

# 比较流量策略和流量形状以限制带宽

## 目录

---

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[规则](#)

[背景信息](#)

[管制与整形](#)

[选择标准](#)

[令牌刷新速率](#)

[流量整形](#)

[流量策略](#)

[最小带宽与最大带宽控制](#)

[相关信息](#)

---

## 简介

本文档介绍流量整形和流量策略之间的功能差异，这两种差异都限制了输出速率。

## 先决条件

### 要求

本文档没有任何特定的要求。

### 使用的组件

本文档不限于特定的软件和硬件版本。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

### 规则

有关文档规则的详细信息，请参阅 [Cisco 技术提示规则](#)。

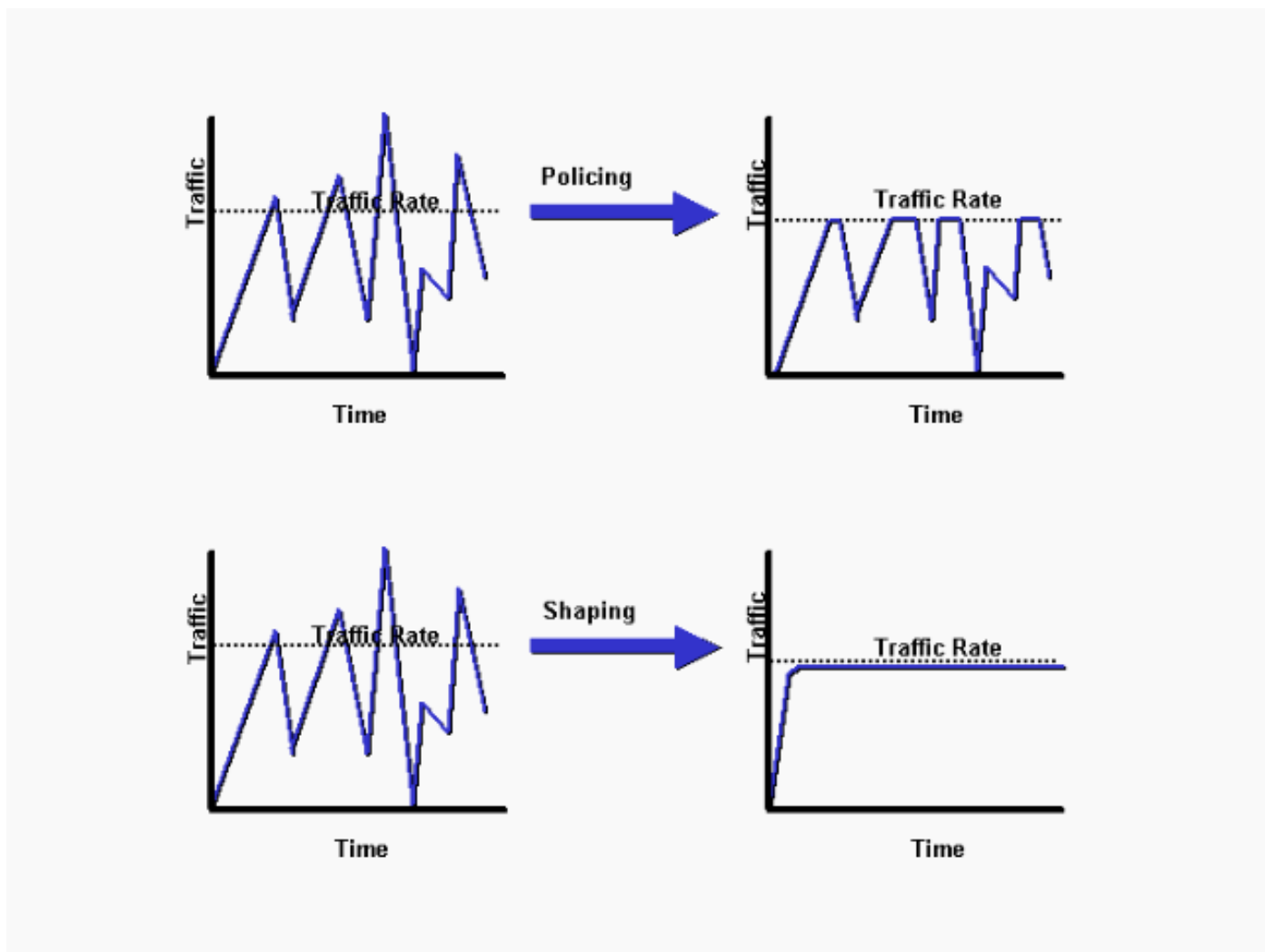
## 背景信息

本文档阐明了流量整形和策略管制之间的功能差异。这两个功能都会限制流量输出速率。两种机制都使用令牌桶作为流量计量器来测量数据包速率。有关令牌桶的详细信息，请参阅[什么是令牌桶](#)

## 管制与整形

流量策略会传播突发流量。当流量速率达到所配置的最大速率时，将丢弃（或重新标记）超额流量。结果显示为带有波峰和波谷的锯齿形输出速率。与策略相比，流量整形在队列中保留超额的数据包，然后安排超额数据包在以后增加的时间里进行传输。流量整形的结果是一个平滑的数据包输出速率。

下图说明两个流量选项之间的主要区别。



### 管制与整形

整形暗示队列和用于缓冲延迟数据包的充足内存的存在，而策略却不是。队列是一个出站概念；离开接口的数据包将进入队列并可进行整形。对接口上的入站流量只能实施策略。启用整形时，请确保有足够的内存。此外，整形需要一种功能，用于调度任何延迟数据包的稍后传输。此调度功能允许您将整形队列组织到不同的队列中。此功能的示例包括基于类的加权公平 Queuing 度(CBWFQ)和低延迟 Queuing (LLQ)。

### 选择标准

下表列出了整形和策略之间的差异，以帮助您选择适当的流量解决方案。

	整形	管制
目标	对超额数据包进行缓冲并排队，使其超过承诺速率。	丢弃（或注释）超过承诺速率的数据包。不进行缓冲。*
令牌刷新速率	在时间间隔的开始处递增。（需要最小间隔数。）	基于公式： $1 / \text{承诺信息速率}$ 连续
令牌值	按比特/秒配置。	按字节配置。
配置选项	<ul style="list-style-type: none"> <li>• 服务质量模块命令行界面 (MQC) 中的 shape 命令可实现基于类的整形。</li> <li>• Frame-relay traffic-shape命令可实现帧中继流量整形(FRTS)。</li> <li>• Traffic-shape命令可实现通用流量整形(GTS)。</li> </ul>	<ul style="list-style-type: none"> <li>• MQC 中的 police 命令可实现基于类的策略。</li> <li>• Rate-limit命令可实现承诺接入速率(CAR)。</li> </ul>
适用于入站	无	Yes
适用于出站	Yes	Yes
突发	至少在八个时间间隔内控制突发并平滑输出速率。使用漏桶来延迟流量以达到平滑效果。	传播突发流量。不进行平滑处理。
优势	由于对超额数据包进行了缓冲，因而降低了丢弃超额数据包的可能性。（数据包最多可缓冲到队列的长度。如果超额流量以较高的速率持续，则可能发生丢包现象。）通常可避免因掉包而进行重新传输。	通过掉包控制输出速率。避免 queuing造成的延迟。
缺点	可能因 queuing而引入延迟，尤其是深度队列。	丢弃超额数据包（配置时），限制 TCP窗口大小，并降低受影响的流量流的整体输出速率。过度激进的突发量大小可能导致过多的丢包并限制整体输出速率，尤其是基于TCP的流量。
可选数据包重新标记	无	是（使用传统的 CAR 功能）。

---

\* 虽然管制不应用缓冲区，但已配置的机 queuing 制适用于需要在等待在物理接口上序列化时排队的已符合条件的数据包。

## 令牌刷新速率

整形和策略之间的主要区别就是补充令牌的速率。整形和策略都使用令牌桶比喻。令牌桶本身没有丢弃或优先级策略。

使用令牌桶功能：

- 

令牌以一定速率放到桶中。

- 

每个令牌都允许源将一定数量的位发送到网络中。

- 

要发送数据包，数据流调整器必须能够从令牌桶中去除等于数据包大小的令牌数量。

- 

如果令牌桶中没有足够的令牌来发送数据包，数据包将等到令牌桶中有足够的令牌（对于整形器），或者丢弃数据包或将其降级（对于监视器）。

- 

桶本身具有指定的容量。如果令牌桶容量已满，则到达的新令牌将被丢弃，并且不能用于未来的数据包。因此在任何时间，信息源能够发送到网络的最大突发传输都是令牌桶大小的组成部分。令牌桶允许突发流量，但会限制突发流量。

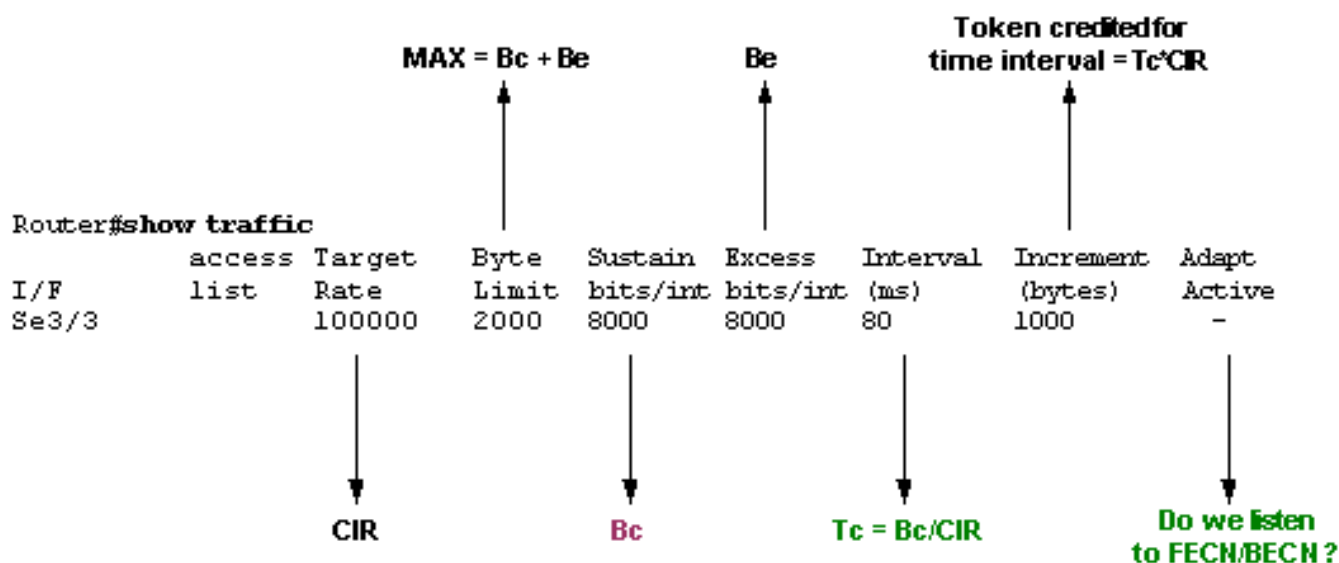
整形以使用每秒位数(bps)值的时间间隔增加令牌桶。整形器使用下一个公式：

$$Tc = Bc/CIR \text{ (in seconds)}$$

在此等式中，Bc 表示承诺突发流量，CIR 表示承诺信息速率。（有关详细信息，请参阅[配置帧中继流量整形。](#)）Tc 值定义了为维持 CIR 的平均速率，要在其间发送 Bc 比特流量的时间间隔（以秒计）。

Tc 的范围在 10 毫秒到 125 毫秒之间。使用Cisco 7500系列上的分布式流量整形(DTS)，最小Tc为4毫秒。路由器会基于 CIR 和 Bc 值在内部计算此值。如果 Bc/CIR 少于 125 毫秒，它将使用该等式计算出的 Tc。如果Bc/CIR大于或等于125毫秒，并且如果Cisco

IOS®确定数据流在更小的时间间隔内更稳定，则使用内部Tc值。使用show traffic-shape命令确定路由器是使用Tc的内部值还是使用在命令行配置的值。show traffic-shape命令的下一个输出示例在[帧中继流量整形的show命令](#)中进行了说明。



*show traffic output*

当超额突发传输(Be)配置为非0值时，成型机则允许令牌存储在桶中，高至Bc+Be。令牌桶可以达到的最大值是 Bc + Be，溢出令牌被丢弃。在桶中具有超过 BC 令牌数的唯一方式是在一个或多个 Tc 期间不要使用所有 BC 令牌。因为令牌桶通过每 Tc 与 BC 令牌被重新补充，您能累计未使用的令牌用于 BC +Be 的最新应用。

相反，基于类的策略和速率会limiting 不断将令牌添加到桶中。具体而言，令牌到达速率的计算如下：

$$(time\ between\ packets < \text{which is equal to } t-t1 > * policer\ rate) / 8\ bits\ per\ byte$$

换句话说，如果数据包以前的到达时间是t1并且当前时间是t，那么根据令牌到达的速率，桶被更新，字节值为 t-t1。



注意：流量监视器使用以字节为单位指定的突发值，并且先前的公式从位转换为字节。

---

以下示例使用8000 bps的CIR（或监视器速率）和1000字节的正常突发量：

```
<#root>
```

```
Router(config)#
```

```
policy-map police-setting
```

```
Router(config-pmap)#
```

```
class access-match
```

```
Router(config-pmap-c)#
```

```
police 8000 1000 conform-action transmit exceed-action drop
```

令牌桶从1000字节开始。如果 450 字节数据包已经到达，那么数据包符合条件，因为这时令牌桶中已经有足够的字节可以使用。数据包将采取符合操作（传输），并从令牌桶中删除450字节（并保留550字节）。如果下一个数据包在0.25秒后到达，将根据以下公式将250个字节添加到令牌桶中：

$$(0.25 * 8000)/8$$

计算结果将 700 字节留在令牌桶中。如果下一个数据包是 800 个字节，数据包将超出并且将采取超出行动(丢弃)。这时不会从令牌桶中取走任何字节。

## 流量整形

Cisco® IOS支持以下流量整形方法：

- 

- [一般流量整形](#)

-

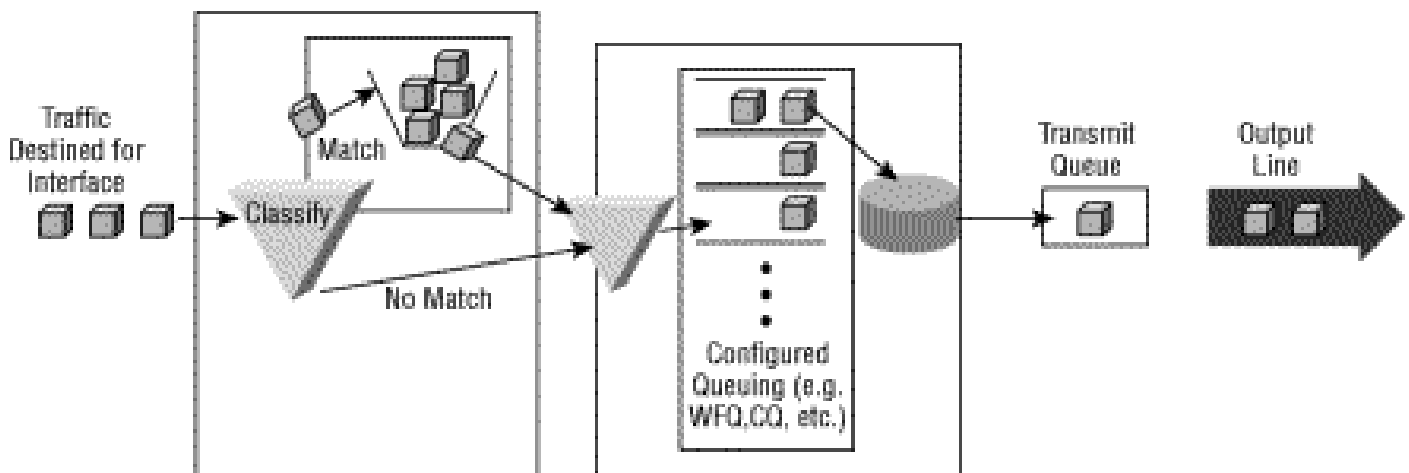
## 帧中继流量整形

- 

### 基于类的整形和分布式基于类的整形

所有流量整形的实施方法是相似的，虽然他们的命令行接口 (CLI) 有些不同，但是他们使用不同类型的队列，来包含和整形被延迟的数据流。思科建议使用模块化QoS CLI配置的基于类的整形和分布式整形。

下图说明QoS策略如何将流量安排到类中，以及如何将超过配置整形速率的数据包排队。



## 流量策略

Cisco IOS支持以下流量管制方法：

- 

### 承诺接入速率

- 

### 基于类的策略

这两种机制在功能上具有重要区别，[比较基于类的策略和承诺接入速率](#)。应用QoS策略时，思科建议使用基于类的策略以及模块化QoS CLI的其他功能。

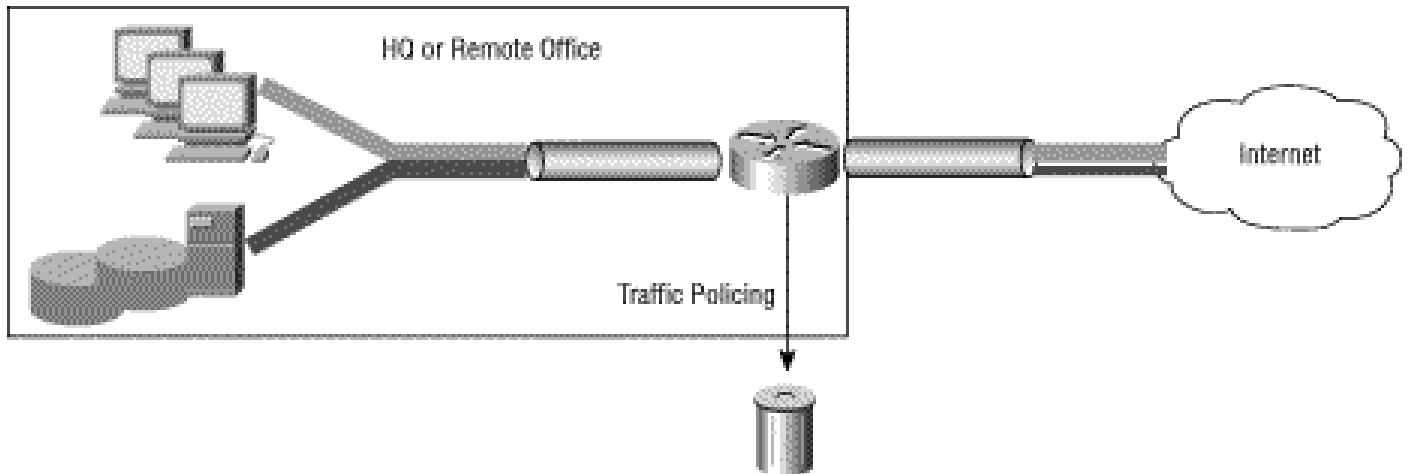
使用**police**命令指定流量等级必须有一个强加的最大速率，如果超出该速率，必须立即采取行动。换句话说，在 **police** 命令下，不能选择缓冲数据包后将其发出，而 **shape** 命令则可以这样做。

另外在使用警管的情况下，令牌桶将确定数据包是否超出或是否与应用的速率一致。在任一情况下，策略都会实施可配置操作，包括



IP优先级或差分服务代码点(DSCP)。

下图说明了在通常应用QoS功能的拥塞点执行流量策略的常见应用。



### 最小带宽与最大带宽控制

**shape** 和 **police** 命令都会将输出速率限制为最大 **kbps** 值。重要的是，这两种机制在拥塞期间都不会提供最低带宽保证。可以使用 **bandwidth** 或 **priority** 命令提供这种保证。

分层策略使用两个服务策略-父策略将QoS机制应用于流量聚合，子策略将QoS机制应用于流量或聚合的子集。子接口和隧道接口等逻辑接口需要分层策略，在父级具有流量功limiting 能，在更低级具有队列。流量功limiting 能会降低输出速率并（可能）造成拥塞，如过 queuing 多数据包所示。

下一个配置不是最理想的，目的是显示在流量聚合（在此情况下级别默认）到最大速率时，limiting police 与shape 命令之间的区别。在此配置中，**police**命令根据数据包的大小和符合及超额令牌桶中剩余的字节数量从子类中发送数据包。（请参阅[流量策略](#)。）结果，由于police功能覆盖priority功能所做的保证，因此不能保证为IP语音(VoIP)和Internet协议(IP)类指定的速率。

然而，如果使用 **shape** 命令，则结果为分层队列系统，并做出所有保证。换句话说，当流入负载超出整形速率的时候，会保证 VoIP 和 IP 类的速率，并且类默认数据流(在子级上)会发生丢包。



**注意：**建议不要使用此配置，并显示此设置以说明在限制流量聚合时**police**与**shape**命令之间的区别。

```
class-map match-all IP
  match ip precedence 3
class-map match-all VoIP
  match ip precedence 5

policy-map child
  class VoIP
    priority 128
  class IP
    priority 1000

policy-map parent
  class class-default
    police 3300000 103000 103000 conform-action transmit exceed-action drop
    service-policy child
```

要使之前的配置有意义，必须通过整形来替换策略。

例如：

```
policy-map parent
  class class-default
    shape average 3300000 103000 0
    service-policy child
```



注意：要了解有关父策略和子策略的详细信息，请参阅[优先级类的QoS子服务策略](#)。

---

相关信息

- [服务质量\(QoS\)技术支持](#)
- [技术支持和文档 - Cisco Systems](#)

## 关于此翻译

思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。