

# 理解和实施多层交换网络中的服务质量【QoS】

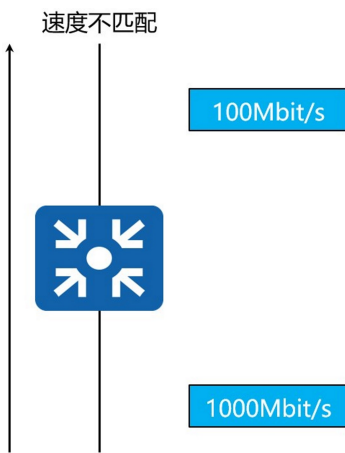
## 一、QoS的概念

- 1、互连网络中共存的数据多种多样，不同的数据对带宽及传输的速度需求亦不相同
- 2、为满足语音、视频及普通的文本文件共享网络，QoS可为不同的数据提供不同的带宽保障
- 3、QoS通过使用分类、标记、流量调节、拥塞管理及拥塞避免这5大组件，为网络中的不同数据提供优质的服务保证
- 4、Huawei VRP路由器及交换机上支持很多QoS特性，包括以下内容：
  - 4.1、控制对资源的使用：控制哪些资源可访问，哪些资源不可访问
  - 4.2、更有效地利用网络资源：通过使用网络分析管理和统计工具，可以知道通信流是如何被处理的，哪些通信流存在延迟、抖动和丢包
  - 4.3、定制服务：让Internet服务提供商能够给客户定制不同的服务等级
  - 4.4、关键任务应用共存：确保对企业来说最重要的关键任务应用能够优先使用网络

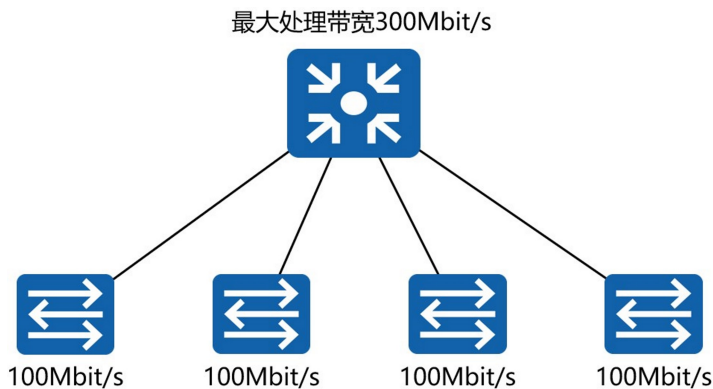
## 二、对QoS的需求

1、即使在有足够带宽的情况下，多种网络设计属性也可能影响网络性能，下列网络设计属性可能导致拥塞：

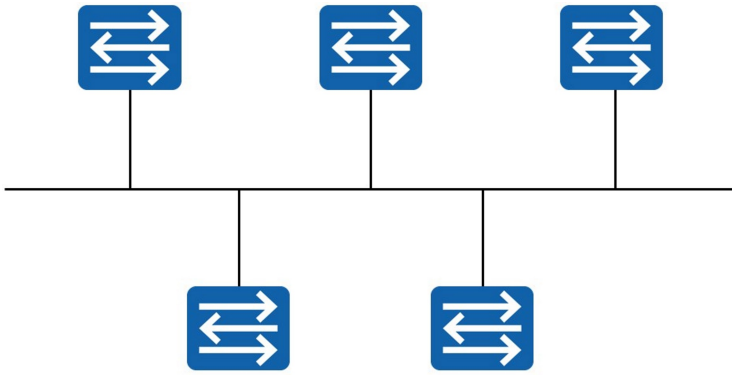
### 1.1、以太网速度不匹配



### 1.2、多对一交换矩阵



### 1.3、聚集



#### 1.4、异常行为

如：网络攻击、黑客入侵、病毒植入等

2、拥塞是影响网络可用性及稳定性的重要因素，导致稳定性差的主要3个因素为：

2.1、延迟【时滞】：数据包到达目标所需的时间

2.2、延迟变化【抖动】：同一个流中不同数据包的延迟之间的差别

2.3、丢包：数据包在信源和目标之间传输的过程中丢包的度量

#### 3、延迟/时滞【Delay】

发送方与接收方之间的端到端延迟【时滞】包括以下两种类型：

3.1、固化网络延迟：包括编解码时间、电/光信号沿介质传输到接收方所需的时间，该延迟为介质的硬件属性，QoS无助于固化延迟

3.2、下述4种延迟为固化延迟：

3.2.1、数据包化延迟：对信号进行分段、取样、编码、数据处理、将数据转换为数据包所需的时间

3.2.2、串行化延迟：将封装在数据帧中的数据包中的比特放到物理介质上所需的时间

3.2.3、传播延迟：通过物理介质传输数据帧中的比特所需的时间

3.2.4、处理延迟：网络设备将数据帧从输入接口取出，将其放入到接收队列，再将其放入到出站接口的输出队列所需的时间

3.3、可变网络延迟：网络状态能够影响数据包从信源传输到目标的总延迟【排队延迟】，该延迟可通过QoS得到控制

排队延迟：数据包在接口的输出队列中等待的时间，只有排队延迟可以通过Huawei VRP中的QoS来进行控制，其它类型的延迟不受QoS配置的直接影

#### 4、抖动【Jitter】

4.1、抖动对网络保持稳定的数据速率具有至关重要的意义

4.2、来自同一个源，去往同一个目的地的同一类型数据包【FEC | 转发等价类】，到达目标网络所花费的时间之差，被称为抖动

4.3、所有终端和Huawei网络设备都使用抖动缓冲区来消除数据包在到达时间上的差别【TCP滑动窗口】

4.4、若后续数据包的到达时间超过了特定阈值，将导致抖动缓冲区欠载

4.5、若大量数据包以过快的速度到达，将造成抖动缓冲区超载

#### 5、丢包【Drop】

1、丢包是多层交换网络中的一个严重的问题

2、过大的延迟将会造成丢包【Huawei VRP设备中，若2s无法收到对端的响应则认为数据包丢失】

3、高速发送接口与低速接收接口直接相连，将导致输出接口队列填满，最终只有丢弃数据包

4、丢包将直接导致网络不可用

### 三、QoS的解决方案

1、通过对通信流进行分类和标记，使得网络设备区分不同的通信流

2、通过流量调节，能够使得通信流适应特定的行为和吞吐量

3、当通信流速率达到特定阈值的时候，丢弃数据包

4、通过对数据包进行调节，输出队列中先传输高优先级的，再传输低优先级的

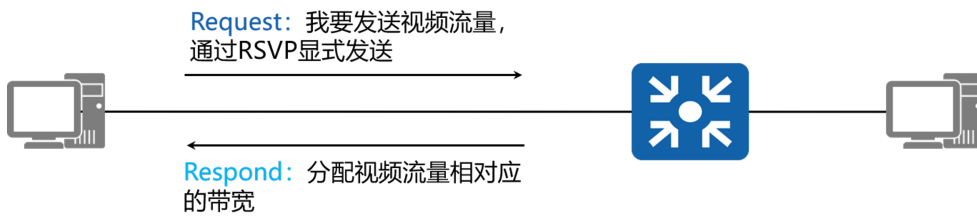
5、通过对数据队列进行管理，避免等待传输的低优先级数据包独占缓冲空间

### 四、QoS的服务模型

QoS体系结构共分为3类服务模型：

## 1、综合服务模型【硬QoS】

- 1.1、在信源发送数据之前，网络中的所有设备已经预先为各类数据流量分配好固定的带宽，保证网络对应用的处理行为是可预测的
- 1.2、综合服务模型通过使用RSVP【资源预留协议】进行实现
- 1.3、在信源与接收方、中间的网络设备上均需要启用RSVP功能



### 1.4、RSVP的工作原理

1.4.1、数据流在发送之前，信源向网络请求特定类型的服务，并将其流量配置文件告诉网络中的每个中间节点，请求网络提供一种能够满足其带宽和延迟要求的服务；在从网络得到确认后，方可开始发送数据

1.4.2、资源预留的过程分为5步：

- 信源发送RSVP Path控制消息，该消息描述了将要被发送的数据的信息
- 每个RSVP路由器收到Path消息后，保存上一跳的IP地址，并继续向下发送
- 接收方接收到RSVP Path消息后，使用RSVP Resv消息向上一跳路由器请求RSVP资源预留
- RSVP路由器确定是否可以满足这些RSVP请求，若可以，则预留出资源，并继续向上一跳路由器发送RSVP请求；若不能，则拒绝
- 信源接收到RSVP请求，说明预留已经就绪，则开始发送数据

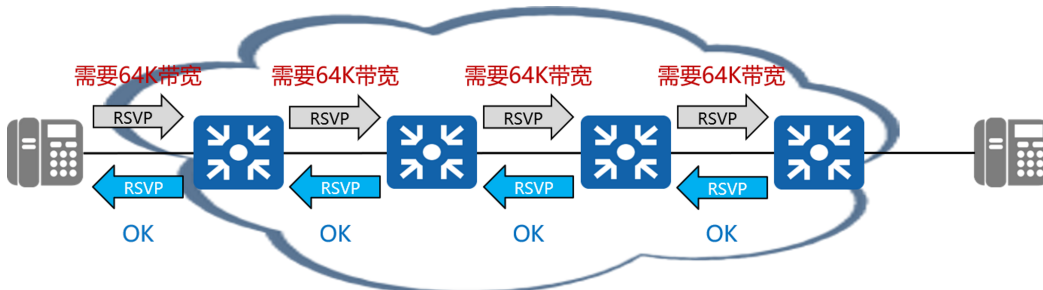
1.4.3、RSVP资源预留方式可分为两大类：

- 独占式：适用于多个信源同时发送的应用程序；针对每种数据清晰的分配资源，不同数据类型不可使用其它数据类型的资源
- 共享式：适用于多个信源不太可能同时发送的应用程序；在无其它数据类型发送时，某一类数据可使用全部资源

1.4.4、RSVP的优缺点：

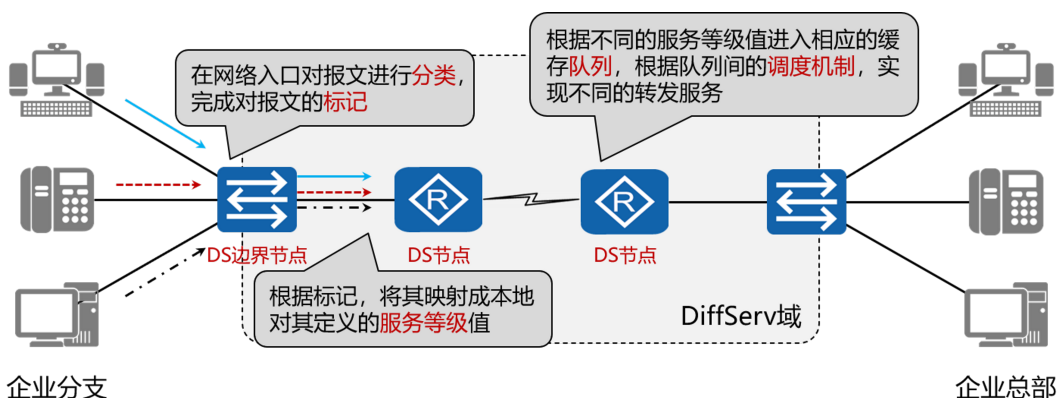
优点：可以为某些特定业务提供带宽、延迟保证

缺点：实现较复杂；当无流量发送时，仍然独占带宽，使用率较低；该方案要求端到端所有节点设备都支持并运行RSVP协议



## 2、区分服务模型【软QoS】

为解决综合服务模型的协议实现复杂性及带宽利用率低等问题，在网络中可部署DiffServ区分服务模型来保证关键业务的通信质量



### 2.1、区分服务的运作流程：

将网络中的流量分成多个类，然后为每个类定义相应的处理行为，使其拥有不同的优先转发、丢包率及时延等

## 2.2、区分服务模型概述：

业务流分类和标记由边界路由器完成；边界路由器可通过多种条件【如报文的源地址及目的地址、ToS域中的优先级、协议类型等】灵活地对报文进行分类；之后对不同类型的报文设置不同的标记字段，从而令其它路由器只需识别报文中的此标记，然后对其进行相应的资源分配和流量控制即可

## 2.3、区分服务是一种基于数据流的QoS模型

2.4、区分服务不同于集成服务，发送数据前不需要显式地通知网络设备

2.5、区分服务通过指定IP数据报头的ToS【服务类型】字段，来为不同的数据分配不同的优先级别

2.6、区分服务使用DSCP【差异化服务编码点】作为QoS优先级描述符

2.7、数据链路层交换机使用二层帧头中的CoS【服务类别，3bit】位；网络层路由器使用三层包头中的【服务类型与优先级，6bit】位

典型的以太网数据包



第二层ISL帧



↑  
3位用于CoS

第二层802.1Q帧



↑  
3位用于CoS【用户优先级】

第三层IPv4数据包



↑  
IP Precedence/DSCP

2.5、数据链路层交换机在封装VLAN时，或使用IEEE 802.1Q，或使用Cisco ISL进行VLAN封装，无论使用哪种封装方式，其均使用3bit表示分类，因此数据链路层交换机最多可标识8种【 $2^3=8$ 】不同的服务级别【0 — 7】

2.6、网络层路由器的IP包头中共有6bit表示分类，因此网络层路由器最多可标识64种【 $2^6=64$ 】不同的服务级别【0 — 63】

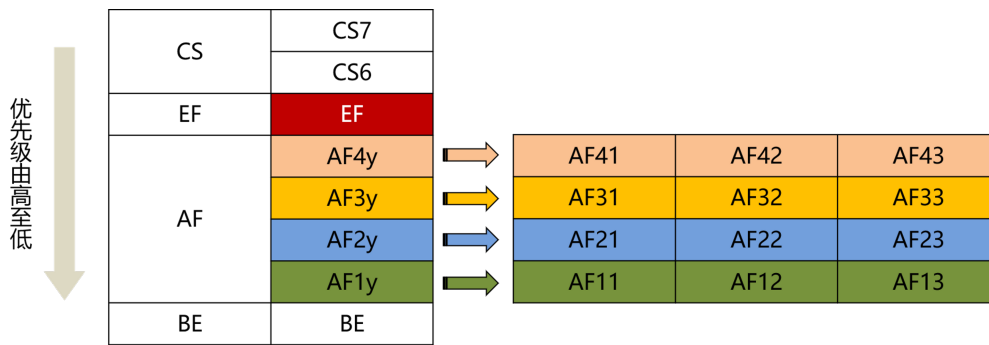
2.7、当交换机将数据帧转发给路由器时，由于交换机仅能标识8种服务级别，无法与路由器的64种服务级别形成一一对应的关系，因此在交换机与路由器传递数据的过程中，路由器仅使用6bit位中的前3bit，故而可与交换机形成一一对应的对应关系

```
[Huawei]display qos map-table dot1p-dscp
Input Dot1p      DSCP
-----
0                0
1                8
2               16
3               24
4               32
5               40
6               48
7               56
[Huawei]
```

## 2.8、DSCP值存在2种表示方式

2.8.1、数字形式：DSCP取值范围为【0 — 63】

2.8.2、关键字表达方式：用关键字标识DSCP值



### 2.8.3、AF【Assured Forwarding | 保证转发】

AF定义了4种类别：AF1y — AF4y；每种类别中有3种丢弃概率【故而共有12种标识】，用户可根据要求的吞吐量，延迟，抖动和丢包率来自行配置

- a、金级【Gold】：占用50%的可用带宽
- b、银级【Silver】：占用30%的可用带宽
- c、铜级【Bronze】：占用20%的可用带宽

不同关键字【即x值】常用于标识不同类型的报文：

AF1y常用来承载普通上网业务

AF2y常用来承载VOD【Video on Demand，视频点播】流量；相对于直播流量而言，VOD对实时性要求不那么强烈，其允许有时延或缓冲

AF3y常用来承载IPTV的直播流量；直播实时性强，需保证其连续性与大吞吐量的带宽

AF4y常用来承载语音的信令【振铃】流量；信令为电话的呼叫控制，用户允许在语音接通前等待少许时间，但绝不允许在通话过程中的语音中断

### 2.8.4、EF【Expedited Forwarding | 快速转发】

EF能够在区分服务域中提供低丢包率，低延迟，低抖动，带宽有保证的端到端服务

EF服务将网络看作【P2P, Point To Point, 点到点】

EF的优先级仅次于优先级6和7，若某类数据被配置为EF，则在该类数据传输没有完成之前，其它数据将无带宽可用，成为饥饿队列

## 2.9、综合服务与区分服务的比较

2.9.1、综合服务必须维护网络中每个流的状态信息，通过根据数据包的类别指定不同的排队策略，流量控制和丢弃参数

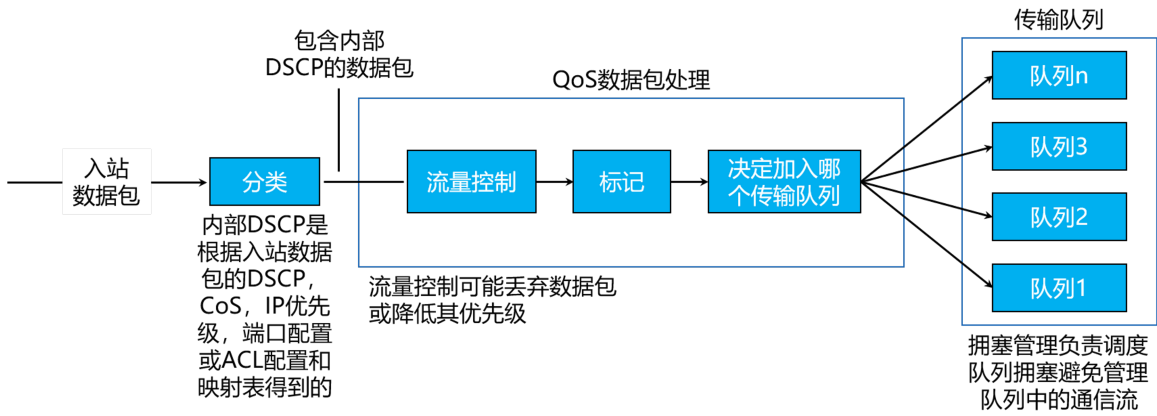
2.9.2、区分服务使用基于类别的策略，不同于集成服务使用基于流的策略，易于管理和维护

	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

3、尽力而为服务模型：尽力而为服务属于没有任何保证的标准连接方式，使用先进先出【FIFO | First In First Out】队列，不存在优先处理的概念，是Huawei VRP设备的默认行为

## 五、VRP QoS基础

1、网络设备对入站数据包进行分类，然后决定将数据包加入队列还是丢弃；在没有使用拥塞避免机制的情况下，相应的队列被填满时，排队机制才会丢弃数据包

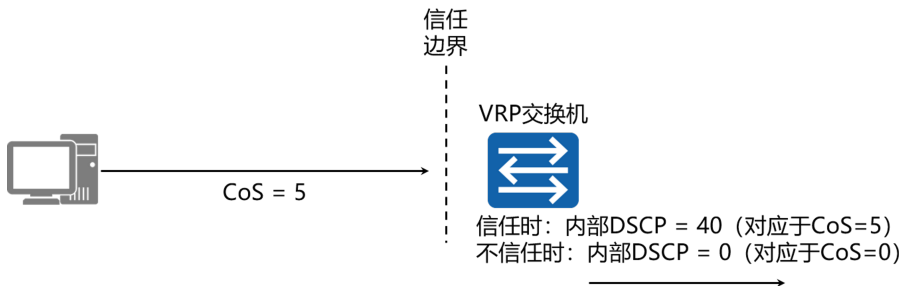


## 2、QoS的5大主要组件：

- 2.1、分类
- 2.2、标记
- 2.3、流量调节：策略和整形
- 2.4、拥塞管理
- 2.5、拥塞避免

## 六、分类

- 1、QoS策略的首要任务是识别需要进行分类的通信流
- 2、分类是能够以特定优先级或预定的参数来区分帧或数据包
- 3、Huawei VRP设备的接口默认模式为【不信任】，因此所有进站数据帧/包的CoS值/DSCP值均为0
- 4、在Huawei VRP设备上，可使用如下分类方法：
  - 4.1、按接口的信任模式
    - 4.1.1、当交换机端口上信任进站数据帧的CoS时，其将根据映射表将进站CoS值映射至DSCP值
    - 4.1.2、当进站接口的QoS配置为不信任时，交换机将所有进站数据包的内部DSCP都设置为0



### 4.2、按简单流分类【使用DSCP, IP优先级或CoS值】

若只想信任特定的进站通信流，可使用简单流分类

步骤：

- 4.2.1、创建一个用户自定义的类别映射表
- 4.2.2、配置类别映射表子句
- 4.2.3、创建一个用户自定义的策略映射表
- 4.2.4、配置策略映射表调用类别映射表
- 4.2.5、将策略映射表应用于接口的进站或出站方向

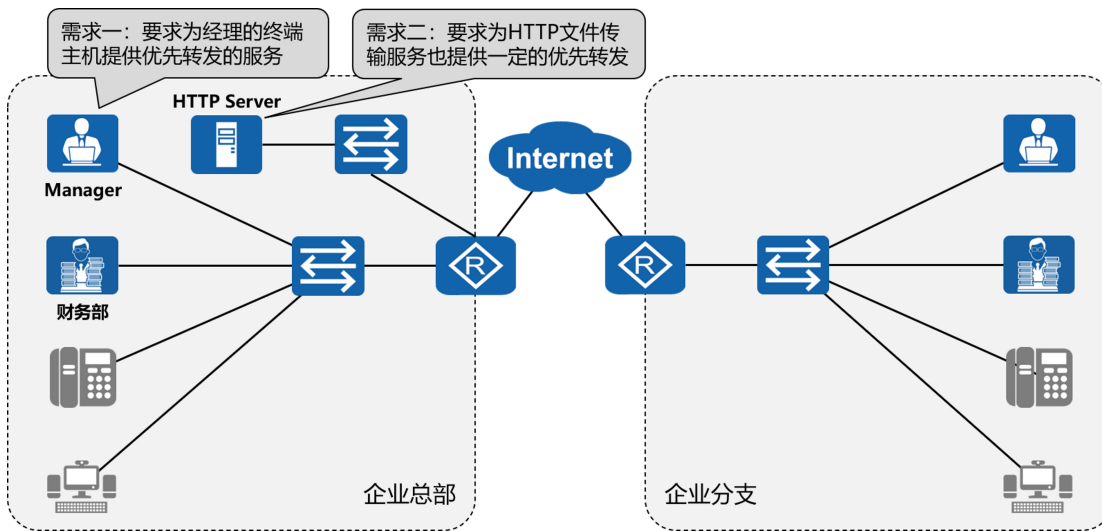
```

<Huawei>system
Enter system view, return user view with Ctrl+Z.
[Huawei]traffic classifier manager
[Huawei-classifier-manager]if-match source-mac 0050-56C0-0004
[Huawei-classifier-manager]quit
[Huawei]traffic behavior manager
[Huawei-behavior-manager]remark 8021p 2
[Huawei-behavior-manager]quit
[Huawei]traffic policy atnet
[Huawei-trafficpolicy-atnet]classifier manager behavior manager
[Huawei-trafficpolicy-atnet]quit
[Huawei]int g0/0/1
[Huawei-GigabitEthernet0/0/1]traffic-policy atnet inbound
[Huawei-GigabitEthernet0/0/1]

```

### 4.3. 按复杂流分类【基于访问控制】

针对简单流分类不够精细化的缺点，提出了复杂流分类



复杂流的匹配依据及匹配项

复杂流分类依据	常用匹配项	具体阐明
链路层复杂流分类	VLAN内/外层Tag的802.1p	各匹配项可任意组合
	源/目的MAC地址	
IP层复杂流分类	IP-Precedence	各匹配项可任意组合
	源/目的IPv4地址	
	TCP/UDP源端口	
	TCP/UDP目的端口	
	协议号	

```

<Huawei>system
Enter system view, return user view with Ctrl+Z.
[Huawei]acl 3001
[Huawei-acl-adv-3001]rule permit tcp source-port eq 80
[Huawei-acl-adv-3001]quit
[Huawei]traffic classifier http
[Huawei-classifier-http]if-match acl 3001
[Huawei-classifier-http]quit
[Huawei]traffic behavior http
[Huawei-behavior-http]remark 8021p 2
[Huawei-behavior-http]quit
[Huawei]traffic policy atnet
[Huawei-trafficpolicy-atnet]classifier http behavior http
[Huawei-trafficpolicy-atnet]quit
[Huawei]int g0/0/1
[Huawei-GigabitEthernet0/0/1]traffic-policy atnet inbound
[Huawei-GigabitEthernet0/0/1]

```

## 七、标记

标记指的是修改入站帧的CoS值或DSCP值，在网络设备对数据进行分类的时候，就已经完成了数据的标记【详见上述配置案例】

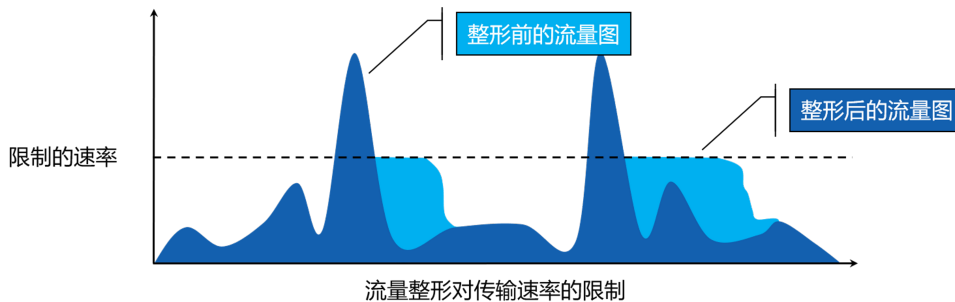
## 八、流量调节

通过配置流量调节，将能够为不同的数据流分配不同的吞吐量，以达到限速的目的

流量调节的方式有2种：

### 1、整形【GTS】

整形测量通信流的速率，并推迟【缓冲】超额通信流的传输，确保通信流的速率不超过指定的值，因此GTS不适合敏感通信流



流量整形的配置步骤：

定义感兴趣流量：

```
[Huawei]acl 3001
[Huawei-acl-adv-3001]rule permit tcp source 172.16.1.254 0 source-port eq 20 destination 192.168.1.10 0
[Huawei-acl-adv-3001]rule permit tcp source 172.16.1.254 0 source-port eq 21 destination 192.168.1.10 0
[Huawei-acl-adv-3001]rule deny ip source any destination any
```

创建并进入传输类别，调用感兴趣流量

```
[Huawei]traffic classifier ftp
[Huawei-classifier-ftp]if-match acl 3001
```

创建并进入传输行为列表，配置GTS

```
[Huawei]traffic behavior atnet
[Huawei-behavior-atnet]gts cir 512 96256
```

创建并进入传输策略，匹配传输类与传输行为

```
[Huawei]traffic policy atnet
[Huawei-trafficpolicy-atnet]classifier ftp behavior ftp
```

进入相应的接口并调用传输策略

```
[Huawei]interface g0/0/1
[Huawei-GigabitEthernet0/0/1]traffic-policy atnet outbound
```

### 2、监管

2.1、监管对超过指定速率的通信流采取特定措施，而不是推迟或缓冲；通常采取的措施是丢弃，但也可以采取其他措施，如信任或标记

2.2、监管机制采用令牌桶算法，这种算法允许通信流突发，而并非限定最大速率

2.3、令牌桶分为3种类型：

2.3.1、单色单桶：只有一个令牌桶，以固定速率向桶内装入令牌，桶内有多少令牌，就可以转发多少数据

2.3.2、单色双桶：存在两个令牌桶，单两个桶以相同的速率装入令牌【华为不支持】

2.3.3、双色双桶：存在两个令牌桶，每个桶可以以不同的速率装入令牌

### 2.4、cir、cbs、pbs的概念

cir：Committed Information Rate | 承诺信息速率

cbs：Committed burst size | 承诺突发尺寸

pbs：Peak burst size | 峰值突发尺寸

默认情况下：



CBS = 200 x CIR  
PBS = 2 x CBS

注：关于令牌桶的详细讲解

### 1、单色单桶：

假设有一套自动出卡系统，包括一台自动出卡机和一个盒子C。自动出卡机以固定的速率（3.6秒出一张通行卡）向盒子中投放通行卡。这个固定的速率，就相当于单色单桶模式中的CIR参数。

CIR (Committed Information Rate)：承诺信息速率，表示向C桶（单桶模式中只有一个令牌桶，称为C桶）中投放令牌的速率，即C桶允许传输或转发报文的平均速率。

如果平均每3.6秒来了不止1辆车，通行卡很快就领完了。这时自动出卡机上的红灯亮起，新到的车辆禁止进入高速公路。这就起到了限定作用，将放行车辆的速率限制在放卡速率范围内。

如果平均每3.6秒来了不到1辆车，盒子里就会有一些积累起来的通行卡。这些累积起来的通行卡可以应付车队（比如车友俱乐部的车队）要进入高速公路这样的情况。我们规定每辆车都要领取通行卡，都由领队的车领取。

假设某时刻来了一个车队共8辆车，但盒子里只有6张通行卡，领队车拿起通行卡一数，发现不够，那么这个车队不允许进入高速，通行卡放回盒子中。

盒子中的通行卡不停累积，总有盒子装满的时候。这个盒子的容积，就相当于单色单桶模式中的CBS参数。

CBS (Committed Burst Size)：承诺突发尺寸，表示C桶的容量，即C桶瞬间能够通过承诺突发流量。相当于盛放通行卡的盒子里最多可以放多少张通行卡，也就是说，该自动出卡系统允许通过的车队中最多可以有多少辆车。

在单色单桶模式中，系统按照CIR速率向C桶中投放令牌。

如果可用令牌的总数量（Tc）小于CBS，则令牌数继续增加。

如果令牌桶已满，则令牌数不再增加。

对于到达的报文（报文大小为B），

如果 $B \leq Tc$ ，报文被标记为绿色，且Tc减少B。

如果 $B > Tc$ ，报文被标记为红色，Tc不减少。

### 2、单色双桶：【华为不支持】

在单色单桶模式中说到，如果平均每3.6秒来了不到1辆车，盒子里就会有一些积累起来的通行卡。如果一直没有车辆过来，盒子中的通行卡不停地累积。盒子的容量是有限的，当盒子中装满通行卡之后，不断投放的通行卡就溢出盒子，会造成浪费。

为了避免这种浪费，我们改进了这个自动出卡系统，在原来的基础上增加一个盒子E（改进后的系统对应单色双桶模式，盒子E对应单色双桶中的E桶）。自动出卡机首先向C盒中投放通行卡。当C盒满了，自动出卡机就向E盒中投放通行卡。

为了保证通行卡有序领取，我们规定先领取C盒中的通行卡。如果C盒中的通行卡不够用，就把卡放回C盒，再从E盒中重新领取通行卡。C盒和E盒中的通行卡不能同时取用。

按照通行卡的取用，可以分为三种情况：

如果C盒中的通行卡够用，绿灯亮，车辆领取C盒中的通行卡后通行。

如果C盒中的通行卡不够用但E盒中的通行卡够用，黄灯亮，从C盒中领取的通行卡要归还，车辆领取E盒中的通行卡后通行。

如果E盒中的通行卡也不够用，红灯亮，车辆禁止通行，从E盒中领取的通行卡要归还。

和单色单桶模式一样，对于车队，有几辆车，就领取几张通行卡。当然，E盒的容量也是有限的。E盒的容量就相当于单色双桶模式中的EBS。

EBS (Excess Burst Size)：超额突发尺寸，表示E桶的容量，即E桶瞬间能够通过的超出突发流量。

假设某时刻来了一个车队有8辆车，但C盒里只有5张通行卡，而E盒中有9张通行卡，那么黄灯亮起，这辆长车从E盒中领取8张通行卡。

在单色双桶模式中，系统按照CIR速率向桶中投放令牌。

如果C桶中可用令牌的总数量（ $T_c$ ）小于CBS，则C桶中令牌数增加。

如果 $T_c$ 等于CBS且E桶中的可用令牌总数量（ $T_e$ ）小于EBS，则C桶中令牌数不增加，E桶中令牌数增加。

如果C桶和E桶中的令牌都已满，则两个桶中的令牌数都不再增加。

对于到达的报文（报文大小为B），

如果 $B \leq T_c$ ，报文被标记为绿色，且 $T_c$ 减少B。

如果 $T_c < B \leq T_e$ ，报文被标记为黄色，且 $T_e$ 减少B， $T_c$ 不减少。

如果 $B > T_e$ ，报文被标记为红色，且 $T_c$ 和 $T_e$ 都不减少。

### 3、双色双桶：

前面说到的自动出卡机，都只有一个口可以输出通行卡。而这里说到的高级自动出卡机，有两个口可以出卡，一个口输出的是通行卡，一个口输出的是服务卡。当然，这里也有两个盒子用于盛放卡，分别是盒C盒和P盒。自动出卡机上的两个口分别以各自固定的速率向两个盒子中投放卡。（这个高级自动出卡机系统对应双色双桶模式，C盒和P盒对应双色双桶模式中的C桶和P桶。）

领取卡的规则和前面单速的情况有所不同。我们规定：

先领取服务卡。如果服务卡不够，把卡放回P盒，红灯亮，车辆禁止通行。

如果服务卡足够但通行卡不够，黄灯亮，服务卡可以取走，车辆禁止通行。

如果服务卡和通行卡都足够，绿灯亮，车辆可以通行，服务卡和通行卡都取走。

自动出卡机向P盒投放服务卡的速率和P盒的容量，就分别相当于双色双桶模式中的PIR和PBS。

PIR（Peak information rate）：峰值信息速率，表示向P桶中投放令牌的速率，即P桶允许传输或转发报文的峰值速率。PIR的值应大于CIR（存在服务卡足够而通行卡不够的情况）。

PBS（Peak Burst Size）：峰值突发尺寸，表示P桶的容量，即P桶瞬间能够通过的峰值突发流量。

按照国家高速公路交通安全法的规定，在高速公路上行驶的车辆，最高时速为120km/h。前面领取的通行卡，保证车辆的时速为100km/h。而服务卡的作用，则是允许车辆时速可以达到120km/h。自动出卡机向P盒投放服务卡的速率就是允许的最高车速，相当于PIR。

在双色双桶模式中，系统按照PIR速率向P桶中投放令牌，按照CIR速率向C桶中投放令牌。

如果P桶中可用令牌的总数量（ $T_p$ ）小于PBS，则P桶中令牌数增加。

如果C桶中可用令牌的总数量（ $T_c$ ）小于CBS，则C桶中令牌数增加。

对于到达的报文（报文大小为B），

如果 $T_p < B$ ，报文被标记为红色，且 $T_c$ 和 $T_p$ 都不减少。

如果 $T_c < B \leq T_p$ ，报文被标记为黄色，且 $T_p$ 减少B， $T_c$ 不减少。

如果 $B \leq T_c$ ，报文被标记为绿色，且 $T_p$ 和 $T_c$ 都减少B。

### 2.5、流量监管的配置步骤：

定义感兴趣流量：

```
[Huawei]acl 3001
```

```
[Huawei-acl-adv-3001]rule permit tcp source 172.16.1.254 0 source-port eq 20 destination 192.168.1.10 0
```

```
[Huawei-acl-adv-3001]rule permit tcp source 172.16.1.254 0 source-port eq 21 destination 192.168.1.10 0
[Huawei-acl-adv-3001]rule deny ip source any destination any
```

创建并进入传输类别，调用感兴趣流量

```
[Huawei]traffic classifier ftp
[Huawei-classifier-ftp]if-match acl 3001
```

创建并进入传输行为列表，配置GTS

```
[Huawei]traffic behavior atnet
[Huawei-behavior-atnet]car cir 512 pir 768 cbs 96256 pbs 160256 green pass yellow discard red discard
```

创建并进入传输策略，匹配传输类与传输行为

```
[Huawei]traffic policy atnet
[Huawei-trafficpolicy-atnet]classifier ftp behavior ftp
```

进入相应的接口并调用传输策略

```
[Huawei]interface g0/0/1
[Huawei-GigabitEthernet0/0/1]traffic-policy atnet outbound
```

## 九、拥塞的产生

### 1、拥塞发生的主要场景：

- 1.1、速率不匹配：数据从高速链路进入设备，再由低速链路转发出去
- 1.2、汇聚：数据从多个接口同时进入设备，由一个没有足够带宽的接口转发出去
- 1.3、多对一交换矩阵：多个同样速率的设备同时向某一台设备发送数据

### 2、拥塞可能会引发一系列的负面影响：

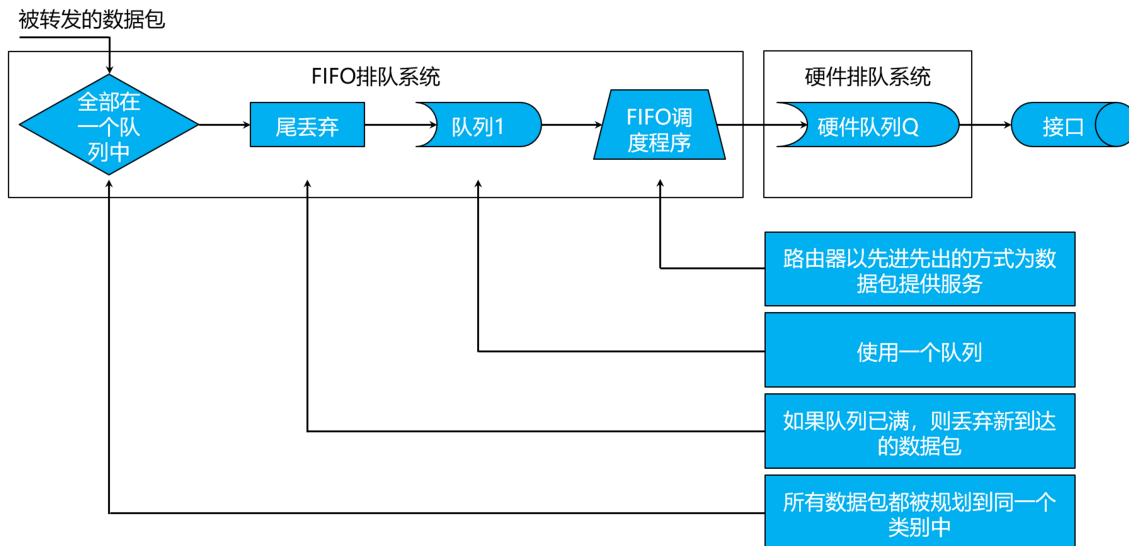
- 2.1、增加了数据传输的延迟与抖动
- 2.2、过高的延迟会引起数据的重传
- 2.3、使网络的有效吞吐率降低，造成网络资源的损害
- 2.4、加剧耗费大量的网络资源，不合理的资源分配甚至可能导致系统陷入资源死锁而崩溃

### 3、拥塞管理实现的方式

- 3.1、根据数据所携带的标记对应的LP【本地优先级】值与队列索引号的对应关系表，来实现将不同的数据进入不同队列
- 3.2、在拥塞发生时使用队列调度来完成不同队列数据的转发顺序

### 4、常用的队列调度算法包括：

- 4.1、FIFO【First In First Out，先进先出队列】
  - 4.1.1、LAN网络接口默认的排队方法是FIFO排队
  - 4.1.2、将所有出站帧加入到同一个队列中，而不考虑其类别
  - 4.1.3、FIFO不使用分类，将所有数据包当作同一种类别
  - 4.1.4、运行Huawei VRP且没有启用QoS的网络设备的默认行为
  - 4.1.5、无需使用任何命令来配置FIFO排队



## 4.2、WFQ【Weighted Fair Queue, 加权公平队列】

4.2.1、WFQ按数据流进行划分，由系统自动完成，无需管理员手工干预；不再按照类别进行划分

4.2.2、系统使用HASH算法自动完成数据入队，尽量将不同特征的流入不同的队列中

4.2.3、出站时，WFQ按流的权重来分配每个流应占有的出口带宽；权重越小，所得带宽越少；权重越大，所得带宽越多

4.2.4、保证相同权重业务之间的公平，体现了不同优先级业务之间的权值

4.2.5、WFQ的设计理念：

- a、为每个数据流创建一个专用的队列，避免某个队列的高饥饿、高延迟和高抖动
- b、在所有数据流之间公平、正确地分配带宽
- c、由于带宽分配的过于公平，为满足某些敏感数据流，引入了加权的概念
- d、WFQ是WAN网络接口带宽小于等于2M时的默认的队列策略

4.2.6、WFQ的优点：

- a、配置简单【无需管理员手工分类】
- b、保证所有的数据流均能获得一定的带宽
- c、大多数厂商均可支持

4.2.7、WFQ的缺点：

- a、流是自动分类的，无法手工干预
- b、缺乏一定的灵活性
- c、受资源限制，当多个流进入同一个队列时无法提供精确服务
- d、无法保证每个流获得的实际资源量
- e、WFQ均衡各个流的延迟与抖动，不适合延迟敏感的业务应用

## 4.3、CBQ【Class-Based Queue, 基于类别的队列】

4.3.1、CBQ是基于类的加权公平队列，是对WFQ功能的扩展，为用户提供了自定义【类】的支持

4.3.2、CBQ并不为所有的流量维护【流】的信息

4.3.3、CBQ可根据DSCP优先级、入接口、IP数据的五元组等规则来对数据进行分类，之后令不同类别的数据进入不同的队列

4.3.4、对于不匹配任何类别的数据，则会匹配进入系统定义的缺省类

4.3.5、CBQ提供三类队列：

- a、EF队列：满足低时延业务
- b、AF队列：满足需要带宽保证的关键数据业务
- c、BE队列：满足不需要严格QoS保证的尽力发送业务

4.3.6、CBQ与WFQ的区别：

	WFQ	CBQ
分类方式	管理员无法控制分类，由HASH算法自行决定	管理员根据类别自行分配流量
带宽的分配	对正常流量处理没问题，但对敏感数据流量太过于【公平】	考虑到公平特性，同时可为敏感数据流量分配高比例带宽

#### 4.4、PQ【Priority Queue，优先级队列】

4.4.1、PQ是为针对关键业务应用而设计的，即需要在拥塞发生时要求优先获得服务以减少响应的延迟

4.4.2、PQ调度机制分为4个队列：高优先队列、中优先队列、正常优先队列及低优先队列，其优先级依次降低

4.4.3、在数据包出站时，PQ首先发送高优先队列中的数据，直至高优先队列中的数据发送完毕，再发送中优先队列中的数据；直至发送完中优先队列的数据，再发送正常优先队列及低优先队列

4.4.4、对于语音应用来说，优先级排队很有用

4.4.5、语音通信流将被加入到高优先队列中；其余非关键业务数据将会在处理关键业务数据的空闲间隙被发送

4.4.6、PQ的缺点：若高优先队列中持续有数据等待发送，则其余优先队列中的数据一直无法被发送，呈现出无带宽可用的现象

#### 4.5、WRR【Weighted Round Robin，加权轮循队列】

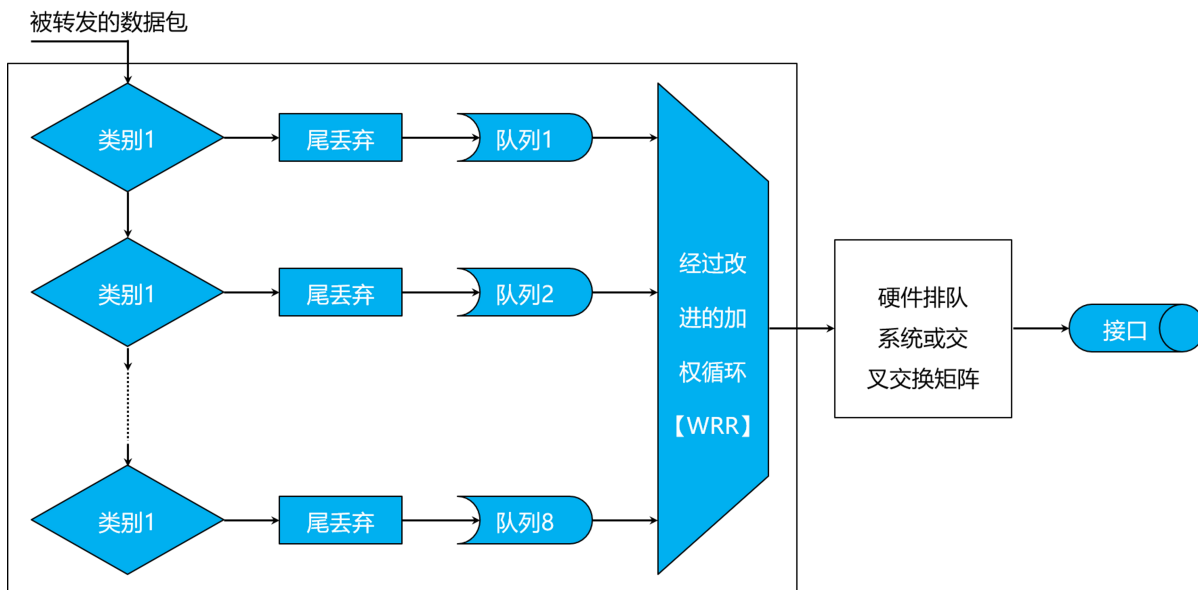
4.5.1、WRR根据每个队列的权重来轮流调度各队列中的数据

4.5.2、为不同通信流类别提供不同服务，为每个出站队列配置权重值，权重值越大，设备给出的优先级越高

4.5.3、WRR在轮循时，每个队列享受的机会和该队列的权重成正比，根据每个队列的权重值进行轮循调度

4.5.4、调度一轮权重值减1，权重值减至0的队列不参与调度；所有队列的权重值减至0时，开始下一轮的调度

4.5.6、每个队列的默认权重值为10



### 十、拥塞避免

1、拥塞避免机制能够预测和避免常见的网络瓶颈点发生拥塞

2、通过使用复杂的算法【而不是简单的尾丢弃算法】来丢弃数据包

3、拥塞避免的3种方法：

#### 3.1、尾丢弃

由于每个队列长度有限，当某一队列被装满时，传统的处理方法是将后续向该队列发送的数据全部丢弃，直至拥塞解除，这种处理方式称为尾丢弃【Tail Drop】



尾丢弃平等对待所有通信流，而不管其内部DSCP值大小

缺点：

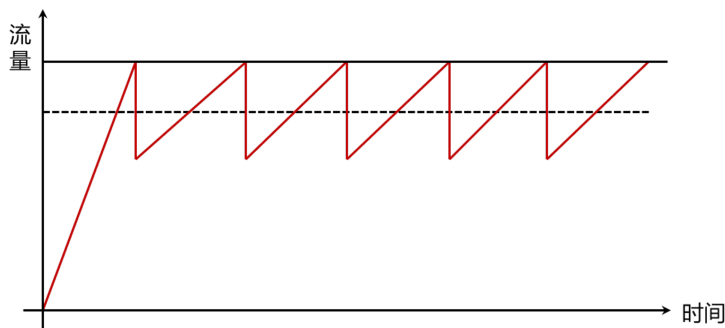
3.1.1、丢弃帧通常会影响到正在进行的TCP会话，随意丢弃TCP会话中的帧将导致并行的TCP会话降低速度，然后再提高速度，形成【锯齿】效果在拥塞点，出站队列中排队的数据包过多将导致延迟和抖动

3.1.2、因为没有丢弃机制，高优先级通信流与尽力而为通信流待遇相同

注：TCP全局同步的解释

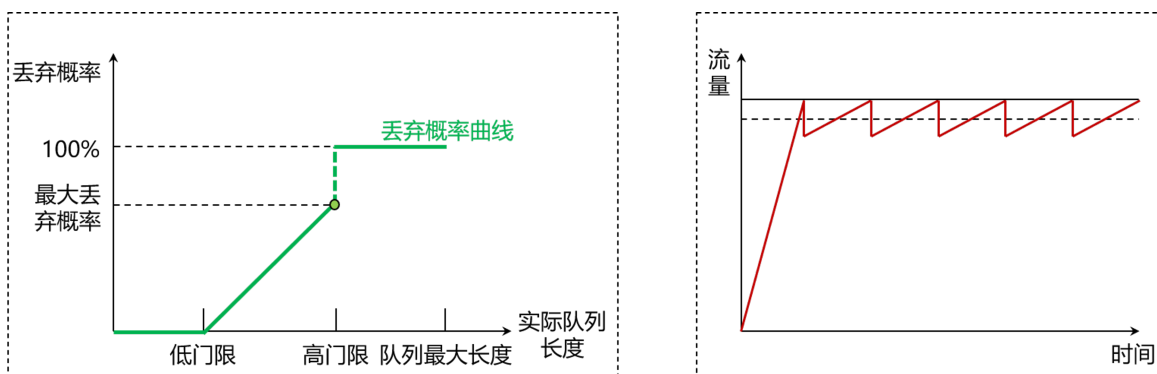
1、对于TCP报文，若大量的数据被丢弃，将造成TCP超时，从而引发TCP慢启动，使得TCP减少报文的发送；当队列同时丢弃多个TCP连接的数据时，将造成多个TCP连接同时进入拥塞避免和慢启动状态以调整并降低流量，这种现象被称为TCP全局同步

2、多个TCP连接发往队列的数据将同时减少，而后又在某个时刻同时出现流量高峰，如此往复，使网络资源利用率低下



### 3.2、RED【随机早期检测】

为避免TCP全局同步，可在队列未满载时先随机丢弃一部分数据，通过预先降低一部分TCP连接的传输速率来尽可能延缓TCP全局同步的到来这种预先随机丢弃数据的行为被称为随机早期检测【RED】



缺点：在丢弃时不考虑数据的优先级，容易造成丢弃了大量的低优先级的重要数据

### 3.3、WRED【加权随机早期检测】

3.3.1、WRED可以弥补尾丢弃的三个缺点，且大幅提高链路带宽利用率

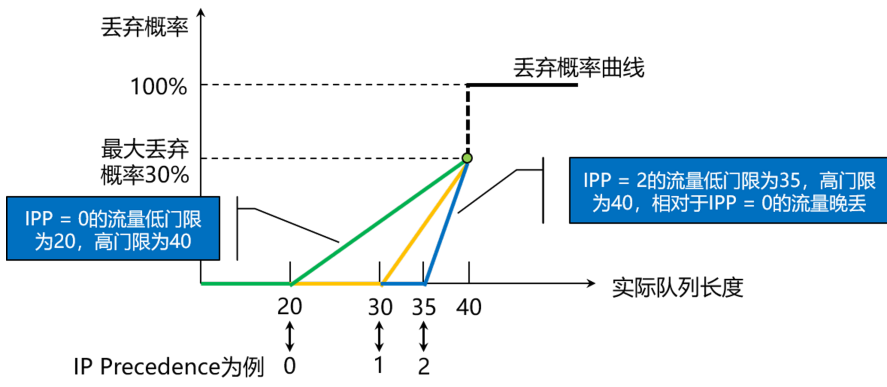
3.3.2、WRED技术可以通过对不同优先级数据包或队列设置相应的丢弃策略，以实现针对不同流量进行区分丢弃

3.3.3、通过使用WRED技术，可实现每一种优先级都能独立设置数据丢包的高门限、低门限及丢包率

3.3.4、数据到达低门限时，开始丢包，到达高门限时丢弃所有数据

3.3.5、随着门限的增高，丢包率不断增加，最高丢包率不超过设置的最大丢包率，直至到达高门限，数据全部丢弃

3.3.6、按照一定的丢弃概率主动丢弃队列中的数据，从一定程度上避免了尾丢弃带来的所有缺点



十一、QoS的配置  
 详细配置见实验手册