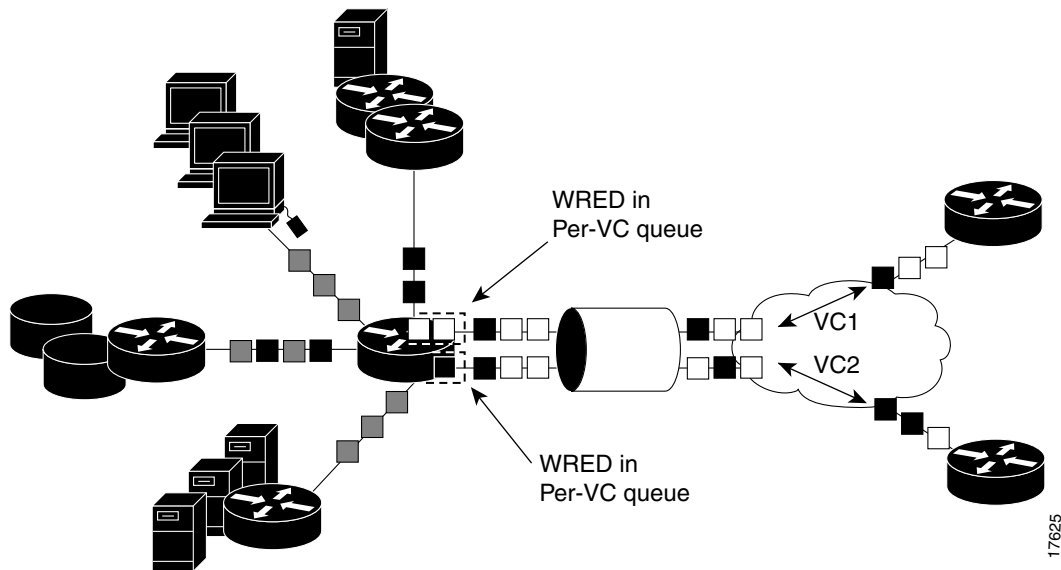
With the IP to ATM CoS feature, all congestion is pushed out of the ATM network (because the router only sends traffic out onto the ATM network at a fixed ATM shaping rate). The congestion is thus pushed back onto the routers on the edge of the ATM network that can perform IP-level service differentiation and congestion management in the form of intelligent discard according to the WRED algorithm, taking into account every IP packet's precedence.

The main characteristics of the IP to ATM CoS Phase 1 feature are as follows:

- It operates on the Cisco 7500 series using the ATM-PA-A3 port adapters plugged into VIP2-50 interface processors.
- It uses a single ATM permanent virtual circuit (PVC) between any pair of ATM-connected routers.
- The ATM-PA-A3 performs ATM traffic shaping to strictly comply with constant bit rate (CBR) or variable bit rate (VBR) ATM Traffic Contract so that the ATM network can ensure that virtually no ATM cells are dropped on the ATM CBR or VBR PVC.
- Per-VC queuing is maintained in the router so that when congestion occurs an IP packet queue will develop for the congested VC.
- The router performs congestion management and service differentiation by implementing the WRED algorithm independently over each per-VC queue.

Figure 2 illustrates IP to ATM CoS Phase 1 operations.

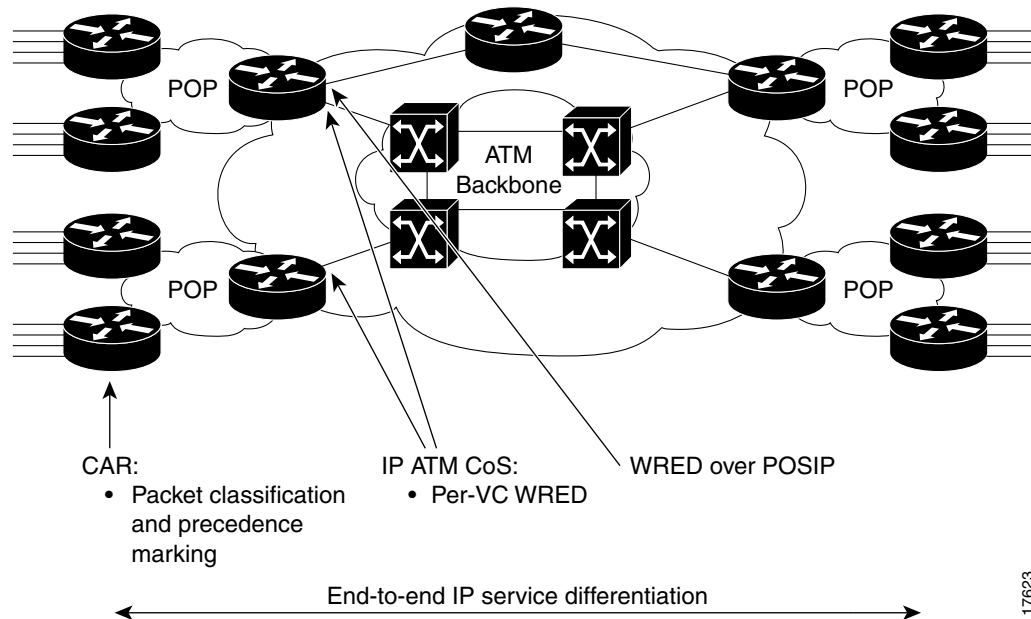
**Figure 2 IP to ATM CoS Phase 1 Operations**



Because WRED is also supported by Cisco IOS software on technologies other than ATM such as POSIP, serial, Ethernet, Fast Ethernet, FDDI, and so forth, a consistent service differentiation can be deployed end-to-end by Internet/intranet service providers across heterogeneous networks comprising any mix of ATM and other transport technologies.

Figure 3 illustrates where IP to ATM CoS Phase 1 belongs in an end-to-end differentiated service deployed over a heterogeneous network.

**Figure 3 IP to ATM CoS Phase 1 Operations in an End-to-End Differentiated Service**



## Looking Beyond IP to ATM CoS Phase 1

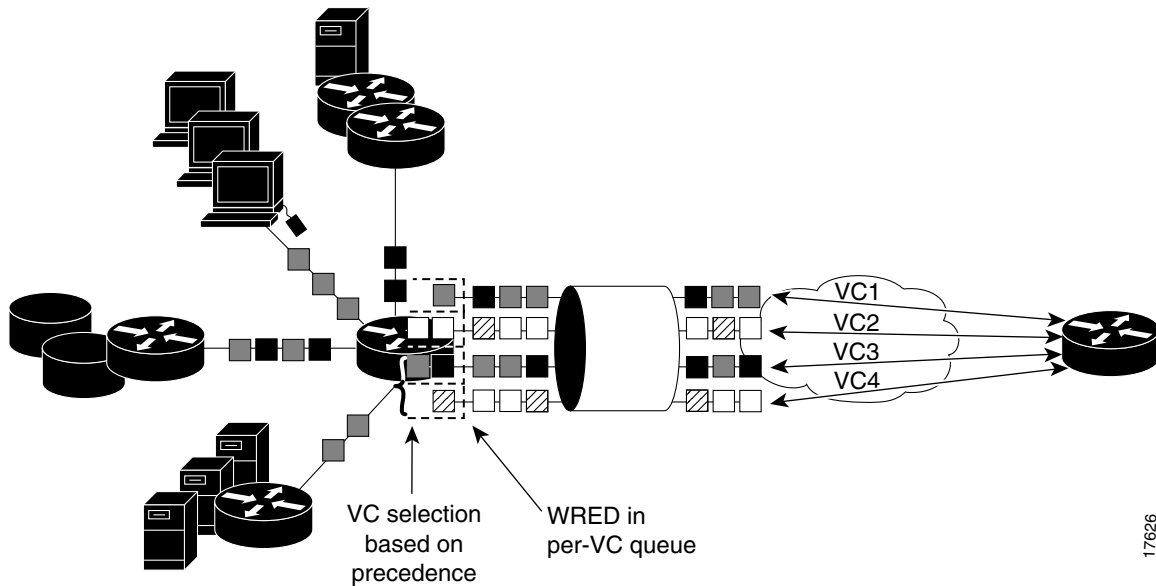
### IP to ATM CoS Phase 2

Phase 2 of IP to ATM CoS will expand the IP to ATM CoS feature to add the following:

- Operation over Cisco 7200 series routers in addition to Cisco 7500 series routers
- Support for multiple ATM PVCs (referred to as a VC bundle) between any pair of ATM-connected routers, each PVC of a bundle having its own ATM traffic class (including available bit rate [ABR]) and ATM traffic parameters
- Service differentiation by flexible distribution of the IP Precedence levels over the different VCs of a VC bundle
- Service differentiation across the precedences transported over the same PVC from a bundle through WRED
- Flexible management of the VC bundle on PVC failure

Figure 4 illustrates IP to ATM CoS Phase 2 operations.

Figure 4 IP to ATM CoS Phase 2 Operations



Through the support of multiple parallel ATM VCs as a VC bundle, Phase 2 of IP to ATM CoS will allow the service provider, where appropriate, to take more advantage of the ATM QoS to provide stronger service differentiation at the IP layer. For instance, the service provider may want to transport the IP traffic belonging to real-time CoS (such as Voice over IP traffic) on an ATM VC with very strict time constraints such as a CBR or VBR-rt PVC, while transporting the nonreal-time traffic over an elastic ATM ABR PVC, enabling the service provider to efficiently make use of the spare capacity in the ATM network. The service provider could also elect to transport the IP traffic for which no commitment is provided (“best-effort” traffic) over an unspecified bit rate (UBR) PVC, the UBR traffic class being effectively the ATM version of a best-effort service.

### 12000 IP to ATM CoS

The IP to ATM CoS feature is planned to also be available on the Cisco 12000 series.

### IETF Diff-Serv

The IETF has created a working group (Diff-Serv) that is standardizing the concept of IP differentiated services.

The charter of this working group and all the relevant Internet drafts can be found on the IETF Web site: <http://www.ietf.org/html.charters/diffserv-charter.html>.

The architecture and concepts standardized by Diff-Serv are identical to those supported by Cisco IOS software, but Cisco IOS software is expected to align with any detailed specifications from Diff-Serv that are ratified. Mapping of Diff-Serv over ATM will be maintained by alignment of the IP to ATM CoS feature for support of standardized differentiated services.

### Tag Switching/MPLS Differentiated Services

Tag switching is a Cisco technology marrying Layer 2 connection-oriented concepts with Layer 3 connectionless concepts and is being standardized by the IETF as MultiProtocol Label Switching (MPLS).

Tag switching/MPLS offers the following benefits:

- Enhanced services (for example, IP VPNs, traffic engineering, QoS, multicast, and so forth)
- Flexibility (for example, independence of underlying technology such as ATM, SDH, fiber, Gigabit Ethernet, introduction of new services and new routing paradigms, and independence of network protocols IPv4, IPv6)
- Scalability (hierarchy of routing, removal of VC meshing and peering issue)
- Investment protection (operates over existing routers and switches)

IP differentiated services are also supported by tag switching/MPLS in a consistent way. The same differentiated architecture applies with packet classification, measurement, and marking being performed on the edge while congestion control and queue management are performed at high performance in the backbone. The main difference is that tag switching first maps the IP Precedence of IP packets into the tag/MPLS CoS field of the tag/MPLS packet and then performs congestion avoidance and queue management based on the tag/MPLS CoS field.

Thus, as large networks migrate to tag switching/MPLS, the IP differentiated services can be very easily maintained and migrated from the current network environment to a full tag/MPLS environment or to a hybrid environment.

## Detailed IP to ATM CoS Phase 1 Operations

### IP to ATM CoS Phase 1 Operations

To support IP differentiated services using IP Precedence signalling in the parts of the network where ATM is the underlying transport technology, the IP to ATM CoS Phase 1 feature relies on the ATM layer QoS to offer a lossless transport of cells through the ATM network. The feature also relies on the router's awareness of high level priority through the IP Precedence field to perform the actual service differentiation exclusively at the IP layer on the edge of the ATM network.

The router uses ATM PVCs as virtual lossless pipes and performs service differentiation at the point of enqueueing packets onto the lossless ATM PVCs. To enable the ATM network to offer lossless ATM PVCs, the IP to ATM CoS Phase 1 feature makes use of ATM PVCs that are individually shaped by the PA-A3 according to their ATM traffic class of CBR, VBR-rt, VBR-nrt, and ATM traffic parameters. The ATM network is then required to offer a virtually lossless service on these shaped VCs.

---

**Note** The IP to ATM CoS Phase 1 feature can also be configured to run on a UBR PVC. However, because lossless operation is required from the ATM network (because the ATM network is not aware of the various classes of differentiated service and would perform indiscriminate loss), configuring IP to ATM CoS on a UBR PVC would only be useful in the very particular case where the ATM network is so underutilized that UBR connections can be assumed to be lossless. If that was the case, WRED should still be activated on the UBR PVC because it is possible that a queue may build up in the router on a UBR PVC.

The IP to ATM CoS feature will also make use of ABR PVCs when ABR PVCs are supported by the PA-A3. When ABR PVCs are supported by the PA-A3, the ATM switches would have the appropriate buffering and the ABR algorithm so that the ABR PVCs would offer negligible ATM loss.

---

---

**Note** If the ATM network could not offer lossless transport despite the ATM shaping on the router, the operation of the IP to ATM CoS feature would be compromised for two reasons. First, because the ATM layer is unaware of the IP Precedence and therefore of the “priority” of the transported packets, ATM cell loss would be indiscriminate and thus would compromise the desired service differentiation. Second, the desired sophisticated dynamic interaction of WRED and TCP for proper congestion avoidance would be affected, which could prevent WRED benefits such as avoiding TCP synchronization and global oscillation.

---

In situations of congestion, the amount of IP traffic received would be superior to the shaping rate of the supporting VC. Consequently, packets for the congested VC would accumulate in the router. The IP-ATM CoS Phase 1 feature maintains a separate logical queue for each VC (that is, a per-VC queue) so that traffic is backed up independently on every VC based on its own level of congestion. Inside each of these per-VC queues, traffic from all the classes of service would thus be competing for the underlying ATM PVC bandwidth. To implement service differentiation, the IP to ATM CoS Phase 1 feature implements WRED independently on each of the per-VC output queues.

The IP to ATM CoS Phase 1 feature is implemented on the Cisco 7500 series routers as a distributed service on the output interface, which means that the IP to ATM CoS feature is actually performed jointly by the VIP and the PA-A3 supporting the output ATM interface without involvement of the Cisco 7500 Route Switch Processor (RSP). All the IP packets destined to the ATM interface are switched normally by input VIP interface processors (or by the RSP if the input interface is supported by an interface processor, for example, xIP, which does not support distributed switching) and are all forwarded onto the output VIP supporting the ATM interface regardless of their CoS. The output VIP and PA-A3 will then jointly enforce congestion management and service differentiation over all the ATM PVCs.

The VIP and the PA-A3 collaborate in the following ways:

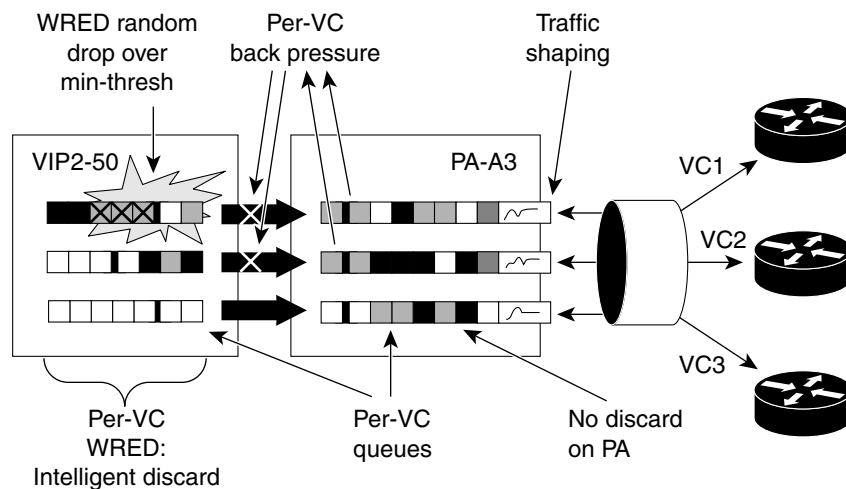
- The PA-A3 transmits ATM cells on each ATM PVC according to the ATM shaping rate.
- The PA-A3 maintains a per-VC first-in, first-out (FIFO) queue for each VC where it stores the packets waiting for transmission onto that VC.
- The PA-A3 provides explicit back pressure to the VIP so that the VIP only transmits packets to the PA when the PA has sufficient buffers available to store the packets, which ensures that the PA-A3 will never need to discard any packets regardless of the level of congestion on the ATM PVCs. When the VIP has packets to transfer to the PA-A3 but is throttled by the PA-A3 back pressure, the VIP will store the packets into per-VC queues (that is, one logical queue for each ATM PVC configured on the ATM interface). The per-VC queue is a FIFO queue storing all packets, in order of arrival, that are to be transmitted onto the corresponding VC.
- The PA-A3 provides back pressure both at the VC level and at the aggregate all-VC level. The per-VC back pressure notifies the VIP to stop sending to the PA-A3 packets destined for that particular VC because the per-VC queue on the PA for that VC has reached a certain occupancy level. The VIP then stops forwarding to the PA any packets destined to that ATM PVC but keeps forwarding to the PA packets destined to other ATM PVCs. The aggregate all-VC back pressure notifies the VIP to stop sending any packets to the PA-A3 when the total number of packets buffered on the PA-A3 for all ATM PVCs reaches a certain level compared to the total available buffering space. The per-VC back pressure prevents unnecessary overconsumption of buffers by a single ATM PVC while the aggregate all-VC back pressure allows dynamic sharing of the PA-A3 buffers across all VCs.
- The VIP then monitors the level of congestion independently on each of its per-VC queues and performs a WRED selective congestion avoidance algorithm independently on each of these queues that enforces service differentiation across the IP classes of service. For each instance of

the per-VC WRED algorithm, the IP to ATM CoS Phase 1 feature computes a separate moving average queue occupancy (expressed in number of packets and taking into account packets of all precedences) and supports a separate set of configurable WRED drop profiles with one profile per precedence.

In summary, the ATM layer functions such as ATM shaping are handled by the PA-A3, while the IP-level service differentiation is performed by the VIP. Through explicit back pressure from the PA to the VIP, the PA operates in a lossless environment and all congestion management and selective drops are performed on the VIP.

Figure 5 illustrates IP to ATM CoS Phase 1 detailed operations.

**Figure 5 IP to ATM CoS Phase 1 Detailed Operations**



The PA-A3 has a pool of buffers on board dedicated to the storage of packets in the PA-A3 per-VC queues.

The VIP2-50 also has a separate pool of buffers dedicated to storage of packets in the VIP2-50 per-VC queues. The number of buffers available for per-VC queueing on the VIP2-50 depends on the amount of SRAM installed on the VIP. With 8 MB of SRAM on board, up to 1085 packets worth of buffers may be available to the IP to ATM CoS feature for per-VC queueing. It must be noted that a per-VC queue will only develop on the VIP for the ATM PVCs on which there is temporary congestion (that is, there is more incoming IP traffic than the egress ATM shaping rate of the corresponding ATM PVC) and that this queue will only remain on the VIP for the duration of the burst.

---

**Note** The amount of SRAM that can be used by the IP to ATM CoS feature for per-VC queueing over the PA-A3 depends on whether another PA is supported on the same VIP2-50. The amount of 1085 packets worth of buffers assumes there is no PA other than the PA-A3 on the VIP2-50.

Note also that the fraction of VIP SRAM that can be used by the Layer 3 services (such as the IP to ATM CoS feature for per-VC buffering) is expected to be increased, which will result in a higher number of available buffers for the same SRAM configuration on the VIP.

---

In addition to ensuring that no drop is performed by the PA and, therefore, that all drops can be subject to WRED's selective intelligent discard, implementation of per-VC queueing both on the PA-A3 and on the VIP ensures a strict isolation of ATM PVCs from the congestion viewpoint. In other words, congestion experienced by one ATM PVC will not spill over other uncongested ATM PVCs, so they will not experience induced packet loss.

---

**Note** When the IP to ATM CoS feature is not activated on an ATM PVC of the ATM interface supported by the VIP2/PA-A3 (that is, WRED was not activated on the PVC by the **random-detect group-name** command), the PA-A3 still performs per-VC queueing. However, the PA-A3 does not provide any back pressure to the VIP for that PVC and the VIP does not perform buffering for that particular PVC. Consequently, when the IP to ATM CoS feature is not used, ATM PVC isolation from the congestion viewpoint is also guaranteed by the PA-A3. However, it must be noted that in this case, the PA-A3 performs indiscriminate discard across IP classes of service and the buffers available on the VIP cannot be used to absorb longer traffic burst.

---

The IP to ATM CoS Phase 1 solution has the following characteristics:

- It is very simple and has minimal operational overhead. The number of ATM PVCs to be created on the ATM network and configured on the routers is kept to the strict minimum (that is, one per pair of ATM-connected routers).
- It is very efficient from the viewpoint of bandwidth allocation and bandwidth sharing. Because a single ATM PVC is shared by all the precedences/CoS, the ATM traffic parameters (for example, SCR/PCR) to be provisioned between any two routers are determined and allocated globally across all the CoS (that is, for all traffic) between the pair of routers. Sharing of a single ATM PVC across all the precedences/CoS allows simpler bandwidth allocation on a global basis rather than on a per-CoS basis or on a per-set-of CoS basis (as would be the case when using multiple parallel ATM PVCs between one pair of routers, as will be supported in Phase 2 of IP to ATM CoS).
- It ensures consistent relative service differentiation regardless of the current relative traffic load for each CoS compared to the expected load. Even if there is less low priority traffic than expected and more high priority than expected, because per-VC WRED operates on a moving average calculated across all the precedences, the high priority traffic will always experience better service (that is, lower drop probability and thus less throttling) than the low priority traffic. With other forms of mapping involving separate bandwidth allocation to the high and low priority traffic (such as would be the case when using multiple parallel ATM PVCs between one pair of routers, as will be supported in Phase 2), then if there is actually less low priority traffic than expected and more high priority traffic than expected, the relative service differentiation may be distorted.
- It does not introduce any packet reordering because the service differentiation relies on WRED performed on a per-VC FIFO queue common to all the precedence levels and because a single ATM PVC is used per next hop.
- Because the service differentiation relies on WRED performed on a per-VC FIFO queue common to all the precedence levels, the different service classes will experience different drop probabilities, but for nondiscarded packets all the CoS will experience the same delay characteristics. The IP to ATM CoS Phase 1 feature does not provide delay differentiation across CoS, which also means that the PVC ATM traffic class must be selected to satisfy the most demanding CoS. For instance, if one CoS is to be used for transport of delay-sensitive traffic, then the ATM PVC used should be CBR or VBR-rt. By contrast, the IP to ATM CoS Phase 2 feature will allow use of a CBR/VBR-rt PVC for the delay-sensitive CoS only while using a

VBR-nrt/ABR/UBR PVC for the rest of the traffic. This would provide stronger delay differentiation across the two types of classes while taking advantage of high ATM statistical gain for nondelay-sensitive traffic.

- Because the service differentiation relies on WRED, the network operator cannot directly control by configuration the exact share of bandwidth to be allocated to each competing CoS in times of congestion. The IP to ATM CoS Phase 1 service differentiation is not targeted at providing strict bandwidth apportionment across classes but rather at providing different drop probabilities across classes. Note that the combination of this differentiated drop probability and TCP's adaptive behavior on packet loss would indirectly result in a TCP flow of higher CoS enjoying a throughput during congestion than a TCP flow from lower CoS.
- Because of the required interaction between WRED selective IP discard and TCP flow control, the IP to ATM CoS Phase 1 service differentiation assumes, as is the case in the Internet and typical intranets, that the majority of traffic is TCP with IETF-compliant adaptive behavior. The IP to ATM CoS service differentiation effectiveness would be compromised in environments where a significant part of the traffic is nonadaptive to loss (such as User Datagram Protocol [UDP]) or where the TCP flow control implemented by traffic sources departs significantly from the IETF-specified behavior.

### Activating WRED on an ATM PVC

Details on the command-line interface (CLI) for activation and configuration of the IP to ATM CoS Phase 1 feature can be found in the feature module for the IP to ATM CoS feature, which is accessible on CCO at:

<http://www.cisco.com/univered/cc/td/doc/product/software/ios111/cc111/ipatmcs1.htm>

With IP to ATM CoS Phase 1, a WRED algorithm can be activated and run independently on each of the VIP-based per-VC queues. The WRED algorithm is activated on a particular PVC by associating a random-detect group to a particular PVC in the **atm pvc** command:

```
atm pvc vcd vpi vci aal-encap [[midlow midhigh] [peak average [burst]] [inarp [minutes]] [oam [seconds] [random-detect [group-name]]]
```

The random-detect group specifies the actual parameters of the WRED algorithm to be performed on the particular PVC. Where the same WRED drop profiles and parameters are to be used by multiple ATM PVCs, they can be associated with the same random-detect group, which facilitates configuration by requiring configuration of the WRED parameter only once, regardless of how many ATM PVCs will be using that same set of parameters.

Of course, whether ATM PVCs are associated with the same random-detect group or not, a separate instance of WRED algorithm always is running independently for each VC, with its own set of variables. In particular, a moving average queue size is computed independently for each per-VC queue on the VIP even when multiple VCs are associated with the same random-detect group. Note, however, that for each per-VC queue a single moving average is computed across all precedence values (that is, taking into account arrival of packets of any CoS).

### Configuring WRED Parameters

The service differentiation is controlled by configuring different WRED drop probability parameters for each precedence for each random-detect group.

To configure the WRED parameters of a random-detect group, use the **random-detect group** command to enter random-detect group configuration mode:

```
random-detect group group-name
```

Subcommands of the random-detect group configuration mode can then be used to configure the WRED parameters for that particular random-detect group:

- To modify the exponential weight factor for calculating the moving average queue length, use the following subcommand:

**exponential weighting constant**

- To modify the packet drop probability parameters individually for each precedence value, use the following subcommand:

**precedence** *precedence min-threshold max-threshold mark-probability-denominator*

A WRED configuration that provisions all precedences with the same thresholds and mark-probability-denominator will display an egalitarian drop behavior and will be in fact performing a standard RED congestion management algorithm. It is only by varying the drop probability parameters of a certain precedence against others that a corresponding priority for that precedence traffic can be provided.

## Switching Mode

The IP to ATM CoS feature is supported by Cisco Express Forwarding (CEF) switching mode and operates in a distributed mode only. Thus Distributed CEF (DCEF) must be enabled on the ATM interface running the IP to ATM CoS feature. DCEF is enabled using the **ip cef distributed** command in global configuration mode.

## VC Type

The IP to ATM CoS feature supports ATM PVCs. ATM switched virtual circuits (SVCs) are not supported.

## Supported ATM Traffic Classes

The IP to ATM CoS feature supports all the ATM traffic classes supported by the PA-A3.

With IP to ATM CoS Phase 1, the PA-A3 supports ATM shaping on ATM PVCs of the following ATM traffic classes:

- CBR
- VBR-rt
- VBR-nrt
- UBR

---

**Note** The IP to ATM CoS feature will also make use of ABR PVCs when ABR PVCs are supported by the PA-A3. When ABR PVCs are supported by the PA-A3, the ATM switches would have the appropriate buffering and the ABR algorithm so that the ABR PVCs would offer negligible ATM loss.

---

ATM shaping on a CBR ATM PVC is achieved on the PA-A3 by configuring the average rate and the peak rate to the same bandwidth when the **atm pvc** command is issued. Configuring them for the same bandwidth will result in an ATM shaping with constant intercell gap, which is characteristic of CBR shaping.

The IP to ATM CoS Phase 1 feature can also be configured to run on a UBR PVC. However, because lossless operation is required from the ATM network (because the ATM network is not aware of the various classes of differentiated service and would perform indiscriminate loss), configuring IP to ATM CoS on a UBR PVC would only be useful in the very particular case where the ATM network is so underutilized that UBR connections can be assumed to be lossless. If that was the case, WRED should still be activated on the UBR PVC because it is possible that a queue may build up in the router on a UBR PVC.

## ATM Shaping

The PA-A3 supports very accurate ATM traffic shaping in accordance with the Generic Cell Rate Algorithm (GCRA) documented in the ATM Forum Traffic Management Specification 4.0. It supports the following shaping parameters:

- SCR = sustainable cell rate, configured by the *average* CLI parameter
- PCR = peak cell rate, configured by the *peak* CLI parameter
- MBS = maximum burst size, configured by the *burst* CLI parameter (in units of 32 cells)

The granularity of the PA-A3 ATM traffic shaping (the difference between the two closest consecutive achievable effective shaping rates) is less than 0.02 percent for PCR around 10 Mbps and is less than 0.25 percent for PCR around 100 Mbps.

The PA-A3 ATM traffic shaping accuracy (the difference between the configured *average* rate and the measured effective SCR) is less than 0.3 percent for SCR rates around 100 Mbps.

It must be noted that the difference between the effective shaping rate and the configured shaping rate may be either positive or negative. In other words, depending on the configured shaping rate, the effective shaping rate actually achieved by the PA-A3 will in some cases be slightly lower than the configured rate and in some cases be slightly higher than the configured rate.

In the cases where the effective shaping rate is higher than the configured rate, it is essential that the ATM PVC be configured throughout the ATM network with a PCR/SCR that is at least as high as the effective shaping rate, rather than just the configured rate. In particular, if the ingress ATM switch performs an ATM policing function, this policing must be configured to accept traffic up to the effective shaping rate of the PA-A3; failure to do so could result in systematic dropping of cells by the policing function, which would compromise the IP to ATM CoS operation that relies on lossless ATM PVCs. Similarly, the ATM switches must allocate a bandwidth at every hop that corresponds to the PA-A3 effective shaping rate; failure to do so could result in ATM switches dropping cells on the ATM PVC because, for example, their per-VC queue would be growing due to an incoming cell rate higher than expected. Such behavior would again compromise the IP to ATM CoS operation. It is thus recommended that values for the ATM PVC traffic parameters (SCR/PCR) for policing, scheduling, and resource allocation be configured on the ATM switches that are slightly higher than the corresponding values configured on the router for shaping; this safety margin will enable the ATM network to guarantee lossless operation at the effective cell rate on the ATM PVC.

## Subinterface Type

The IP to ATM CoS feature supports both multipoint and point-to-point subinterfaces.

## Encapsulations

The IP to ATM CoS Phase 1 feature supports aal5snap and aal5mux encapsulations.

## Address Resolution

The IP to ATM CoS Phase 1 feature supports both static map (with the **map-group** and **map-list** commands) and ATM Inverse ARP (with the **inarp** keyword in the **atm pvc** command) for the next hop protocol address resolution.

When the **atm pvc** command is issued with the **inarp** keyword, Cisco IOS software will send an InARP request as per RFC 1293 for the protocols configured in the interface. It will also respond to InARP requests from the remote end if that protocol is configured on that interface. The requests are responded to only if the request is from the same network as the router and all other requests are silently discarded. The InARP mappings are aged out and periodic requests are issued to refresh the map table.

## Monitoring of PVC Status

The ATM OAM features supported in Cisco IOS software can be used for monitoring the status of the PVC when deploying the IP to ATM CoS feature.

In conjunction with the IP to ATM CoS Phase 1 feature, Cisco IOS software supports the following ATM OAM features for monitoring the status of ATM PVCs:

- The router can detect reception of AIS and RDI OAM.
- The router will generate RDI on reception of AIS.
- The router can generate OAM F5 loopback cells regularly if the **atm pvc** command is issued with the **oam** keyword. If OAM response cells are missed (indicating the lack of connectivity), the system console displays a debug message indicating the failure of the PVC, provided the **debug atm errors** command is enabled. If you suspect that a PVC is faulty, enabling OAM cell generation and the **debug atm errors** command allows you to monitor the status of the PVC on the console display. Note that when it detects missing OAM response cells the router will not turn the ATM PVC status to inactive.

---

**Note** A Cisco IOS feature called PVC Management was introduced in Cisco IOS Release 11.3 T and is carried through in Cisco IOS Release 12.0 where the IP to ATM CoS Phase 2 feature was introduced. The PVC Management feature includes the ability to perform a full continuity check by permanently controlling the return of loopback OAM cells and using this information to automatically determine if an ATM PVC is active or inactive.

The PVC Management feature with the IP to ATM CoS Phase 2 feature will provide immediate notification to the routing layer of any change in the PVC status and whatever triggered the PVC status change, such as detection of AIS/RDI or continuity checking through OAM loopback cells. Immediate notification will result in optimum rerouting speed in case of PVC failure.

---

Detection of AIS or RDI by the router will result in the ATM PVC being recognized as inactive by the ATM layer. However, this information is not explicitly communicated to the routing layer. Consequently rerouting will take place through detection of PVC failure by the routing protocol's polling/keepalive mechanisms. With the IP to ATM CoS Phase 1 feature, rerouting time on ATM PVC failure is thus determined entirely by the routing protocol.

## WRED and Non-IP Traffic

With the IP to ATM CoS Phase 1 feature, non-IP packets are subjected to the WRED algorithm and treated exactly like IP packets with precedence 0. In other words, non-IP packets are taken into account when computing the WRED average queue length and a drop decision is made for every non-IP packet using the WRED drop profile of IP precedence 0 traffic (that is, precedence 0's min-threshold, max-threshold, and mark-probability-denominator).

It must be noted that when the WRED configuration is set up for decreasing drop probabilities as the precedence increases, non-IP traffic will be experiencing the lowest QoS enforced by WRED. In other words, non-IP traffic will be as likely to be dropped as IP best-effort traffic and is more likely to be dropped than any IP packet with non-0 precedence.

If non-IP traffic is set up for a QoS different than IP's best-effort traffic (for example, if the non-IP traffic experiences a higher QoS than IP best-effort), then IP best-effort traffic could be transported on the IP backbone using a precedence value different from 0 (for example, precedence 2). The WRED discard profile could then be configured and tuned separately for non-IP traffic (by configuring precedence 0's WRED parameters) and for IP best-effort (by configuring precedence 2's WRED parameters). Note also that even if the default configuration of WRED defines decreasing drop probability profiles as the precedence increases, the drop profile parameters of each precedence can be configured without any relative order so that precedence 0's discard profile can be configured to offer the highest QoS if this is deemed desirable for non-IP traffic.

---

**Note** In order for IP best-effort traffic to be transported on the IP backbone, a consistent policy for use of precedence values for different types of traffic throughout the network is required. In particular, precedence marking of IP traffic on the edge of the network (for example, marking of incoming IP Precedence through the Cisco IOS CAR feature or through policy routing) should be configured accordingly; in this example, incoming best-effort traffic would be marked with precedence 2 on every device performing packet classification and marking on the edge of the network.

---

## WRED and Multicast/Broadcast

The IP to ATM CoS feature is compatible with the Cisco IOS pseudobroadcasting that performs router-based replications of multicast/broadcast packets over ATM PVCs that have been declared as part of the pseudobroadcast domain by the **broadcast** keyword in the **map-list** statement.

The IP to ATM CoS Phase 1 feature enforces the same precedence-based service differentiation on IP multicast/broadcast packets as on IP unicast packets. An IP multicast/broadcast packet is treated by per-VC WRED exactly like a unicast IP packet of the same precedence.

Non-IP multicast packets are subject to per-VC WRED and subject to the IP precedence 0 discard profile.

## Protection of Routing and Network Control Protocols

Ensuring integrity of routing protocol operations when deploying differentiated services over ATM is of paramount importance. It is not very useful for a router to allocate all resources to important data traffic if this allocation results in loss of important routing traffic and thus affects routing stability. Consequently, the IP to ATM CoS scheduling and discarding mechanisms implemented to prioritize important data make special provisions to guarantee that routing protocols have appropriate access to resources in order to maintain routing operations while still offering the required QoS to important data traffic.

The following sections describe several mechanisms that cooperate with the IP to ATM CoS Phase 1 feature to achieve very strong protection of routing protocols while offering IP differentiated services:

- Per-VC queueing and back pressure
- Precedence marking and selective discard
- WRED/discard-bypass for critical routing packets
- Selective packet discard on input

A case study also follows that illustrates a design for protection of routing protocols in a congested network.

The same mechanisms are also used to protect a number of network control messages that are necessary for network stability and integrity (such as ATM Inverse ARP).

### Per-VC Queueing and Back Pressure

Per-VC queueing and back pressure mechanisms were described earlier. In the context of routing protection, per-VC queueing ensures that a congested VC does not result in a shortage of buffers on another VC, therefore preventing the congestion from spilling across VCs; routing traffic in uncongested VCs cannot be subject to drop because of congestion on another VC. Back pressure ensures that the VIP is notified of an increase or decrease of output VC queues on the PA before overflow or underflow occurs in that output VC queue. This means that no indiscriminate drop will occur on the PA-A3 and that the WRED selective discard mechanism running on the VIP can take into account the fact that a packet may be an important routing packet and thus should not be discarded.

This mechanism applies to IP and to non-IP routing protocols.

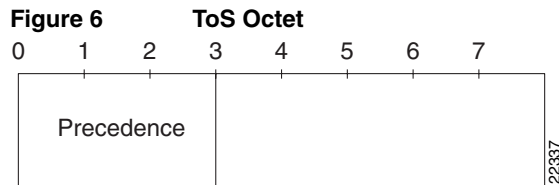
### Precedence Marking and Selective Discard

As per the IP specification RFC 791, routing and network control packets may be flagged in the network to facilitate preferential treatment of those packets by routers. The packets are flagged by setting the Precedence field inside the IP type of service (ToS) octet to particular values (precedences 6 and 7). Figure 6 shows the placement of the Precedence field in the ToS octet and lists the corresponding precedence values. The network control and internetwork control precedences (7 and 6) are used in many networks running on Cisco routers, allowing easy identification of routing traffic so that it can be protected (for example, by Cisco IOS output scheduling mechanisms or Cisco IOS selective packet discard), which contributes to network stability in case of intense route fluctuations.

---

**Note** RFC 1349, *Type of Service in the Internet Protocol Suite*, updates RFC 791 with respect to specification of the ToS field coded in bit 3-6 of the ToS octet. It does not modify RFC 791 with respect to the Precedence field.

---



Precedence  
 111- Network Control  
 110- Internetwork Control  
 101- Critic/ECP  
 100- Flash Override  
 011- Flash  
 010- Immediate  
 001- Priority  
 000- Routine

Routing packets generated by Cisco routers are marked with precedence 6. At the time these packets are enqueued, and consequently have gone through the WRED algorithm, the routing packets will experience a selective discard with very low loss probability (based on the WRED parameters for precedence 6).

We recommend you not use precedence 6 and 7 for any data traffic (that is, data traffic should only use precedences 0 to 5) so that routing traffic can always be easily identified and treated accordingly.

In the case where the network administrator modifies the WRED parameters from their default values for an ATM VC, we strongly recommend you ensure that the WRED parameters yield a very low loss probability for packets with precedence 6 and 7.

This mechanism applies to IP routing protocols only.

### WRED/Discard-Bypass for Critical Routing Packets

The IP and non-IP routing protocol packets generated by the router and that are most critical to the stability of the routing protocol (such as link-state peer/adjacency keepalives/hellos) are flagged inside the router. Those packets are then treated with extreme care at any buffering stage and given the highest priority in terms of discard. In particular, these packets bypass the WRED discard algorithm and are never dropped. Consequently, even if the WRED parameters were not properly tuned to the network environment or if the IP traffic was not backing off as expected in response to the WRED discards, the critical routing packets would still be transmitted and the routing protocol would stay up.

This mechanism applies to IP routing protocols and non-IP routing.

### Selective Packet Discard on Input

The three mechanisms just described for protection of routing traffic ensure that routing packets are protected from discard on output from the router toward the ATM network. Cisco IOS software also includes a feature called selective packet discard (SPD) ensuring protection of routing on input on a router. When enabled, the SPD feature ensures that, in situations where the router is receiving more packets than it can process, the router will selectively discard nonrouting packets and protect incoming routing packets to guarantee integrity of routing exchange and stability of the routing protocol.

This feature is particularly useful when cache-based switching modes are in use on the router because in case of major topology changes in the network, cache invalidation would result in a surge of packets requiring process switching, which in turn increases CPU load and packet backlog on input for processing if packets arrive faster than the CPU can handle them. In these particular

situations, input congestion could result in lost routing packets. Because the IP to ATM CoS feature operates with CEF/DCEF switching, topology changes do not trigger cache invalidation and thus do not result in CPU load increase nor in input congestion. SPD is thus not expected to be required when the IP to ATM CoS feature is used.

However, SPD might still be useful to protect routing protocol integrity in very particular situations where the router CPU is heavily loaded (for instance because of a malicious attack whereby a large amount of traffic is addressed to the router CPU itself). In those situations, SPD would ensure that routing packets arriving on the router would not be discarded regardless of the volume of incoming traffic addressed to the router.

## VIP Requirements

The IP to ATM CoS Phase 1 feature requires a VIP2-50 model for the VIP interface processor.

Because the IP to ATM CoS feature maintains a per-VC queue for each VC on the VIP, these queues are used simultaneously on all the VCs to absorb temporary bursts and to store early congestion traffic until the sources react to the WRED drops and adapt their transmission rate. We recommend use of a larger SRAM configuration (8 MB) when multiple ATM PVCs are expected to encounter congestion simultaneously.

## IP to ATM CoS Phase 1 Performance and Limitations

### Throughput

The IP to ATM CoS Phase 1 feature operates at a high rate. Details regarding throughput are provided in the following paragraphs.

#### Throughput over DS3/E3 interface

When an ATM DS3/E3 interface is used, activating the IP to ATM CoS Phase 1 feature does not reduce the no-drop rate achieved by the VIP2-50/PA-A3, which remains at the interface line rate with typical Internet packet size distribution such as IMIX traffic, even with a high number of ATM VCs (for example, 200 PVCs).

---

**Note** The no-drop rate is the highest throughput (in packets per second) achieved by the router onto the ATM interface without any packet drop introduced by the router.

---

---

**Note** The “Internet MIX” or IMIX is a well known packet mix representative of Internet traffic that includes packets with the following packet length distribution: 40-byte IP datagrams (58 percent), 552-byte IP datagrams (33 percent), and 1500-byte IP datagrams (9 percent). In other words, for every 12 packets, 7 are with 40-byte IP payload (padded to 46-byte payload on Ethernet), 4 with 552-byte IP payload, and 1 with 1500-byte IP payload. Therefore, IMIX traffic is also referred to as traffic with 7:4:1 distribution.

---

## Throughput over STM-1/OC-3 Interface

When an STM-1/OC-3 interface is used with IMIX traffic, activating the IP to ATM CoS Phase 1 feature does not significantly reduce the maximum no-drop rate achieved by the VIP2-50/PA-A3 on the ATM STM-1/OC-3 interface.

Activating the IP to ATM CoS Phase 1 feature on a small number of ATM PVCs will not reduce the no-drop rate with IMIX traffic. The IP to ATM CoS Phase 1 feature then achieves a no-drop rate throughput of 88 percent of the STM-1/OC-3 ATM line rate.

Activating the IP to ATM CoS Phase 1 feature simultaneously over 50 ATM PVCs will degrade the no-drop rate with IMIX traffic by 2 percent. The IP to ATM CoS Phase 1 feature then achieves a no-drop rate throughput of 84 percent of the STM-1/OC-3 ATM line rate.

Activating the IP to ATM CoS Phase 1 feature simultaneously over 200 ATM PVCs will degrade the no-drop rate with IMIX traffic by 10 percent. The no-drop rate achieved by the STM-1/OC-3 ATM line rate will be 76 percent.

---

**Note** More significant performance degradation could be observed in particular situations where the traffic contains an extraordinarily high percentage of very short packets.

---

The difference between the no-drop rate and the saturation throughput under persistent extreme overload is smaller when the IP to ATM CoS Phase 1 feature is activated than when it is not. This means that even in abnormal situations where the traffic sources do not back off as expected in reaction to WRED discards and they keep sending traffic significantly in excess of the throughput achievable through the network, the IP to ATM CoS Phase 1 feature does not introduce additional risks of performance collapse.

---

**Note** The saturation rate is the throughput of packets transmitted onto the interface when significantly more than the no-drop rate is sent onto the router (and thus some packets get discarded). Note that the saturation rate is typically lower than the no-drop rate.

---

Consequently, activating the IP to ATM CoS Phase 1 feature to take advantage of service differentiation and congestion avoidance is not expected to introduce a risk of compromising the network performance such as in reducing the no-drop rate or in creating a bottleneck that did not exist prior.

## Number of PVCs with WRED

It is possible to configure per-VC WRED simultaneously over any number of ATM PVCs existing on the ATM interface. Regardless of the number of ATM PVCs configured with per-VC WRED and regardless of the level of simultaneous congestion on those PVCs, the back pressure mechanism of the PA-A3 will ensure that packets never need be indiscriminately dropped by the PA-A3, but rather that the necessary drops be performed on the VIP2-50 running the WRED algorithm.

---

**Note** Configurations with per-VC WRED operating simultaneously on up to 200 ATM PVCs have been extensively tested.

---

The following must be noted:

- The WRED algorithm inherently requires significant packet buffers in the queue it manages to be effective (it must be able to absorb the normal burst of traffic that sources may send until they have the time to react to the WRED loss).
- The IP to ATM CoS Phase 1 feature runs one such WRED algorithm for each ATM PVC simultaneously over its own per-VC queue.
- The amount of packet memory on the VIP2-50 that can be used for the per-VC queues managed by WRED is finite.

Consequently, if excessive congestion was occurring on a very high number of ATM PVCs at the very same point in time, the IP to ATM CoS Phase 1 feature could run out of VIP buffers for performing its selective random drop with full effectiveness on all VCs.

---

**Note** For instance, a VIP2-50 with 8-MB SRAM provides 1085 packet buffers available for IP to ATM CoS per-VC queueing on which WRED operates. If 100 ATM PVCs were configured and if all VCs were experiencing excessive congestion simultaneously (as could be simulated in test environments where non-TCP flow controlled source would be used), then on average each PVC would have about 10 packets worth of buffering, which may be too short for WRED to operate successfully.

---

VIP2-50 devices with large SRAM are thus strongly recommended in designs with a high number of ATM PVCs running per-VC WRED and that can experience congestion simultaneously.

The higher the number of configured active PVCs, the lower their SCR would need to be, and consequently the shorter the queue required by WRED to operate on the PVC. Thus, as is the case when using the default WRED profiles of the IP to ATM CoS Phase 1 feature, configuring lower WRED drop thresholds when per-VC WRED is activated on a very large number of low-speed congested ATM PVCs would minimize the risk of buffer shortage on the VIP.

---

**Note** Buffer shortage on the VIP does not result in any malfunction. In the case of buffer shortage on the VIP, the IP to ATM CoS Phase 1 feature simply degrades to FIFO tail drop during the period of buffer shortage (that is, the same drop policy that would occur if the IP to ATM CoS feature were not activated on this PVC).

---

Full dynamic sharing of the VIP buffers across all per-VC queues allows maximum efficiency in the use of the available buffers. The buffers can be used at any time by any per-VC queue needing to absorb a burst of traffic. This dynamic sharing ensures that any PVC experiencing congestion at any point in time will have access to any required number of available buffers.

It must also be noted that the no-drop rate slightly decreases as the number of ATM PVCs increases. Refer to the “Throughput” section for information on the no-drop rate measured for various numbers of PVCs over IMIX traffic.

## Single PA-A3 per VIP2-50

The IP to ATM CoS Phase 1 feature allows a single PA-A3 per VIP2-50.

## IP to ATM CoS Phase 1 Monitoring

### Total Drops on Interface: show interface atm

To learn the total number of packets that have been dropped on output on an ATM interface, use the **show interface atm** command.

```
7513-1-31# show interface atm 11/0/0
ATM11/0/0 is up, line protocol is up
Hardware is cyBus ENHANCED ATM PA
MTU 4470 bytes, sub MTU 4470, BW 149760 Kbit, DLY 80 usec, rely 255/255, load 1/255
Encapsulation ATM, loopback not set, keepalive not set
Encapsulation(s): AAL5 AAL3/4
4096 maximum active VCs, 8 current VCCs
VC idle disconnect time: 300 seconds
Last input 00:00:00, output 00:00:00, output hang never
Last clearing of "show interface" counters never
Queueing strategy: fifo
Output queue 0/40, 801870 drops; input queue 0/75, 0 drops
30 second input rate 19000 bits/sec, 1 packets/sec
30 second output rate 22000 bits/sec, 2 packets/sec
4072808 packets input, 757905093 bytes, 0 no buffer
Received 0 broadcasts, 0 runts, 0 giants
0 input errors, 0 CRC, 0 frame, 0 overrun, 1225 ignored, 0 abort
3672469 packets output, 260485571 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
0 output buffers copied, 0 interrupts, 0 failures
7513-1-31#
```

“Output” drops indicate the total aggregate of all drops on output (VIP WRED discard, VIP discard because of VIP buffer shortage, discard on the PA, and so forth).

### Drops on the PA: show atm vc

To learn the number of packets that have been dropped by the PA on output on an ATM interface on a per-VC basis, use the **show atm vc** command.

Note that in normal operations with the IP to ATM CoS feature, there should be no drops on the PA because the PA provides back pressure to the VIP so that congestion is pushed back from the PA to the VIP, which can perform WRED selective discard on a per-VC basis.

```
7513-1-31# show atm vc 100
ATM11/0/0.100: VCD: 100, VPI: 5, VCI: 100, etype:0x0, AAL5 - LLC/SNAP, Flags: 0xC0030
PeakRate: 16000, Average Rate: 8000, Burst Cells: 512, VCmode: 0x0
OAM DISABLED, InARP frequency: 15 minute(s)
InPkts: 135, OutPkts: 134, InBytes: 9416, OutBytes: 9256
InPRoc: 135, OutPRoc: 0, Broadcasts: 133
InFast: 0, OutFast: 0, InAS: 0, OutAS: 1
InPktDrops: 0, OutPktDrops: 0
CrcErrors: 0, SarTimeOuts: 0, OverSizedSDUs: 0
OAM F5 cells sent: 0, OAM cells received: 0
Status: ACTIVE
7513-1-31#
```

“OutPktDrops” keeps track of the packet drops performed by the PA because of the per-VC queue on the PA being congested and running out of buffers.

## WRED Intelligent Drops on the VIP: show queueing interface atm

To learn the number of packets that have been selectively dropped by the WRED algorithm on the VIP on a per-VC basis, use the **show queueing interface atm** command.

```
7513-1-31# show queueing interface atm 11/0/0.103
VC 5/103 -
  ATM11/0/0.103 queue size 46
    packets output 1346100, drops 134315, nobuffer drops 0
  WRED: queue average 44
    weight 1/512, max available buffers 1021

  Precedence 0: 40 min threshold, 81 max threshold, 1/10 mark weight
    1344366 packets output, drops: 134304 random, 10threshold
  Precedence 1: 45 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 2: 50 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 3: 55 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 4: 60 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 5: 65 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 6: 70 min threshold, 81 max threshold, 1/10 mark weight
    1734 packets output, drops: 1 random, 0 threshold
  Precedence 7: 75 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
```

“Random” and “threshold” drops indicate for each precedence how many packets have been discarded by the WRED algorithm:

- **Random:** When the average queue length was between min-threshold and max-threshold. In this case the discard was a random decision with a probability varying from zero (if average queue length is just above the min-threshold) to mark-weight (if average queue length is just below the max-threshold). Of course, this counter only counts the packets for which the random decision was actually to drop the packet.
- **Threshold:** When the average queue length was above the max-threshold. In this case the discard was a deterministic decision and all the packets for which the average queue length is calculated to be beyond the max-threshold are discarded.

## Buffer-Shortage Drops on the VIP: show queueing interface atm

In some exceptional situations, the output VIP could have no buffers left to store a packet that is switched to this output VIP from the RSP or from an input VIP. Consequently, the VIP will need to indiscriminately drop that packet regardless of its precedence.

Such an exceptional situation could occur as a result of heavy congestion combined with misconfiguration of the WRED parameters. As an example, if the exponential weighting constant has been reconfigured from the default value to an excessively large value, then the WRED algorithm is slow to react to congestion (because the moving average increases only slowly as the instantaneous queue fills up) and WRED may not start intelligent discard early enough, so that the bursts keep filling the buffers.

Such situations should be avoided because these drops indiscriminately affect high precedence traffic.

Drops on the VIP due to buffer shortage can be monitored through the **show queueing interface atm** command through the “nobuffer drop” counter.

```
7513-1-31# show queueing interface atm 11/0/0.103
VC 5/103 -
  ATM11/0/0.103 queue size 46
    packets output 1346100, drops 134315, nobuffer drops 0
  WRED: queue average 44
    weight 1/512, max available buffers 1021

  Precedence 0: 40 min threshold, 81 max threshold, 1/10 mark weight
    1344366 packets output, drops: 134304 random, 10 threshold
  Precedence 1: 45 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 2: 50 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 3: 55 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 4: 60 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 5: 65 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 6: 70 min threshold, 81 max threshold, 1/10 mark weight
    1734 packets output, drops: 0 random, 1 threshold
  Precedence 7: 75 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
```

“Nobuffer drops” indicate how many packets have been dropped indiscriminately by the VIP because no buffer was available at that time to accept the packet when it was handed over to the output VIP by the RSP or the VIP that received the packet. Because the VIP drops the packet without being able to run the IP to ATM CoS feature, and in fact without even looking at the packet at all, such packets are dropped irrespective of the moving average queue occupancy for the particular VC and irrespective of the packet precedence.

## Instantaneous WRED Operation: show queueing interface atm

The **show queueing interface atm** command can also be used to learn how many total buffers are available in the VIP buffers for the WRED algorithm for a particular VC, and to learn the instantaneous value of the real buffer queue length and the instantaneous WRED average queue length.

```
7513-1-31# show queueing interface atm 11/0/0.103
VC 5/103 -
  ATM11/0/0.103 queue size 46
    packets output 1346100, drops 134315, nobuffer drops 0
  WRED: queue average 44
    weight 1/512, max available buffers 1021

  Precedence 0: 40 min threshold, 81 max threshold, 1/10 mark weight
    1344366 packets output, drops: 134304 random, 10 threshold
  Precedence 1: 45 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 2: 50 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 3: 55 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 4: 60 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 5: 65 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 6: 70 min threshold, 81 max threshold, 1/10 mark weight
    1734 packets output, drops: 1 random, 0 threshold
  Precedence 7: 75 min threshold, 81 max threshold, 1/10 mark weight
    (no traffic)
```

“Queue size” indicates the number of packets buffered on the VIP in the per-VC queue corresponding to that VC at the time the **show** command was run.

“Queue average” indicates the instantaneous value of the moving average queue length computed by WRED for that particular VC at the time the **show** command was run.

“Max available buffers” indicates the number of packet buffers that are available on the VIP and can be dynamically used by all the IP to ATM CoS per-VC queues.

## IP to ATM CoS Phase 1 Fine-Tuning

### Where WRED Fine-Tuning Can (and Cannot) Help

As discussed earlier, the main goal of WRED is to achieve service differentiation across various CoS identified by the IP Precedence. Thus, WRED enables the network operator to achieve different performance and QoS objectives for each CoS. For example, the objectives for each CoS can include different levels of parameters such as the drop probability and the bandwidth achieved through the network by a TCP-compliant flow; these parameters could be defined in various ways in time (for example, average over some long-term period, average across the busy hour), in various ways in topology (for example, across all users, across some specific user population), and in various statistical ways (across all the traffic, across a representative set of flows).

---

**Note** Because per-VC WRED relies on a single FIFO queue, the IP to ATM CoS Phase 1 feature does not offer delay differentiation across classes and cannot be used by network operators to achieve different packet transit delay objectives across different CoS. Note, however, that because of WRED's inherent congestion avoidance behavior, the IP to ATM CoS Phase 1 feature is expected to improve the average transit delay across all CoS in congested networks.

---

A number of network-wide inputs can be used by the network operator to determine whether WRED enforces the required level of service for each precedence. These inputs include the results of a per-precedence network performance measurement campaign, the information provided by a per-precedence service level agreement monitoring tool, or the WRED per-precedence discard statistics provided by the Cisco IOS **show queueing interface atm** command.

If these inputs reveal that the performance and QoS experienced by some classes of service is below their respective objectives, while the performance and QoS experienced by other classes of service is above their respective objectives, it is most likely that WRED fine-tuning can help. WRED's service differentiation can be tuned to reallocate network resources across the classes of service so that the underserved CoS can be more favored by WRED at the expense of the overserved CoS. See the section "WRED Fine-Tuning to Reallocate Resources across CoS" for more recommendations on this fine-tuning.

If these inputs reveal that the performance and QoS are below the respective objectives for all classes of service, then the level of traffic is probably simply too high compared to the resources available in the network to satisfy all CoS objectives. WRED fine-tuning likely will not help. Either the objectives should be loosened for some classes of service or more resources need to be deployed in the network.

---

**Note** As discussed earlier, WRED's service differentiation relies on the dynamic interaction of WRED's probabilistic discard and how transport protocol's flow control adapts to the discards performed by WRED. More specifically, WRED has been designed to operate best in an environment where traffic is dominated by TCP traffic implementing the IETF specified TCP flow control mechanisms. Consequently, in special situations where there would be a significant proportion of nonadaptive traffic (such as a network carrying a high proportion of multimedia or Voice over IP traffic running over UDP), it is possible that even with fine-tuning, WRED may not be able to enforce the required relative performance and QoS objectives. Such situations are not expected to be encountered in an environment carrying current Internet and intranet traffic.

---

## Considerations on WRED Fine-Tuning

Because of the dynamic nature of RED and WRED and their complex interactions with transport level flow control mechanisms (such as TCP flow control), fine-tuning WRED to achieve specific IP service differentiation objectives in particular operating conditions is a delicate exercise and great caution is recommended. We recommend you start operations or testing with the default WRED settings (or from configurations close to the WRED default settings) and fine-tune from there. Modifications to WRED parameters should be tested and validated under a vast range of network conditions before being deployed in a large network.

The default WRED settings take into account the following:

- The experience developed in the Internet research community on RED parameter setting
- Configuration of related parameters (such as shaping parameters of the ATM VC on which WRED is run)
- A different discard profile per precedence (the higher the precedence, the better the default corresponding service)

For these reasons we believe that the default WRED settings will provide a robust base to start from when deploying IP service differentiation over ATM (see the section “Default WRED Configuration”).

Because higher rate ATM PVCs must cope with larger bursts, we recommend you not apply the same WRED profile (that is, the same random-detect group) to all VCs if they have vastly different rates. However, fine-tuning of WRED parameters for each different ATM PVC rate is neither required nor practical. We therefore suggest you make extensive use of the ability to associate the same random-detect group to multiple VCs and only define a few random-detect groups, each being applied to all the VCs belonging to a range of VC rates (for instance the “slow VCs,” the “medium VCs,” and the “fast VCs”). Then it is only necessary to fine-tune a few random-detect groups to achieve fine-tuning across all VCs.

For each random-detect group, it is expected that to achieve particular service differentiation objectives in a particular environment, only minor modifications of per-precedence parameters from the default values would be required (see the sections “Relative per-Precedence WRED Fine-Tuning” and “Common per-Precedence WRED Fine-Tuning”). We do not expect this configuration to normally require major modifications of the per-precedence parameters from the default values, nor do we expect that modification of the WRED exponential weighting to be required (see the section “Exponential Weighting Constant WRED Fine-Tuning”).

General recommendations on RED setting can be found in the paper by Sally Floyd defining RED gateways (Random Early Detection Gateways): <http://www.aciri.org/floyd/red.html>.

Additional brief discussions on RED parameter settings are also available from Sally Floyd’s web site at <http://www.aciri.org/floyd/REDparameters.txt>.

## Default WRED Configuration

Using the **atm pvc ... [random-detect [group\_name]]** command to associate to an ATM PVC a random-detect group without a group name, or with a group name whose WRED parameters have not been further configured through the **random-detect group group-name** command, results in a set of default WRED parameters being applied to the ATM PVC.

The default values allocate the same max-thresholds and the same mark-probability to all the precedences. However, the default min-threshold is different for every precedence: the higher the precedence, the higher the min-thresholds. Consequently the default WRED configuration offers an increasingly “better” service to higher precedences.

As indicated previously, these default values take into account the ATM shaping parameters associated with the VC. To accommodate for larger bursts that can appear at higher rates, the higher the VC shaping rate the larger the default min- and max-thresholds. For instance, with a 10-kbps ATM, here are the default WRED parameters applied to the VC in a particular router:

```

nf-7505-1# show run
interface ATM1/1/0.47 point-to-point
  atm pvc 47 0 47 aal5snap 10 10 1 random-detect wredgroup1

nf-7505-1# show queueing red

VC 0/47 -
random-detect group default:

exponential weight 9
precedence    min-threshold    max-threshold    mark-probability
-----
0:            20                40                1/10
1:            22                40                1/10
2:            24                40                1/10
3:            26                40                1/10
4:            28                40                1/10
5:            30                40                1/10
6:            32                40                1/10
7:            34                40                1/10
    
```

In comparison, here are the default WRED parameters applied by the same router to a VC shaped at 9 Mbps of SCR and 10 Mbps of PCR:

```

nf-7505-1# show run
interface ATM1/1/0.49 point-to-point
  atm pvc 49 0 49 aal5snap 10000 9000 100 random-detect wredgroup3
nf-7505-1# show queueing red

VC 0/49 -
random-detect group default:

exponential weight 9
precedence    min-threshold    max-threshold    mark-probability
-----
0:            72                144               1/10
1:            81                144               1/10
2:            90                144               1/10
3:            99                144               1/10
4:            108               144               1/10
5:            117               144               1/10
6:            126               144               1/10
7:            135               144               1/10
    
```

## WRED Fine-Tuning to Reallocate Resources across CoS

### Relative per-Precedence WRED Fine-Tuning

In most environments the per-precedence WRED fine-tuning described in this section will be the only fine-tuning required to achieve specific IP service differentiation objectives (such as per-precedence service level agreements with different packet loss commitments).

In addition, in most environments we expect only minor modifications over the per-precedence WRED default values to be required.

Per-precedence WRED fine-tuning refers to adjustment to the three WRED parameters configurable on a per-precedence basis to shuffle resource allocation across CoS. Those per-precedence parameters are min-threshold, max-threshold, and mark-probability.

For instance, consider a situation where the default (or current) WRED parameters are the following:

```

nf-7505-1# show queuing red

VC 0/49 -
random-detect group default:

exponential weight 9
precedence    min-threshold    max-threshold    mark-probability
-----
0:            72                144              1/10
1:            81                144              1/10
2:            90                144              1/10
3:            99                144              1/10
4:            108               144              1/10
5:            117               144              1/10
6:            126               144              1/10
7:            135               144              1/10
    
```

The network administrator may want to configure the following:

- An even more aggressive discard for the precedence 0 traffic (such as best-effort traffic) to provide a better level of services to all other precedences. This may be accomplished by the following:
  - By reducing the max-threshold from 144 to 133 so that the best-effort traffic gets deterministically discarded at a point where a significant proportion of traffic on other precedences would still be accepted.
  - By increasing its drop probability from 1/10 to 1/8 (reduce the mark-probability-denominator from 10 to 8) so that best-effort traffic always experiences an even higher proportion of discarded traffic than the other precedences in any congestion situation.
- An even less aggressive discard of precedence 6 and 7 traffic (network control traffic such as routing) to ensure even higher protection of routing traffic over any other traffic by similarly increasing the max-threshold and reducing the discard probability of precedences 6 and 7.

The updated WRED parameters for the corresponding random-detect group would then look like the following:

```

nf-7505-1# show queueing red

VC 0/49 -
random-detect group wredgroup3:

exponential weight 9
precedence    min-threshold    max-threshold    mark-probability
-----
0:            72                133              1/8
1:            81                144              1/10
2:            90                144              1/10
3:            99                144              1/10
4:            108               144              1/10
5:            117               144              1/10
6:            135               155              1/100
7:            135               155              1/100
    
```

### Common per-Precedence WRED Fine-Tuning

Common per-precedence WRED fine-tuning refers to the action of modifying the per-precedence WRED parameters to make the discard more aggressive across all precedences (or less aggressive across all precedences) rather than to modify these parameters to change the relative allocation of resources among the precedences as described previously.

Common per-precedence WRED fine-tuning may be required in situations where a significant number of drops occur on the VIP because of buffer shortage (see the earlier section “Instantaneous WRED Operation: show queueing interface atm”) and some classes of service experience lower QoS than their corresponding objectives.

---

**Note** Improper configuration of the exponential weighting constant also could be responsible for VIP buffer shortage—improper configuration could result in the average queue not tracking congestion fast enough and thus not starting random drops until all buffers are actually exhausted. However, fine-tuning of the exponential weighting constant, which is discussed in the next section, usually should not be required. In normal situations, VIP buffer shortage is more likely to be properly addressed by common adjustment of the per-precedence drop profile than by fine-tuning the exponential weighting constant.

---

When a VIP is experiencing a buffer shortage, you can modify the WRED parameters to enforce more aggressive discard across all the precedences by reducing the min- and max-thresholds for all (or a significant subset of) precedences (and perhaps also by increasing the mark weight). Such a common per-precedence WRED fine-tuning should be useful in an environment where a high number of ATM PVCs are running WRED and are simultaneously experiencing congestion.

---

**Note** A delicate balance exists when modifying WRED parameters to enforce more aggressive discard across all the precedences. Configuring discard parameters that are excessively aggressive across all precedences would result in an unnecessary reduction in performance and QoS levels across all precedences.

---

## Exponential Weighting Constant WRED Fine-Tuning

We recommend you not modify the exponential weighting constant. We do not expect the fine-tuning guidelines provided in this section to be commonly required. Modifying the exponential weighting constant should only be performed with great care and if you have very clearly determined that your applications would benefit from the changed values.

The WRED average queue size computation is based on the previous average and the current size of the queue. The formula is as follows:

$$\text{average} = (\text{old\_average} * (1 - 1/2^n)) + (\text{current\_queue\_size} * 1/2^n)$$

where  $n$  is the exponential weighting constant specified in the **exponential weighting constant** subcommand of the **random-detect group** command. Thus, the higher the factor, the more dependent the average is on the previous average.

The default exponential weighting constant value of 9 was chosen based on the best current RED expertise available.

For high values of  $n$ , the previous average becomes more significant in the moving average calculation. A large factor smooths out the peaks and lows in queue length. The average queue size is unlikely to change very quickly, avoiding drastic swings in size. The WRED process will be slow to start dropping packets, but it may continue dropping packets for a time after the actual queue size has fallen below the minimum threshold. The slow-moving average will accommodate temporary bursts in traffic.

If the value of  $n$  gets too high, WRED will not react to congestion. Packets will be transmitted or dropped as if WRED were not in effect because there will be a shortage of buffers before the WRED algorithm starts performing selective discard across the CoS. The result could be indiscriminate drops across all CoS, which in turn may prevent the network from achieving the performance and QoS objectives of the most stringent CoS. The result also could be the undesirable TCP synchronization phenomenon sometimes observed with FIFO queueing. In this situation, many TCP sources back off simultaneously when there is a heavy drop caused by buffer shortage. They then reopen their windows gradually at the same time, until eventually there is another buffer shortage, and they back off together again.

In situations where there is a significant number of drops due to buffer shortage on a VIP (see the earlier section “Buffer-Shortage Drops on the VIP: show queueing interface atm”), it may not be obvious whether modification of the per-precedence WRED parameters is required across all or most precedences (see the section “Common per-Precedence WRED Fine-Tuning”) or whether modification of the exponential weighting constant is required. The following suggestions are provided based on monitoring the WRED instantaneous operations at critical times (see the earlier section “Instantaneous WRED Operation: show queueing interface atm”).

If the queue size is often high while the average queue remains low compared to the min-thresholds configured for the precedences carrying the bulk of the traffic, then the average queue length likely is not tracking the congestion fast enough and reducing the exponential weighting to allow the moving average to adjust faster to congestion is advised (that is, increase instantaneous queue size).

If the queue size is often very low while the average queue remains high beyond min- and max-thresholds for the precedences carrying the bulk of the traffic, then the average queue length likely is not tracking the congestion disappearance fast enough. Reducing the exponential weighting to allow the moving average to adjust faster to congestion is advised (that is, increase instantaneous queue size).

For low values of  $n$ , the average queue size closely tracks the current queue size. The resulting average may fluctuate quickly with changes in the traffic levels. In this case, the WRED process responds quickly to congestion (that is, to small increases in queue length) and starts dropping packets as soon as small bursts of traffic are detected.

If the value of  $n$  gets too low, WRED will overreact to temporary traffic bursts and drop traffic unnecessarily, essentially preventing WRED from accepting normal traffic bursts for which it actually has physical buffering capacity. This overreaction will result in unnecessarily reducing the performance and QoS observed across all CoS.

### WRED Fine-Tuning with High Numbers of Congested ATM PVCs

As described earlier in the section “IP to ATM CoS Phase 1 Performance and Limitations,” the VIP may not have enough buffers to maintain a long per-VC queue for all the VCs simultaneously in the special cases where a very high number of ATM PVCs are configured with per-VC WRED and where a very high number of them can simultaneously experience congestion. In this particular situation, adjusting the WRED drop profile so that WRED starts dropping earlier on shorter average queue occupancy (that is, reducing min-threshold and also perhaps max-threshold) may be beneficial in allowing faster WRED reaction on all VCs.

## IP to ATM CoS Phase 1 Case Studies

### Case Study 1: Protection of IP Routing Traffic in a Congested Network

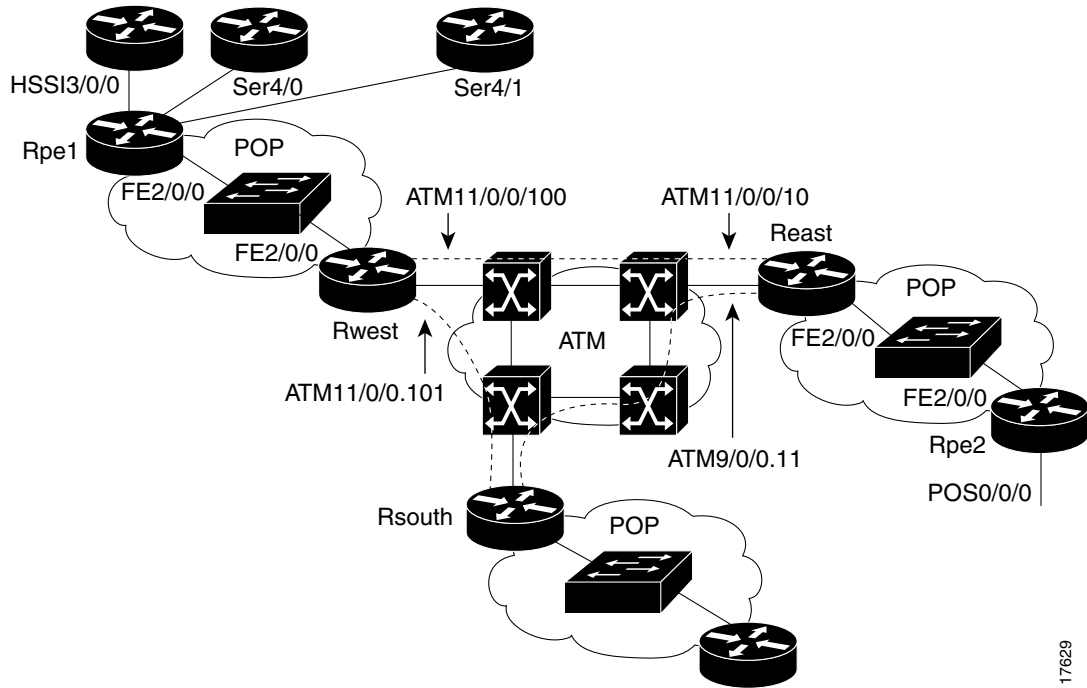
This case study presents an IP network design where a single level of service is supported for data traffic but where routing traffic is protected over user data in case of congestion.

The main design goals are as follows:

- Offer a single level of service for all user/data traffic
- Ensure that even in case of congestion routing traffic is protected so that routing stability is maintained

Figure 7 illustrates the network topology for this case study.

Figure 7 Network Topology for Case Study 1



17629

In this design, the data traffic is transported as precedence 0. We strongly recommend you re-mark all traffic coming from outside of the network with precedence 0 to ensure that packets are not injected inadvertently or maliciously into the network with their precedence set to 6 or 7, in which case they would be “stealing” high-quality service and would be threatening the routing traffic integrity. In this design CAR is used on the edge of the network to perform re-marking of external traffic.

IP routing protocol messages are always generated by Cisco routers with precedence 6, which does not require any configuration action.

As described earlier, routing protocols are protected through per-VC queueing and back pressure, through selective discard, and through WRED/discard-bypass for critical routing packets. In this case study, the IP to ATM CoS Phase 1 selective discard mechanism of per-VC WRED is configured by setting the WRED discard profiles on every VC so that precedences 6 and 7 will experience a much lower drop probability than the data traffic.

## Rpe1 Configuration

Rpe1 is a provider edge router. It is a router from the service provider network that sits on the edge and directly connects to routers on the customer premises. Every interface that receives packets from outside the service provider network (that is, from customer premises routers) performs re-marking of the precedence to 0.

```

ip cef distributed
!
interface Hssi3/0/0
  description Edge interface connecting to a CPE
  ip address 192.168.10.1 255.255.255.0
  rate-limit input 10000000 24000 24000
    conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
!
interface serial4/0
  description Edge interface connecting to a CPE
  ip address 192.168.50.1 255.255.255.0
  rate-limit input 1000000 24000 24000
    conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
  encapsulation hdlc
  ip route-cache
...
!
interface FastEthernet2/0/0
  description Interface connecting Rpe1 to the intra-POP switch
  ip address 17.1.1.1 255.255.255.0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable

```

## Rpe2 Configuration

Rpe2 is a provider edge router. It is a router from the server provider network that sits on the edge and directly peers with an edge router from another server provider. The interface that receives packets from the other server provider network performs re-marking of the precedence to 0.

```

!
ip cef distributed
!
interface POS0/0/0
  description Edge interface connecting to another Service provider
  ip address 192.168.100.1 255.255.255.0
  rate-limit input 10000000 24000 24000
    conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
  ip route-cache distributed
  no keepalive
  clock source internal
  no cdp enable
!
interface FastEthernet2/0/0
  description Interface connecting Rpe2 to the intra-POP switch
  ip address 17.2.1.1 255.255.255.0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
!

```

## Rwest Configuration

Rwest is a backbone router. It is a router from the server provider network that interconnects its POP with other remote POPs through the ATM WAN cloud. The IP to ATM CoS Phase 1 feature is activated on the two ATM PVCs connecting Rwest respectively to Reast and to Rsouth. The ATM PVCs are to be shaped as CBR PVCs respectively at 10 Mbps and 1 Mbps. Because these two PVCs have shaping rates that differ by an order of magnitude, different WRED profiles are configured on the two VCs, with larger discard thresholds on the faster PVC to accommodate for normal temporary bursts that would naturally be larger (in number of packets) on the faster VC (assuming similar Round Trip Time [RTT]).

```

!
ip cef distributed
!
interface ATM11/0/0
  description ATM interface to the ATM WAN cloud
  no ip address
  no ip route-cache optimum
  ip route-cache distributed
  no ip mroute-cache
  no keepalive
!
interface ATM11/0/0.100 point-to-point
  description ATM PVC to Reast
  ip address 17.0.0.1 255.255.255.0
  atm pvc 100 5 100 aal5snap 10000 10000 10 inarp
  random-detect fast-wred-profile
!
interface ATM11/0/0.101 point-to-point
  description ATM PVC to Rsouth
  ip address 17.0.1.1 255.255.255.0
  atm pvc 101 5 101 aal5snap 1000 1000 10 inarp
  random-detect slow-wred-profile
!
interface FastEthernet2/0/0
  description Interface connecting Rwest to the intra-POP switch
  ip address 17.1.1.2 255.255.255.0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
!
random-detect-group slow-wred-profile
  precedence 0 20 40 10
  precedence 1 20 40 10
  precedence 2 20 40 10
  precedence 3 20 40 10
  precedence 4 20 40 10
  precedence 5 20 40 10
  precedence 6 35 60 100
  precedence 7 35 60 100
!
random-detect-group fast-wred-profile
  precedence 0 72 144 10
  precedence 1 72 144 10
  precedence 2 72 144 10
  precedence 3 72 144 10
  precedence 4 72 144 10
  precedence 5 72 144 10
  precedence 6 130 200 100
  precedence 7 130 200 100
!

```

## Rsouth Configuration

Rsouth is a backbone router. It is a router from the server provider network that interconnects its POP with other remote POPs through the ATM WAN cloud. The IP to ATM CoS Phase 1 feature is activated on the two ATM PVCs connecting Rsouth respectively to Reast and to Rwest. The ATM PVCs are to be shaped as CBR PVCs respectively at 9 Mbps and 1 Mbps. Because these two PVCs have shaping rates that differ by almost an order of magnitude, different WRED profiles are configured on the two VCs, with larger discard thresholds on the faster PVC to accommodate for normal temporary bursts that would naturally be larger (in number of packets) on the faster VC (assuming similar RTT).

Rsouth thus has a similar configuration to Rwest.

## Reast Configuration

Reast is a backbone router. It is a router from the server provider network that interconnects its POP with other remote POPs through the ATM WAN cloud. The IP to ATM CoS Phase 1 feature is activated on the two ATM PVCs connecting Reast respectively to Rsouth and to Rwest. The ATM PVCs are to be shaped as CBR PVCs respectively at 9 Mbps and 10 Mbps. Because these two PVCs have shaping rates that are of the same order of magnitude and quite close, the same WRED profile is configured on the two VCs.

```

!
ip cef distributed
!
interface ATM9/0/0
  description ATM interface to the ATM WAN cloud
  no ip address
  no ip route-cache optimum
  ip route-cache distributed
  no ip mroute-cache
  no keepalive
!
interface ATM9/0/0.10 point-to-point
  description ATM PVC to Rwest
  ip address 17.0.0.2 255.255.255.0
  atm pvc 10 5 100 aal5snap 10000 10000 10 inarp
  random-detect fast-wred-profile
!
interface ATM9/0/0.11 point-to-point
  description ATM PVC to Rsouth
  ip address 17.0.2.2 255.255.255.0
  atm pvc 11 5 101 aal5snap 9000 9000 10 inarp
  random-detect fast-wred-profile
!
interface FastEthernet2/0/0
  description Interface connecting Reast to the intra-POP switch
  ip address 17.2.1.2 255.255.255.0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
!

```

```

random-detect-group fast-wred-profile
precedence 0 72 144 10
precedence 1 72 144 10
precedence 2 72 144 10
precedence 3 72 144 10
precedence 4 72 144 10
precedence 5 72 144 10
precedence 6 130 200 100
precedence 7 130 200 100
!

```

### Examples of Show Commands

The **show queuing interface** command confirms that per-VC WRED performs selective packet drops on data traffic in a situation of congestion but does not perform any drops of routing packets. It also confirms that there were no drops due to buffer shortage on the VIP.

```

Rwest# show queuing interface atm 11/0/0.100
VC 5/100 -
  ATM11/0/0.100 queue size 83
    packets output 134757, drops 17402, nobuffer drops 0
  WRED: queue average 82
    weight 1/512, max available buffers 1021

  Precedence 0: 72 min threshold, 144 max threshold, 1/10 mark weight
    133023 packets output, drops: 17402 random, 0 threshold
  Precedence 1: 72 min threshold, 144 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 2: 72 min threshold, 144 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 3: 72 min threshold, 144 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 4: 72 min threshold, 144 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 5: 72 min threshold, 144 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 6: 130 min threshold, 200 max threshold, 1/100 mark weight
    1734 packets output, drops: 0 random, 0 threshold
  Precedence 7: 130 min threshold, 200 max threshold, 1/100 mark weight
    (no traffic)

```

The **show atm vc** command confirms that no packets are being dropped on the PA.

```

Rwest# show atm vc 100
ATM11/0/0.100: VCD: 100, VPI: 5, VCI: 100, etype:0x0, AAL5 - LLC/SNAP, Flags: 0xC0030
PeakRate: 10000, Average Rate: 10000, Burst Cells: 10, VCmode: 0x0
OAM DISABLED, InARP frequency: 15 minute(s)
InPkts: 1350, OutPkts: 1340, InBytes: 94160, OutBytes: 92560
InPRoc: 1350, OutPRoc: 0, Broadcasts: 1330
InFast: 0, OutFast: 0, InAS: 0, OutAS: 1
InPktDrops: 0, OutPktDrops: 0
CrcErrors: 0, SarTimeOuts: 0, OverSizedSDUs: 0
OAM F5 cells sent: 0, OAM cells received: 0
Status: ACTIVE
Rwest#

```

The **eigrp log-neighbor-change** command is used in the Enhanced IGRP router configuration. Because no neighbor-change messages are sent to the syslog (and to the console), it confirms that, regardless of the congestion level on the data traffic, Enhanced IGRP neighbors remain visible to each other from the routing perspective and that the Enhanced IGRP routing thus remains stable.

```
!
router eigrp 209
  eigrp log-neighbor-change
!
```

The **show ip eigrp neighbor detail** command is used to confirm that Enhanced IGRP routing messages are not lost and thus do not require retransmission.

```
Rwest# show ip eigrp neighbor detail
IP-EIGRP neighbors for process 209
H  Address                Interface  Hold  Uptime  SRTT  RTO  Q  Seq
                               (sec)          (ms)    Cnt Num
0  17.0.0.2                ATM11/0/0  14    17:11:35  1    200  0  1027
Version 11.1/1.0, Retrans: 1, Retries: 0
```

## Case Study 2: ISP Offering Premium Differentiated Service

Before the availability of IP QoS features, ISPs needed to focus on a single market segment characterized by a particular trade-off between level of performance and price. For instance, some ISPs focused on low-cost services (standard) often sought by residential customers, while some ISPs focused on high-performance services (premium) often sought by business customers. These services needed to remain on separate physical networks because they each required very different traffic engineering practices: The high-performance service required a network with no or low overbooking to achieve high performance and the low-cost service required high overbooking in the network to achieve low cost.

The IP QoS features now enable an ISP to support both (or even multiple) standard and premium CoS on the same infrastructure by activating congestion management or scheduling mechanisms that operate independently for each CoS.

This case study presents an IP network design where a premium CoS is supported in addition to the standard best-effort CoS. The premium CoS may be marketed, for instance, to business users. The premium CoS is not marketed as a strictly guaranteed service (that is, with strictly guaranteed performance parameters) but rather is marketed as an engineered best-effort service just like “high end/overprovisioned” Internet backbones are marketed today. WRED, and per-VC WRED where the underlying transport technology is ATM, is used to enforce the service differentiation by more aggressively dropping standard traffic in case of congestion so that standard traffic is in effect more throttled in case of congestion than the premium service. This selective dropping results in a premium IETF-compliant TCP flow enjoying superior throughput over a standard IETF-compliant TCP flow. In case of congestion, an end user communicating using the standard service will experience the congestion while an end user communicating using the premium service will have a similar “experience” as if the user were using a lightly loaded network.

Packets that are part of the standard CoS are marked/re-marked with precedence 0 when received from the customer and enter the server provider network using the CAR feature. This marking ensures that traffic from a standard customer will be treated with the standard CoS.

Packets that are part of the premium CoS are marked/re-marked with precedence 4 when received from the customer and enter the server provider network using the CAR feature. This marking ensures that traffic from a premium customer will be treated with the premium CoS.

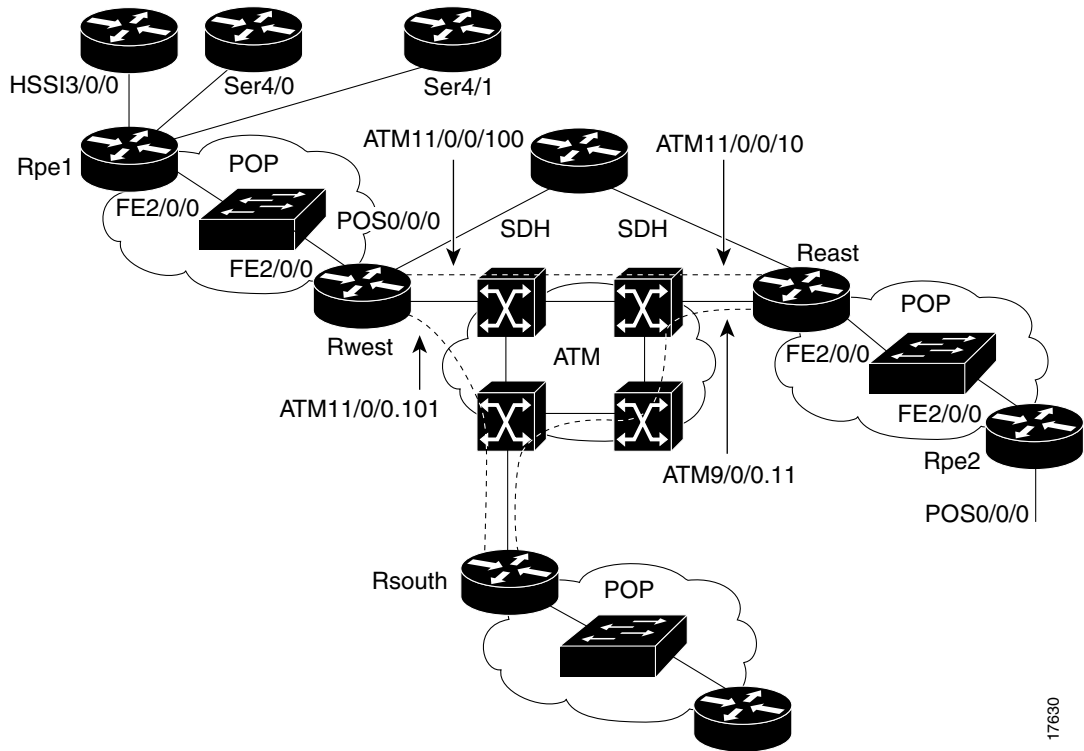
The service provider also uses CAR to sell and enforce a subrate premium service, a service where the amount of premium traffic to be sent by the customer is limited to a bandwidth that is lower than the physical interface bandwidth. For instance, the service provider can sell 5 Mbps of premium service to a customer connected at 34 Mbps through a High-Speed Serial Interface (HSSI).

QoS Policy Propagation via BGP (QPPB) is used to ensure that all traffic destined to a premium customer is marked with precedence 4 when it enters the service provider network (regardless of whether it comes from a premium or standard customer) and is thus treated with the premium CoS. This configuration of QPPB ensures that a premium customer benefits from high performance both on the traffic it sends and on the traffic it receives.

Routing traffic is treated inside the server provider network as a separate class that is not visible nor accessible to user traffic and that enjoys a lower drop probability to increase routing protection. Routing traffic is identified using precedence 6.

Figure 8 illustrates the network topology for this case study.

**Figure 8 Network Topology for Case Study 2**



17630

## Rpe1 Configuration

Rpe1 is a provider edge router. It is a router from the server provider network that sits on the edge and directly connects to routers on the customer premises. Every interface that receives packets from outside the service provider network (that is, from customer premises routers) performs re-marking of the precedence to 0 for packets belonging to the standard CoS and to 4 for packets belonging to the premium CoS.

```

!
ip cef distributed
!
interface serial4/0
description Edge interface connecting to a Standard Customer
ip address 192.168.50.1 255.255.255.0
rate-limit input 1000000 24000 24000
conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
encapsulation hdlc
! activate QPPB on this interface
bgp-policy destination ip-prec-map
ip route-cache
!
interface serial4/1
description Edge interface connecting to a Premium Customer
ip address 192.168.60.1 255.255.255.0
rate-limit input 1000000 24000 24000
conform-action set-prec-transmit 4 exceed-action set-prec-transmit 4
encapsulation hdlc
! activate QPPB on this interface
bgp-policy destination ip-prec-map
ip route-cache
!
interface Hssi3/0/0
description Edge interface connecting to a Customer
with 5Mb/s of Premium
ip address 192.168.70.1 255.255.255.0
! traffic below 5 Mb/s is marked as Premium (Precedence 4).
! traffic beyond 5 Mb/s is marked as Standard (Precedence 0).
! Only considers traffic which hasn't been ear-marked as going to
! a Premium (ie traffic has qos-group 0)
rate-limit input qos-group 0 5000000 24000 24000
conform-action set-prec-transmit 4 exceed-action set-prec-transmit 0
! activate QPPB on this interface (for Precedence and qos-group setting)
bgp-policy destination ip-prec-map
bgp-policy destination ip-qos-map
no ip route-cache optimum
ip route-cache distributed
no keepalive
no cdp enable
...
!
interface FastEthernet2/0/0
description Interface connecting Rpe1 to the intra-POP switch
ip address 17.1.1.1 255.255.255.0
no ip route-cache optimum
ip route-cache distributed
no keepalive
no cdp enable
!
router bgp 210
! tell BGP to take into account the "receive_community" route-map
table-map receive_community
! tell BGP to advertise prefixes with communities as defined in
! the "mark_community" route-map
neighbor 210.210.14.1 remote-as 210

```

```
neighbor 200.200.14.1 update-source Loopback0
neighbor 210.210.14.1 route-map mark_community out
neighbor 210.210.14.1 send-community
!
ip bgp-community new-format
!
! access lists for all the Premium customers on Rpe1
access-list 1 permit 192.168.60.0 0.0.0.255
access-list 2 permit 192.168.70.0 0.0.0.255
!
! advertise this Premium Customer with Community Premium (ie 210:4)
route-map mark_community permit 10
match ip address 1
set community 210:4
!
! advertise this Premium Customer with Community Premium (ie 210:4)
route-map mark_community permit 10
match ip address 2
set community 210:4
!
! advertise all other Customers with Community Standard (ie 210:0)
route-map mark_community permit 30
set community 210:0
!
ip community-list 1 permit 210:4
!
! configure QPPB to set Precedence to 4 on packets destined to an
! address whose BGP advertisement has been received with Community 210:4
route-map receive_community permit 10
match community 1
set ip precedence 4
set ip qos-group 4
!
! configure QPPB to set Precedence to 0 otherwise
route-map receive_community permit 20
set ip precedence 0
set ip qos-group 0
!
```

### Rwest Configuration

Rwest is a backbone router. It is a router from the server provider network that interconnects its POP with other remote POPs through the ATM WAN cloud. The IP to ATM CoS Phase 1 per-VC WRED feature is activated on the two ATM PVCs connecting Rwest respectively to Reast and to Rsouth. The ATM PVCs are to be shaped as VBR PVCs with a PCR of 12 Mbps and an SCR of 8 Mbps. Because these two PVCs have identical shaping rates, the same WRED profile is configured on the two VCs.

The IP QoS WRED feature is activated on the POSIP link to Rnorth. Use of WRED over POSIP and over ATM allows for consistent end-to-end support.

```

!
ip cef distributed
!
interface POS0/0/0
  description SDH/POSIP interface connecting to the SDH WAN cloud
  ip address 17.10.0.1 255.255.255.0
  random-detect
  ip route-cache distributed
  no keepalive
  clock source internal
  no cdp enable
!
interface ATM11/0/0
  description ATM interface to the ATM WAN cloud
  no ip address
  no ip route-cache optimum
  ip route-cache distributed
  no ip mroute-cache
  no keepalive
!
interface ATM11/0/0.100 point-to-point
  description ATM PVC to Reast
  ip address 17.0.0.1 255.255.255.0
  atm pvc 100 5 100 aal5snap 12000 8000 512 inarp
  random-detect fast-wred-profile
!
interface ATM11/0/0.101 point-to-point
  description ATM PVC to Rsouth
  ip address 17.0.1.1 255.255.255.0
  atm pvc 101 5 101 aal5snap 12000 8000 512 inarp
  random-detect fast-wred-profile
!
interface FastEthernet2/0/0
  description Interface connecting Rwest to the intra-POP switch
  ip address 17.1.1.2 255.255.255.0
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
!
random-detect-group fast-wred-profile
  precedence 0 72 120 10
  precedence 1 72 120 10
  precedence 2 72 120 10
  precedence 3 72 120 10
  precedence 4 110 144 30
  precedence 5 110 144 30
  precedence 6 144 200 100
  precedence 7 144 200 100
!

```

## Examples of Show Commands

The **show queueing interface** command confirms that per-VC WRED performs selective packet drops with lower drop probability on the premium traffic than on the standard traffic in a situation of congestion. It also confirms that it does only exceptional dropping of routing packets and that it does not drop any packets due to buffer shortage on the VIP.

```
Rwest# show queueing interface atm 11/0/0.100
VC 5/100 -
  ATM11/0/0.100 queue size 63
    packets output 150460, drops 29071, nobuffer drops 0
  WRED: queue average 75
    weight 1/512, max available buffers 1021

  Precedence 0: 72 min threshold, 120 max threshold, 1/10 mark weight
    133023 packets output, drops: 17402 random, 11343 threshold
  Precedence 1: 72 min threshold, 120 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 2: 72 min threshold, 120 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 3: 72 min threshold, 120 max threshold, 1/10 mark weight
    (no traffic)
  Precedence 4: 110 min threshold, 144 max threshold, 1/30 mark weight
    15703 packets output, drops: 211 random, 112 threshold
  Precedence 5: 110 min threshold, 144 max threshold, 1/30 mark weight
    (no traffic)
  Precedence 6: 144 min threshold, 200 max threshold, 1/100 mark weight
    1734 packets output, drops: 3 random, 0 threshold
  Precedence 7: 144 min threshold, 200 max threshold, 1/100 mark weight
    (no traffic)
```

The **show atm vc** command confirms that no packets are being dropped on the PA.

```
Rwest# show atm vc 100
ATM11/0/0.100: VCD: 100, VPI: 5, VCI: 100, etype:0x0, AAL5 - LLC/SNAP, Flags: 0xC0030
PeakRate: 12000, Average Rate: 8000, Burst Cells: 512, VCmode: 0x0
OAM DISABLED, InARP frequency: 15 minute(s)
InPkts: 1350, OutPkts: 1340, InBytes: 94160, OutBytes: 92560
InPProc: 1350, OutPProc: 0, Broadcasts: 1330
InFast: 0, OutFast: 0, InAS: 0, OutAS: 1
InPktDrops: 0, OutPktDrops: 0
CrcErrors: 0, SarTimeOuts: 0, OverSizedSDUs: 0
OAM F5 cells sent: 0, OAM cells received: 0
Status: ACTIVE
Rwest#
```

## Case Study 3: European ISP Offering Premium Internet Access over Transatlantic ATM Link

As the price of intercontinental bandwidth (bandwidth between Europe/Asia and the United States) remains high compared to domestic bandwidth, this international bandwidth is often a scarce resource for European/Asian IP server providers. It is typically the most narrow bottleneck for Internet access from Europe or Asia. Large ISPs acting as network service providers (NSPs) for smaller ISPs or offering Internet access to business customers see a key business opportunity in generating direct revenue from this asset and want to control how the transatlantic bandwidth is distributed among competing users. In other words, they want to change their business model and reflect the cost of provisioning international bandwidth to the downstream customer (local ISPs, academic customers, business customers) based on the level of bandwidth overbooking that each customer is contracting for and experiencing.

To succeed, a European/Asian ISP can offer multiple levels of differentiated services on Internet access. In that service, an end user can contract a premium Internet access. Grouping the customers that have the same level of requirements for Internet access makes it possible for the ISP to apply more generous resource provisioning and lower overbooking ratio that is appropriate for that pool of premium customers and offer them the required cost/performance point. Of course, the service could be enhanced further to offer more than two CoS for Internet access.

Because a very high proportion of Internet content is pulled from servers located in the United States and because the most stringent bottleneck for Internet access from Europe and Asia is often the transatlantic link, it is often sufficient to apply IP QoS mechanisms on that transatlantic link and only in the direction from the United States. There is typically little discard in the domestic network because it is usually far less congested or not congested at all.

This case study presents a European ISP network design supporting two levels of international access to the Internet (standard and premium) where the international bandwidth is provided through an ATM connection. In this environment, the service differentiation is provided through the IP to ATM CoS feature activated on a router on the U.S. side of the transatlantic link.

The main design goals are as follows:

- Offer two levels of international Internet access by performing separate congestion management for each CoS over the most stringent bottleneck (the international ATM connection from the United States to Europe)

---

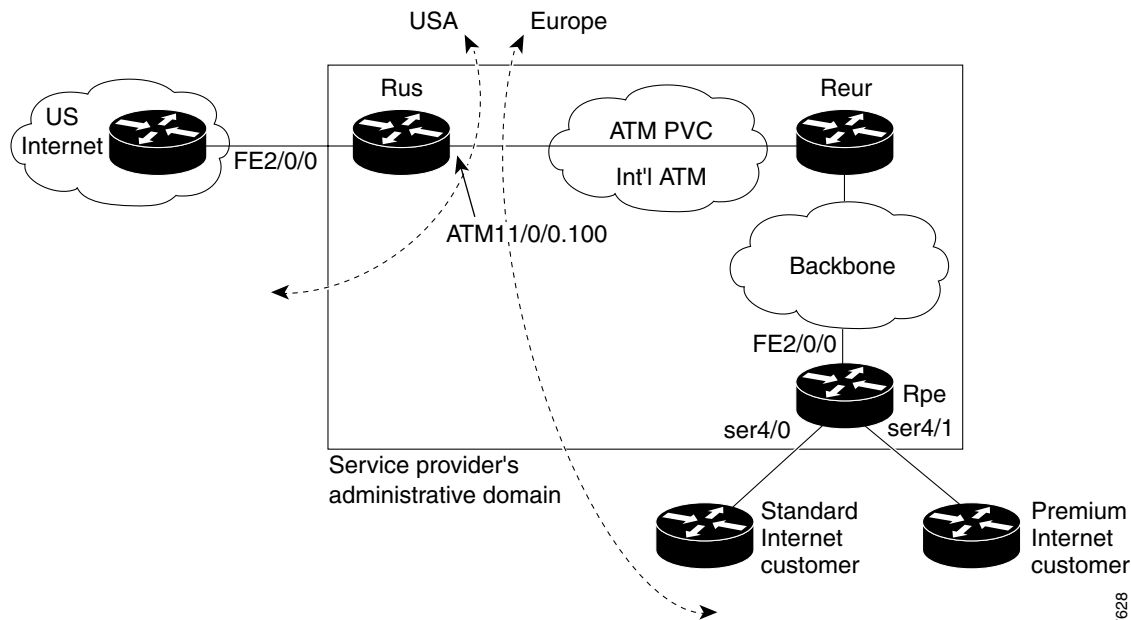
**Note** If the bandwidth in the direction of Europe to the United States is also a significant bottleneck, then the IP to ATM CoS Phase 1 feature should also be applied on Reur (see Figure 9) over the ATM PVC toward the United States, allowing service differentiation to also be enforced on traffic traveling from Europe to the United States. The precedence of packets from Europe to the United States would then be marked by CAR when the packets enter the ISP network from the end customers' sites. Such a situation may occur if the international ATM server provider offers ATM PVCs with asymmetrical bandwidth; the European ISP could use an ATM PVC with a high bandwidth from the United States to Europe and a lower bandwidth from Europe to the United States to reflect the asymmetry of Internet traffic.

---

- Ensure that, even in the case of congestion routing, traffic is protected so that routing stability is maintained

Figure 9 illustrates the network topology for this case study.

Figure 9 Network Topology for Case Study 3



17628

### Rus Configuration

Rus is an interconnect router sitting in the United States. It supports the interconnection to the U.S. Internet network provider on one side and supports the ATM international connection to the European backbone.

Rus can be operated by one of the following:

- The European server provider.
- The U.S. Internet network provider but with some configuration access being given to the European ISP so that the European ISP can configure the service differentiation features (for example, IP to ATM CoS Phase 1 and QPPB) and parameters.
- Entirely by the U.S. Internet network provider with a special peering agreement with the European ISP that includes an agreed service differentiation configuration (for example, IP to ATM CoS Phase 1 and QPPB). QPPB can then be used to ensure that the configuration on Rus is totally independent of whichever of the European ISPs happens to be standard and premium. The European ISP configures its customers in the standard and premium services. QPPB allows BGP to dynamically notify Rus about which customers are premium and which customers are standard so that the configuration in Rus is independent of the changes/additions/deletions of standard/premium customers in the European ISP network.

QPPB is used so that Rus is dynamically notified through BGP about which prefixes correspond to standard international Internet access customers and which prefixes correspond to premium international Internet access customers. QPPB is also used to set the precedence to 0 in packets destined to the standard customers and to 2 in packets destined to the premium customers.

The IP to ATM CoS Phase 1 per-VC WRED feature is also activated on Rus over the ATM PVC toward Reur in Europe. The ATM PVC is shaped as a CBR PVC with a PCR of 10 Mbps, and the IP to ATM CoS phase 1 per-VC WRED feature performs the service differentiation on traffic destined to the standard international Internet access customers and destined to the premium international Internet access customers because the packets respectively have their precedences set to 0 and 2.

```
interface FastEthernet2/0/0
  mac-address 0001.0002.0001
  ip address 192.168.1.1 255.255.255.0
  ! remark precedence of all incoming traffic to 0
  rate-limit input 1000000 24000 24000 conform-action set-prec-transmit 0
    exceed-action set-prec-transmit 0
  ! activate QPPB on this interface
  bgp-policy destination ip-prec-map
  no ip route-cache optimum
  ip route-cache distributed
  no keepalive
  no cdp enable
  !
interface ATM11/0/0
  no ip address
  no ip route-cache optimum
  ip route-cache distributed
  no ip mroute-cache
  no keepalive
  !
interface ATM11/0/0.100 point-to-point
  ip address 17.0.0.1 255.255.255.0
  atm pvc 100 5 100 aal5snap 10000 10000 10 inarp
    random-detect internet-access
  !
router bgp 210
  ! tell BGP to take into account the "receive_community" route-map
  table-map receive_community
  neighbor 210.210.14.1 remote-as 210
  neighbor 200.200.14.1 update-source Loopback0
  !
  ip bgp-community new-format
  !
  ip community-list 1 permit 210:4
  !
  ! configure QPPB to set Precedence to 2 on packets destined to an
  ! address whose BGP advertisement has been received with Community 210:2
  route-map receive_community permit 10
    match community 1
    set ip precedence 2
  !
  ! configure QPPB to set Precedence to 0 otherwise
  route-map receive_community permit 20
    set ip precedence 0
  !
random-detect-group internet-access
  precedence 0 72 120 10
  precedence 1 72 120 10
  precedence 2 110 144 30
  precedence 3 110 144 30
  precedence 4 110 144 30
  precedence 5 110 144 30
  precedence 6 144 200 100
  precedence 7 144 200 100
  !
```

## Rpe Configuration

Rpe is a provider edge router. It is a router from the server provider network that sits on the edge and directly connects to routers on the customer premises. Every interface that receives packets from outside the service provider network (that is, from customer premises routers) performs re-marking of incoming traffic to precedence 0 on all packets from all customers because our design assumes that the only service differentiation is performed over the international connection for traffic traveling from the United States. However, prefixes of customers subscribing to the premium international Internet access are advertised by Rpe with a community distinguishing them as such so that Rus can then mark the precedence to 2 of packets destined to these customers through QPPB.

```
!
ip cef distributed
!
interface serial4/0
description Edge interface connecting to a Standard Internet Customer
ip address 192.168.50.1 255.255.255.0
! remark all incoming traffic to precedence zero
rate-limit input 100000 2400 2400
conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
encapsulation hdlc
ip route-cache
!
interface serial4/1
description Edge interface connecting to a Premium Internet Customer
ip address 192.168.60.1 255.255.255.0
! remark all incoming traffic to precedence zero
rate-limit input 100000 2400 2400
conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
encapsulation hdlc
ip route-cache
!
interface FastEthernet2/0/0
description Interface connecting Rpe to the intra-POP switch
ip address 17.1.1.1 255.255.255.0
no ip route-cache optimum
ip route-cache distributed
no keepalive
no cdp enable

!
router bgp 210
neighbor 210.210.20.1 remote-as 210
neighbor 210.210.20.1 update-source Loopback0
! tell BGP to advertise prefixes with communities as defined in
! the "mark_community" route-map
neighbor 210.210.20.1 route-map mark_community out
neighbor 210.210.20.1 send-community
!
ip bgp-community new-format
!
! access lists for all the Premium Internet Customers on Rpe
access-list 5 permit 192.168.60.0 0.0.0.255
!
! advertise this Premium Customer with Community Premium (ie 210:2)
route-map mark_community permit 10
match ip address 5
set community 210:2
!
! advertise all other Customers with Community Standard (ie 210:0)
route-map mark_community permit 30
set community 210:0
!
```

## Case Study 4: Service Differentiation for Enhanced Multimedia in Enterprise Networks

This case study presents an enterprise network where the IP to ATM CoS Phase 1 feature is used to prioritize multimedia traffic (multimedia training distributed by a server running the Cisco IP/TV application) over other data traffic (FTP) to enhance the quality of the multimedia application and maximize effectiveness of the employees relying on these applications.

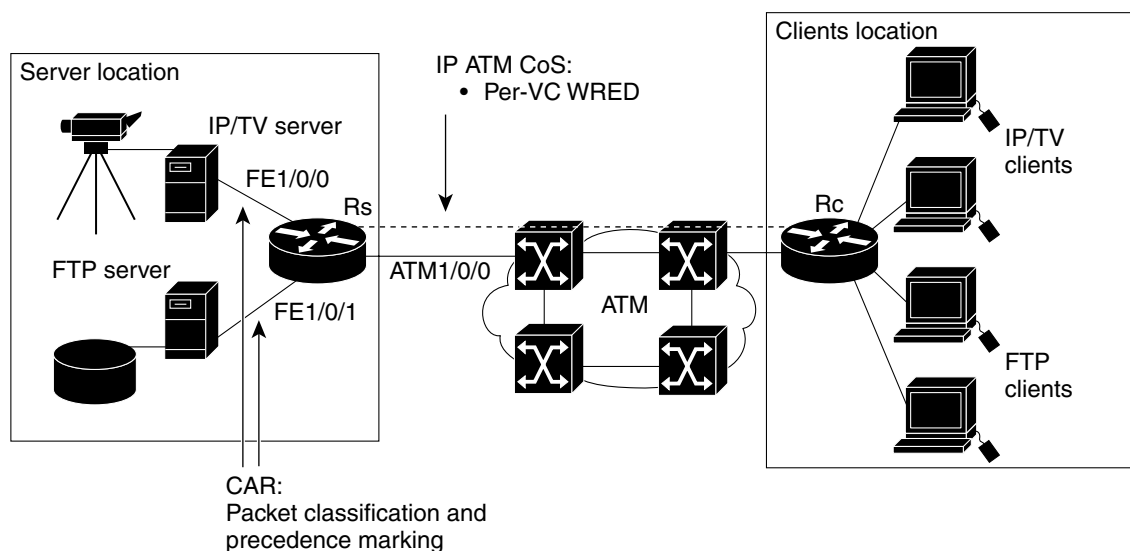
Clients in one location access various applications onto servers in a remote location. These applications include on-demand training video, live broadcast training multimedia sessions, and heavy FTP transfers. Connectivity between the locations is provided through an ATM PVC between the two site routers. The ATM PVC bandwidth is a possible bottleneck.

The main design goals are as follows:

- Ensure that in case of congestion of the ATM PVC the multimedia applications over the FTP traffic are prioritized so that the multimedia application users experience proper quality for their audio, video, and whiteboard streams without being affected by the congestion
- Ensure that the FTP traffic can use the bandwidth unused by the multimedia applications
- Ensure that in case of congestion the FTP traffic is dynamically and effectively throttled to adapt to the remaining bandwidth unused by the multimedia applications without impact on the multimedia applications

Figure 10 illustrates the network topology for this case study.

**Figure 10 Network Topology for Case Study 4**



To achieve these goals, CAR is used on the router Rs connecting the servers to mark the precedence of incoming traffic depending on the server it is coming from: CAR marks as high priority (precedence 4) all the traffic sent by the IP/TV server that is the multimedia traffic, and CAR marks as low priority all the traffic coming from the FTP server. Then the IP to ATM CoS Phase 1 feature is used on Rs to prioritize the multimedia traffic by enforcing a more aggressive WRED discard on the low priority traffic so that in case of congestion, through TCP's adaptive flow control, the FTP transfers get throttled back to the bandwidth left over by the multimedia applications on the ATM PVC.

---

**Note** Because the IP to ATM CoS Phase 1 feature relies on WRED probabilistic discard, this design only operates efficiently when the low priority traffic is adaptive and implements strict flow control (such as TCP traffic as assumed in this case study). If the low priority traffic does not dynamically adapt to loss as expected from TCP (for example, because the TCP implementation does not back off as recommended by the IETF or because there is a significant volume of non-TCP traffic), then the low priority traffic will not adapt to the bandwidth left over by the multimedia applications and may keep sending at a higher rate than can be accepted over the ATM PVC. In this case, the IP to ATM CoS Phase 1 feature often will operate in tail drop mode for the low priority traffic with high levels of drops on the low priority traffic.

---

In this case study, the video streams transferred by the IP/TV server are MPEG1 at 1.5 Mbps each, and the ATM PVC between the two locations is a CBR PVC at 5 Mbps. Thus, when two clients are simultaneously pulling a different video stream in the remote location, the FTP traffic is left with about 2 Mbps of bandwidth. When a third client starts pulling a different video stream in the remote location, the FTP traffic must adjust to the reduced remaining bandwidth inside the ATM PVC of about 500 kbps.

Note that because the Cisco IP/TV application uses IP multicast, when multiple clients are following the same multimedia channel (for example, live broadcast training multimedia session), the replication can be performed in the remote client site by the router supporting the multiple clients (Rc). Thus only 1.5 Mbps of bandwidth is consumed by this session over the ATM PVC regardless of the number of clients simultaneously following that session.

## Rs Configuration

Rs is the router on the server site.

```

!
ip cef distributed
!
interface FastEthernet1/0/0
description interface connecting the IP/TV Server
ip address 192.168.70.1 255.255.255.0
! mark incoming traffic as High Priority (Precedence 4)
rate-limit input 10000000 24000 24000
    conform-action set-prec-transmit 4 exceed-action set-prec-transmit 4
no ip route-cache optimum
ip route-cache distributed
no keepalive
no cdp enable
!
interface FastEthernet1/0/1
description interface connecting the FTP Server
ip address 192.168.80.1 255.255.255.0
! mark incoming traffic as Low Priority (Precedence 0).
rate-limit input 10000000 24000 24000
    conform-action set-prec-transmit 0 exceed-action set-prec-transmit 0
no ip route-cache optimum
ip route-cache distributed
no keepalive
no cdp enable
!
interface ATM11/0/0
description ATM interface to the ATM WAN cloud
no ip address
no ip route-cache optimum
ip route-cache distributed
no ip mroute-cache
no keepalive
!
interface ATM11/0/0.100 point-to-point
description ATM PVC to Rc
ip address 192.168.100.1 255.255.255.0
atm pvc 100 5 100 aal5snap 12000 8000 512 inarp
random-detect client-server-wred-profile
!
!
random-detect-group client-server-wred-profile
precedence 0 50 120 10
precedence 1 50 120 10
precedence 2 50 120 10
precedence 3 50 120 10
precedence 4 110 144 30
precedence 5 110 144 30
precedence 6 144 200 100
precedence 7 144 200 100
!

```

## Glossary

<b>Table 1</b>	<b>Glossary</b>
ABR	Available Bit Rate
ATM	Asynchronous Transfer Mode
CAR	Committed Access Rate
CBR	Constant Bit Rate
CCO	Cisco Connection Online
CEF	Cisco Express Forwarding
CLI	Command-Line Interface
DCEF	Distributed Cisco Express Forwarding
ISP	Internet Service Provider
NSP	Network Service Provider
PA	Port Adapter
PCR	Peak Cell Rate
POSIP	Packet over SONET/SDH Interface Processor
PVC	Permanent Virtual Circuit
QoS	Quality of Service
QPPB	QoS Policy Propagation via BGP
RED	Random Early Detection as proposed by Floyd and Jacobson
RSP	Route Switch Processor
RTT	Round Trip Time
SCR	Sustainable Cell Rate
ToS	Type of Service
UBR	Unspecified Bit Rate
VBR	Variable Bit Rate
VBR-nrt	Variable Bit Rate-nonreal time
VBR-rt	Variable Bit Rate-real time
VIP	Versatile Interface Processor
WFQ	Weighted fair queueing: Queueing algorithm that provides bandwidth allocations to queues of different classification
WRED	Weighted Random Early Detection: The latter of two Cisco implementations of the RED algorithms and model. Differs from RED primarily in that different “knobs” are provided for each IP Precedence value.