

## HCRSE114-QoS 原理

QoS ( Quality of Service ) 服务质量

MQC ( Modular QoS Command-Line Interface ) 模块化 QoS 命令行

PQ ( Priority queuing ) 优先级队列

SLA ( Service-Level Agreement ) 服务等级协议

DSCP ( Differentiated Services Code Point ) 差分服务代码点

WFQ ( Weighted Fair Queuing ) 加权公平排队

WRED ( Weighted Random Early Detection ) 加权随机先期检测

MQC 三个要素：

流分类 ( traffic classifier )、流行为 ( traffic behavior ) 和流策略 ( traffic policy )。

影响网络通信质量的因素：带宽、时延、抖动、丢包率

### QoS 四大组件

报文分类和标记，流量管理 ( 流量监管、流量整形和接口限速 )，拥塞管理，拥塞避免

### QoS 服务模型：

尽力而为的服务模型 ( Best Effort Services Model )

综合服务模型 ( Integrated Services Model )

区分服务模型 ( Differentiated Services Model )

	优点	缺点
尽力而为服务模型	实现机制简单	对不同业务流不能进行区分对待
综合服务模型	可提供端到端QoS服务，并保证带宽、延迟	需要跟踪和记录每个数据流的状态，实现较复杂，且扩展性较差，带宽利用率较低
区分服务模型	不需跟踪每个数据流状态，资源占用少，扩展性较强；且能实现对不同业务流提供不同的服务质量	需要在端到端每个节点都进行手工部署，对人员能力要求较高

服务模型

服务模型，是指一组实现端到端 QoS 保证的方式，包括 Best Effort、IntServ 和 DiffServ 三种服务模型。

Best Effort 模型 尽力而为的服务模型

Best Effort 模型（即尽力而为模型）是一个单一的服务模型，也是最简单的服务模型。应用程序可以在任何时候，发出任意数量的报文，而且不需要事先获得批准，也不需要通知网络。Best Effort 模型中，网络尽最大的可能性来发送报文，但对时延、可靠性等性能不提供任何保证。

Best Effort 模型是 Internet 的缺省服务模型，它适用于绝大多数网络应用，如 FTP、E-Mail 等，它通过先入先出（FIFO）调度方式来实现。

IntServ 模型 综合服务模型

IntServ ( Integrated Service ) 模型是一个综合服务模型，它的特点是在发送报文前要先向网络提出申请。这个请求是通过信令来完成的，一个实例是资源预留协议 RSVP ( Resource Reservation Protocol )。应用程序首先通过 RSVP 信令通知网络它的 QoS 需求 ( 如时延、带宽、丢包率等指标 )。在收到资源预留请求后，传送路径上的网络节点实施许可控制 ( Admission control )，验证用户的合法性并检查资源的可用性，决定是否为用户预留资源。

一旦认可并为应用程序的报文分配了资源，则只要应用程序的报文控制在流量参数描述的范围内，网络节点将承诺满足应用程序的 QoS 需求。预留路径上的网络节点可以通过执行报文的分类、流量监管、低延迟的排队调度等行为，来满足对应用程序的承诺。IntServ 模型常与组播应用结合，适用于需要保证带宽、低延迟的实时多媒体应用，如电视会议、视频点播等。

当前，采用 RSVP 协议的 IntServ 模型定义了两种业务类型：

**保证型服务** ( Guaranteed Service )

提供保证的带宽和时延限制来满足应用程序的要求。如 VoIP ( Voice over IP ) 应用可以预留 10M 带宽和要求不超过 1 秒的时延。

**负载控制型服务** ( Controlled-Load Service )

保证即使在网络过载 ( overload ) 的情况下，仍能对报文提供类似 Best Effort 模型在未过载时的服务质量——即在网络拥塞的情况下，保证某些应用程序报文的低时延和低丢包率需求。

可以提供端到端的 QoS 投递服务是 IntServ 模型的最大优点。IntServ 模型的最大缺点是可扩展性不好。网络节点需要为每个资源预留维护一些必要的软状态 ( Soft State ) 信息；在与组播应用相结合时，还要定期地向网络发资源请求和路径刷新信息，以支持组播成员的动态加入和退出。

上述操作要耗费网络节点较多的处理时间和内存资源。在网络规模扩大时，维护的开销会大幅度增加，对网络节点特别是核心节点线速处理报文的性能造成严重影响。因此，IntServ 模型不适宜于在流量汇集的骨干网上大量应用。

### DiffServ 模型 区分服务模型

为了在 Internet 上针对不同的业务提供有差别的服务质量，IETF 定义了 DiffServ ( Differentiated Service ) 模型。

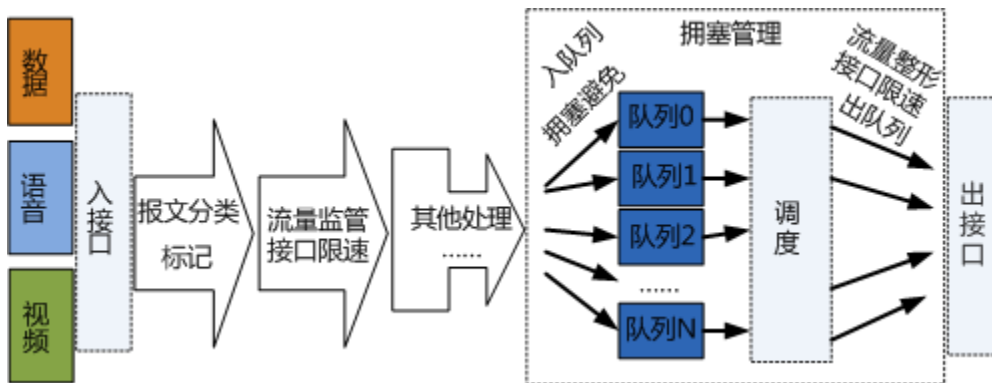
DiffServ 模型是一种多服务模型，它可以满足不同的 QoS 需求。与 IntServ 模型不同，应用程序在发出报文前，通过设置报文头部的优先级字段，向网络中各设备通告自己的 QoS 需求，而不需要通知途经的网络设备为其预留资源。

DiffServ 模型中，网络不需要为每个流维护状态，它根据每个报文携带的优先级来提供特定的服务。可以用不同的方法来指定报文的 QoS，如 IP 报文的优先级 ( IP Precedence )，报文的源地址和目的地址等。网络通过这些信息来进行报文的分类、流量整形、流量监管和队列调度。

DiffServ 模型一般用来为一些重要的应用提供端到端的 QoS。通常在配置 DiffServ 模型后，边界设备通过报文的源地址和目的地址等信息对报文进行分类，对不同的报文设置不同的优先级，并标记在报文头部。而其他设备只需要根据设置的优先级来进行报文的调度。

### QoS 的度量指标：带宽、时延和抖动、以及丢包率

DiffServ 模型包含了**四大组件**，通过这四大组件实现端到端的 QoS：



入接口的数据进来，根据类型的不同，进行简单流或复杂流的分类

可以基本流或接口进行流量监管，把超过的流量丢弃

接着查一些路由表，MAC 表进行转发

进入设备后，会进入队列，0-7 个队列，同时也会进行拥塞避免的处理，主动丢弃部分报文

去出时进行拥塞管理，通过调度算法安排报文的转发次序

在出去之前，根据下行接口的情况，看是否需要进行流量整形，缓存一下空闲再发送

### 1. 优先级映射（报文分类和标记）

要实现差分服务，首先需要将报文分为不同的类别。类别确定好了，设备才能针对性地提供服务。

### 2. 流量管理（流量监管、流量整形和接口限速）

令牌桶：单速单桶，单速双桶，双速双桶

限速：通过配置接口发送报文速率占接口带宽的百分比，实现对接口发送报文速率的限制

限制用的是单速单桶

流量监管：单速双桶或双速双桶，三色

是将流量限制在特定的带宽内。当业务流量超过额定带宽时，

超过的流量将被丢弃。这样可以防止个别业务或用户无限制地占用带宽。

**流量整形：单速单桶，双色**

是一种主动调整流的输出速率的流控措施，使流量比较平稳地传送给下游设备，避免不必要的报文丢弃和拥塞。流量整形通常在接口出方向使用。

**接口限速：单速单桶，双色**

是对一个接口上发送或者接收全部报文的总速率进行限制。当不需要对报文类型进行进一步细化分类而要限制通过接口全部流量的速率时，接口限速功能可以简化配置。

**流量监管与流量整形的区别**

在进行报文流量控制时，流量监管是对超过流量限制的报文进行丢弃；而流量整形则将超过流量限制的报文缓存在队列中，等待链路空闲的时候再发送。

### 3. 拥塞管理

拥塞管理是在网络发生拥塞时，通过一定的调度算法安排报文的转发次序，保证网络可以尽快恢复正常。拥塞管理通常在接口出方向使用。

### 4. 拥塞避免

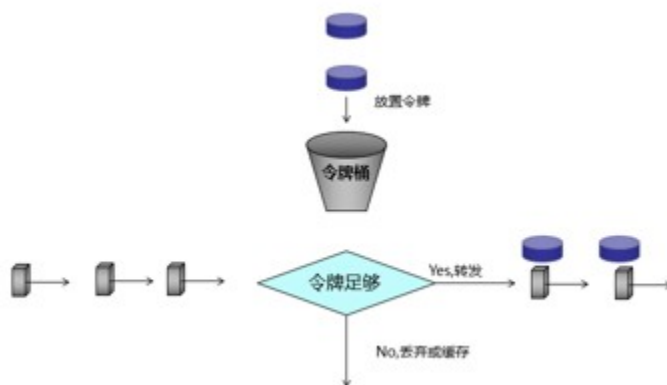
拥塞避免可以监视网络资源（如队列或内存缓冲区）的使用情况。在拥塞有加剧的趋势时，主动丢弃报文，避免网络拥塞继续加剧。拥塞避免通常在接口入方向使用。

综上所述，报文分类是基础，是有区别地实施服务的前提，流量监管、流量整形和接口限速主要用于预防拥塞，拥塞管理和

拥塞避免是用来解决拥塞。

## 令牌桶算法

要实现流量的控制，必须有一种机制可以对通过设备的流量进行度量。令牌桶是目前最常采用的一种流量测量方法，用来评估流量速率是否超过了规定值。



当数据流到达设备时首先会根据数据的大小从令牌桶中取出与数据大小相当的令牌数量用来传输数据。也就是说要使数据被传输必须保证令牌桶里有足够多的令牌，如果令牌数量不够，则数据会被丢弃或缓存。这就可以限制报文的流量只能小于等于令牌生成的速度，达到限制流量的目的。

## 单速率令牌桶结构

traffic behavior toisp

car cir 1024 cbs 192512 pbs 320512 green pass yellow pass red discard

单速率令牌桶主要由如下三个参数构成

CIR ( Committed Information Rate ) : 承诺信息速率

CBS ( Committed Burst Size ) : 承诺突发尺寸



EBS ( Extended burst size ) :

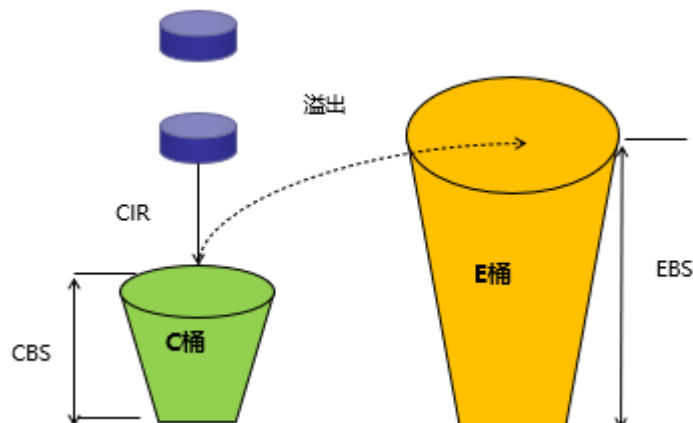
超额突发尺寸

CIR：承诺信息速率，单位是 bit/s，表示向令牌桶中投放令牌的速率。

CBS：承诺突发尺寸，单位为 bit，用来定义在部分流量超过 CIR 之前的最大突发流量，即为令牌桶的容量（深度）。承诺突发尺寸必须大于报文的最大长度（最大时一个分组可以领取桶中的全部令牌）。CBS 越大，表示所允许的突发量越大。

EBS ( Extended burst size )：超额突发尺寸，用来定义在所有流量超过 CIR 之前的最大突发量。

对于单速率标记算法的实现，华为路由器采用双桶结构。



双桶结构由两个桶实现，为方便将两个令牌桶称为 C 桶和 E 桶。C 桶容量为 CBS，E 桶容量为 EBS，总容量是 CBS+EBS。如果不允许有突发流量，EBS 则设置成 0。

当 EBS≠0 时，称为单速双桶。当 EBS=0，E 桶的令牌数始终为 0，相当于只使用了一个令牌桶——C 桶，这种情况也称为单速单桶。单速单桶中，报文只标记为绿、红两色。

绿色通过，红色丢弃

单速率流量评估规则



### 单速率令牌添加方式

单速率标记算法中，初始状态时两桶是满的。先往 C 桶中添加令牌，等 C 桶满了，再往 E 桶中添加令牌（E 桶的令牌用做以后临时超过 CIR 的突发流量），当两桶都被填满时，新产生的令牌将会被丢弃。

### 单速率流量评估规则

当报文到来后，直接与桶中的令牌数相比较，如果有足够的令牌就转发（通常用一个令牌关联一个比特的转发权限），如果没有足够的令牌则丢弃或缓存。

为方便用  $T_c$  和  $T_e$  表示桶中的令牌数量， $T_c$  和  $T_e$  初始化等于 CBS 和 EBS。

令牌桶算法的评估结果都是为报文打上红、黄、绿三种颜色的标记，所以称为“三色标记”。QoS 会根据报文的颜色做相应的处理

### 色盲模式（不需要原有报文已被标记的颜色）

对于单速单桶：

如果报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，  
如果报文长度超过 C 桶中的令牌数  $T_c$ ，报文被标记为红色， $T_c$  值不变。

对于单速双桶：

如果报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，  
如果报文长度超过 C 桶中的令牌数  $T_c$  但不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，  
如果报文长度超过 E 桶中的令牌数  $T_e$ ，报文被标记为红色，但  $T_c$  和  $T_e$  不变。

查看附件中的图片：单速双桶

- 1.假设第 1 个到达的报文是 1500bytes。检查 C 桶发现令牌数大于数据包的长度，所以数据包被标为绿色，C 桶减少令牌 1500bytes，还剩 500bytes，E 桶令牌数量保持不变。
- 2.假设 1ms 之后到达第 2 个报文 1500bytes。在此间隔内，C 桶新增令牌 =  $CIR \times 1ms = 1000bits = 125bytes$ ，加上 C 桶原来剩余的令牌 500bytes，此时 C 桶共有 625bytes，检查发现 C 桶内令牌数量不够。检查 E 桶发现有足够令牌，因此报文标记为黄色，E 桶减少令牌 1500bytes，剩余 500bytes，C 桶剩余 625bytes 保持不变。
- 3.假设又过 1ms 后到达第 3 个报文 1000bytes。在此间隔内，C 桶新增令牌 125bytes，加上 C 桶原来剩余的令牌 625bytes，此时 C 桶共有 750bytes，检查发现 C 桶内令牌数量不够。检查 E 桶发现令牌数量也不够，因此报文被标记为红色，C 桶、E 桶令牌数不变。
- 4.假设又过 20ms 后到达第 4 个报文 1500bytes。在此间隔内，C 桶新增令牌 =  $CIR \times 20ms = 20000bits = 2500bytes$ ，加上 C 桶原来剩余的令牌 750bytes，C 桶此时令牌数为 3250bytes。而 CBS = 2000bytes，因此溢出的 1250bytes 添加到 E 桶，此时 E 桶有 1750bytes。由于 C 桶中令牌数大于报文长度，报文标记为绿色，C 桶减少令牌 1500bytes，剩余 500bytes，E 桶令牌数量保持不变。

色敏模式下，（要查看原有报文已被标记的颜色）

对于单速单桶：

如果报文已被标记为绿色但报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ；

如果报文已被标记为绿色且报文长度超过 C 桶中的令牌数  $T_c$ ，则报文被标记为红色， $T_c$  保持不变；

如果报文已被标记为黄色或红色，都直接将报文标记为红色， $T_c$  保持不变。

对于单速双桶：

如果报文已被标记为绿色且报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，

如果报文已被标记为绿色且报文长度超过 C 桶中的令牌数  $T_c$  但不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，

如果报文已被标记为黄色但报文长度不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，

如果报文已被标记为黄色且报文长度超过 E 桶中的令牌数  $T_e$ ，则报文被标记为红色，且  $T_e$  保持不变，

如果报文已被标记为红色，直接将报文标记为红色， $T_c$  和  $T_e$  不变。

### 双速率令牌

CIR ( Committed Information Rate ) : 承诺信息速率

CBS ( Committed Burst Size ) : 承诺突发尺寸

PIR ( Peak Information Rate ) : 表示峰值信息速率

PBS ( Peak Burst Size ) : 表示峰值突发尺寸

### 单速

CIR ( Committed Information Rate ) : 承诺信息速率

CBS ( Committed Burst Size ) : 承诺突发尺寸

EBS ( Extended burst size ) : 超额突发尺寸

CIR：承诺信息速率，表示端口允许的信息流平均速率，即保证能够通过速率。

CBS：承诺突发尺寸，用来定义在部分流量超过 CIR 之前的最大突发流量。即瞬间能够通过的承诺流量，即第一个令牌桶的深度（假定该桶为 C 桶）。承诺突发尺寸必须不小于报文的最大长度。

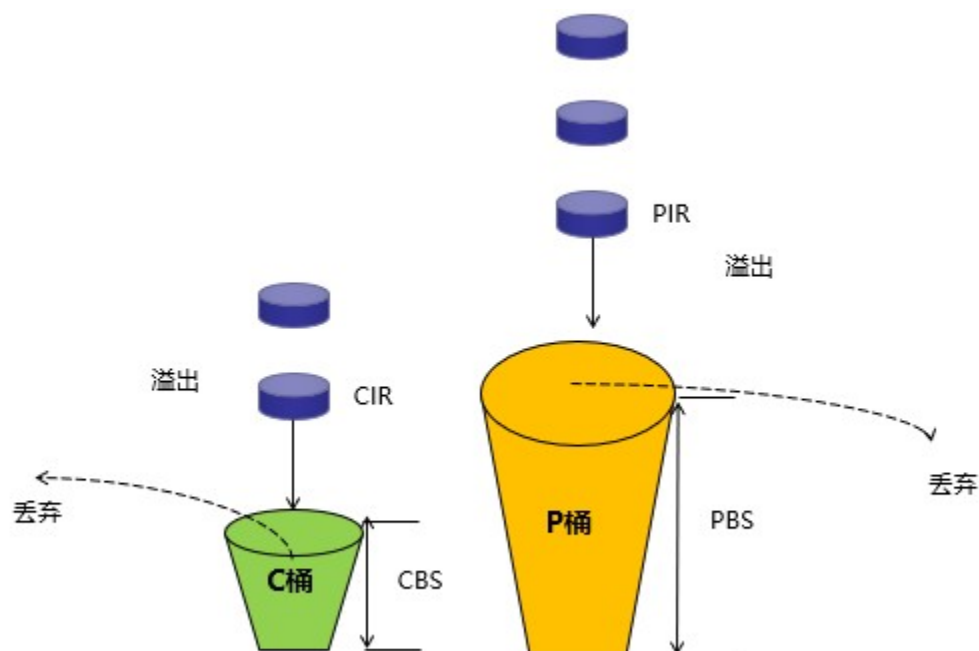
PIR：表示峰值信息速率，表示端口允许的突发流量的最大速率。该值必须不小于 CIR 的设置值。

PBS：表示峰值突发尺寸，用来定义每次突发所允许的最大流量尺寸。即瞬间能够通过的峰值流量，即第二个令牌桶的深度（假定该桶为 P 桶）

双速率三色标记算法业界都使用两个令牌桶，但它关注的是速率的突发，所以不像单速率三色标记算法那样把第一个桶中未使用的令牌放到第二个桶中，而是使用两个独立的令牌桶，存在两个令牌填充速率。为方便将两个令牌桶称为 C 桶和 P 桶，C 桶容量为 CBS，令牌填充速率为 CIR，P 桶容量为 PBS，令牌填充速率为 PIR。说明：“双速率”是指该算法中两个令牌桶中的令牌填充速率不同。

双速率令牌添加方式

初始状态时两桶是满的。往 C 桶和 P 桶分别以 CIR 和 PIR 的速率填充令牌。因这两个令牌桶是相互独立的，当其中一个桶被填满时，这个桶新产生的令牌将会被丢弃，另一个桶继续填充令牌。



### 双速率流量评估规则

双速率三色标记算法关注的是**速率的突发**，首先评估的是数据流的速率是否符合规定的突发要求，其规则是先比较 P 桶，再比较 C 桶。

双速率三色标记算法也有色盲模式和色敏模式两种。为方便使用  $T_c$  和  $T_p$  表示桶中的令牌数量， $T_c$  和  $T_p$  初始化等于 CBS 和 PBS。

#### 色盲模式：

如果报文长度不超过 P 桶中，也不超过 C 桶中的令牌数  $T_c$ ，报文被标记为绿色，且  $T_p = T_p - B$ ， $T_c = T_c - B$ 。

如果报文长度不超过 P 桶中的令牌数  $T_p$  但超过 C 桶中的令牌数  $T_c$ ，则报文被标记为黄色，且  $T_p = T_p - B$ ，

如果报文长度超过 P 桶中的令牌数  $T_p$ ，则报文被标记为红色，且  $T_c$  和  $T_p$  保持不变，

#### 色敏模式下：

如果报文已被标记为绿色且报文长度超过 P 桶中的令牌数  $T_p$ ，

则报文被标记为红色，且  $T_p$  和  $T_c$  不变。

如果报文已被标记为绿色且报文长度不超过 P 桶中的令牌数  $T_p$  但超过 C 桶中的令牌数  $T_c$ ，则报文被标记为黄色，且  $T_p = T_p - B$ ， $T_c$  不变。

如果报文已被标记为绿色且报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_p = T_p - B$ ， $T_c = T_c - B$ 。

如果报文已被标记为黄色，则只比较 P 桶，如果报文长度超过 P 桶中的令牌数  $T_p$ ，则报文被标记为红色，且  $T_p$  和  $T_c$  不变。

如果报文已被标记为黄色，且报文长度不超过 P 桶的令牌数，则报文被标记为黄色，且  $T_p = T_p - B$ ， $T_c$  不变。

如果报文已被标记为红色，直接将报文标记为红色， $T_c$  和  $T_p$  不变。

双速双桶模式允许流量速率突发，

当用户的流量速率小于配置的 CIR 时，报文被标记为绿色；

当用户的流量大于 CIR 而小于 PIR 时，报文被标记为黄色；

当用户的流量大于 PIR 时，报文被标记为红色。

1.第 1 个到达的报文假设是 1500bytes。检查发现报文长度不超过 P 桶也不超过 C 桶，所以数据包被标为绿色，C 桶和 P 桶都减少令牌 1500bytes，C 桶还剩 500bytes，P 桶还剩 1500bytes。

2.假设 1ms 后到达第 2 个报文 1800bytes。在此间隔内，P 桶新增令牌 =  $PIR \times 1ms = 2000bit = 250bytes$ ，加上 P 桶原来剩余的令牌 1500bytes，此时 P 桶共有 1750bytes，小于报文长度。C 桶新增令牌 =  $CIR \times 1ms = 1000bits = 125bytes$ ，加上 C 桶原来剩余的令牌 500bytes，此时 C 桶共有 625bytes。报文标记为红色，P 桶、C 桶令牌数不变。

3.假设又过 1ms 后到达第 3 个报文 1000bytes。在此间隔内，

P 桶新增令牌 250byte，加上 P 桶原来剩余的令牌 1750bytes，此时 P 桶共有令牌 2000bytes，大于报文长度。再检查 C 桶，C 桶新增令牌 125bytes，加上 C 桶原来剩余的令牌 625bytes，此时 C 桶共有 750bytes，仍然小于报文长度。因此报文被标记为黄色，P 桶减少令牌 1000bytes，剩余 1000bytes，C 桶令牌不变。

4.假设又过 20ms 之后到达报文 1500bytes。在此间隔内，P 桶新增令牌 =  $PIR \times 20ms = 40000bits = 5000bytes$ ，超过 P 桶容量 PBS，因此 P 桶令牌数 = PBS = 3000bytes，溢出的令牌丢弃。这样 P 桶有 2000bytes，大于报文长度。此时 C 桶增加令牌 =  $CIR \times 20ms = 20000bits = 2500bytes$ ，超过 C 桶容量 CBS，因此 C 桶令牌数 = CBS = 2000bytes，溢出的令牌丢弃。C 桶此时令牌数 2000bytes，大于报文长度。报文被标记为绿色，P 桶减少令牌 1500bytes，剩余 1500bytes；C 桶减少令牌 1500bytes，剩余 500bytes。



## 前言

- 随着网络技术的飞速发展，互联网中的业务越来越多样化。除了传统的WWW、E-Mail、FTP应用外，企业还尝试在Internet上拓展新业务，比如IP电话、电子商务、多媒体游戏、远程教学、远程医疗、可视电话、电视会议、视频点播、在线电影等。企业用户也希望通过VPN技术，将分布在各地的分支机构连接起来，开展一些事务性应用，比如访问公司的数据库或通过Telnet管理远程设备。
- 网络的普及，业务的多样化，使互联网流量激增，产生网络拥塞，转发时延增加，严重时还会产生丢包，导致业务质量下降甚至不可用。所以，要在IP网络上开展这些实时性业务，就必须解决网络拥塞问题。QoS技术就是在这种背景下发展起来的。本章我们将学习QoS高级特性。



## QoS MQC配置

- MQC概念：
  - 模块化QoS命令行MQC（Modular QoS Command-Line Interface）是指通过将具有某类共同特征的报文划分为一类，并为同一类报文提供相同的服务，对不同类的报文提供不同的服务。
- MQC三要素：
  - MQC包含三个要素：流分类（traffic classifier）、流行为（traffic behavior）和流策略（traffic policy）。

## 流分类

- 二层流分类规则
  - 源MAC地址、目的MAC地址、VLAN、MPLS EXP、基于二层的协议字段、ACL 4000~4999匹配的字段等。
- 三层流分类规则
  - IP DSCP、IP precedence、IP协议类型、RTP端口号、ACL 2000~3999匹配的字段、ACL6 2000~3999匹配的字段等。
- 流分类用来定义一组流量匹配规则，以对报文进行分类。

## 流分类配置

- **traffic classifier classifier-name [ operator { and | or } ]**
  - **and**表示流分类中各规则之间关系为“逻辑与”，指定该逻辑关系后：
    - 当流分类中有ACL规则时，报文必须匹配其中一条ACL规则以及所有非ACL规则才属于该类。
    - 当流分类中没有ACL规则时，则报文必须匹配所有非ACL规则才属于该类。
  - **or**表示流分类各规则之间是“逻辑或”，即报文只需匹配流分类中的一个或多个规则即属于该类。  
缺省情况下，流分类中各规则之间的关系为“逻辑或”
- **traffic classifier classifier-name [ operator { and | or } ]**
- **classifier-name**：指定流分类名称。
- **Operator**:指定流分类下各规则之间的逻辑运算符。如果没有指定 operator，则各规则之间缺省为逻辑“或”的关系。
- **And**:指定流分类下各规则之间是逻辑“与”的关系。
- **Or**:指定流分类下各规则之间是逻辑“或”的关系。指定该逻辑关系后，报文只需匹

配流分类下的一个或多个规则就属于该类。

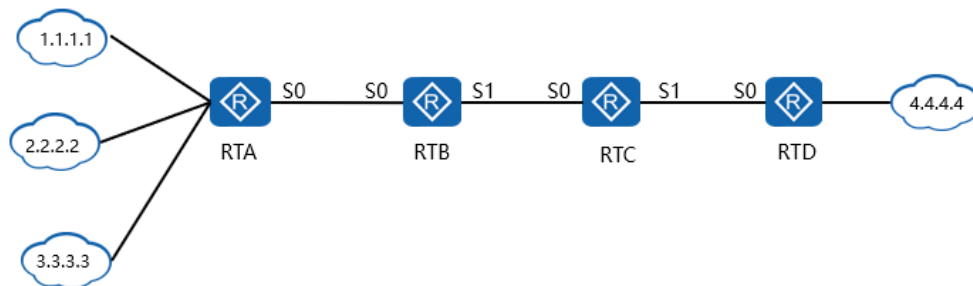
## 流行为配置

- MQC流行为动作有：
  - 报文过滤
  - 重标记优先级
  - 重定向
  - 流量监管
  - 流量统计
- 通过命令**traffic behavior behavior-name**可创建一个流行为

## 流策略配置

- 流量策略把流量类和流量动作关联
  - 执行命令**traffic policy policy-name**, 创建一个流策略并进入流策略视图
  - 执行命令**classifier classifier-name behavior behavior-name [ precedence precedence-value ]**

## MQC配置示例



- 在RTA上将流量分成3类：AF11、AF21、EF
- RTB上将AF21流量重新标记成AF22
- 这是一个基于类的 QoS 的配置示例。在 RTA 上执行流分类；在 RTB 上执行带宽限制、重新标记等策略。

## MQC配置-RTA配置

```
acl number 2001
 rule 0 permit source 1.1.1.1 0
acl number 2002
 rule 0 permit source 2.2.2.2 0
acl number 2003
 rule 0 permit source 3.3.3.3 0

traffic classifier Class_af11 operator and
if-match acl 2001
traffic classifier Class_af21 operator and
if-match acl 2002
traffic classifier Class_ef operator and
if-match acl 2003 //流分类

traffic behavior Behavior_af11
 remark dscp af11
traffic behavior Behavior_af21
 remark dscp af21
traffic behavior Behavior_ef
 remark dscp ef //流行为

traffic policy Mark
 classifier Class_af11 behavior Behavior_af11
 classifier Class_af21 behavior Behavior_af21
 classifier Class_ef behavior Behavior_ef //流策略

interface Serial0
 ip address 12.12.12.1 255.255.255.252
 traffic-policy Mark outbound
```

- RTA 上主要是执行流量分类。根据源地址的不同，分别把流量标记为 AF11、AF21 和 EF。

## MQC配置-RTB配置

```
traffic classifier Class_af21 operator and
if-match dscp af21 //流分类

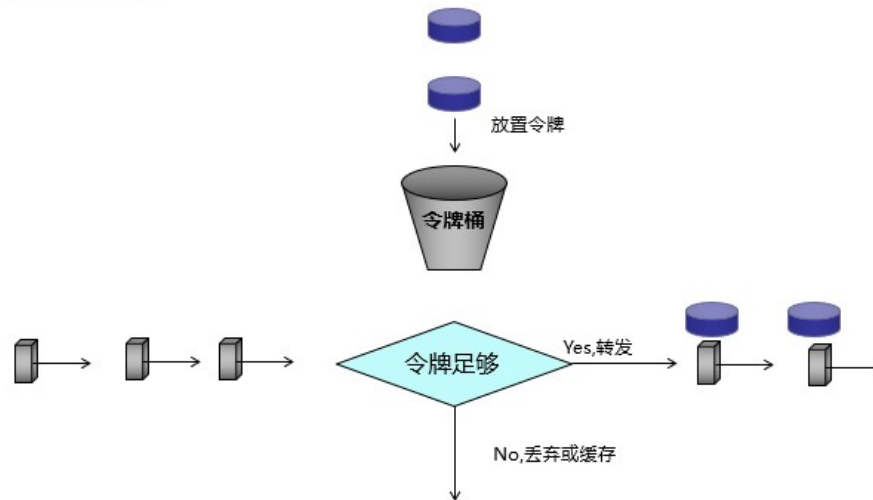
traffic behavior Behavior_af21
 remark dscp af22 //流行为

traffic policy dscp
 classifier Class_af21 behavior Behavior_af21 //流策略

interface Serial0
 ip address 23.23.23.1 255.255.255.252
 traffic-policy dscp inbound
```

- 在 RTB 上根据数据包的标记执行不同的 QoS 策略。

## 令牌桶算法



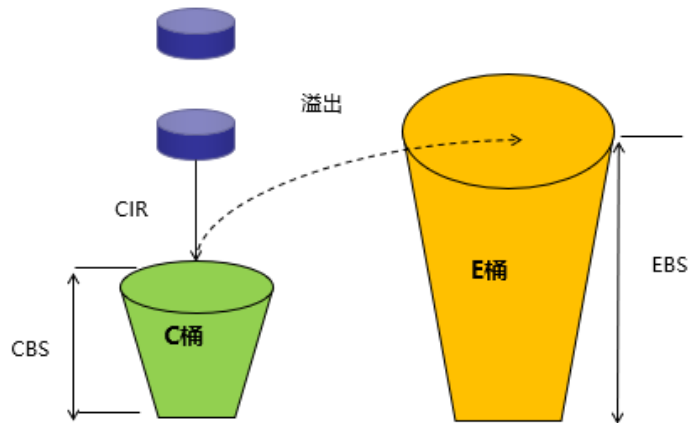
- 要实现流量的控制，必须有一种机制可以对通过设备的流量进行度量。令牌桶是目前最常采用的一种流量测量方法，用来评估流量速率是否超过了规定值。
- 当数据流到达设备时首先会根据数据的大小从令牌桶中取出与数据大小相当的令牌数量用来传输数据。也就是说要使数据被传输必须保证令牌桶里有足够多的令牌，如果令牌数量不够，则数据会被丢弃或缓存。这就可以限制报文的流量只能小于等于令牌生成的速度，达到限制流量的目的。

## 单速率令牌桶参数

- 单速率令牌桶主要由如下三个参数构成
  - CIR (Committed Information Rate)：承诺信息速率
  - CBS (Committed Burst Size)：承诺突发尺寸
  - EBS (Extended burst size)：超额突发尺寸

## 单速率令牌桶结构

- 双桶结构由两个桶实现，为方便将两个令牌桶称为C桶和E桶



- 对于单速率标记算法的实现，华为路由器采用双桶结构。
- 双桶结构由两个桶实现，为方便将两个令牌桶称为 C 桶和 E 桶。C 桶容量为 CBS，E 桶容量为 EBS，总容量是 CBS+EBS。如果不允许有突发流量，EBS 则设置成 0。
- 当 EBS≠0 时，称为单速双桶。当 EBS=0，E 桶的令牌数始终为 0，相当于只使用了一个令牌桶——C 桶，这种情况也称为单速单桶。单速单桶中，报文只标记为绿、红两色。
- CIR CBS EBS 的单位：
- CIR ( Committed Information Rate )：承诺信息速率，单位是 bit/s，表示向令牌桶中投放令牌的速率。
- CBS ( Committed Burst Size )：承诺突发尺寸，单位为 bit，用来定义在部分流量超过 CIR 之前的最大突发流量，即为令牌桶的容量 ( 深度 )。承诺突发尺寸必须大于报文的最大长度 ( 最大时一个分组可以领取桶中的全部令牌 )。CBS 越大，表示所允许的突发量越大。
- EBS ( Extended burst size )：超额突发尺寸，用来定义在所有流量超过 CIR 之前的最大突发量。

# 单速率流量评估规则

- 色盲模式下
  - 单速单桶
  - 单速双桶
- 色敏模式下
  - 单速单桶
  - 单速双桶
- **单速率令牌添加方式**
- 单速率标记算法中，初始状态时两桶是满的。先往 C 桶中添加令牌，等 C 桶满了，再往 E 桶中添加令牌（E 桶的令牌用做以后临时超过 CIR 的突发流量），当两桶都被填满时，新产生的令牌将会被丢弃。
- **单速率流量评估规则**
- 当报文到来后，直接与桶中的令牌数相比较，如果有足够的令牌就转发（通常用一个令牌关联一个比特的转发权限），如果没有足够的令牌则丢弃或缓存。
- 为方便用  $T_c$  和  $T_e$  表示桶中的令牌数量， $T_c$  和  $T_e$  初始化等于 CBS 和 EBS。
- 色盲模式下，在对到达报文（假设报文大小为  $B$ ）进行评估时，遵循以下规则：
  - 对于单速单桶：
    - 如果报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，
    - 如果报文长度超过 C 桶中的令牌数  $T_c$ ，报文被标记为红色， $T_c$  值不变。
  - 对于单速双桶：
    - 如果报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，
    - 如果报文长度超过 C 桶中的令牌数  $T_c$  但不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，
    - 如果报文长度超过 E 桶中的令牌数  $T_e$ ，报文被标记为红色，但  $T_c$  和  $T_e$  不变。
- 色敏模式下，在对到达报文（假设报文大小为  $B$ ）进行评估时，遵循以下规则：
  - 对于单速单桶：
    - 如果报文已被标记为绿色但报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ；
    - 如果报文已被标记为绿色且报文长度超过 C 桶中的令牌数  $T_c$ ，则报文被标记为红色， $T_c$  保持不变；
    - 如果报文已被标记为黄色或红色，都直接将报文标记为红色， $T_c$  保持不变。
  - 对于单速双桶：
    - 如果报文已被标记为绿色且报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_c = T_c - B$ ，

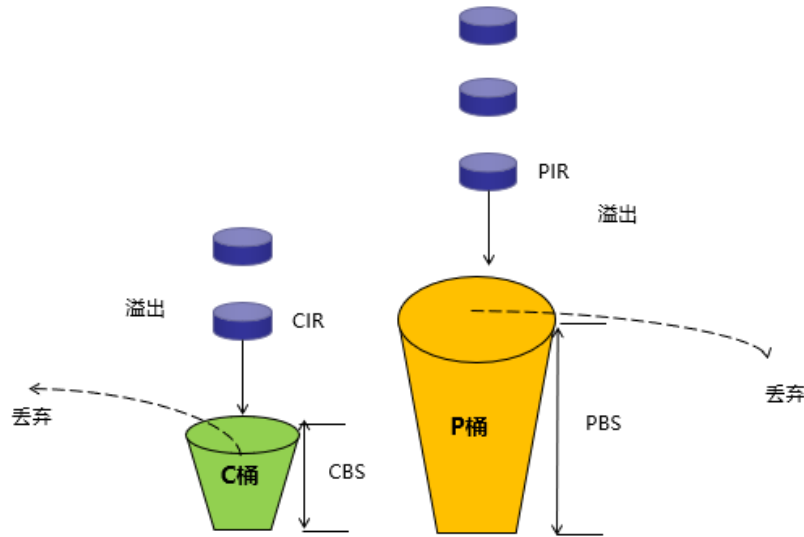
- 如果报文已被标记为绿色且报文长度超过 C 桶中的令牌数  $T_c$  但不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，
- 如果报文已被标记为黄色但报文长度不超过 E 桶中的令牌数  $T_e$ ，则报文被标记为黄色，且  $T_e = T_e - B$ ，
- 如果报文已被标记为黄色且报文长度超过 E 桶中的令牌数  $T_e$ ，则报文被标记为红色，且  $T_e$  保持不变，
- 如果报文已被标记为红色，直接将报文标记为红色， $T_c$  和  $T_e$  不变。

## 双速率令牌桶参数

- CIR (Committed Information Rate)：承诺信息速率
  - CBS (Committed Burst Size)：承诺突发尺寸
  - PIR (Peak Information Rate)：表示峰值信息速率
  - PBS (Peak Burst Size)：表示峰值突发尺寸
- 
- CIR (Committed Information Rate)：承诺信息速率，表示端口允许的信息流平均速率，单位是 bit/s。
  - CBS (Committed Burst Size)：承诺突发尺寸，用来定义在部分流量超过 CIR 之前的最大突发流量，单位为 bit。承诺突发尺寸必须不小于报文的最大长度。
  - PIR (Peak Information Rate)：表示峰值信息速率，表示端口允许的突发流量的最大速率，单位是 bit/s。该值必须不小于 CIR 的设置值。
  - PBS (Peak Burst Size)：表示峰值突发尺寸，用来定义每次突发所允许的最大流量尺寸。



## 双速率令牌桶结构(1)



- 双速率三色标记算法业界都使用两个令牌桶，但它关注的是速率的突发，所以不像单速率三色标记算法那样把第一个桶中未使用的令牌放到第二个桶中，而是使用两个独立的令牌桶，存在两个令牌填充速率。为方便将两个令牌桶称为 C 桶和 P 桶，C 桶容量为 CBS，令牌填充速率为 CIR，P 桶容量为 PBS，令牌填充速率为 PIR。
- 说明：“双速率”是指该算法中两个令牌桶中的令牌填充速率不同。
- **双速率令牌添加方式**
- 初始状态时两桶是满的。往 C 桶和 P 桶分别以 CIR 和 PIR 的速率填充令牌。因这两个令牌桶是相互独立的，当其中一个桶被填满时，这个桶新产生的令牌将会被丢弃，另一个桶继续填令牌。
- **双速率流量评估规则**
- 双速率三色标记算法关注的是速率的突发，首先评估的是数据流的速率是否符合规定的突发要求，其规则是先比较 P 桶，再比较 C 桶。
- 双速率三色标记算法也有色盲模式和色敏模式两种。为方便用  $T_c$  和  $T_p$  表示桶中的令牌数量， $T_c$  和  $T_p$  初始化等于 CBS 和 PBS。
- 色盲模式下，在对到达报文（假设数据包大小为  $B$ ）进行评估时，遵循以下规则：
  - 如果报文长度超过 P 桶中的令牌数  $T_p$ ，则报文被标记为红色，且  $T_c$  和  $T_p$  保持不变，
  - 如果报文长度不超过 P 桶中的令牌数  $T_p$  但超过 C 桶中的令牌数  $T_c$ ，则报文被标记为黄色，且  $T_p = T_p - B$ ，
  - 如果报文长度不超过 C 桶中的令牌数  $T_c$ ，报文被标记为绿色，且  $T_p = T_p - B$ ， $T_c = T_c - B$ 。
- 色敏模式下，在对到达报文（假设报文大小为  $B$ ）进行评估时，遵循以下规则：
  - 如果报文已被标记为绿色且报文长度超过 P 桶中的令牌数  $T_p$ ，则报文被标记为红色，且  $T_p$  和  $T_c$  不变。

- 如果报文已被标记为绿色且报文长度不超过 P 桶中的令牌数  $T_p$  但超过 C 桶中的令牌数  $T_c$ ，则报文被标记为黄色，且  $T_p = T_p - B$ ， $T_c$  不变。
- 如果报文已被标记为绿色且报文长度不超过 C 桶中的令牌数  $T_c$ ，则报文被标记为绿色，且  $T_p = T_p - B$ ， $T_c = T_c - B$ 。
- 如果报文已被标记为黄色，则只比较 P 桶，如果报文长度超过 P 桶中的令牌数  $T_p$ ，则报文被标记为红色，且  $T_p$  和  $T_c$  不变。
- 如果报文已被标记为黄色，且报文长度不超过 P 桶的令牌数，则报文被标记为黄色，且  $T_p = T_p - B$ ， $T_c$  不变。
- 如果报文已被标记为红色，直接将报文标记为红色， $T_c$  和  $T_p$  不变。
- CIR PIR CBS PBS 的单位：
- **cir** *cir-value* 指定承诺信息速率，即保证能够通过速率。整数形式，取值范围是 0 ~ 4294967295，单位是 kbit/s。
- **pir** *pir-value* 指定峰值速率，即最大能够通过速率。整数形式，取值范围是 0 ~ 4294967295，单位是 kbit/s。参数 *pir-value* 的值不应小于已经配置的 *cir-value* 的值。
- **cbs** *cbs-value* 指定承诺突发尺寸 (Committed Burst Size)，即瞬间能够通过承诺流量，即第一个令牌桶的深度 (假定该桶为 C 桶)。整数形式，0 ~ 4294967295，单位是 byte。CBS 的取值必须大于已经配置的 CIR 的值。缺省值与配置的 *cir-value* 有关。
- **pbs** *pbs-value* 指定超出突发尺寸 (Peak Burst Size)，即瞬间能够通过峰值流量，即第二个令牌桶的深度 (假定该桶为 P 桶)。整数形式，0 ~ 4294967295 单位是 byte。缺省值与 *pir-value* 有关。

## QoS现网应用-规划部署

- 原则上在业务接入端口入方向实施分类、标记和限速；业务侧出方向实施流量整形（如果业务侧端口接入多个不同等级的业务，则业务侧出方向还要实施队列调度和拥塞避免）；网络侧出方向实施队列调度和拥塞避免。
- 在部署PQ队列时，建议根据业务需求对PQ队列进行限速。
- 建议在网络边缘（接近用户和业务源）做流分类（复杂流分类），以减少后期分类的复杂度；在汇聚层以上设备只做简单流分类，以降低设备处理负担。
- 原则上低优先级业务先丢包，多丢包；高优先级业务后丢包，少丢包；实时业务尾丢弃；信令协议报文不丢包。

## QoS现网应用-业务优先级规划

- 管理控制信息：如网管、设备间通信协议等，对网络的正常运行至关重要，应按照最高优先级保证；
- 语音流：要求低时延、低抖动、低丢包率的EF业务级别；
- 视频会议、可视电话等实时视频流：要求低时延、低抖动、低丢包率的EF业务级别；
- 办公业务：对网络时延抖动要求较低，一般作为AF、BE业务处理；

## QoS现网应用-流量监管与整形

- 考虑到网管系统可能发生病毒感染，而网管流量是最高优先级，采用PQ方式调度，为避免对语音、视频业务的冲击，必须对从网管服务器进入网络的流量进行监管。
- 企业办公业务，为了防止个别办公业务占用过多带宽，或由于感染病毒而发起的DDoS流量攻击，而造成其它业务服务质量下降，应该根据SLA对每个接入的流量进行流量监管(CAR)，为了充分利用网络带宽和提高客户满意度，超过约定速率的流量不直接丢弃，而是降低为最低优先级，进行尽力而为的转发。

## QoS现网应用-队列调度与丢包策略

- 骨干网边缘路由器只需要在入方向使能简单流分类，骨干网路由器按照报文自身携带的优先级进行队列调度就可以了。在IP与MPLS的边界，一般配置IP DSCP/precedence与MPLS EXP的映射。通常管理控制消息、语音、视频会议和可视电话为高优先级，进行优先转发；其余企业业务可以选择WFQ加权轮循调度。
- 对于语音和视频电话业务等PQ调度业务，选择尾丢弃策略，只有当队列缓存耗尽时，才会发生丢包，最大限度保证业务的无丢弃转发。对于其他WFQ调度业务，为避免全局的TCP同步现象，采用WRED方式。较高优先级的业务设置较高的WRED丢弃阈值，较低优先级的业务设置较低的丢弃阈值，保证低优先级业务较大概率被丢弃。