

# Internet Processor II ASIC: 速率限制和通信策略特性

Juniper 网络公司，爱立信公司，2000 年

内容提要.....	2
速率限制.....	2
通信整形.....	2
采用漏桶算法的通信平滑工具.....	3
采用令牌桶算法的突发整形工具.....	3
通信策略.....	4
硬通信策略和软通信策略应用.....	6
<b>INTERNET PROCESSOR II ASIC：速率限制和通信策略特性.....</b>	<b>7</b>
INTERNET PROCESSOR II ASIC：通信策略设计目标.....	8
在所有接口上为软硬通信策略提供支持.....	9
改善软硬通信策略工具的细密度.....	9
INTERNET PROCESSOR II ASIC：软硬通信策略运行模型.....	10
分组分类程序.....	11
流量监视程序.....	11
转发操作.....	12
<b>INTERNET PROCESSOR II ASIC：通信策略配置实例.....</b>	<b>12</b>

## 内容提要

随着用户对Internet服务需求的提高，服务供应商必须确定共享的网络资源在不同的用户、使用者和应用之间的分配方式。如果服务供应商不能管理通信进入核心传输网络的流量和速度，那么服务供应商就很难管理为每个用户提供的基本服务水平。速率限制和通信策略工具对成功的网络运营非常关键，因为它们允许服务供应商决定哪些通信进入其网络，网络接纳这些通信的流量和速率，以及在网络拥塞程度随着通信负载变化时，路由器的每站丢弃行为。

本文介绍了Juniper网络公司的M系列Internet主干路由器是怎样实现速率限制工具的，以及JUNOS™ Internet软件和Internet Processor II™ ASIC是怎样支持速率限制工具的。它简要介绍了速度限制机制，描述了服务供应商可以怎样使用这些工具，为分层服务提供支持，同时增强网络的性能和稳定性。本文把重点放在JUNOS软件和Internet Processor II ASIC支持的速率限制的具体设计目标、运行模型和配置实例上。由于这些机制与服务等级（CoS）特性密切相关，且与CoS特性交互，因此本文假设读者已经从整体上理解了CoS概念，以及M系列路由器支持CoS的方式。

## 速率限制

开通率低的统计多路复用网络经常会导致长期拥塞。在刚刚开通接入电路时，用户和服务供应商会就链路上的数据传输平均速率（及突发性）达成一致。只要用户履行了自己的那部分合同，并根据服务水平协议（SLA）传输分组，服务供应商就会努力及时地传送这些分组。及时传送并仅仅不是传统数据应用的要求，它对新兴的实时应用也同样关键，如IP语音（VoIP）和桌面会议，因为这些应用都不能容忍网络拥塞带来的延迟和抖动。

为了支持所有用户的各种不同SLA，关键在于服务供应商要调节通信流量，保护网络核心中共有的资源，确保每个用户不会消耗高于其平均份额的带宽。为此，服务供应商需要多种工具，确定每个用户是否遵守SLA，以及在用户试图把配置文件之外的通信注入网络时应该采取哪些措施。

保护服务供应商网络核心中的共享资源有两种方法：

- 通信整形
- 通信策略

## 通信整形

通信整形主要是通过队列头使用整形器来对分组进行排序，从而减少网络拥塞的概率。这些工具平滑分组流，调节网络接纳的通信速率和流量。有两种基本的通信整形速率限制工具：

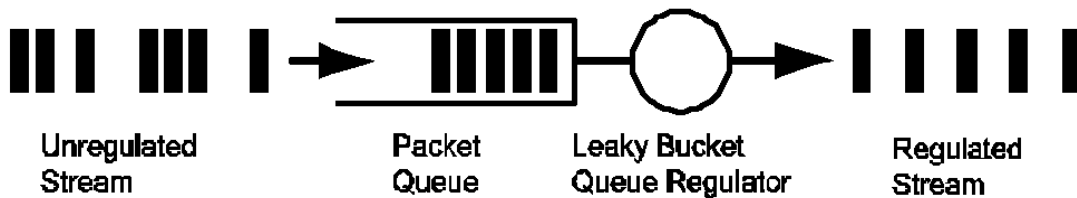
- 通信平滑工具消除了突发性，为网络提供一条稳定的通信流。这种速率限制功能通常使用漏桶算法实现。
- 长期平均通信速率整形工具，允许预先确定的突发尺寸，为网络提供一条调节突发性的通信流。这种速率限制功能通常使用令牌桶算法实现。

每种工具都支持不同的速率限制功能，可以生成具有不同特点的输出流。

### 采用漏桶算法的通信平滑工具

漏桶速率限制算法把突发的分组流转化成常用的空间相等的分组流。图一说明了漏桶速率算法的运行方式。

Figure 1: Traffic-smoothing Using a Leaky Bucket Algorithm



图一 采用漏桶算法的通信平滑技术

Unregulated Stream: 未调节的通信流

Packet Queue: 分组队列

Leaky Bucket Queue Regulator: 漏桶队列调节器

Regulated Stream: 调节的信息流

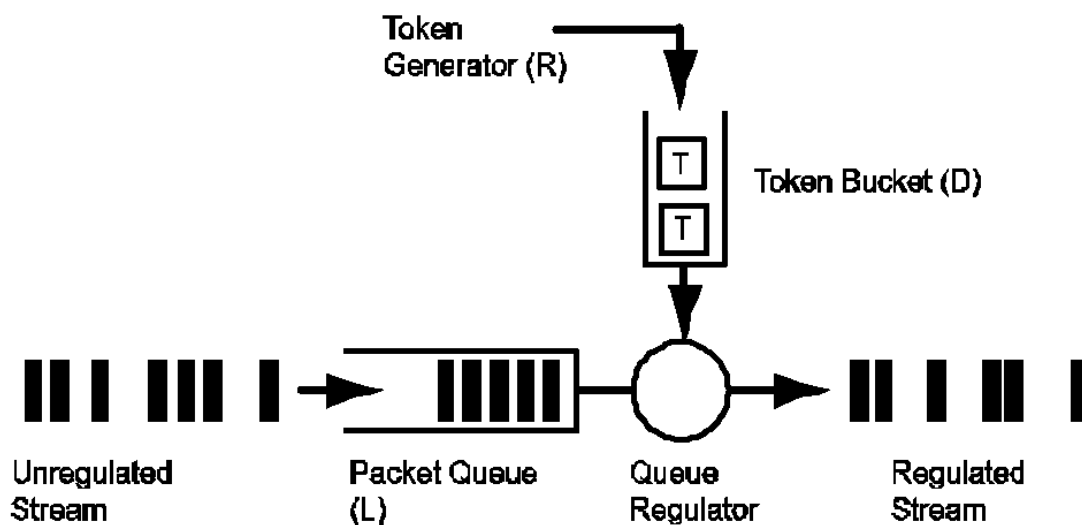
通过漏桶速率限制算法，未调节的分组流被放入漏桶队列调节器控制的分组队列中。如果这条分组流代表的分组超过队列可以存储的分组，那么多出的分组将被丢弃。在分组到达队列头时，它们以队列调节器配置决定的恒定速率转发到网络中。

您可以使用漏桶速率限制工具管理数据流，这样分组不会以高于网络可以吸收或愿意吸收的速率转发到网络中。分组队列的长度（或深度）限制着分组在端到端路径中的这个通信整形器上可能导致的延迟数量。但是，如果网络其余部分未能充分配置，以支持提供的负载，那么分组可能会在下行站上引发额外的延迟。

### 采用令牌桶算法的突发整形工具

令牌桶速率限制算法强制实现长期平均传输速率，同时允许约定的突发速率。在这种方法中，令牌桶用来管理队列调节器，队列调节器则控制着分组流进入网络的速率。图二说明了令牌桶速率限制算法的运行方式。

Figure 2: Burst Shaping Using a Token Bucket Algorithm



图二 采用令牌桶算法的突发整形技术

Token Generator (R): 令牌生成器 (R)

Token Bucket: 令牌桶

Unregulated stream: 未调节的流

Packet Queue (L): 分组队列 (L)

Queue Regulator: 队列调节器

Regulated stream: 调节后的流

令牌生成器以每秒 $R$ 个令牌的速度不断生成令牌，并把它放在深度为 $D$ 个令牌的令牌桶中。假设每个令牌获准传输固定数量的字节，如果令牌桶装满，新生成的令牌将被丢弃。同时，未调节的分组流会到达，被放入最大长度为 $L$ 的分组队列中。如果分组提供的分组数超过队列可以存储的数量，超过的分组将被丢弃。

在决定能否把由 $P$ 尺寸的令牌组成的分组转发到网络中时，队列调节器必须考虑许多因素：

- 如果令牌桶是满的，分组被转发到网络中， $P$ 令牌从桶中清除。
- 如果令牌桶是空的，分组在队列头上等待，直到 $P$ 令牌被生成，并被放在桶中。当桶逐渐包含 $P$ 个令牌时，许多令牌将从桶中清除，分组被发送到网络中。
- 最后，假设令牌桶只是部分装满，含有 $T$ 个令牌。如果 $P$ 小于等于 $T$ ， $P$ 个令牌将从桶中清除，分组被转发到网络中。如果 $P$ 大于 $T$ ，那么分组必须等待其余的 $P$ 减去 $T$ 个令牌，然后可以发送到网络中。当桶中包含所需的 $P$ 个令牌时，许多令牌将从桶中清除，分组被转发到网络中。

您可以使用令牌桶速率限制工具，调节长期平均传输速率，同时允许预先确定的突发尺寸。令牌生成器的速率决定着长期平均通信速率，令牌桶的深度决定着整形器允许的最大突发尺寸，分组队列的长度决定着分组在这个通信整形器上可以引发的延迟数量。

## 通信策略

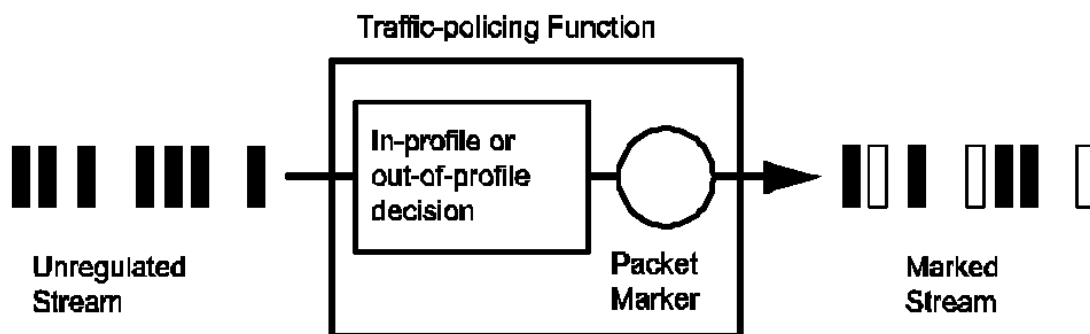
通信策略工具允许检查用户通信流量，丢弃或标出超过SLA的分组。通信策略功能使用令牌

桶算法，但分组队列被换成分组丢弃或分组标记功能。如果策略功能确定某个分组位于配置文件中，那么网络可以接纳该分组。

如果策略功能确定分组位于配置文件之外，那么该网络将被立即丢弃（硬策略），或网络不会接纳该分组，而是标为位于配置文件之外（软策略）。

标记位于配置文件之外的通信(图三)允许从执行通信策略功能的路由器下行方向的各站上，以不同方式处理配置文件之内的分组和配置文件之外的分组。例如，您可以把一台通信策略接入路由器配置成通过改变其丢弃优先级来标记分组，这样核心路由器可以在拥塞过程中，为该分组提供更高的丢弃优先级，同时继续交付配置文件之内的通信。

**Figure 3: Soft Traffic-policing Function**



图三 软通信策略功能

Traffic Policing Function: 通信策略功能

Unregulated Stream: 未调节的流

In-profile or out-of-profile decision: 决定其位于配置文件之内还是位于配置文件之外

Packet Marker: 分组标记器

Marked Stream: 标记后的流

在使用通信策略程序标记分组时，您必须认真确保策略程序能够保持分组在流量中的排列顺序。通过提高其丢弃优先级来标记分组，可以提高核心路由器在网络拥塞过程中丢弃分组的概率。分组排列顺序是可以保持的，因为在网络中的每一站，流量中的所有分组都会被赋予相同的队列。

应该避免根据通信策略程序的配置文件之内或配置文件之外的决策，把分组分配给不同的传输队列，以此来标记分组的作法。把一条通信流量分隔到多个队列中，可能会导致分组重新排序，这可能会给流量的吞吐量性能带来负面影响。在分开的流量流经网络核心中的多个站（每个站具有不同的队列）时，会提高分组重新排序的概率。

采用通信整形工具的速率限制技术具有重要意义，在通信进入网络时，它可以精确地控制通信速率和通信流量，但是这些工具不支持动态反馈机制，而动态反馈机制允许其适应和利用网络核心中临时提供的容量。采用通信整形工具的速率限制技术较通信整形提供了一个更加灵活的解决方案，因为它允许资源保护方案适应不断变化的网络条件。

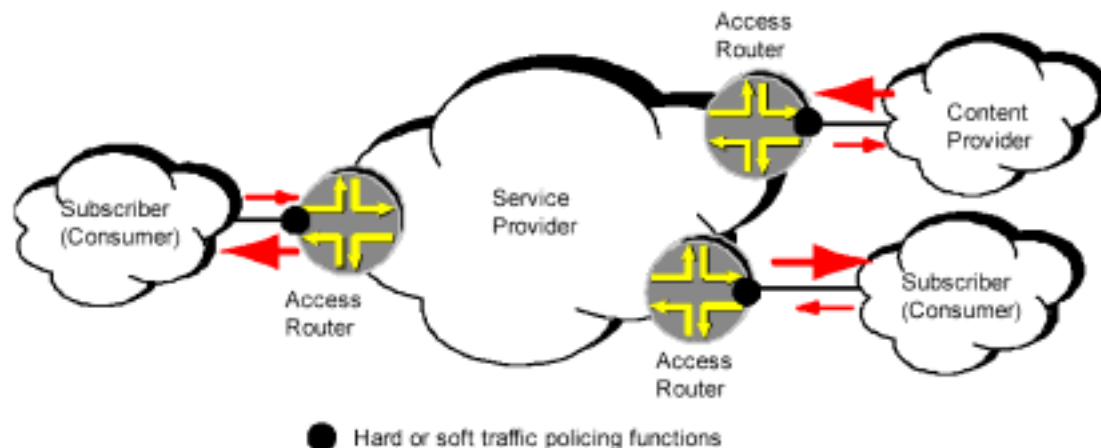
硬通信策略（丢包）类似于通信整形，而软通信策略（分组标记）则是一种完全不同的方法。它能够把丢包的决策延迟到位于网络内部的系统上，允许只在网络核心真正存在拥塞时才丢

弃通信。在带宽充足的时期，软通信策略允许服务供应商为所有用户提供更高水平的服务，同时仍能够保护共享的资源，在带宽缺乏时满足用户SLA。

### 硬通信策略和软通信策略应用

硬通信策略和软通信策略工具可以部署在网络边缘，提供分层用户服务，保护共享的资源。图四说明了可以怎样使用这样工具，限制进入网络核心的用户通信流量，及管理为用户或其它服务供应商提供的带宽。

Figure 4: Hard and Soft Traffic Policing Applications



图四 硬通信策略和软通信策略应用

- Access Router: 接入路由器
- Subscriber (Consumer): 用户 (消费者)
- Access Router: 接入路由器
- Service Provider: 服务供应商
- Access Router: 接入路由器
- Content Provider: 内容供应商
- Subscriber (Consumer): 用户 (消费者)

由于Internet应用的基本特点，您经常需要在网络边缘处理不对称的通信流量。例如，接入Web的内容消费者要求的拉动（下载）带宽通常是推送（上载）带宽的10倍。另一方面，内容供应商需要的推送带宽要超过拉动带宽。通信策略为您提供了多种工具，以精确地控制带宽，根据具体的用户或应用要求提供分层SLA。

下面的检查清单确定了采用硬通信策略或软通信策略的速率限制技术在网络边缘使用的几种应用。

- 通信策略可以用来在T1（或E1）接口上开通低速率接入电路。例如，通过在用户推送方向中提供256 Kbps带宽，在用户拉动方向中提供512 Kbps带宽，您可以把通信策略程序配置成支持不对称的通信负载。
- 可以在逻辑接口上使用通信策略，对一条通信流中的所有分组应用单一的速率限制。例如，您可以使用通信策略，根据为用户的逻辑接口提供的带宽数据创建分层VPN服务。
- 可以根据一个逻辑接口内部每种类别的SLA，使用通信策略限制流量速率。例如，在用户DS3（45-Mbps）接入电路上，您可以把2 Mbps的通信标为金牌DiffServ编码点值，把

5 Mbps的通信标为银牌DiffServ编码点值,而把无限制的通信流量标为铜牌DiffServ编码点值。

- 可以使用通信策略,根据逻辑接口内部的第四层配置文件限制流量速率。例如,您可以允许不受限制的用户HTTP通信流量,但把NNTP通信速率限制为某一带宽门限。
- 可以使用通信策略限制ICMP通信速率,防止拒绝服务(DoS)攻击。例如,smurf攻击可能导致大量的ICMP回声响分组被传送到攻击目标上。这可能会导致严重的网络拥塞或中断。您可以与其它分组过滤工具一起使用通信策略技术,使网络资源免受这些类型的DoS攻击。
- 可以在逻辑接口内部使用分组过滤(没有速率限制),根据第四层配置文件对通信进行排队。例如,您可以把SMTP通信分配给低优先级队列,把VoIP通信分配给高优先级队列。
- 上述方法的任意组合。

在评估在服务供应商网络的不同区域中部署通信策略工具的适当性时,需要考虑许多因素:

- 实现方案的分组分类工具的精细度是多少?如果通信策略机制全面集成了细密的分组过滤工具,应该考虑在详细检查包头的基础上,对通信流量进行分类和调节。
- 实现通信策略特性对性能有什么影响?如果通信策略功能在硬件中执行,可以预计其将给路由器转发性能带来可以预测的影响。这允许在网络中部署通信策略,而不会出现通常与传统基于软件的解决方案相关的性能降级。

### Internet Processor II ASIC: 速率限制和通信策略特性

Juniper Networks®已经发展了M系列路由器,通过频频推出不同版本的JUNOS软件,支持速率限制、软硬通信策略和CoS特性:

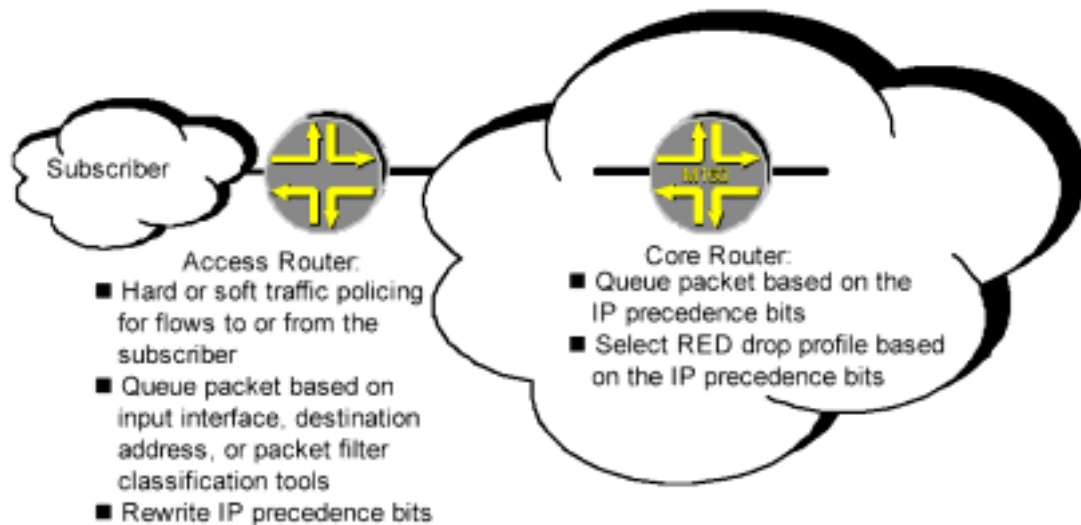
- JUNOS软件3.3版(1999年8月)推出对核心路由器CoS特性的支持。这版软件支持在M20™和M40™ Internet主干路由器上配置基于硬件的CoS特性。它提供了丰富的CoS功能,包括输入第二层通信策略、输出第二层速率限制、配置输出分组传输队列的长度、定义为某个输出队列分配哪些分组、使用加权循环(WRR)算法为每个队列调度传输优先级、使用随机早期检测(RED)算法管理每个队列的拥塞。所有这些机制都在硬件中实现,确保在系统管理实现CoS特性时,提供稳定的、可预测的转发性能。
- JUNOS软件3.4版(1999年12月)扩展了我们的CoS功能,支持接入路由器接入相关特性。这版软件允许接入路由器在包头的IP优先位中对分组的优先级编码。网络核心中的路由器可以读取IP优先位的值,确定接入路由器为每个分组分配的输出队列和RED丢弃优先级。当M系列路由器作为接入路由器部署时,它可以根据输入接口、输入接口利用率或信宿IP地址等项目,确定每个分组的分组队列和RED丢弃优先级。在制订这一决策后,接入路由器可以在把分组传输到网络核心之前,在IP优先位中写入一个适当的值,来指明分组的排队和RED丢弃优先级。
- JUNOS软件4.0版(2000年4月)没有提供新的速率限制、通信策略或CoS特性。但是,它推出了对M160™ Internet主干路由器和Internet Processor II ASIC的支持。Internet Processor II ASIC支持40 Mbps的转发速率,同时提供了丰富的分组处理特性(如过滤、计数、取样、记录和负载均衡),使服务供应商能够在迅速变化的环境中极具竞争力地构建和管理其网络发展。JUNOS软件对Internet Processor II ASIC的支持是一个关键步

骤，因为它打下了良好的基础，允许我们增强通信策略工具的细密度，同时在全系列M系列PIC中为其提供支持。

- JUNOS软件4.1版（2000年8月）扩展了Internet Processor II ASIC的功能，在所有输入接口和输出接口上支持软硬通信策略。这一版本允许M系列路由器定义分组过滤器，由分组过滤器调用通信策略程序，丢弃分组或在分组通知记录标记丢包优先权（PLP）位。
- JUNOS软件4.2版（2000年10月）扩展了Internet Processor II ASIC的功能，允许M系列路由器定义分组过滤器（没有速率限制），根据第四层配置文件把分组分配到输出传输队列上。

图五说明了可以以何种方式在接入核心路由器上实现我们基于硬件的通信策略和CoS特性，提供了一个完善的方法，在网络中支持通信策略和差别化服务。

Figure 5: JUNOS CoS Features on Access and Core Routers



图五 接入和核心路由器上的JUNOS CoS特性

Subscriber: 用户

Core Router: 核心路由器

Queue packet based on the IP precedence bits : 根据IP优先位对分组进行排队

Select RED drop profile based on the IP precedence bits : 根据IP优先位选择RED丢弃配置文件

Hard or soft traffic policing for flows to or from the subscriber : 为用户的通信流量提供硬或软通信策略

Queue packet based on input interface, destination address, or packet filter classification tools : 根据输入接口、信宿地址或分组过滤分类工具对分组进行排队

Rewrite IP precedence bits : 重写IP优先位

Access Router : 接入路由器

### Internet Processor II ASIC : 通信策略设计目标

JUNOS软件4.1版和4.2版推出了软硬通信策略特性，增强Internet Processor II ASIC支持的丰富的分组处理功能。我们的开发工作有两个主要的设计目标：

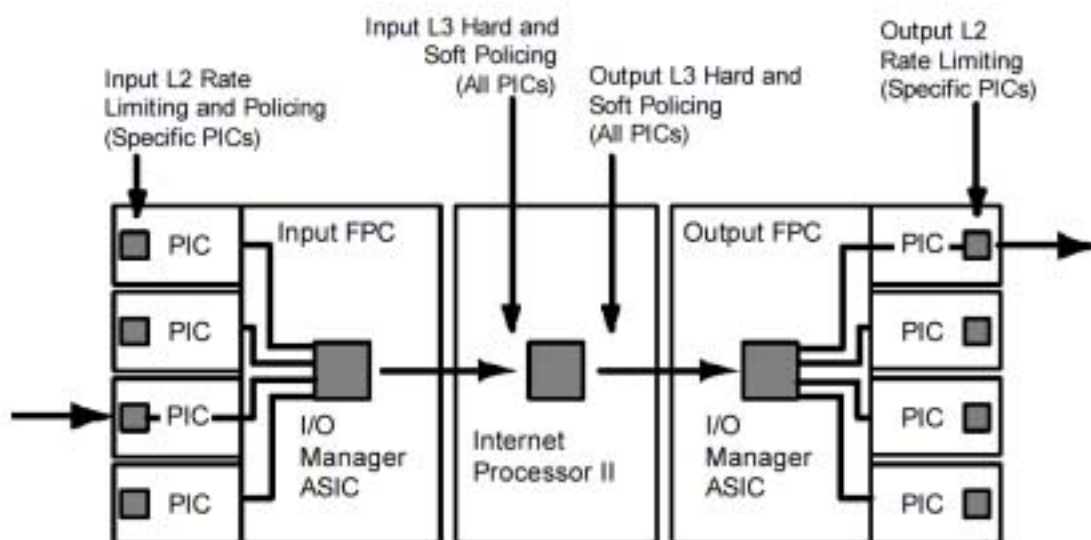
- 在所有接口上为软硬通信策略提供支持
- 增强我们软硬通信策略工具的细密度

## 在所有接口上为软硬通信策略提供支持

在JUNOS软件4.0版中，第二层速率整形和通信策略功能使用令牌桶算法，由介质专用ASIC在每个PIC上实现。但是，只有数量有限的PIC支持这些第二层功能，因此在所有接口上没提供这些特性。由于T1、E1、E3、快速以太网和千兆位以太网接口是接入路由器上部署最广泛的接口，因此绝对需要在这些接口上支持基于硬件的通信策略。

为了实现这一设计目标，JUNOS软件4.1版和4.2版采用在Internet Processor II ASIC（参见图六）上执行的第三层软硬通信策略工具。在集中化资源上实现这些特性，如Internet Processor II ASIC，允许在流经系统的所有分组（输入和输出）上执行第三层软硬通信策略，而不管机箱中安装的是哪种PIC。通过简单地升级到Internet Processor II ASIC，您可以简便地在所有接口上增加基于硬件的通信控制功能，而不必更换一块线路卡。

**Figure 6: Supporting Hard and Soft Traffic Policing on All Interfaces**



图六 在所有接口上支持软硬通信策略

Input L2 Rate Limiting and Policing (Specific PICs)：输入第二层速率限制和策略（具体PIC）

Input L3 Hard and Soft Policing (All PICs)：输入第三层软硬策略（所有PIC）

Output L3 Hard and Soft Policing (All PICs)：输出第三层软硬策略（所有PIC）

Output L2 Rate Limiting (Specific PICs)：输出第二层速率限制（具体PIC）

Input FPC：输入FPC

Output FPC：输出FPC

## 改善软硬通信策略工具的细密度

在JUNOS软件4.0版中，每个PIC上的介质专用ASIC都支持每条信道仅一个第二层速率限制程序或一个通信策略程序。每条信道仅有一个速率限制程序，这一基本限制是指只能对一个接口上的所有通信适应单一的门槛。尽管这种功能对核心路由器已经足够了，但是接入路由器要求更高的细密度水平，以支持部署新的创收型分层服务和实施SLA。

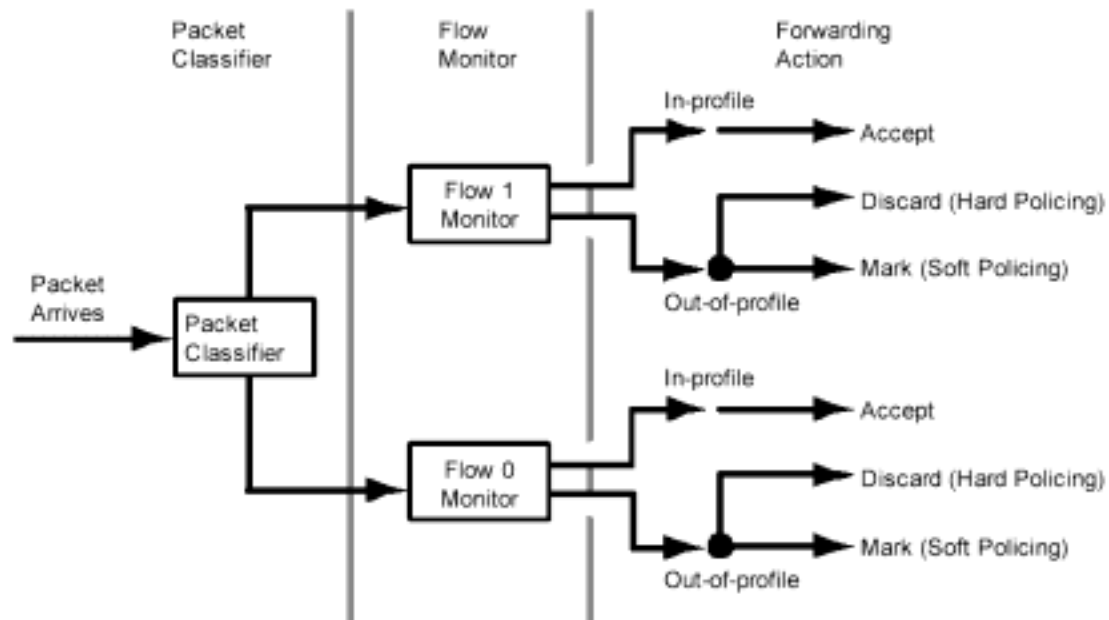
为了实现这一设计目标，第三层软硬通信策略工具现在已经全面集成Internet Processor II ASIC的分组过滤功能。把软硬通信策略与分组过滤集成起来，允许您在详细检查每个分组通知记录的基础上，精确地识别和调节通信流量。您可以根据输入接口、信源和信宿IP地址、IP协议字段、信源和信宿UDP和TCP端口、DiffServ字节等等，把分组成成不同的通信流量。此外，在过滤器条件then子句中执行的过滤器措施可以修改分组通知记录中包含的信息，从而影响下行路由器处理标记分组的方式。

### Internet Processor II ASIC：软硬通信策略运行模型

Internet Processor II ASIC和JUNOS软件4.1版和4.2版支持的第三层软硬通信策略工具使用三种不同的处理组件执行任务（图七）：

- 分组分类程序
- 流量监视程序
- 转发操作

Figure 7: Traffic Policing Operational Model



图七：通信策略运行模型

Packet Classifier：分组分类程序

Flow Monitor：流量监视程序

Forwarding Action：转发操作

Packet Arrives: 分组到达

Packet Classifier：分组分类程序

Flow 1 Monitor: 流量1监视

In-profile：在配置文件之内

Accept：接收

Discard (Hard Policing)：丢弃（硬策略）

Out-of-profile：在配置文件之外

Mark (Soft Policing)：标记（软策略）

Flow 0 Monitor: 流量0监视  
 In-profile : 在配置文件之内  
 Accept : 接收  
 Discard (Hard Policing) : 丢弃 (硬策略)  
 Out-of-profile : 在配置文件之外  
 Mark (Soft Policing) : 标记 (软策略)

### 分组分类程序

分组分类程序检查分组流中的分组，把它们划分成策略同等类别 (PEC)。一个PEC是一个分组分类程序进行相同处理的分组集合。例如，分组分类程序根据分组的IP信宿地址和DiffServ字节进行分类。为了实现简化性，假设分组分类表只含有三种规则，如表一所示：

**Table 1: Example Packet Classification Table**

Rule	IP Destination Address	DiffServ Byte	PEC
2	192.100/16	Expedited Forwarding PHB (101110)	2
1	208.197.169/24	Assured Forwarding 2 Medium Drop PHB (010100)	1
0	*	*	0

**表一：分组分类表实例**

规则	IP信宿地址	地址DiffServ字节	PEC
2	192.100/16	加速转发PHB ( 101110 )	2
1	208.197.169/24	有保证的转发2中等丢弃PHB ( 010100 )	1
0	*	*	0

在本例中，考虑一对分组，分组A和分组B。假设分组A具有IP信宿地址192.100.5.13，DiffServ字节值为101110。假设分组B具有IP信宿地址192.100.34.102，DiffServ字节值为101110。分组A和分组B在同一个PEC (PEC2) 中，因为它们根据规则2同时匹配分类表。

考虑第三个分组，分组C具有IP信宿地址208.167.169.55，DiffServ字节值为010100。分组A和分组C不在同一个PEC中，因为分组A与规则2匹配，分组C与规则1匹配。

规则0是默认规则，与分组分类表中没有和任何其它规则匹配的每个分组相匹配。

### 流量监视程序

流量监视程序维护给定PEC可以给网络增加的负载限制有关的状态。流量监视程序为每个PEC跟踪两个关键参数，即平均带宽和最大突发尺寸。每种限制的配置值可以在每个PEC中有所不同。每个PEC的流量监视程序在每个分组上执行计费操作，确定分组是位于配置之内还是位于配置文件之外。

## 转发操作

为PEC分配的每个分组的转发操作取决于分组是位于配置文件之内还是位于配置文件之外。

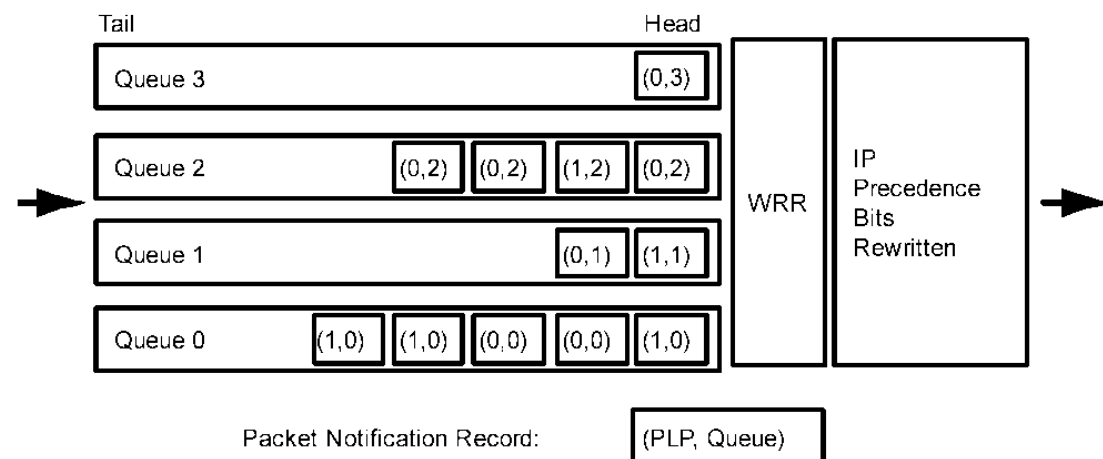
- 如果分组位于配置文件之内，它可以简单地被接纳和转发，而不会受到速率限制程序或通信策略程序的影响。
- 如果分组位于配置文件之外，它可以简便地被丢弃或进行标记。分组的标记方式是，设置分组通知记录中的PLP位，使其能够由通信策略功能下行方向的分组处理引擎进行处理。下行可以指本地路由器中的进程，也可以指网络中分组路径上的后续路由器。

## Internet Processor II ASIC：通信策略配置实例

如果您不熟悉M系列路由器支持的基于硬件的CoS特性，我们将简要讨论其主要机制，及其与软硬通信策略支持的关系。

每条通信流都具有最多四个输出传输队列，编号为0到3，每个队列获得通信流可用队列缓冲器大小的配置百分比（图八）。M系列路由器通过读取分组通知记录中承载的两个队列优先位的值（00，01，10，11），为每个分组确定输出队列。

**Figure 8: M-Series Output Transmission Queues**



图八 M系列输出传输队列

Tail：尾

Head：头

Queue 3：队列3

Queue 2：队列2

Queue 1：队列1

Queue 0：队列0

IP Precedence Bits Rewritten：重写的IP优先位

Packet Notification Record: 分组通知记录

(PLP, Queue)：( PLP ， 队列 )

WRR算法为输出队列服务，RED算法则一直监视每个队列的拥塞情况。如果分组位于配置文件之外，RED算法可以使用较分组位于配置文件之内时更加积极的丢弃概率。通过读取分

组通知记录中承载的丢包优先权（PLP）位，M系列路由器确定每个分组是位于配置文件之内（PLP=0）还是位于配置文件之外（PLP=1）。

最后，执行策略功能的路由器可以配置成写入PLP位，使本地路由器选择具体的RED丢弃配置文件。此外，执行策略功能的路由器可以配置成在传输分组前，在IP优先位中重写RED丢弃优先位。通过在传输分组前重写IP优先位，策略路由器可以命令下行路由器，在其对标记的分组进行排队时，使用具体的RED丢弃配置文件。