

# 第7章

## QoS服务

# 网络服务

---



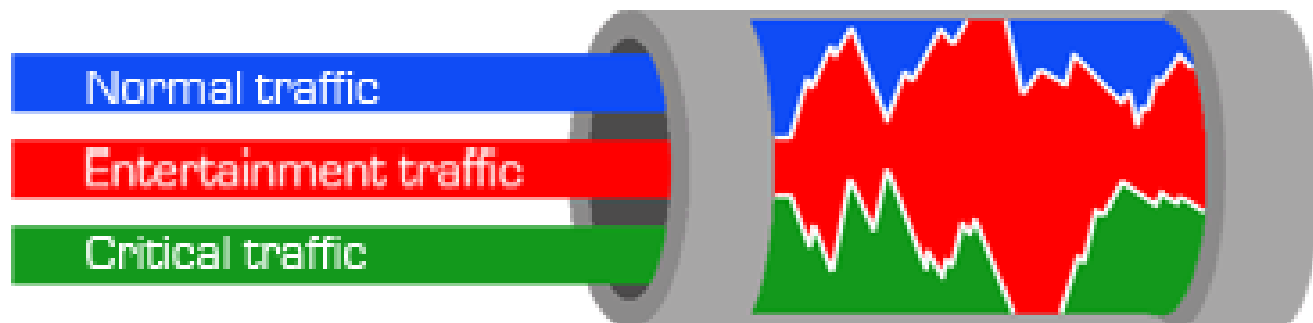
## QoS技术

# 服务质量★

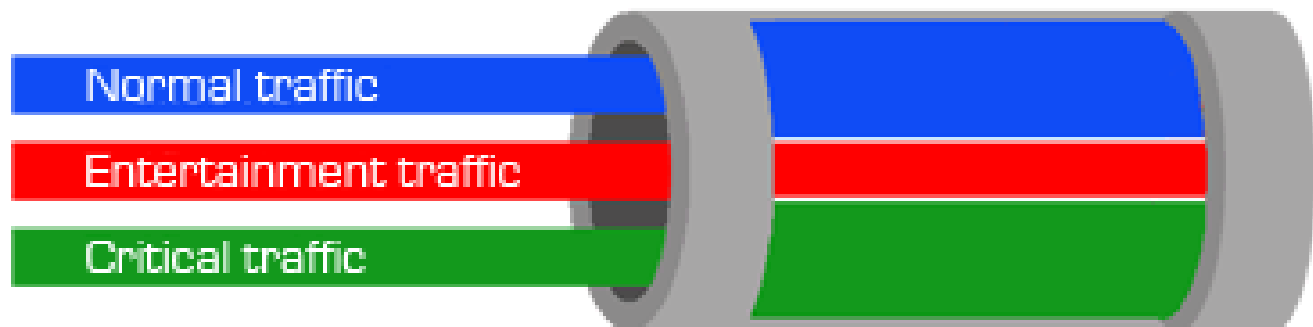
- QoS (Quality of Service) 即服务质量。
- 对于网络业务，服务质量包括：
  - ◆ 传输的带宽 (Bandwidth)、
  - ◆ 传送的时延 (Delay)、
  - ◆ 时延抖动 (Jitter)
  - ◆ 数据的丢包率 (Loss)、

# 带宽 (Bandwidth)

Bandwidth Use without Qos control



Bandwidth Use with QoS control



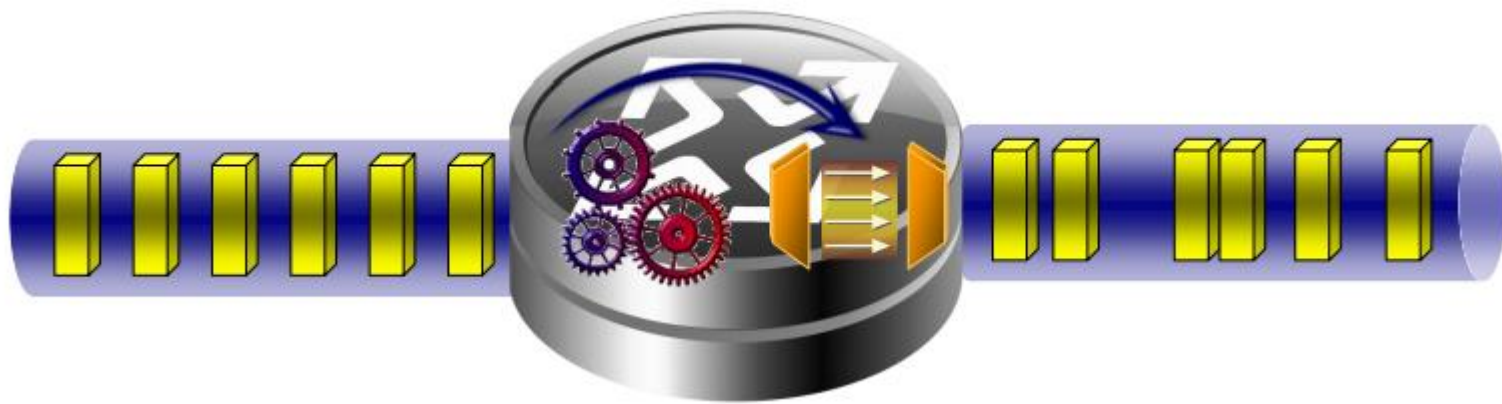
# 时延 (Delay)

- 端到端的时延由传输时延、处理时延和队列时延组成。



# 抖动 (Jitter)

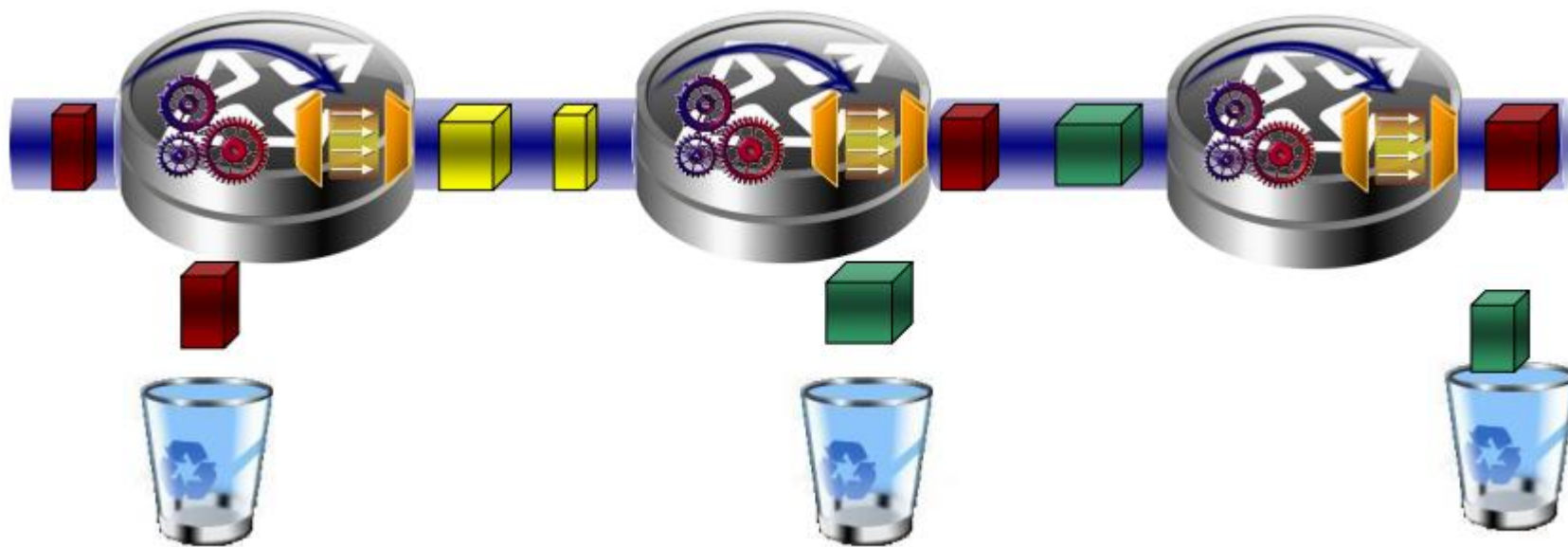
- 抖动是由于属于同一个流的数据包的端到端**时延不相等**造成的。



**抖动**的大小跟**时延**的大小直接相关

# 丢包 (Loss)

- 在很多时候, 丢包一般是因为**队列满**造成的, 在队列满的时候, 一般采用**尾丢弃** (把最后到达的包丢弃) 来进行丢包。



# QoS服务模型简介★

## ■ 通常QoS提供以下三种服务模型：

### ◆ Best-Effort service

（尽力而为服务模型）

### ◆ Integrated service

（综合服务模型，简称Int-Serv）

### ◆ Differentiated service

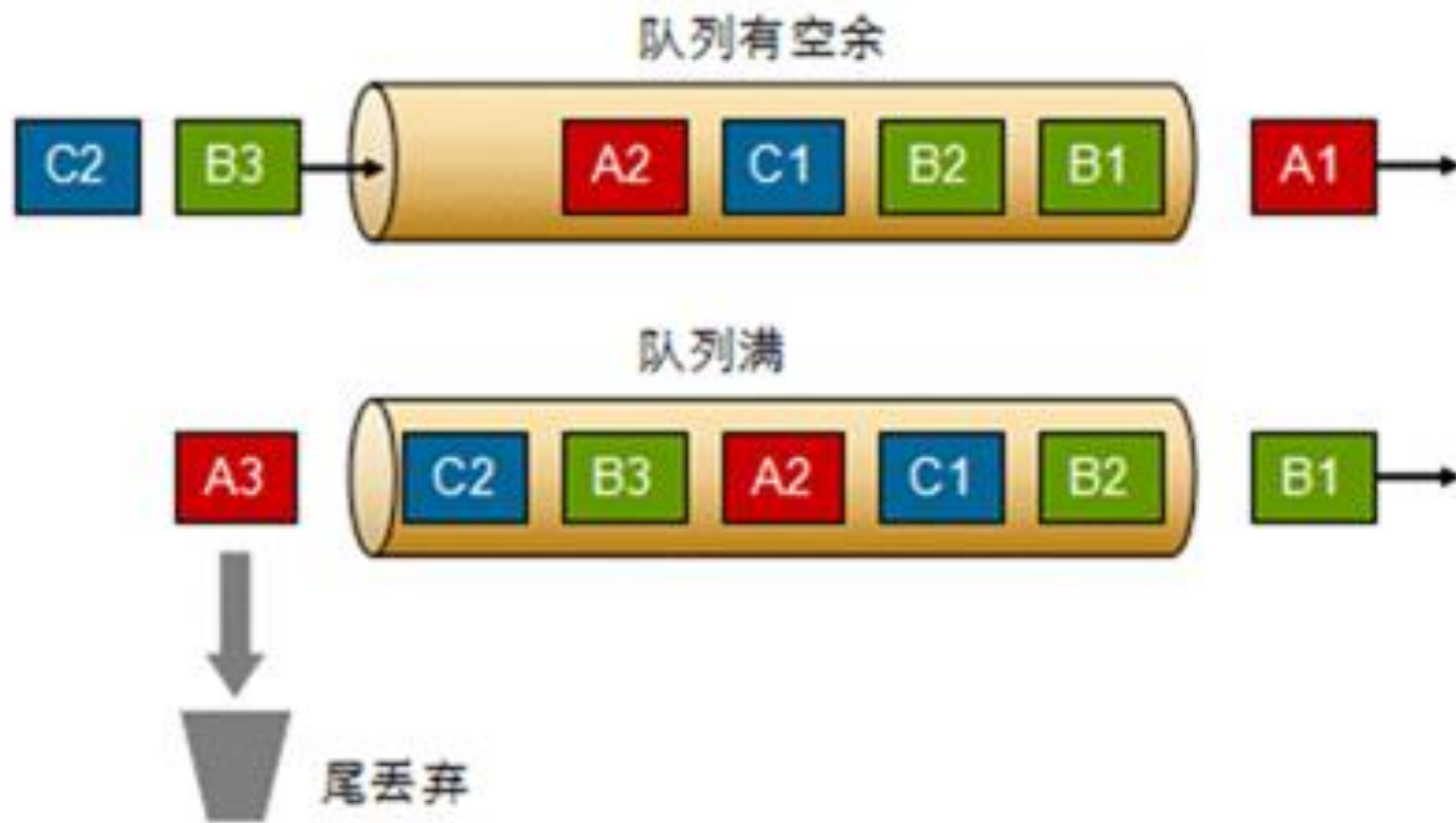
（区分服务模型，简称Diff-Serv）



# 1. Best-Effort服务模型

- Best-Effort, **尽力而为**的服务模型：  
网络尽最大的可能性来发送报文。
  - ◆ 对时延、可靠性等性能**不提供**任何保证。
  - ◆ Best-Effort服务模型是网络的**缺省服务**模型，通过FIFO队列来实现。
  - ◆ 提供的带宽远大于实际需求时  
(Overprovisioning) 可以保证服务质量。

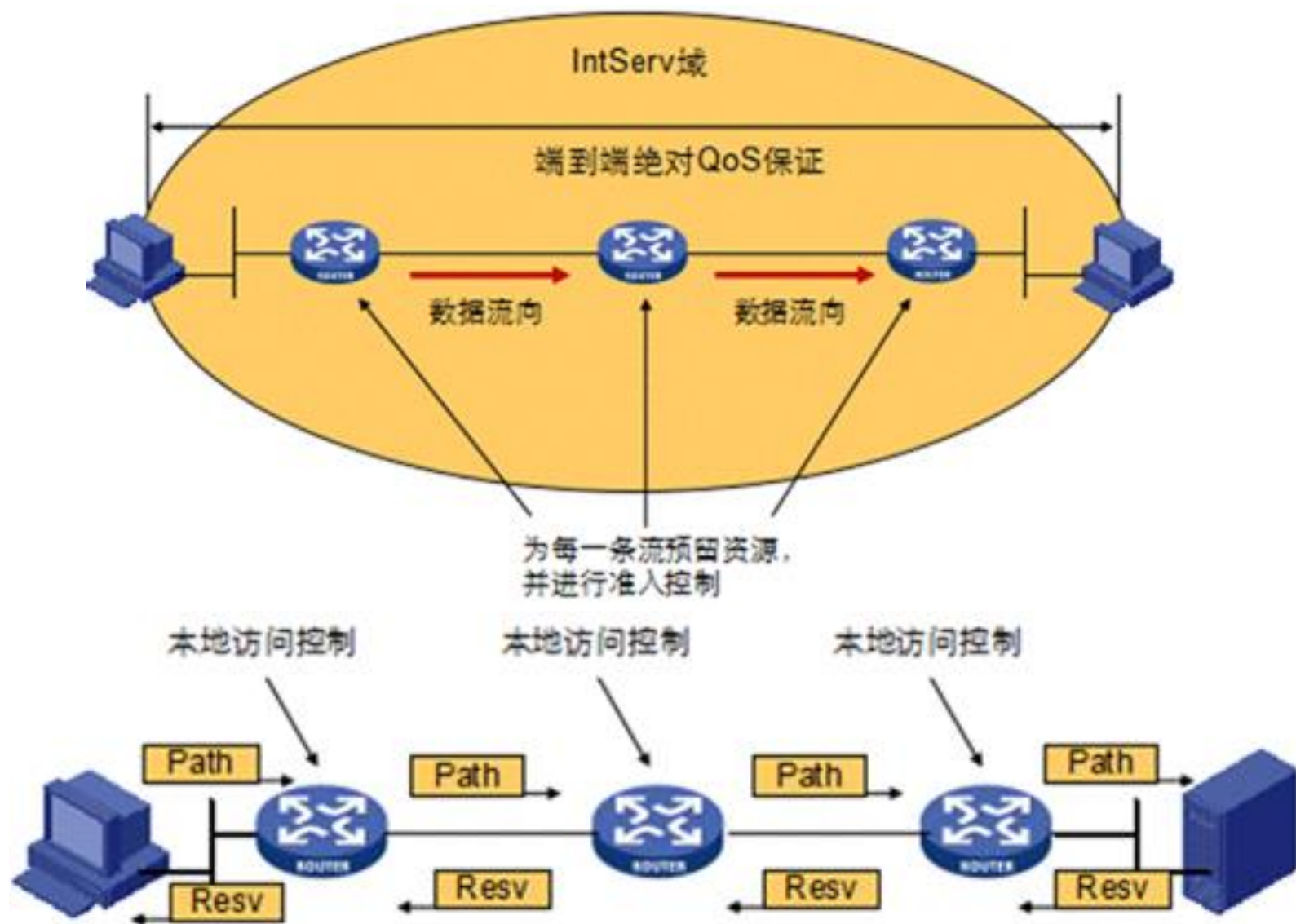
# 图示



## 2. Int-Serv服务模型

- **Int-Serv (Integrated Service), 集成服务模型**：需要向网络申请特定的服务，网络确定完成资源预留后使用。
  - ◆ 要求所有路由器和终端设备均支持QoS。
  - ◆ **资源预留协议 (Resource Reservation Protocol, RSVP)**
  - ◆ 可扩展性差，难以在Internet核心网络实施。

# 图示



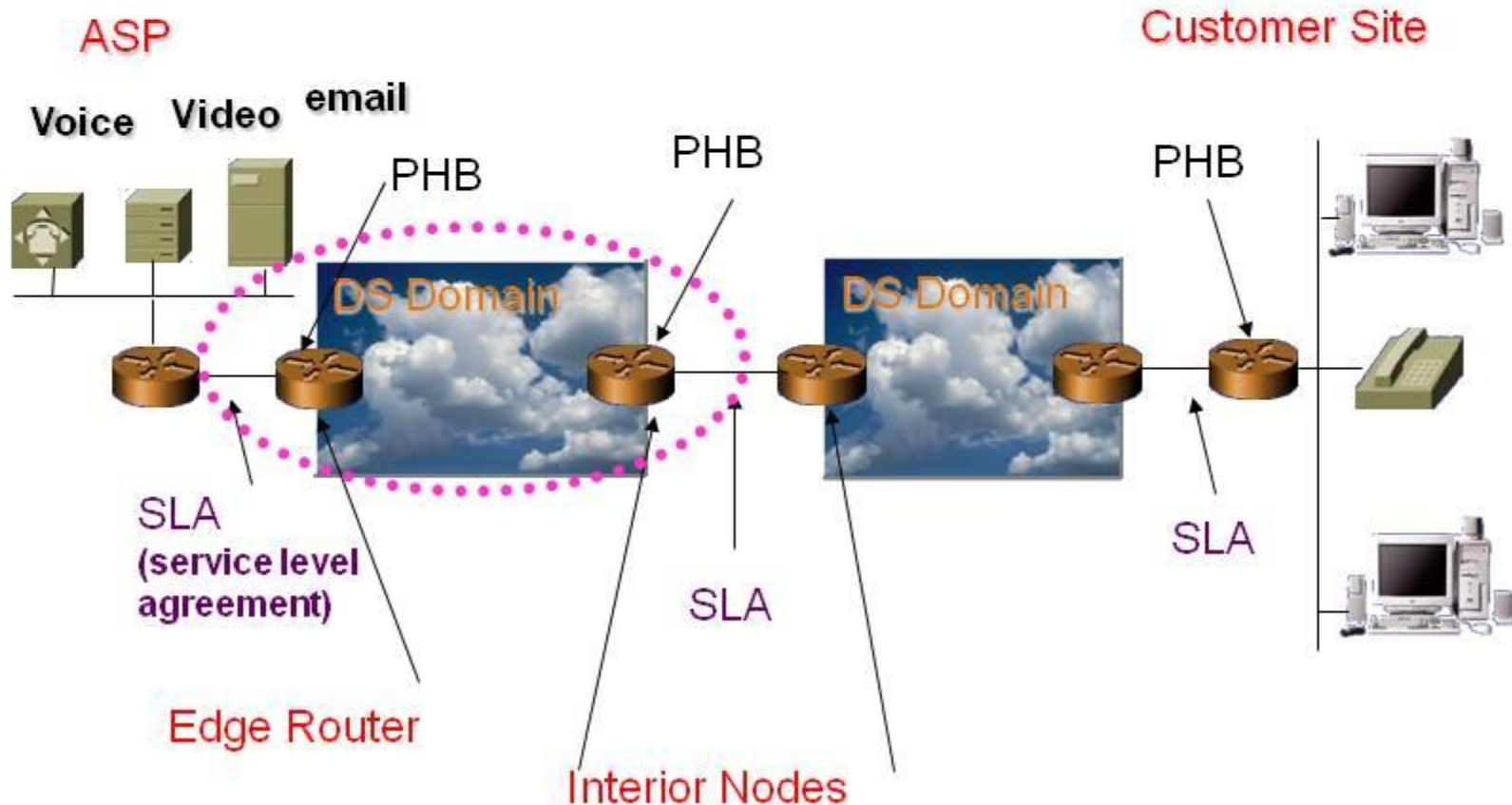
### 3. Diff-Serv服务模型

- Diff-Serv (Differentiated Service) , 差分服务模型。通过报文头部的QoS参数信息, 告知网络节点它的QoS需求。
  - ◆ DS 边界节点, 进行流量控制并设置相关标记
  - ◆ DS 内部节点, 根据相关标记实现流分类和QoS服务。

# DiffServ Operation

- Each ISP configures its own routers to match the service that it offers, and each ISP has its own DiffServ Domain.

⑩per-hop behaviors (PHBs)



# 网络服务

---



## QoS实现

# QoS组件

■ DiffServ模型有如下四个QoS组件：

◆ **流分类和标记** (classification and marking)

- 将数据包分为不同的类别，称为**流分类**，流分类并不修改原来的数据包。
- 将数据包设置为不同的优先级称为**标记**，而标记会修改原来的数据包。



# QoS组件

- DiffServ模型有如下四个QoS组件：
  - ◆ 流分类和标记 (classification and marking)
  - ◆ 流量监管和整形 (Policing and Shaping) :
    - 将超过的流量**丢弃**的技术称为流量监管
    - 将超过的流量**缓存**的技术称为流量整形。

# QoS组件

- DiffServ模型有如下四个QoS组件：
  - ◆ 流分类和标记 (classification and marking)
  - ◆ 流量监管和整形 (Policing and Shaping) :
  - ◆ 拥塞管理 (Congestion management)
    - 在网络发生拥塞时，将报文放入队列中缓存，并采取某种调度算法安排报文的转发次序。

# QoS组件

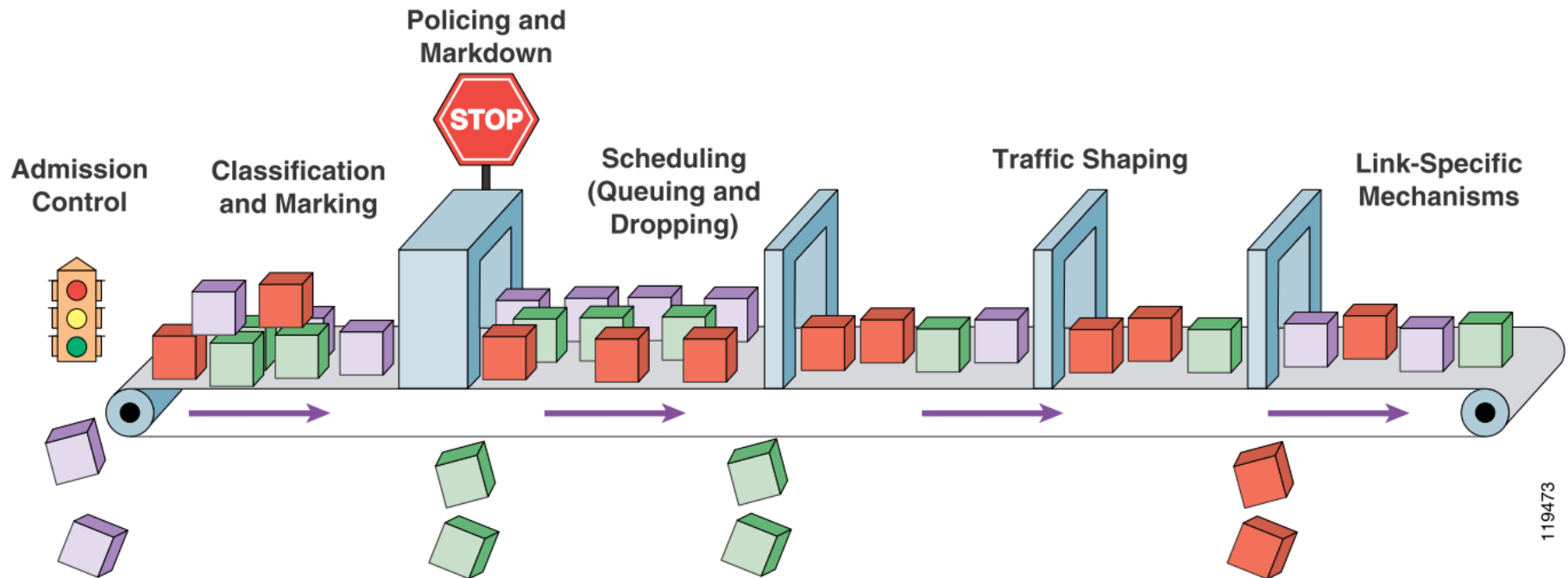
## ■ DiffServ模型有如下四个QoS组件：

- ◆ 流分类和标记 (classification and marking)
- ◆ 流量监管和整形 (Policing and Shaping) :
- ◆ 拥塞管理 (Congestion management)
- ◆ 拥塞避免 (Congestion avoidance) :

- 监督网络资源的使用情况，当发现拥塞有加剧的**趋势**时采取**主动丢弃**报文的策略，通过调整流量来解除网络的过载。

# QoS 处理过程

Figure 1-1 The Cisco QoS Toolset



119473

⑩ 流分类: Classification and Marking

⑩ 流量监管: Policing and Markdown

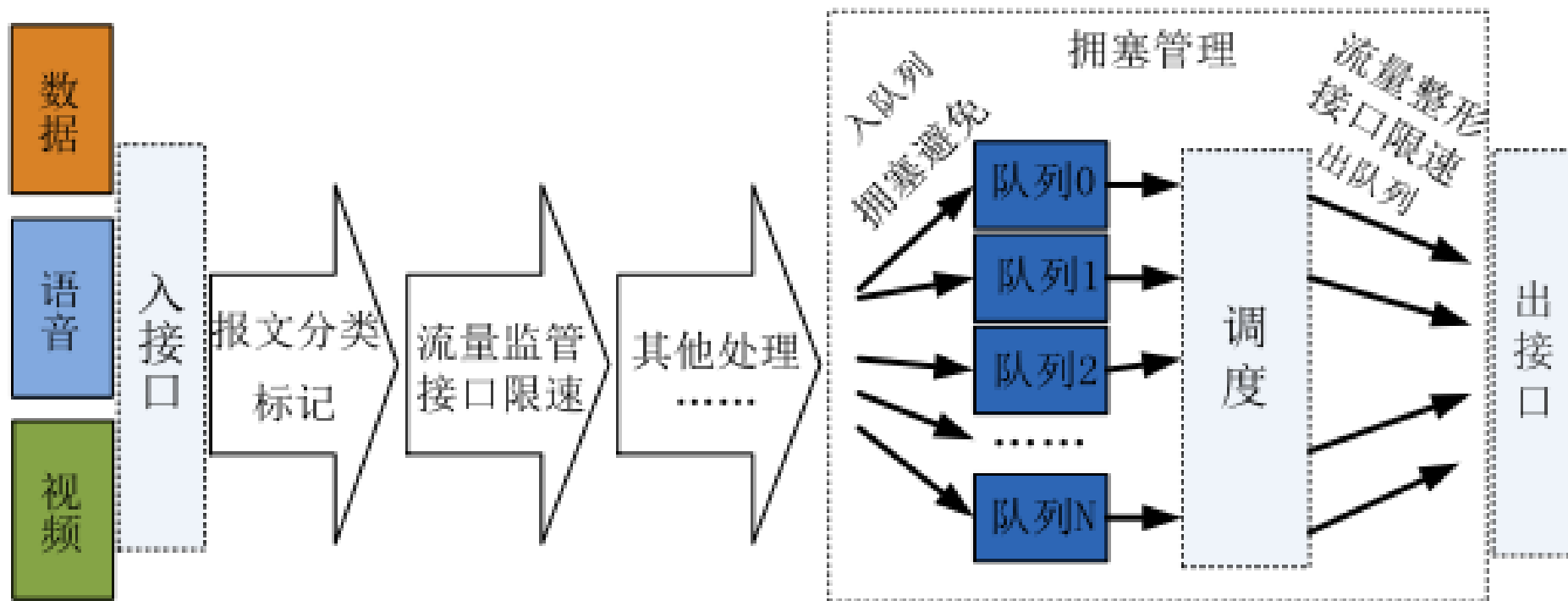
⑩ 拥塞管理: Scheduling

⑩ 流量整形: Traffic shaping

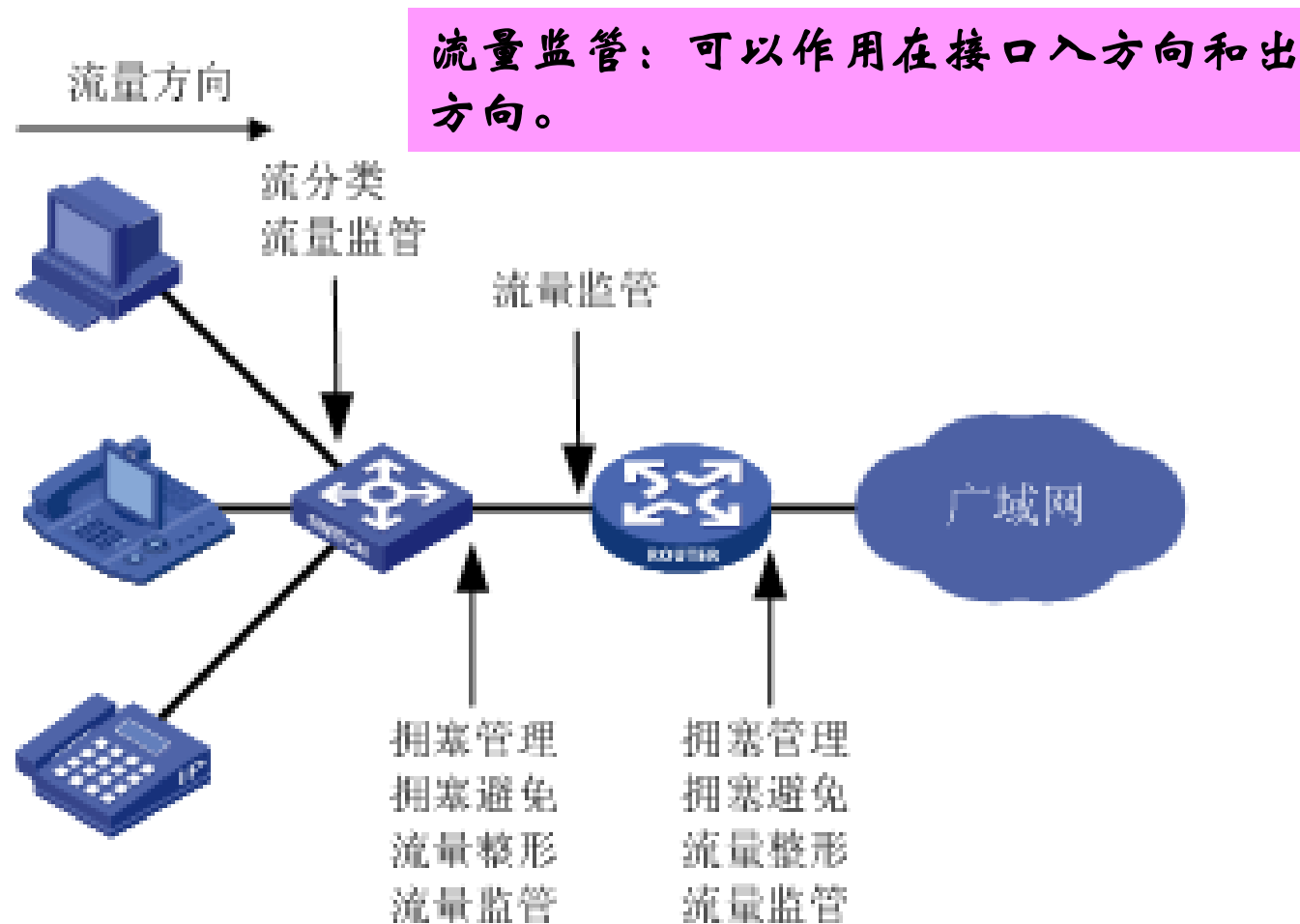
⑩ 拥塞避免: Congestion avoidance

# 流程

图 1-1 QoS 技术处理流程



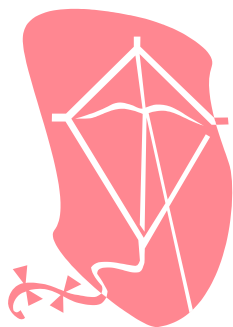
# QoS技术在网络中的位置



# QoS服务

---

## 流量分类



# 流量分类

- **流量分类**是按照一定的规则识别符合某类特征的报文。
- 按照分类规则参考信息的不同，分为：
  - ◆ **简单流分类**是根据报文的优先级
    - 如IP报文头中的DSCP/IP- PRE信息，MPLS报文头中的MPLS EXP信息，二层报文头中的802.1P信息来分类的。
  - ◆ **复杂流分类**是根据更多的报文信息
    - 如报文优先级、源IP、目的IP、源MAC、目的MAC、协议类型等信息来分类的。



# VLAN 优先级 (L2)

Figure 1-10

Figure 1-10 802.1Q/p CoS Bits



⑩ Class of service (CoS)

Three Bits for User Priority  
(802.1p CoS)

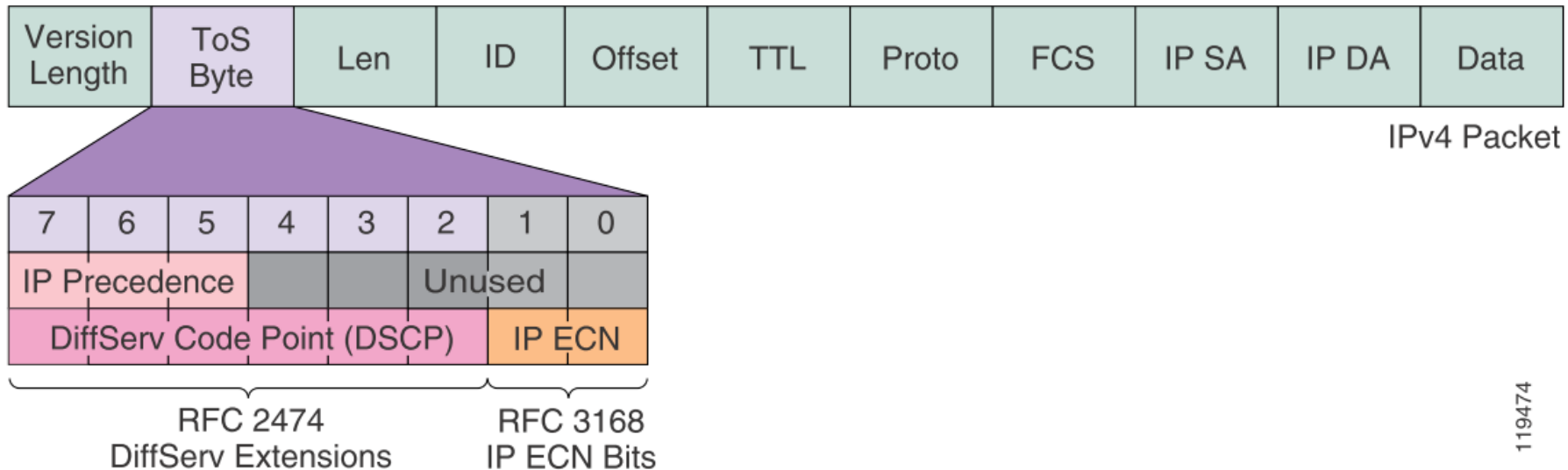
Class of service (CoS)

22660

- ⑩ 802.1Q协议，即Virtual Bridged Local Area Networks 协议，主要规定了VLAN的实现
- ⑩ 802.1p协议定义了优先级的概念。

# IP 优先服务位 (L3)

Figure 1-2 The IP ToS Byte (DSCP and IP ECN)



119474

⑩ IP precedence: IP优先级字段

⑩ Differentiated Services Code Point (DSCP):  
差分服务代码点

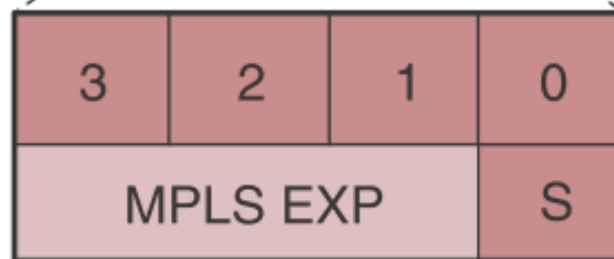
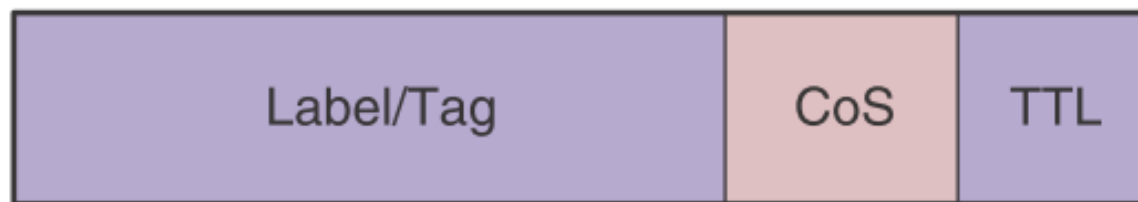
# MPLS 优先级(L2.5)

Figure 1-11

Figure 1-11 MPLS EXP Bits

Label/Tag:  
20 bits

Time to Live (TTL):  
8 bits



⑩ MPLS EXP

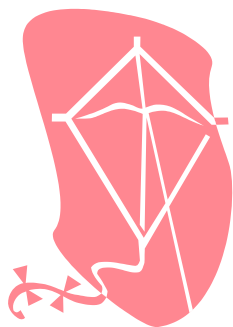
MPLS Experimental (CoS): 3 bits  
Bottom of stack indicator (S): 1 bit

226607

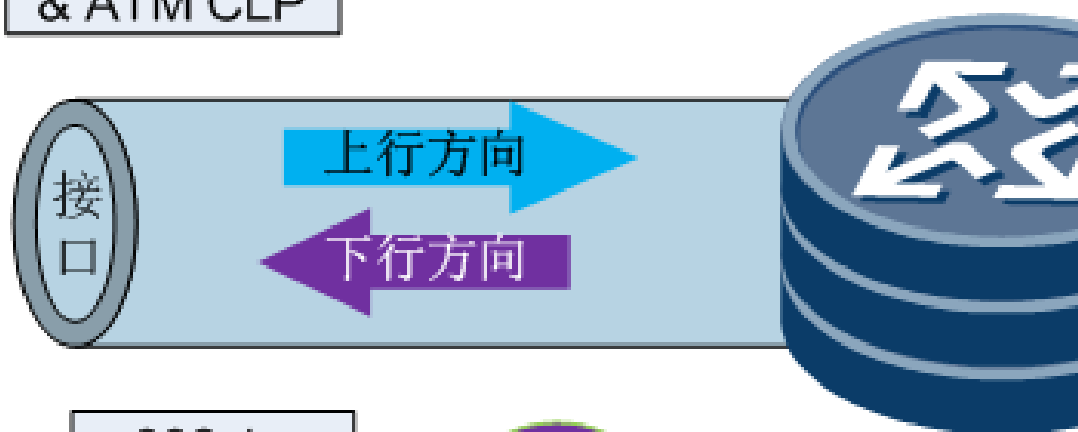
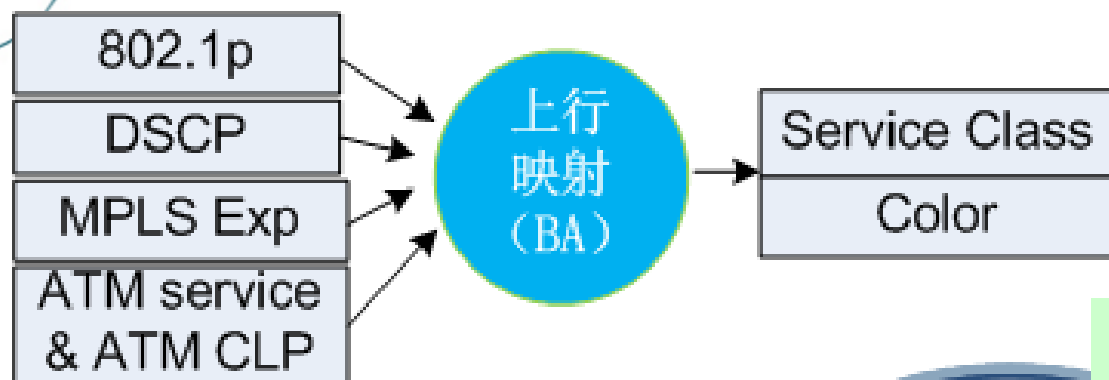
# QoS服务

---

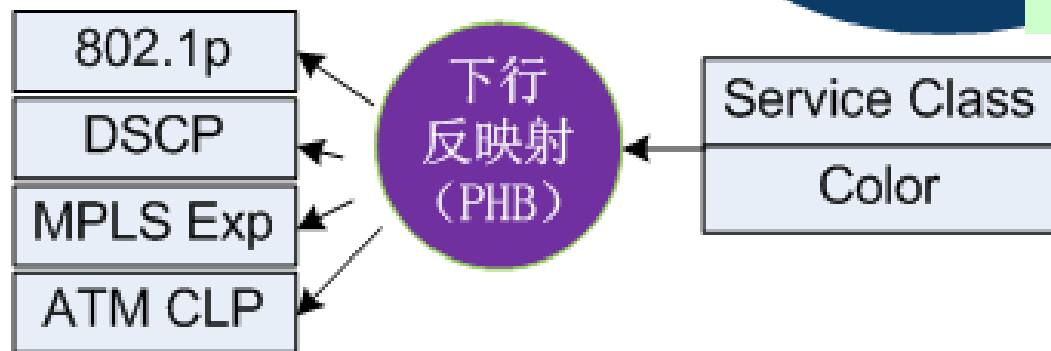
## 流量标记



# 分类和标记



⑩ 需要将报文携带的QoS优先级统一映射到设备内部的服务等级Service Class (也叫做调度优先级PHB) 和丢弃优先级 (也叫颜色Color),



# 链路中分类标识变更

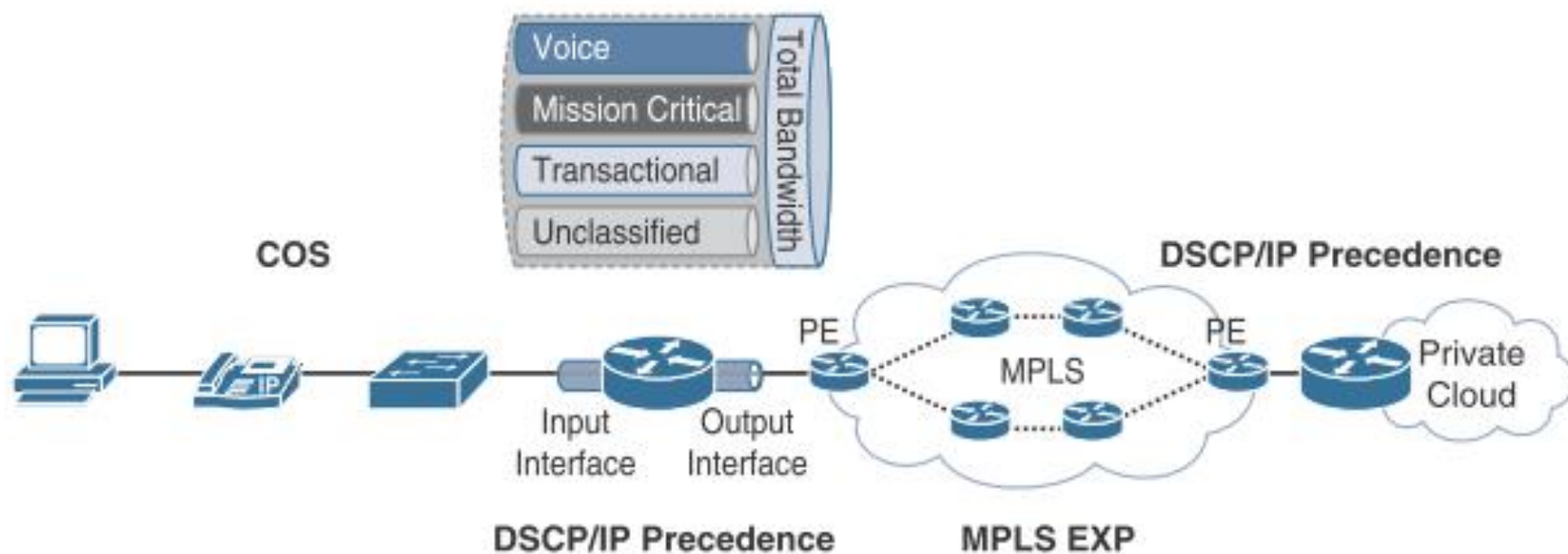
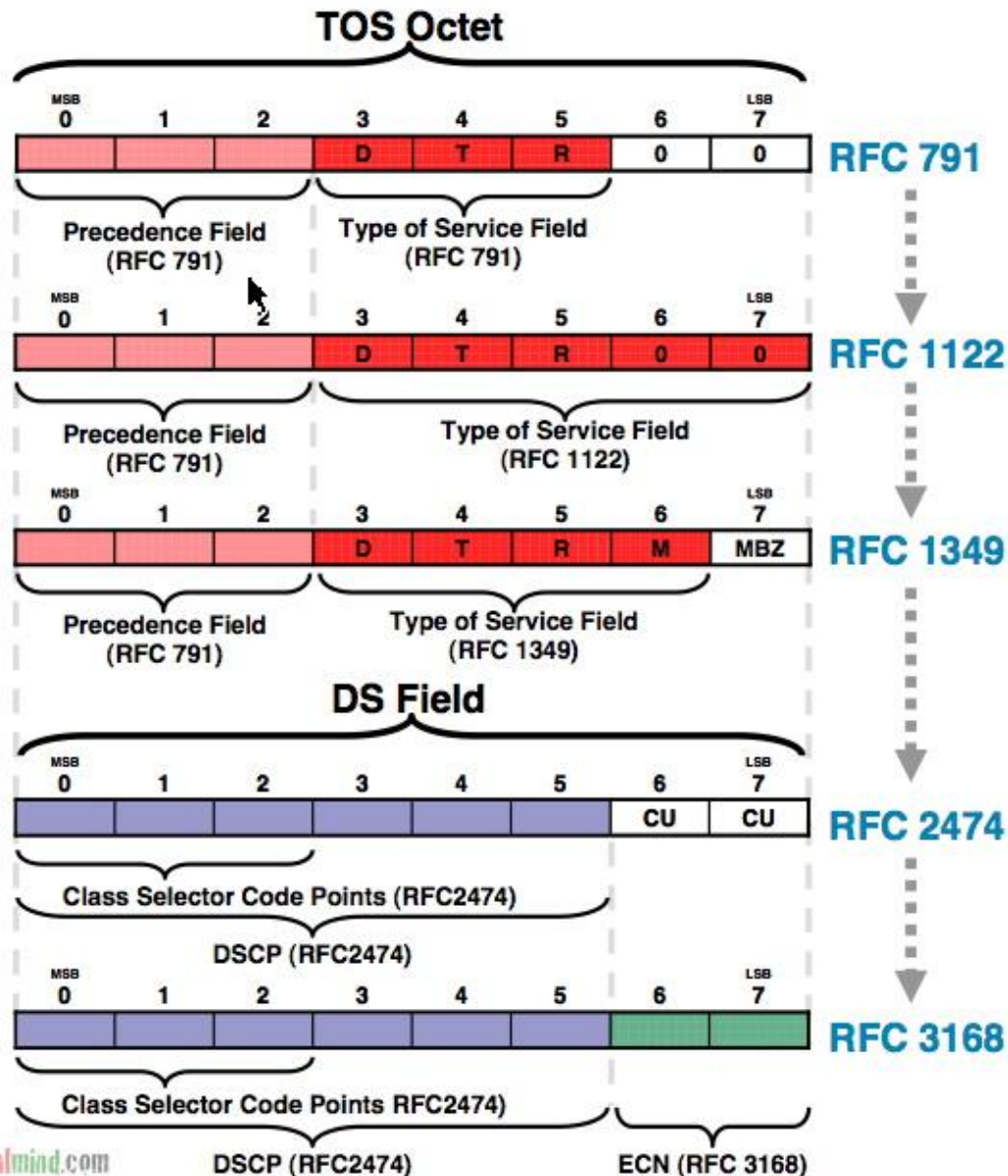


Figure 15-1 QoS Traffic Descriptors

# Evolution to the DS Field



- The Differentiated Services (DS) field has obsoleted both
  - The Type of Service Octet in IPv4
  - The Traffic Class field in IPv6
- 3 least significant bits are the class selector codepoints (CS)
  - Functionally equivalent to the precedence field

# PHB

- PHB (Per-Hop Behaviors), PHB是DS节点作用于数据流的行为。
  - ◆ 网络管理员可以配置DSCP到PHB的映射关系。
  - ◆ 如果报文未定义DSCP到PHB的映射, 则DS节点将选择采用缺省PHB
  - ◆ 每个DS节点必须支持缺省PHB。

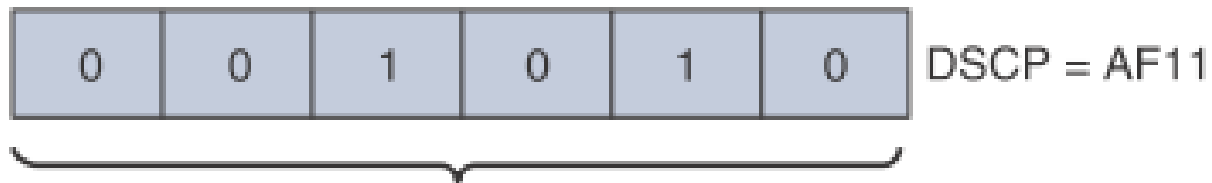


# PHB的分类★

## ■ PHB 的分类:

- ◆ Default PHB, Best-Effort (BE) 尽力而为, 000000, 只关注可达。
- ◆ Class-Selector PHB (CS), 类选择码, XXX000, 和 IP precedence 兼容
- ◆ Expedited Forwarding PHB (EF), 加速转发, 101110, DS节点发出的信息流速率在任何情况下必须获得等于或大于设定的速率。
- ◆ Assured Forwarding PHB (AfxY), 确保转发, XXX YY0, XXX对应IP Precedence(1-4); YY表示丢弃优先级(1-3)

# Afxy

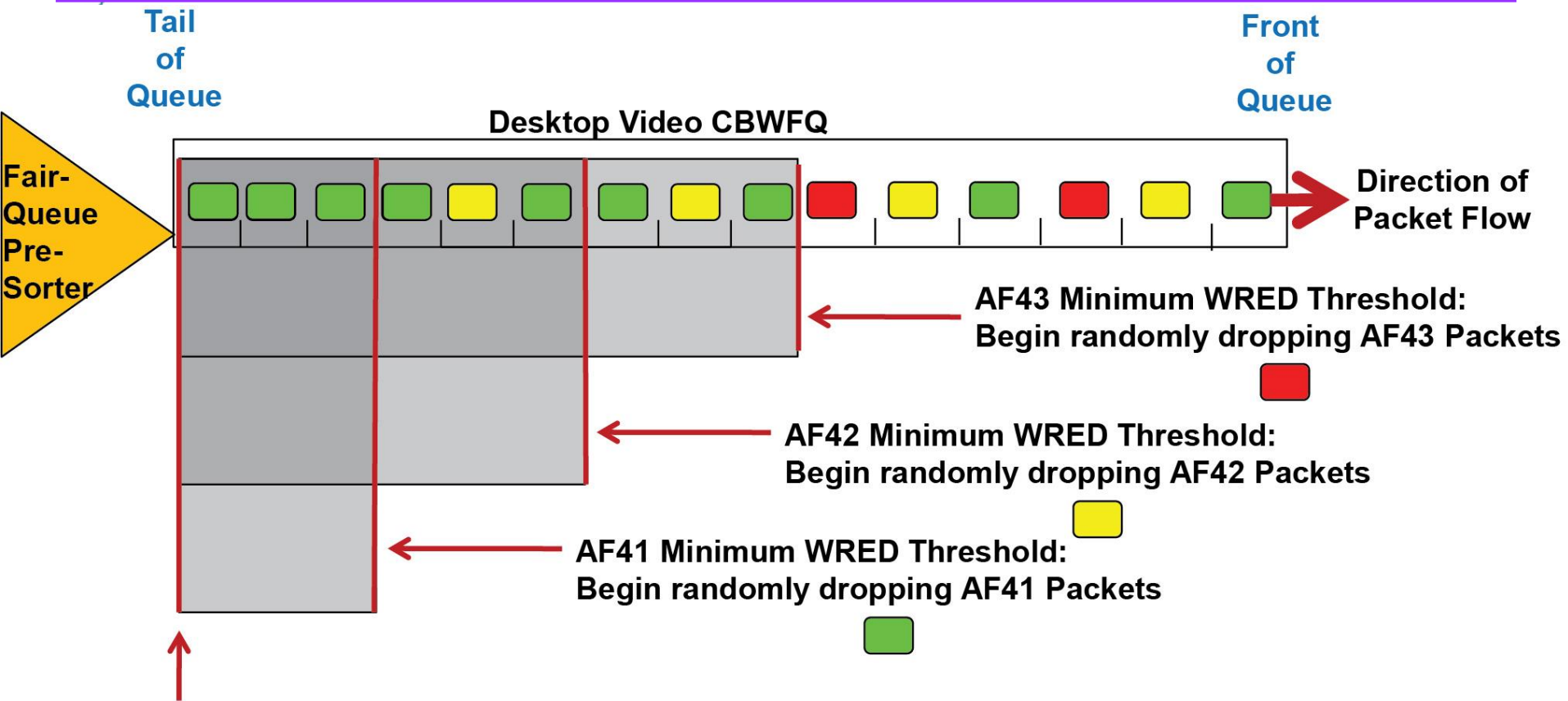


Class	Value		
AF1	001	dd	0
AF2	010	dd	0
AF3	011	dd	0
AF4	100	dd	0

Drop Probability (dd)	Value	AF Value
Low	01	AF11
Medium	10	AF12
High	11	AF13

**Figure 15-7** *AF/IPPrec Dropping Probabilities*

# 丢包示意



Maximum WRED Thresholds for AF41, AF42, and AF43 are set to the tail of the queue in this example.

348829

# PHB & DSCP

## Per-Hop Behaviors (PHB)

## DiffServ Code Points (DSCP)

Expedited Forwarding

EF

<sup>46</sup>  
101110

Assured Forwarding

Low Drop  
Pref

Med Drop  
Pref

High Drop  
Pref

Class 1

AF11

AF12

AF13

<sup>10</sup> 001010    <sup>12</sup> 001100    <sup>14</sup> 001110

Class 2

AF21

AF22

AF23

<sup>18</sup> 010010    <sup>20</sup> 010100    <sup>22</sup> 010110

Class 3

AF31

AF32

AF33

<sup>26</sup> 011010    <sup>28</sup> 011100    <sup>30</sup> 011110

Class 4

AF41

AF42

AF43

<sup>34</sup> 100010    <sup>36</sup> 100100    <sup>38</sup> 100110

Best Effort

BE

<sup>0</sup>  
000000

**Figure 3-1** *DSCP Encoding Scheme*

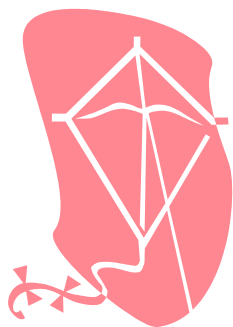
# PHB & DSCP

Application	L3 Classification			L2
	IPP	PHB	DSCP	CoS
Routing	6	CS6	48	6
Voice	5	EF	46	5
Video Conferencing	4	AF41	34	4
Streaming Video	4	CS4	32	4
Mission-Critical Data	3	AF31	26	3
Call Signaling	3	CS3	24	3
Transactional Data	2	AF21	18	2
Network Management	2	CS2	16	2
Bulk Data	1	AF11	10	1
Best Effort	0	0	0	0
Scavenger	1	CS1	8	1

# QoS服务

---

## 流量监管与整形



# 流量监管和整形

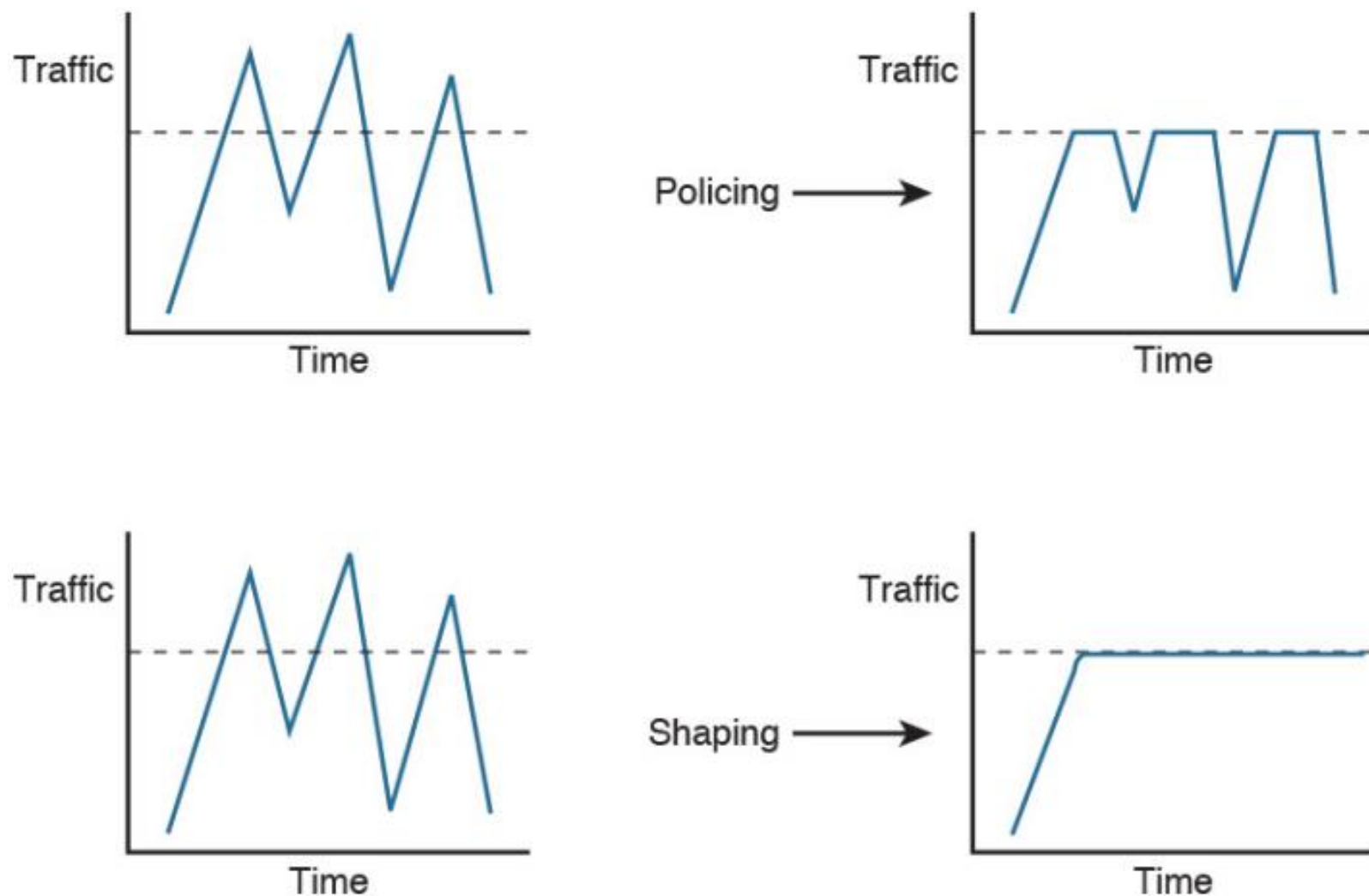


Figure 14-25 Policing and shaping of traffic

# 流量监管与整形



## 流量监管

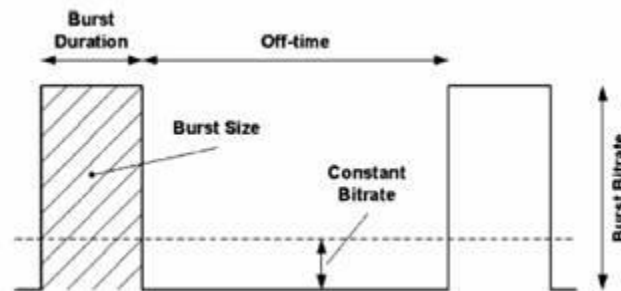
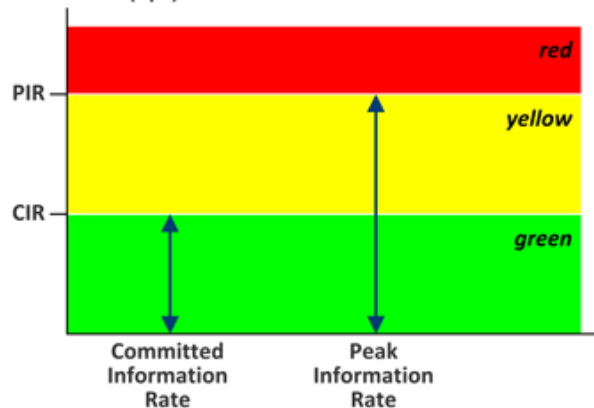


# SLA★

- 运营商之间都签有**服务水平协议 (SLA)**，其中包含每种业务流的流量参数：
  - ◆ 承诺速率CIR (Committed Information Rate)、
  - ◆ 峰值速率PIR (Peak Information Rate)、
  - ◆ 承诺突发尺寸CBS (Committed Burst Size)、
  - ◆ 峰值突发尺寸PBS (Peak Burst Size)

Two Rates & Three Colors

Information Rate (bps)



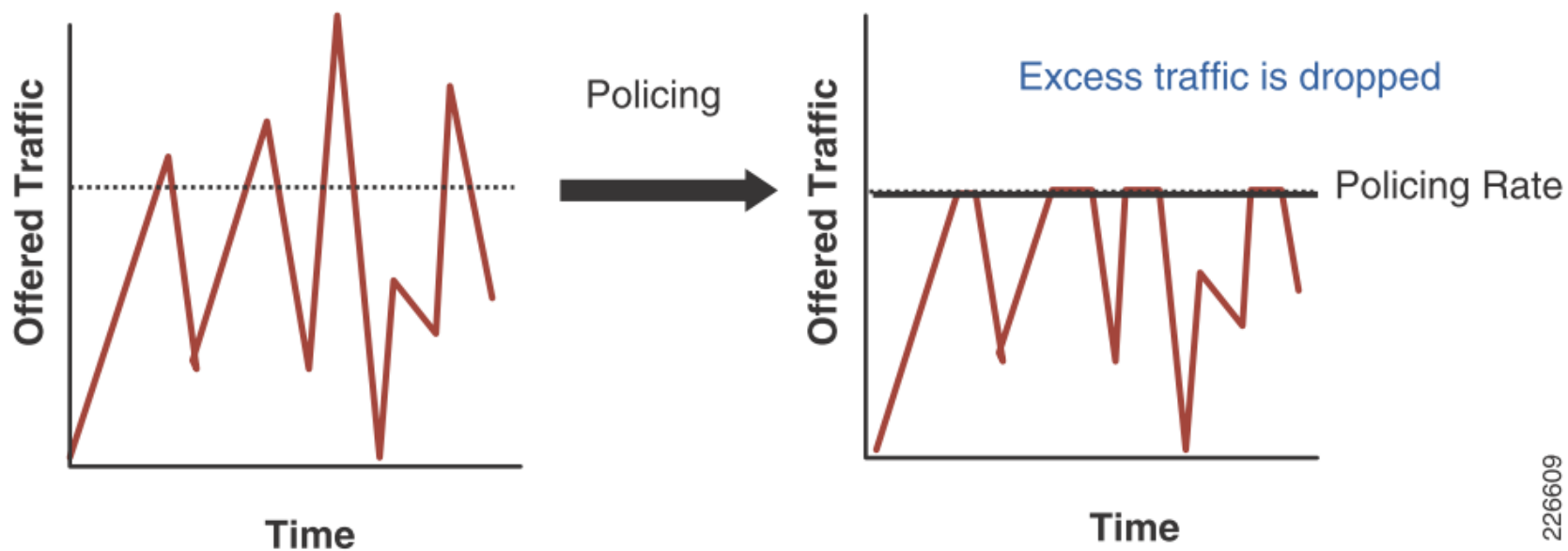
# 流量监管★

- **流量监管TP (Traffic Policing)** 是指对进入设备的流量进行监控，确保其没有滥用网络资源。
- 对超出SLA约定的流量报文可指定给予
  - ◆ pass (通过)、
  - ◆ drop (直接丢弃)
  - ◆ markdown (降级)，
    - 降级是指降低服务等级 (Service Class)，或者是提高丢弃等级 (Color)，

⑩ 转发 (pass)  
⑩ 丢弃 (discard)  
⑩ 重标记 (remark)

# 流量监管—丢弃

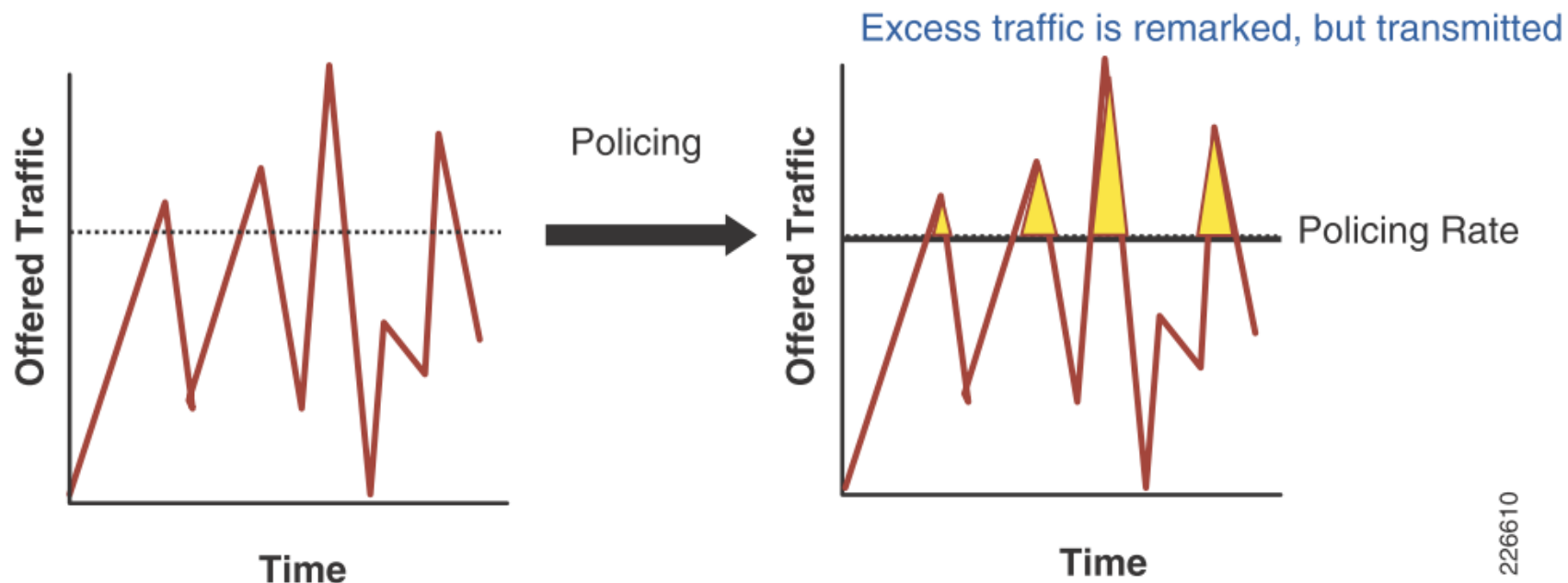
Figure 1-13 A Policing Action



226609

# 流量监管—标记

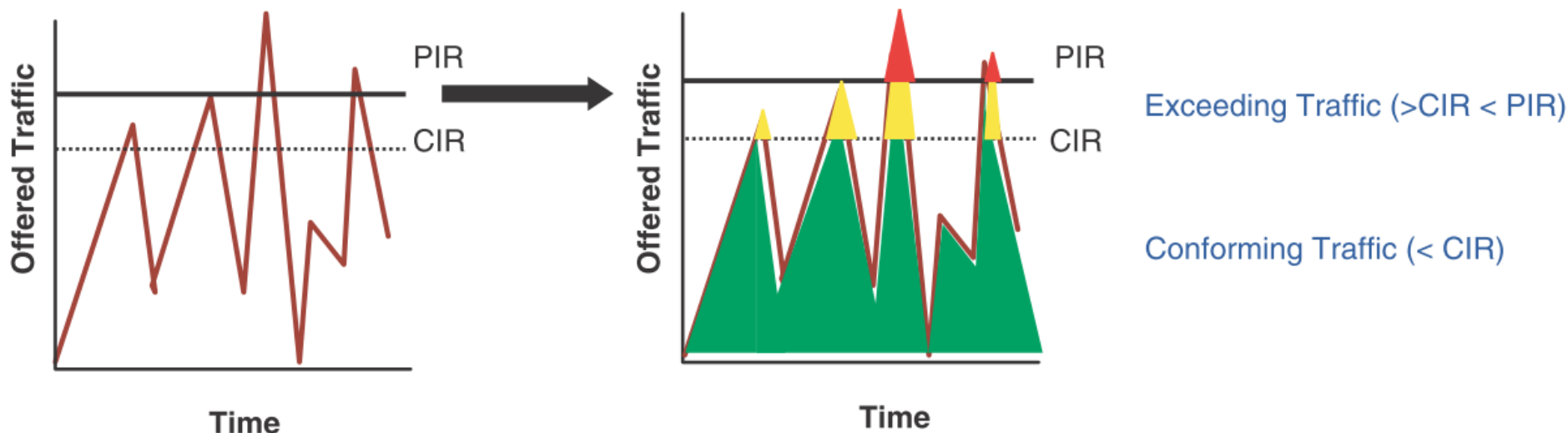
Figure 1-14 A Policer as a Marker



226610

# 流量监管—三色标记

Figure 1-15 A Dual-Rate Policer as a Three-Color Marker



A RFC 2698 “Two Rate Three Color Marker” can:

- mark *conforming* traffic to one value (such as AF31)
- remark *exceeding* traffic to another value (such as AF32)
- remark *violating* traffic to yet another value (such as AF33)

⑩ **PIR** (Peak Information Rate) – 最高信息率

⑩ **CIR** (Committed Information Rate) – 承诺信息率

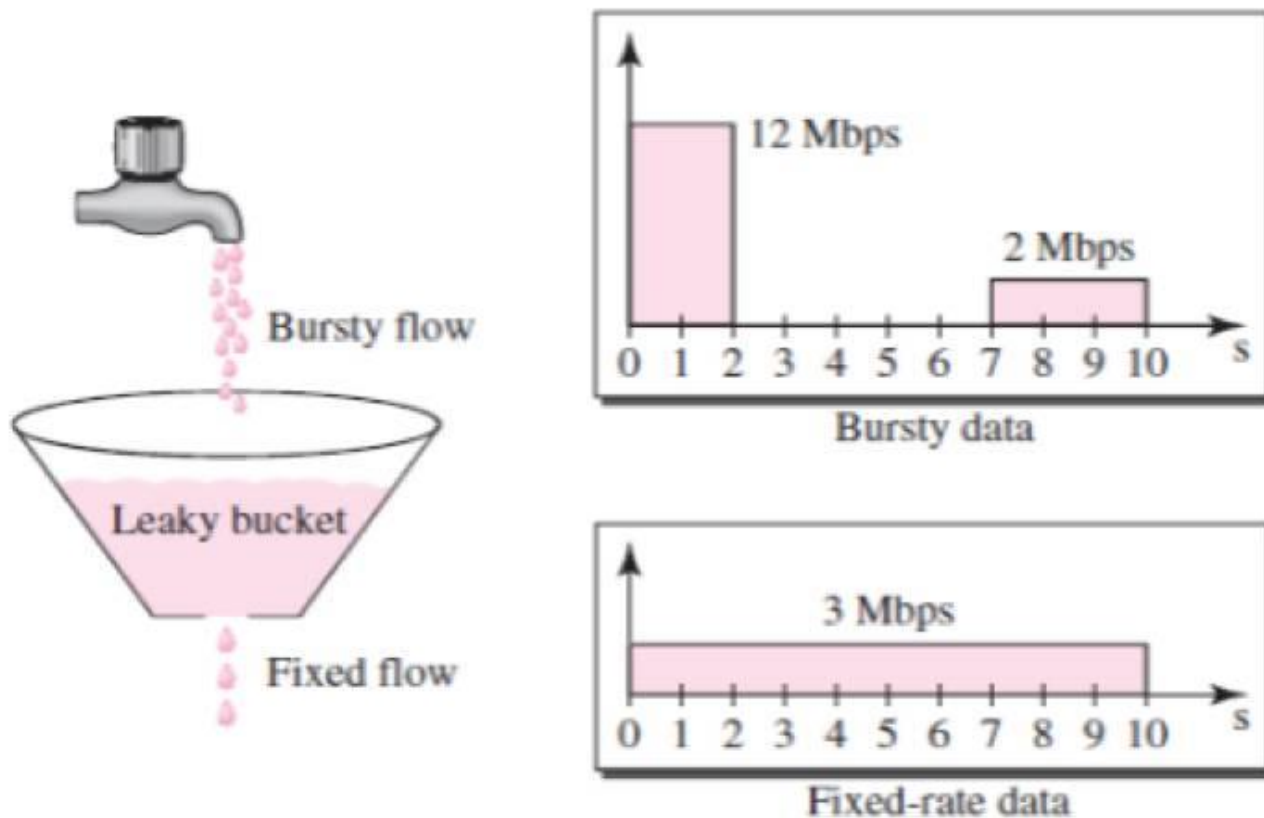
# 流量监管与整形



## 流量整形

# 流量整形

- **流量整形**是对输出报文的速率进行控制，使报文以均匀的速率发送出去。



# 流量监管和整形的相同点

- 流量监管和流量整形的相同点：
  - ◆ 作用都是监控网络流量。
  - ◆ 都是用令牌桶算法评估流量速率。
  - ◆ 主要都用于网络边缘。



# 流量监管和整形的差异点★

## 流量监管

丢弃超额流量或将超额流量重标记为低优先级量

不需要额外的内存资源，不会带来延迟和抖动

丢包可能引发重传

可以重标记流

## 流量整形

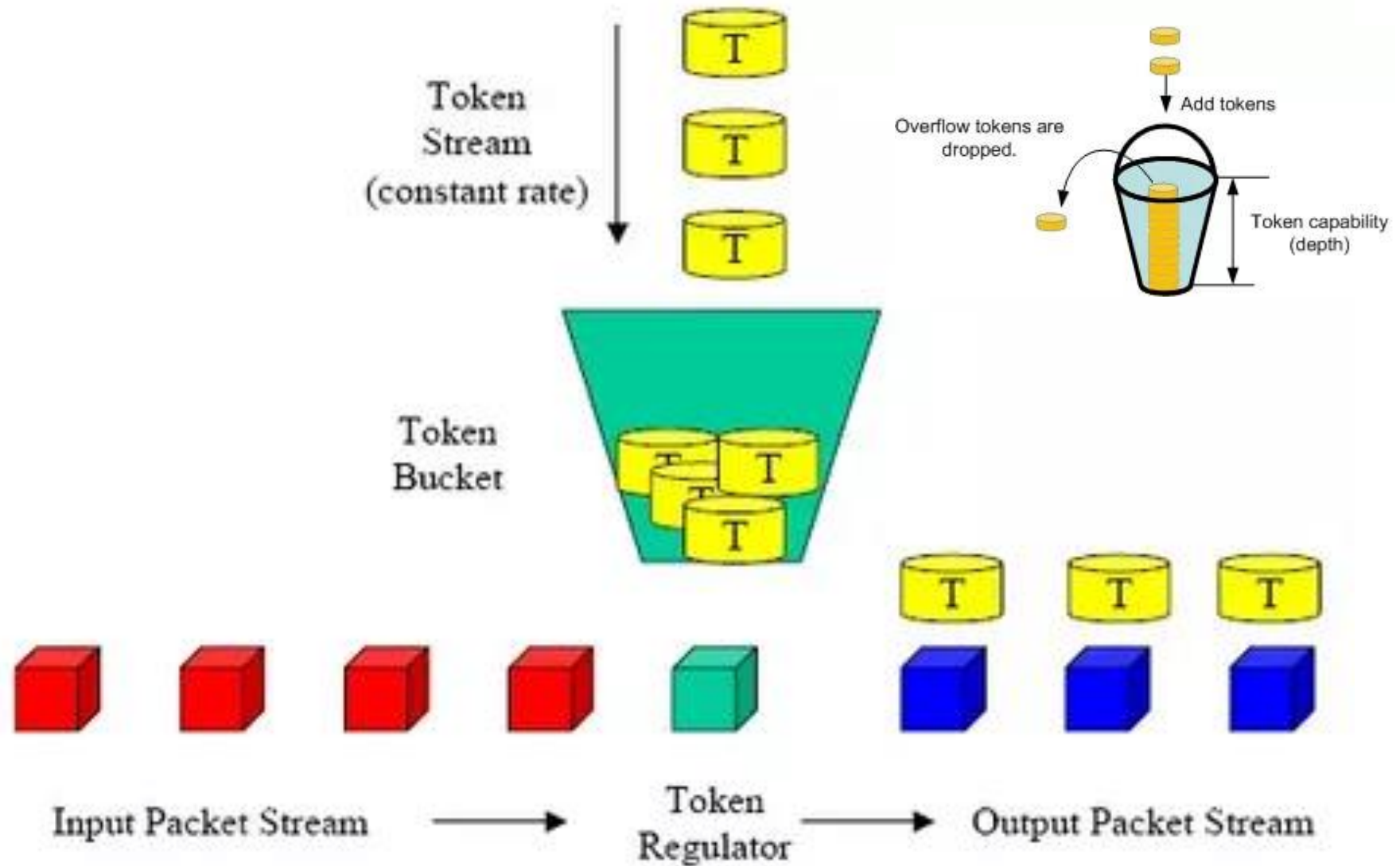
缓存超出策略/协定SLA规定的超额流量

需要内存去缓存超额流量，可能会带来延迟和抖动

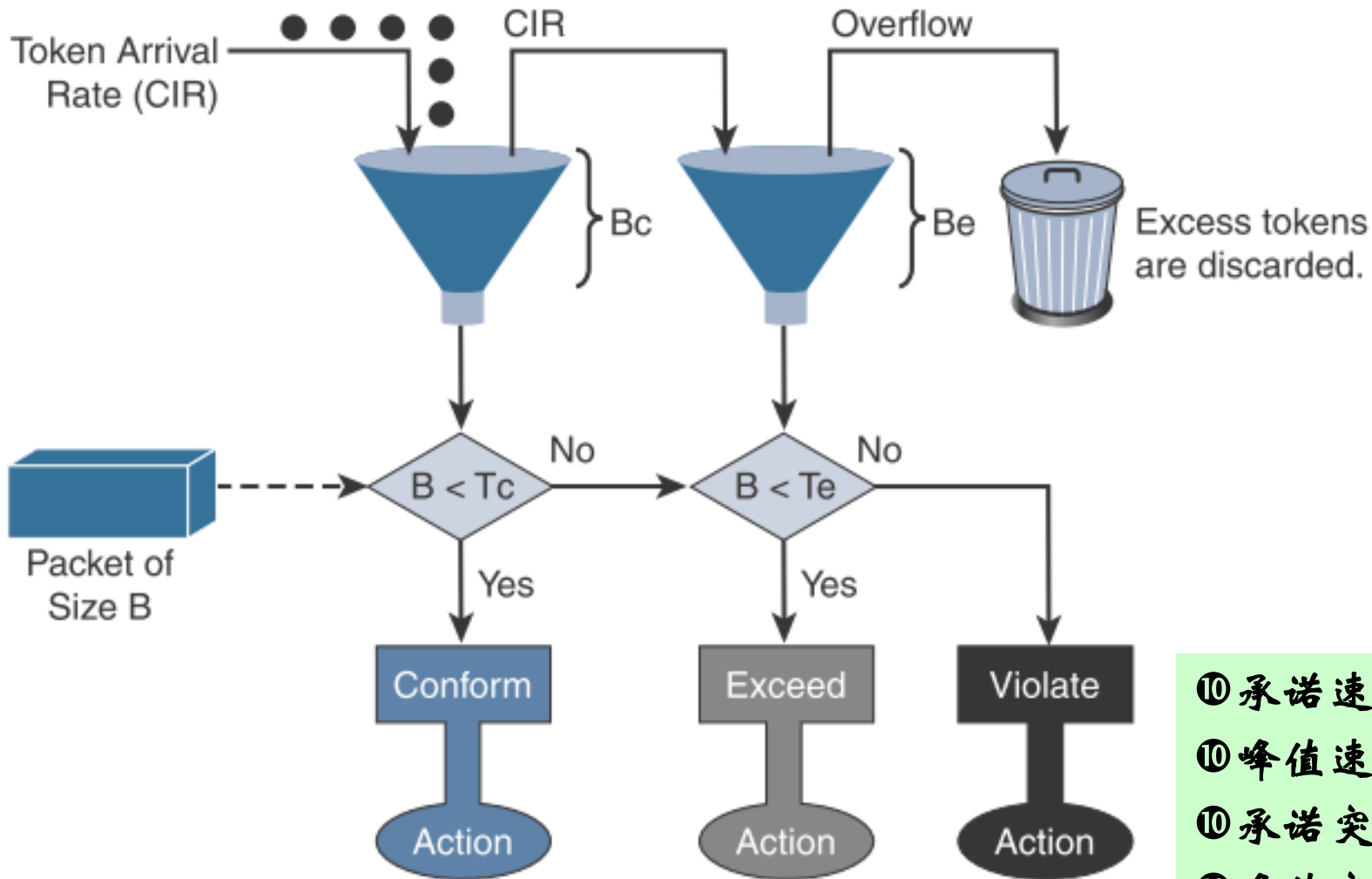
较少的丢包，因而较少导致重传

不能重标记

# 单桶单速★



# 双桶单速★



- ⑩ 承诺速率 CIR
- ⑩ 峰值速率 PIR
- ⑩ 承诺突发尺寸 CBS
- ⑩ 峰值突发尺寸 PBS

Figure 15-15 Dual-token Bucket

# 演示



CIR = 1Mbit/s

First packet



Packet(B) = 1500 bytes



Bucket C

CBS = 2000 bytes

Tc(0) = 2000 bytes



Bucket E

EBS = 2000 bytes

Te(0) = 2000 bytes

# 双桶双速★

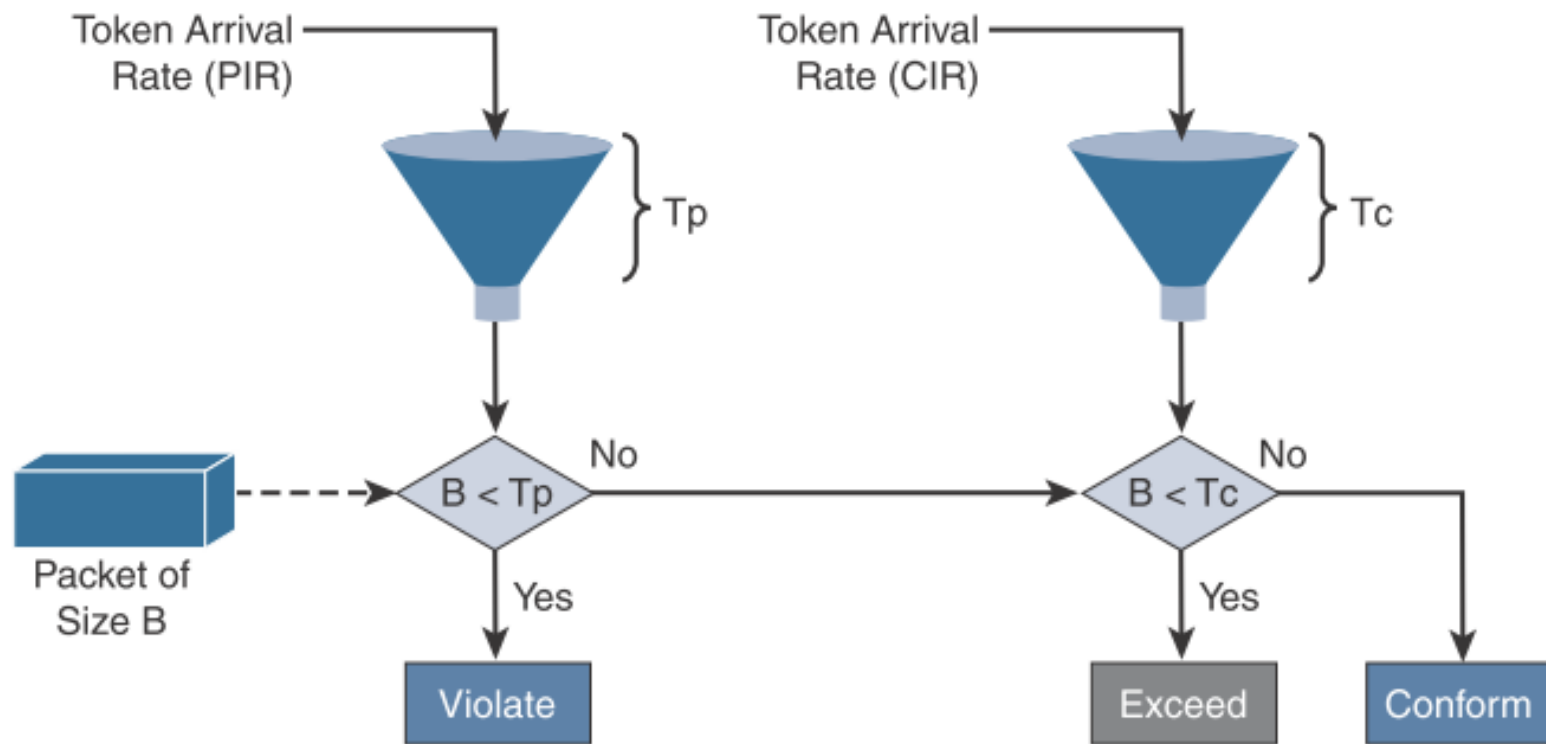


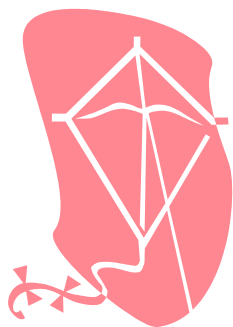
Figure 15-16 Two-rate Three-color Policer

- ⑩ 承诺速率 CIR
- ⑩ 峰值速率 PIR
- ⑩ 承诺突发尺寸 CBS
- ⑩ 峰值突发尺寸 PBS

# QoS服务

---

## 拥塞控制



# 拥塞的管理和对策★

## ■ 拥塞的管理和对策

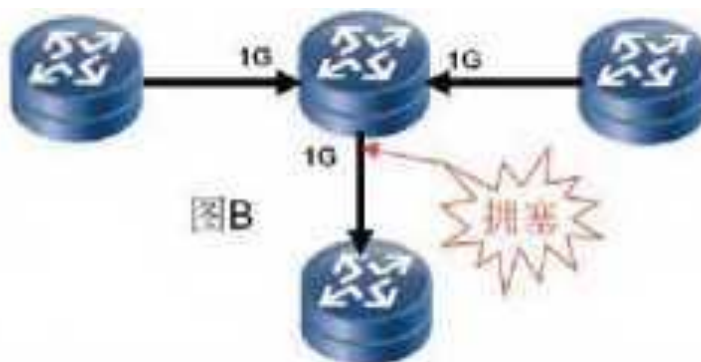
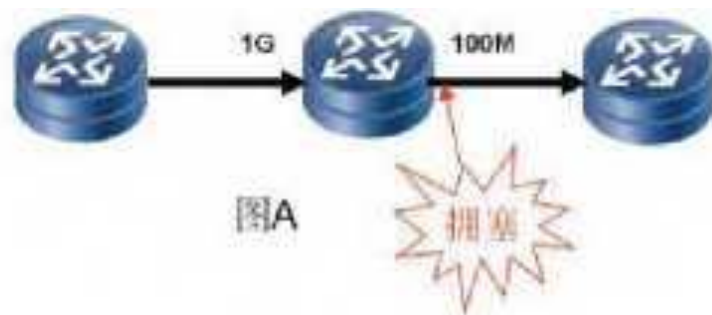
◆ 解决有限的网络资源与用户需求间的矛盾。

## ■ 通常用如下两种策略来缓解网络拥塞：

◆ **拥塞管理**：指网络在**发生**拥塞时，如何进行管理和控制。

◆ **拥塞避免**：在拥塞有加剧的**趋势**时，通过调整网络的流量来解除网络过载的流量控制机制。

# 拥塞的出现





# 拥塞控制

---



## 拥塞管理

# 拥塞管理

- **拥塞管理**：指网络在发生拥塞时，如何进行管理和控制。
  - ◆ 处理的方法是使用**队列**技术，将从一个接口发出的所有报文放入多个队列
  - ◆ 使用**队列调度算法**对各个队列进行处理。

# 队列调度算法★

## ■ 常用的队列调度机制如下：

- ◆ FIFO: First In First Out, 先进先出队列
- ◆ RR: Round Robin, 轮询队列
- ◆ WRR: Weight Round Robin, 按权重轮询队列
- ◆ PQ: Priority Queuing, 优先级队列
- ◆ CQ: Custom Queuing, 自定义队列
- ◆ WFQ: Weighted Fair Queuing, 加权公平队列

# FIFO

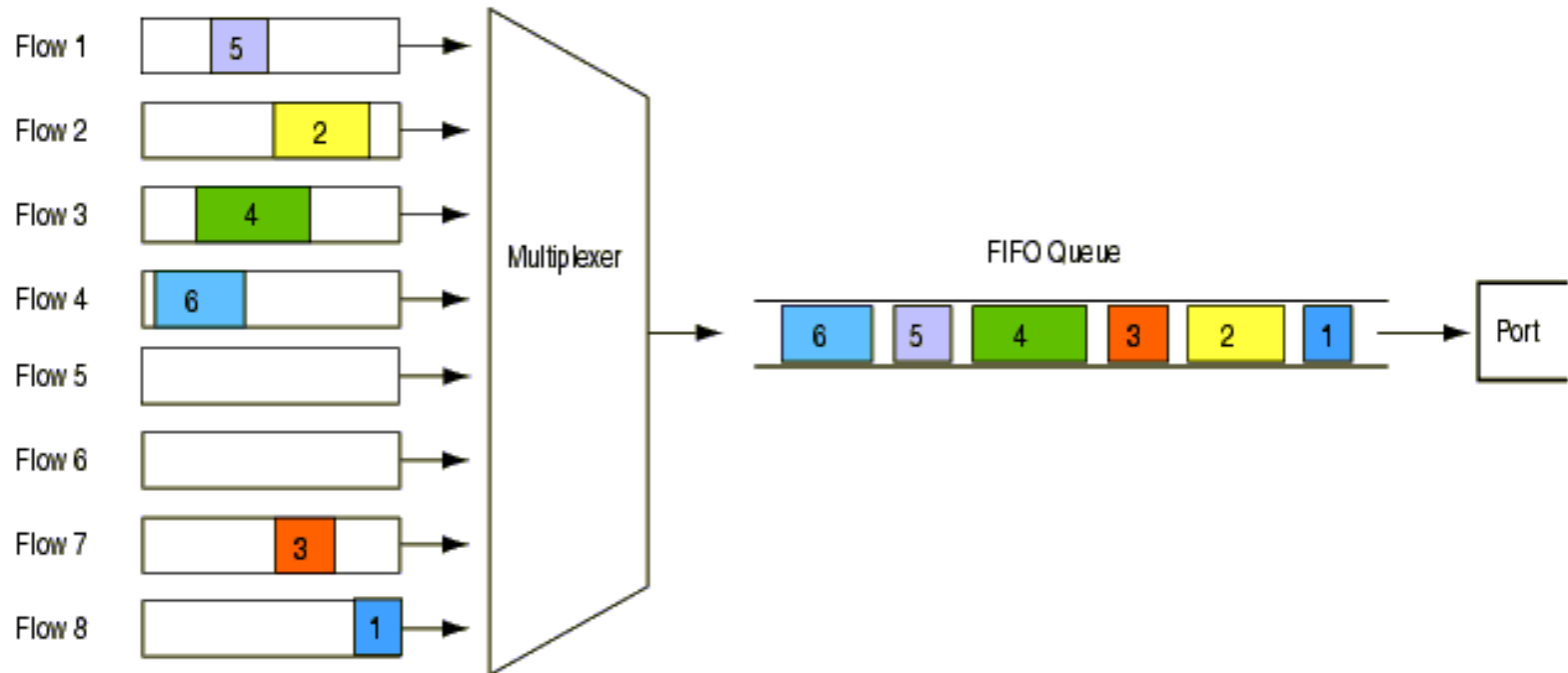
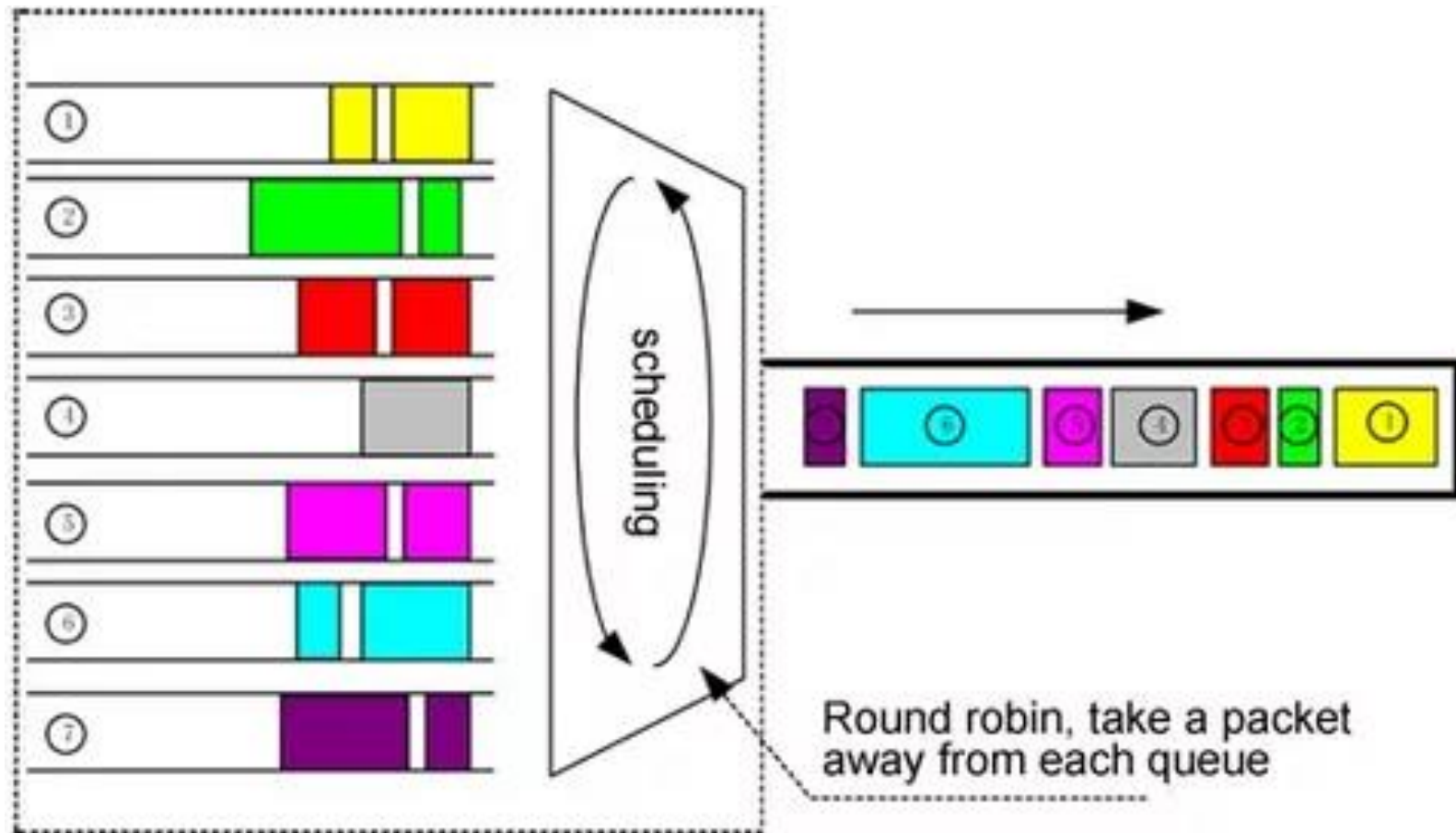


Figure 2.2.1

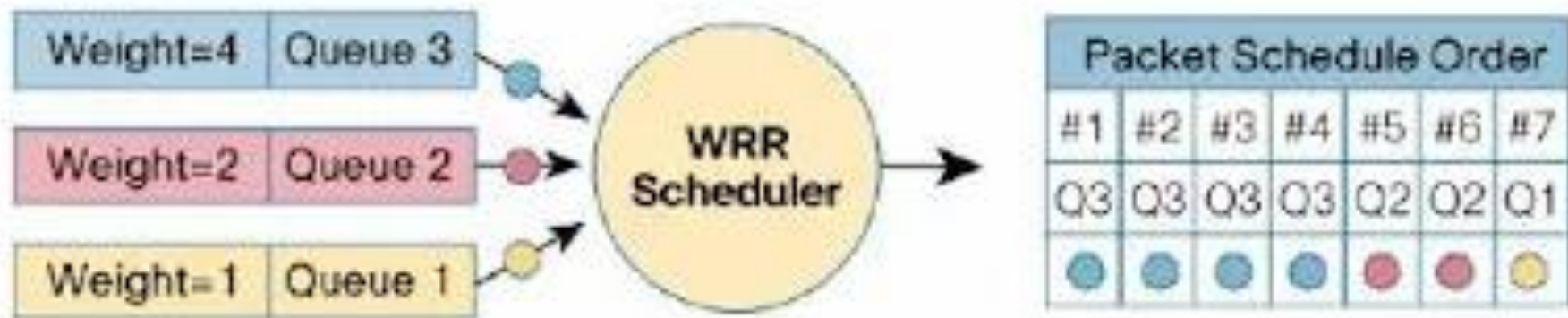
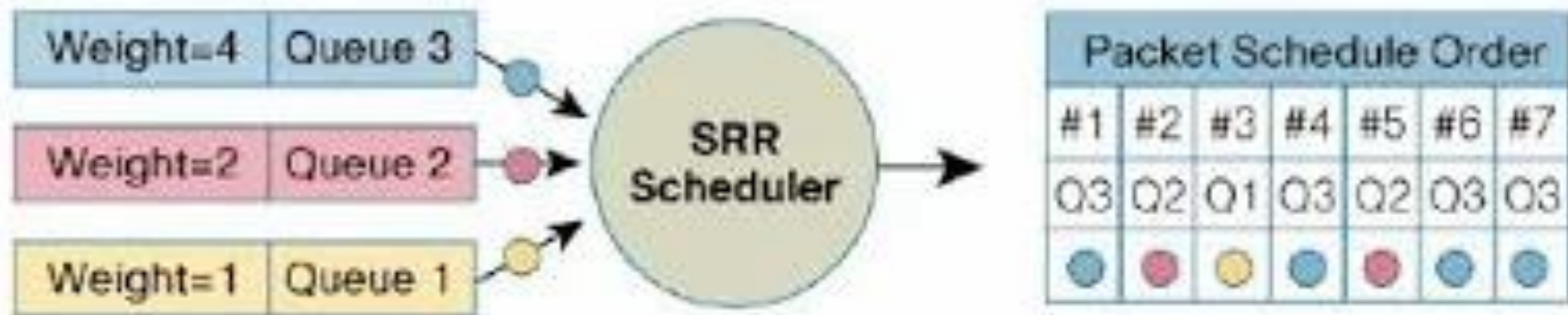
简单，不公平

# Round Robin



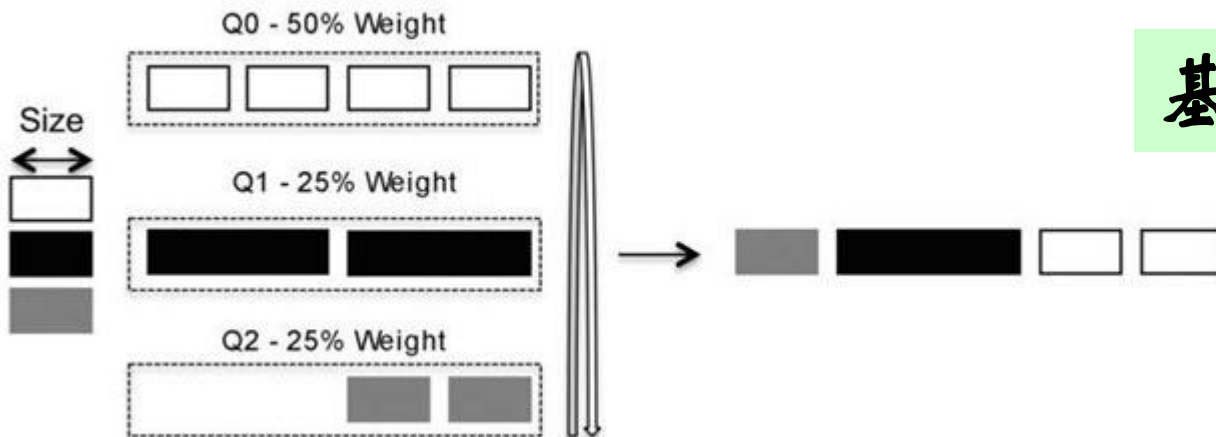
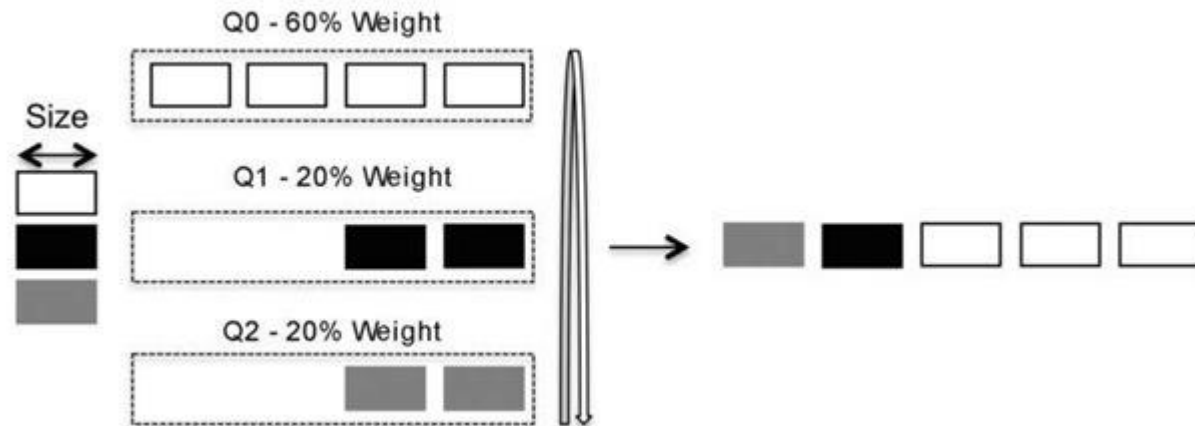
无权重，受包长影响，时延抖动较大

# Weight Round Robin



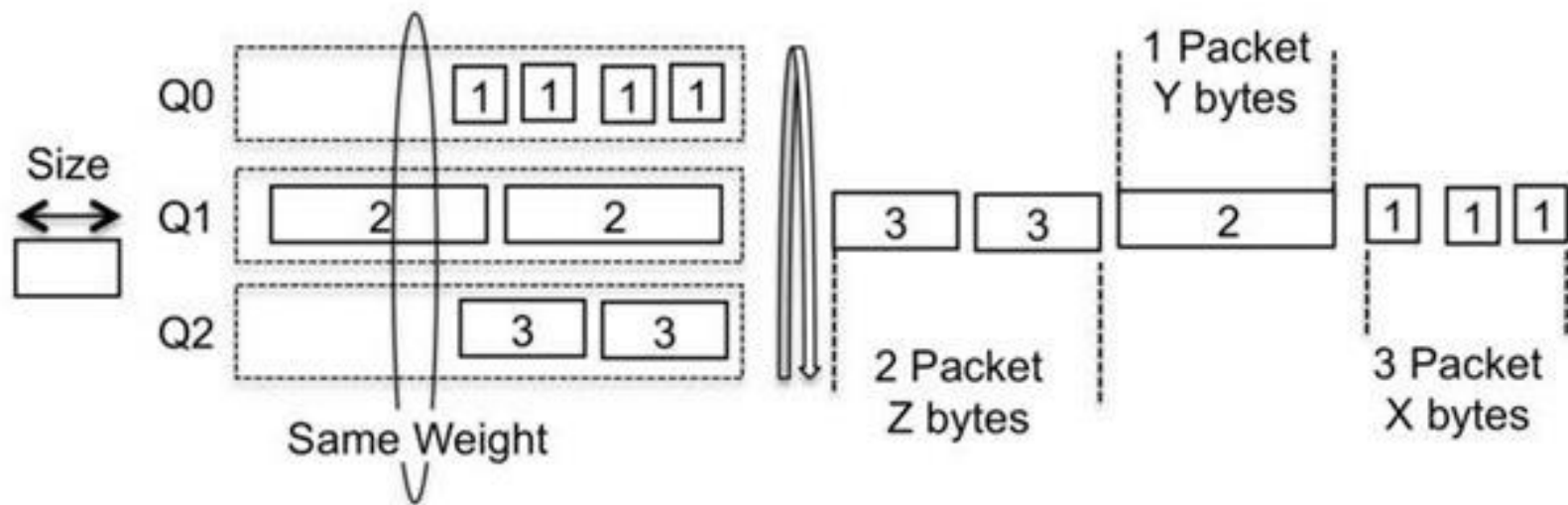
受包长影响，时延抖动较大

# WRR的不足



基于包

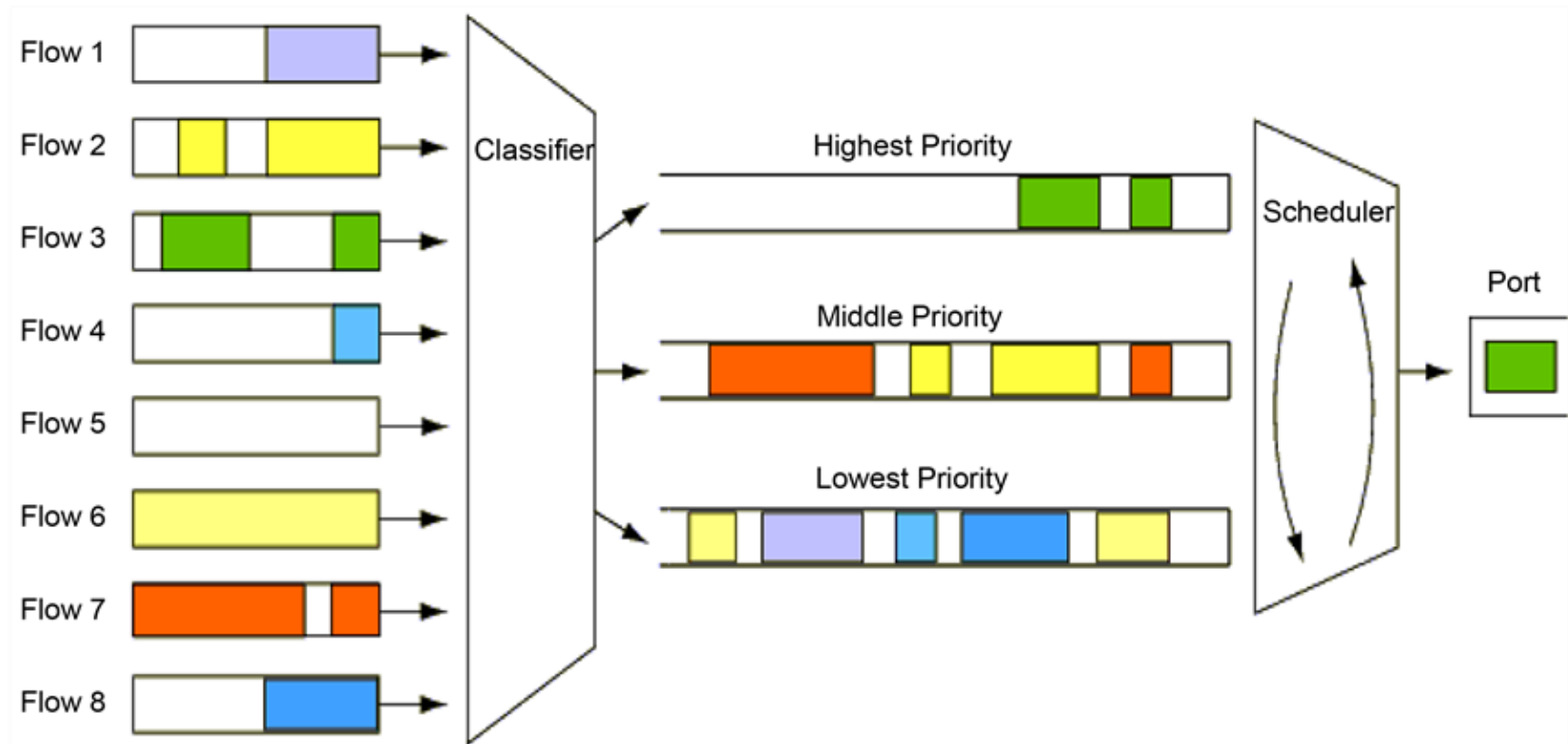
# WFQ



基于字节

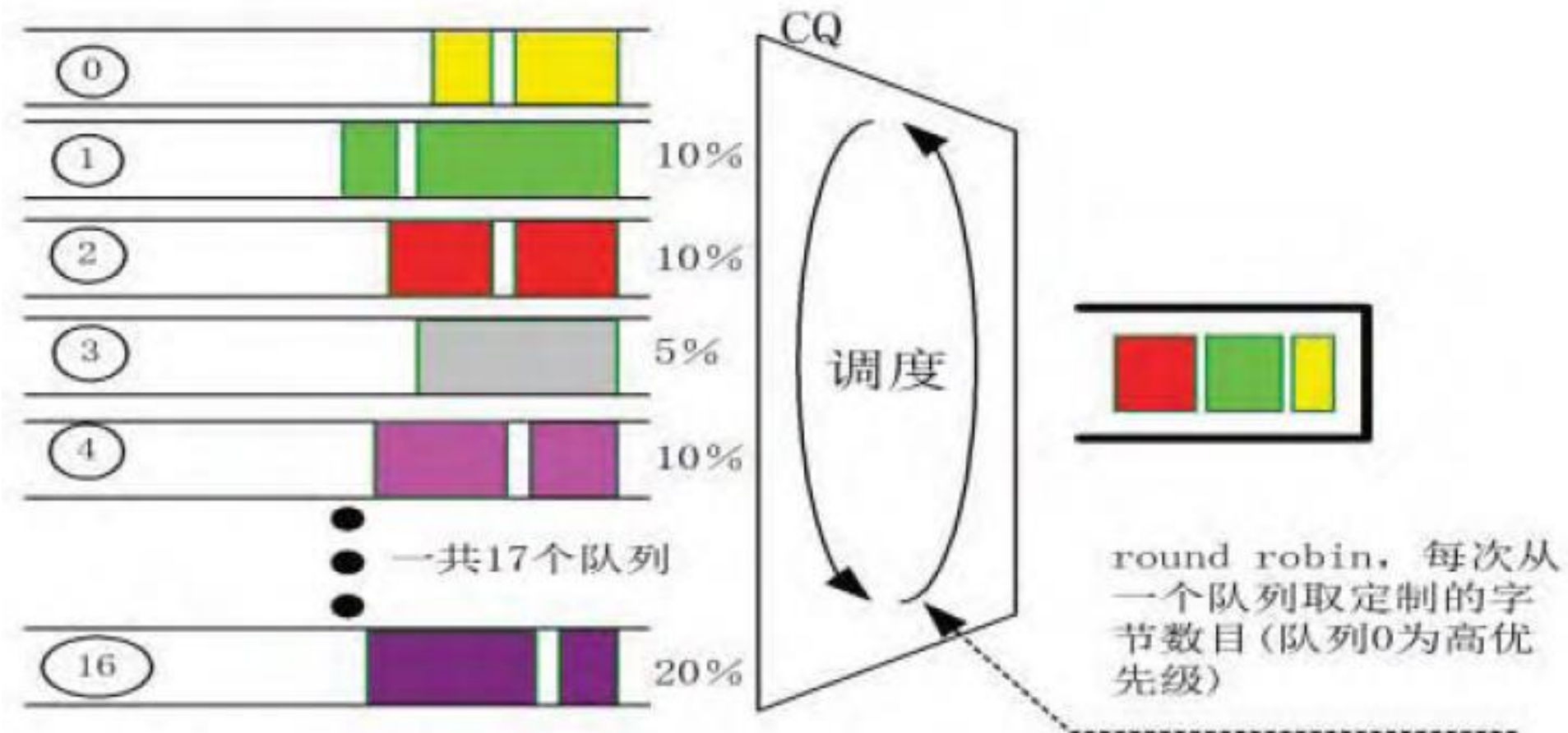


# PQ



低优先级流量可能无法获得服务，长时延

# CQ



# 对比

队列技术	调度的时延/抖动（在速率低的时候明显，速度绝对高的时候可忽略）	公平性
FIFO	差	无
RR	差	依赖包长
WRR	差	依赖包长
PQ	高优先级队列的时延控制非常好	无
CQ	配置字节数小的时候，带宽分配不准确，当差配置字节数大的时候，时延抖动比较大	一般
WI-Q	时延控制较好，抖动小	好

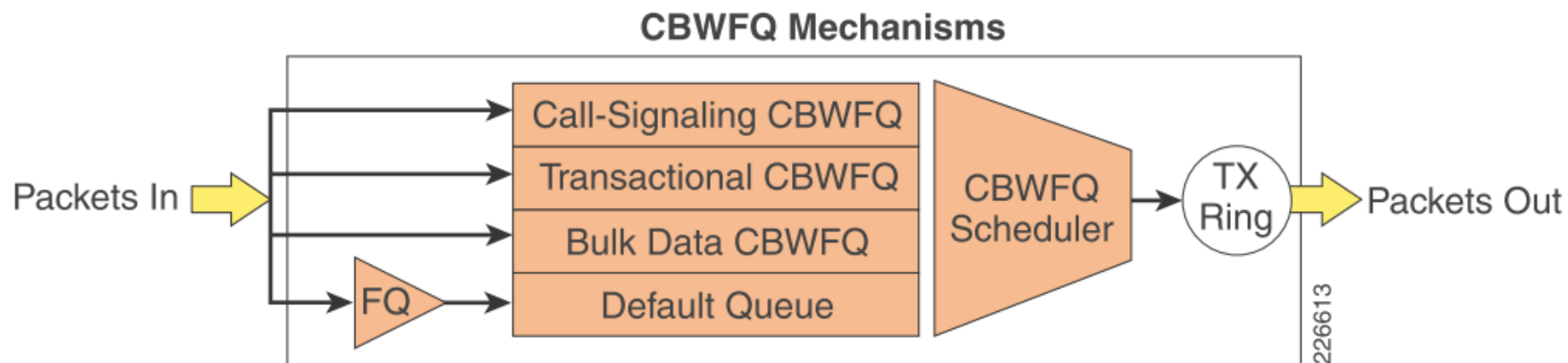
# 拥塞控制



## 队列构成实例

# CBWFQ

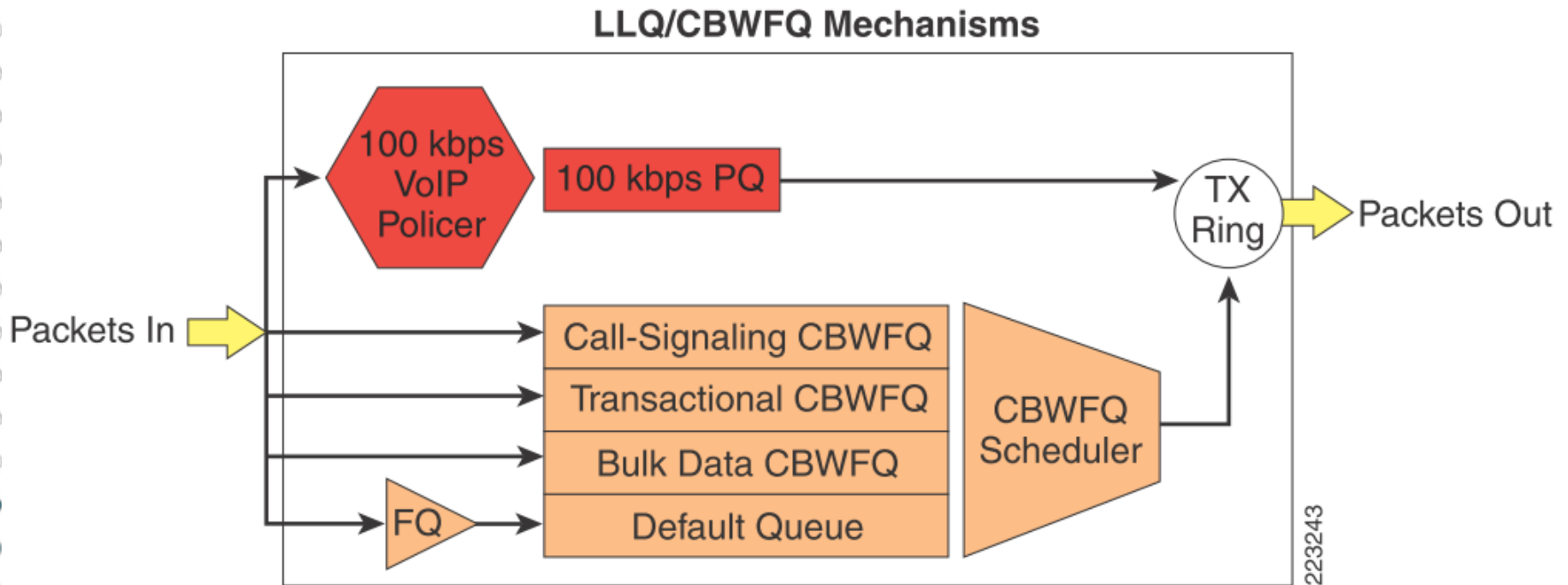
**Figure 1-17 CBWFQ Operation**



基于类的加权公平队列(class-based weighted fair queuing)

# LLQ/CBWFQ

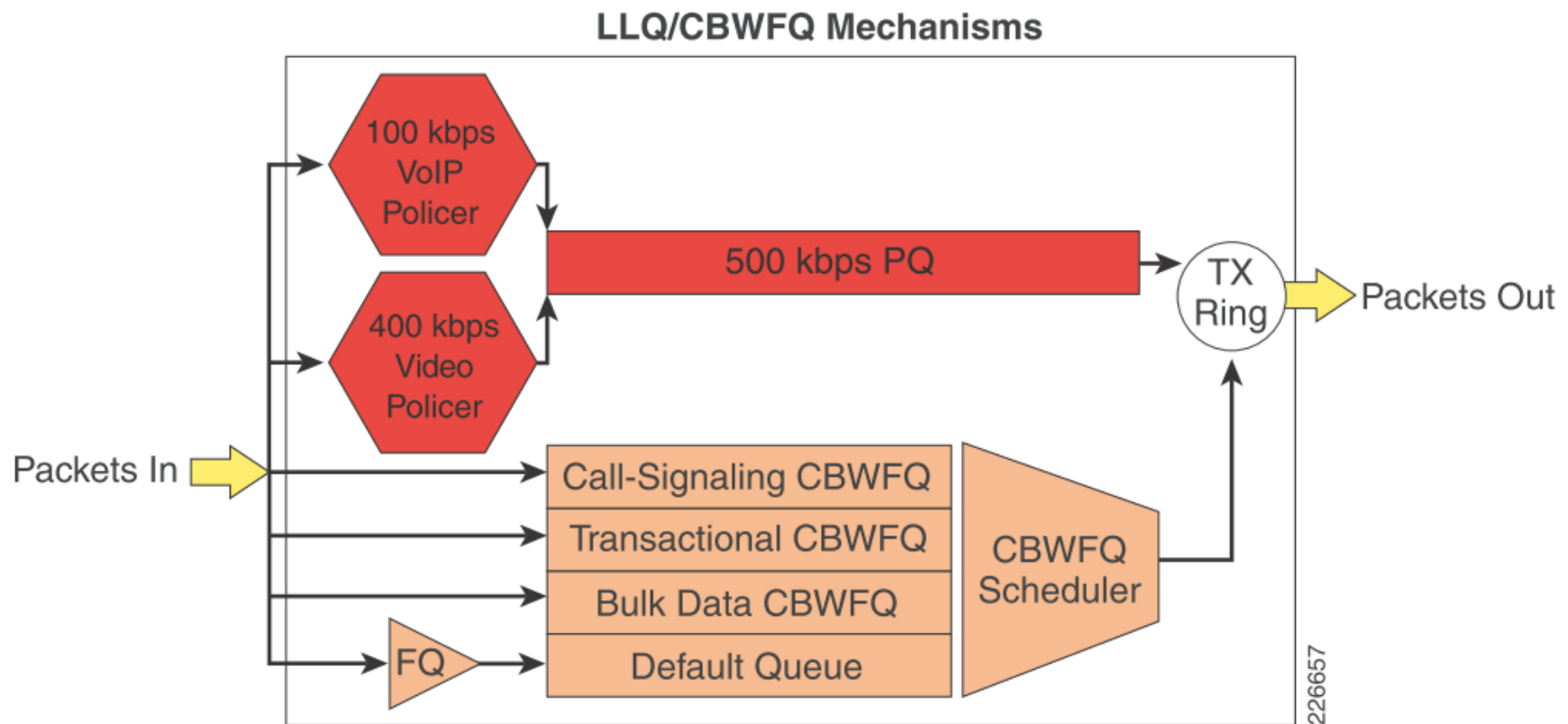
Figure 1-18 LLQ/CBWFQ Operation



⑩LLQ (Low Latency Queueing, 低延迟队列)

# Dual-LLQ/CBWFQ

**Figure 1-19** *Dual-LLQ LLQ/CBWFQ Operation*

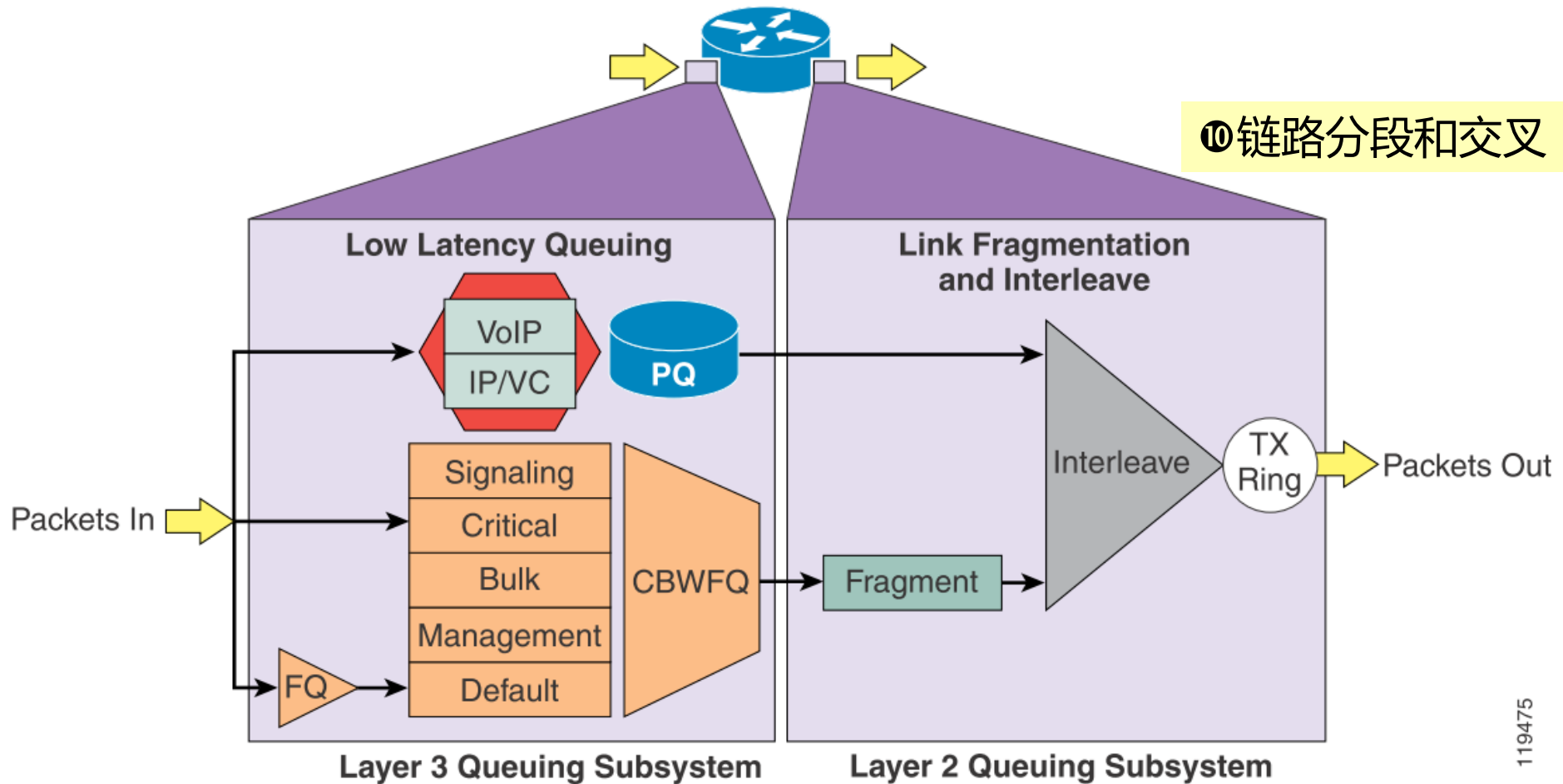


# Link Fragmentation and Interleave

Figure 1-3

LLQ/CBWFQ Operation

⑩链路分段和交叉



119475



# 拥塞控制

---



## 拥塞避免

# 拥塞避免简介

- 拥塞避免（Congestion Avoidance），在拥塞产生或有加剧的趋势时**主动丢弃报文**，通过调整网络的流量来解除网络过载。
- 设备在丢弃报文时，需要与**源端的流量控制**动作（比如TCP流量控制）相配合，调整网络的流量到一个合理的负载状态。

# 尾部丢弃

■ **尾部丢弃 (Tail-Drop)** 的方法。当队列的长度达到最大值后，所有新到来的报文都将被丢弃。

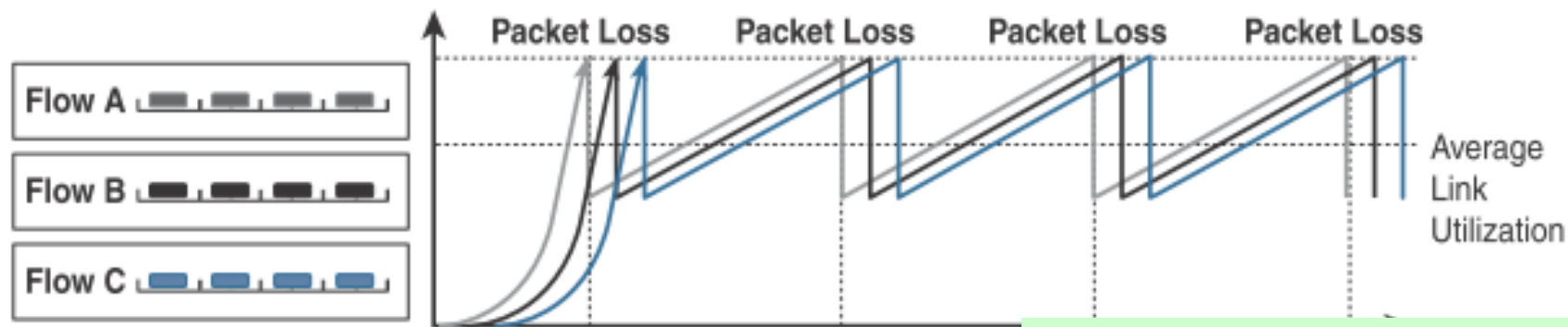


Figure 15-22 Suboptimal Bandwidth Utilization and Tail Drop

⑩ TCP全局同步

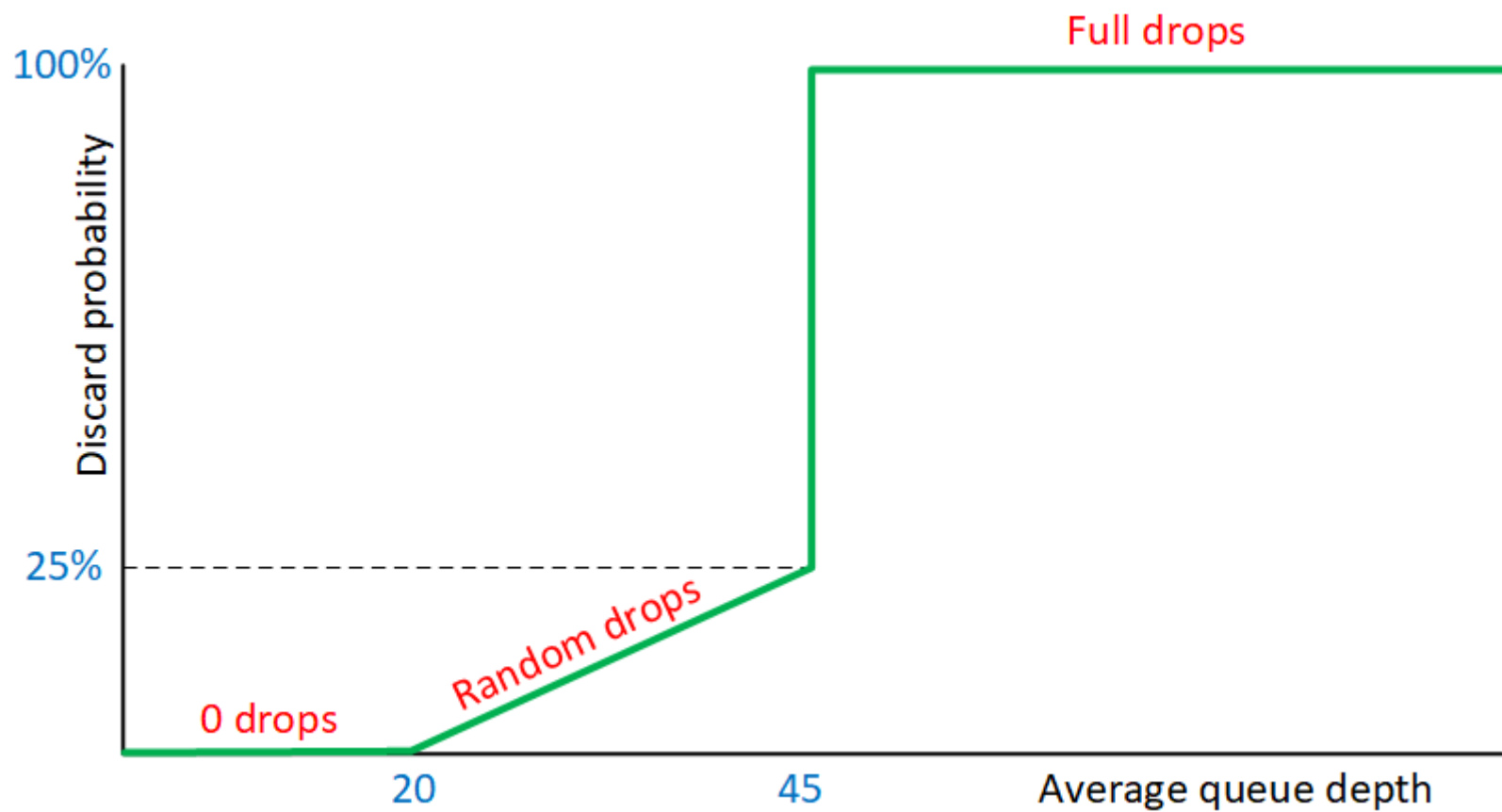
⑩ TCP饥饿

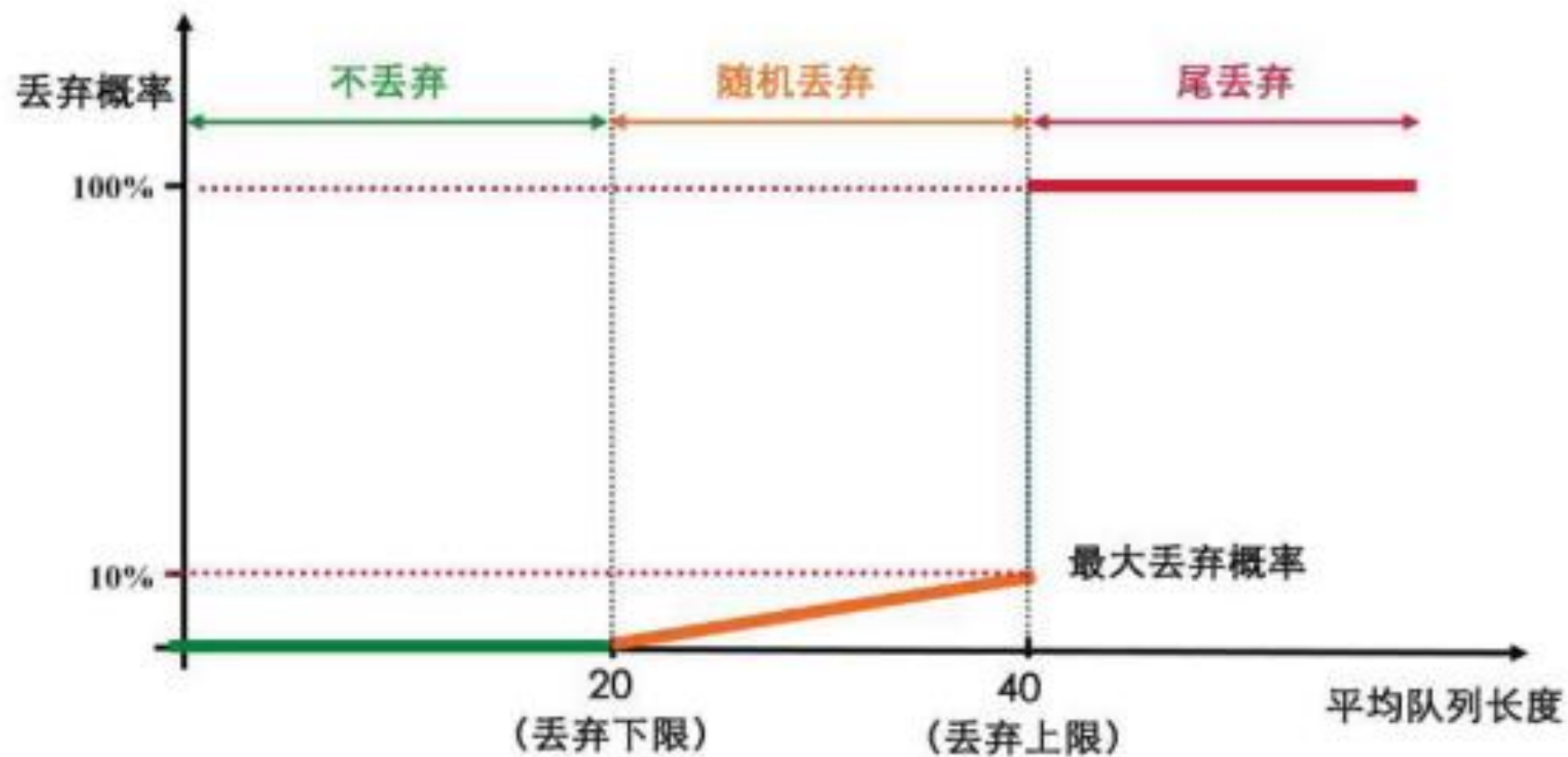
⑩ 高延时和高抖动

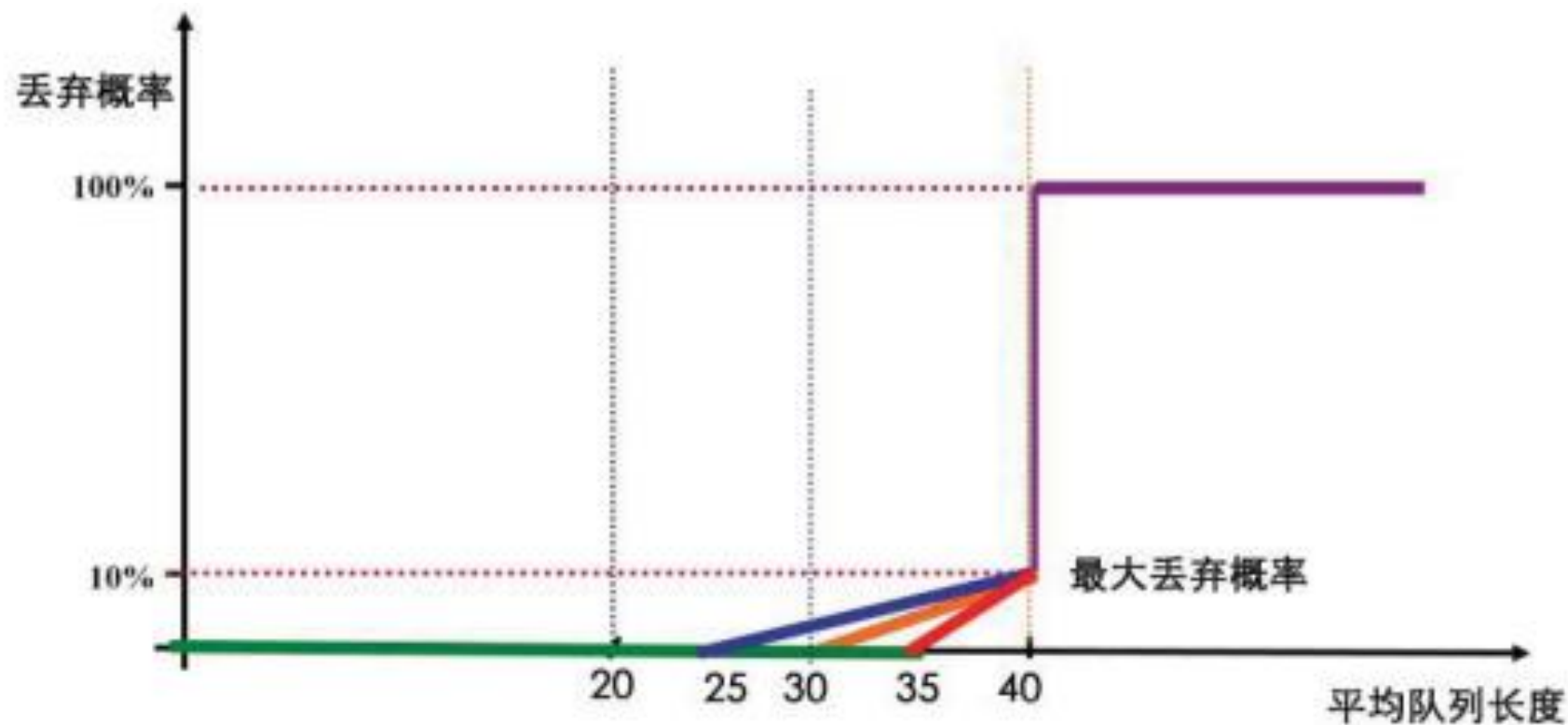
⑩ 无差别的丢弃

# RED与WRED

- 为避免TCP全局同步现象，可使用
  - ◆ RED (Random Early Detection) , 随机早期检测
  - ◆ WRED (Weighted Random Early Detection) , 加权随机早期检测。

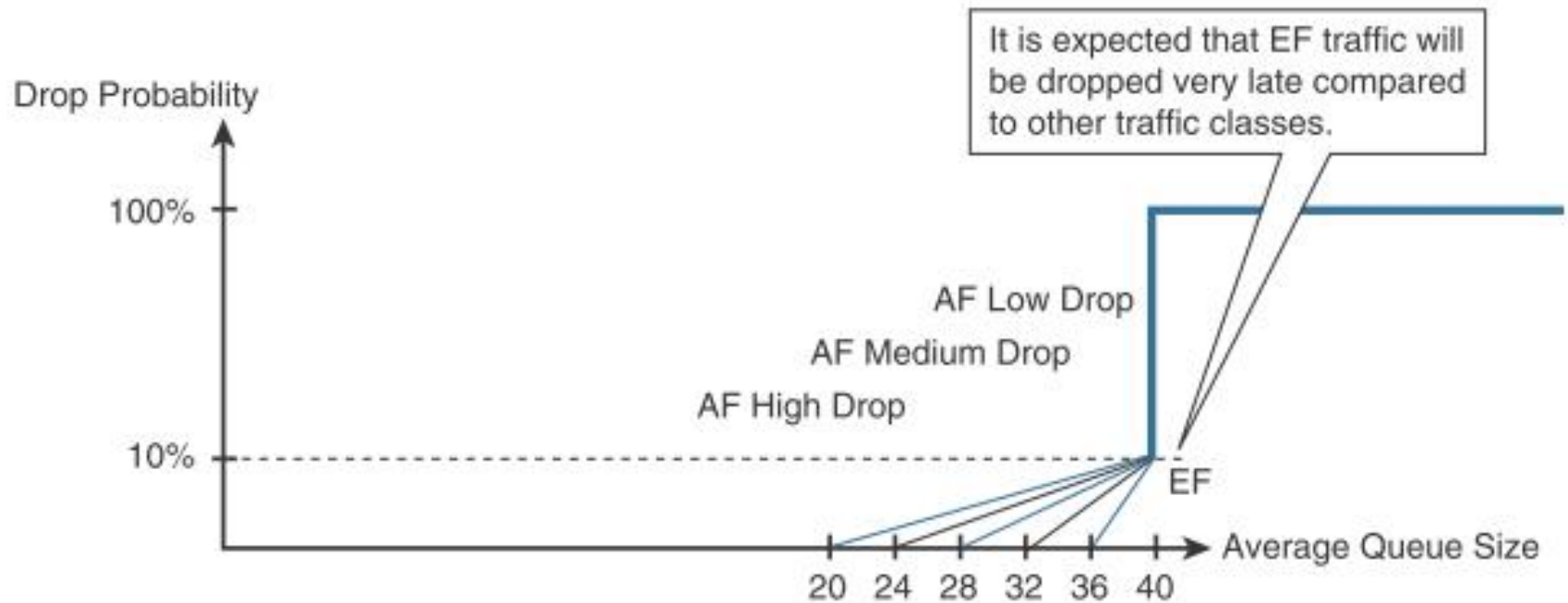






- AF21
- AF22
- AF23

# WRED

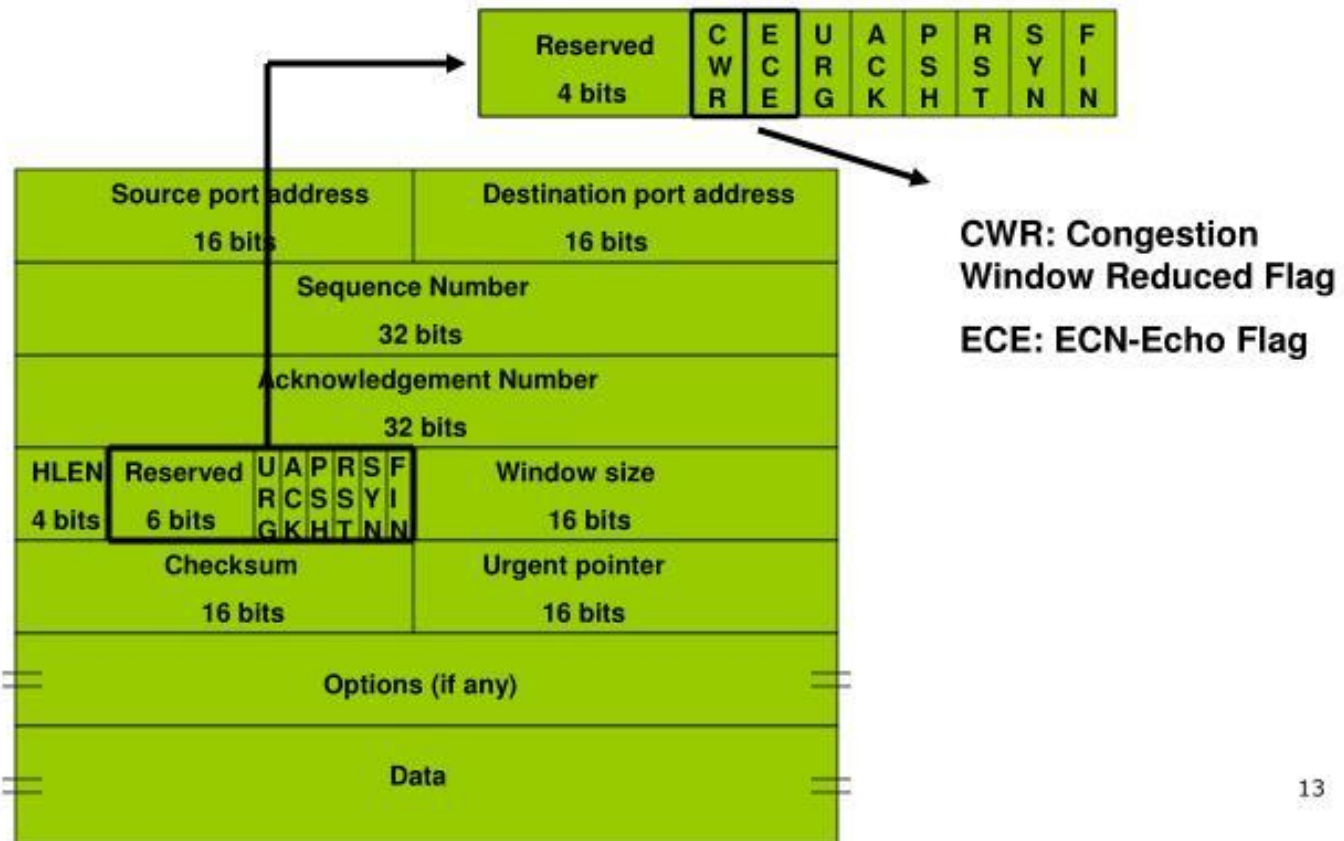


**Figure 15-26** Cisco IOS DSCP-based WRED Dropping Probability

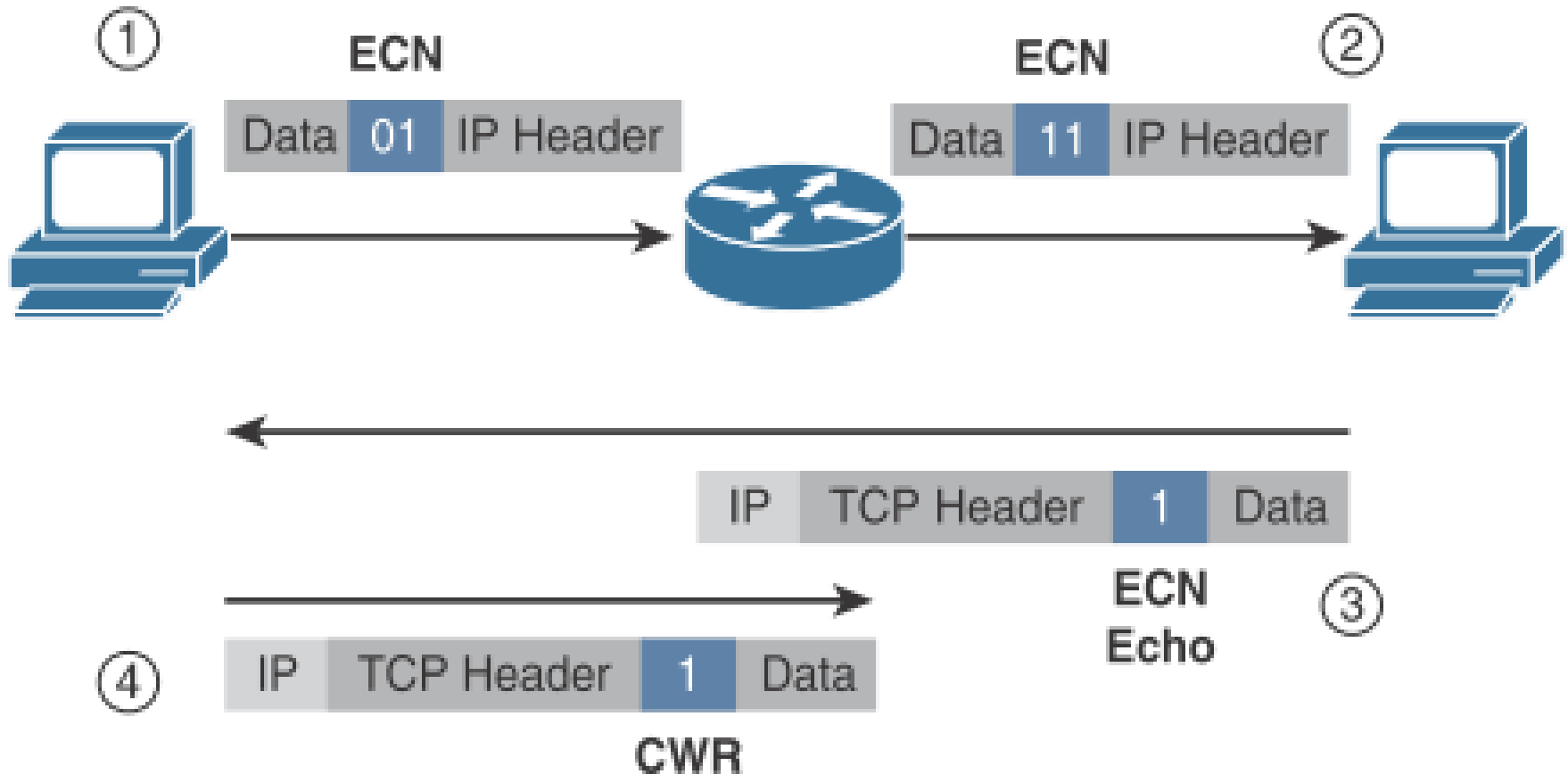


# TCP包头

## ECN bits in TCP header



# ECN 机制

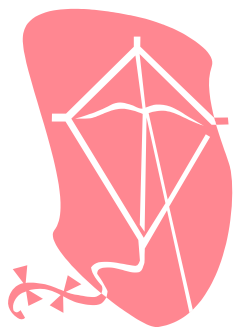


**Figure 15-28** *ECN Operation*

# QoS服务

---

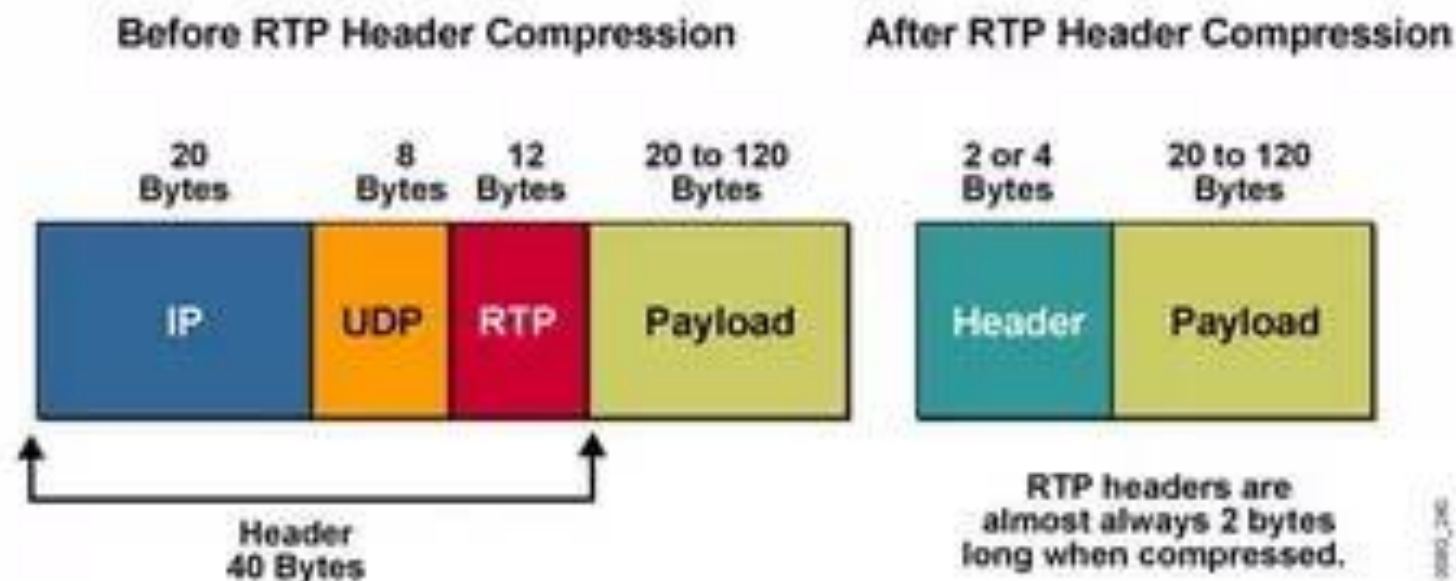
## 其它技术



# 链路专用管理

- 在为语音流量提供QoS时，广域网链路可以使用一下技术：
  - ◆ Compressed Real-time Protocol (CRTP)
  - ◆ Link Fragmentation and Interleaving (LFI)
  - ◆ Frame Relay Fragmentation (FRF)

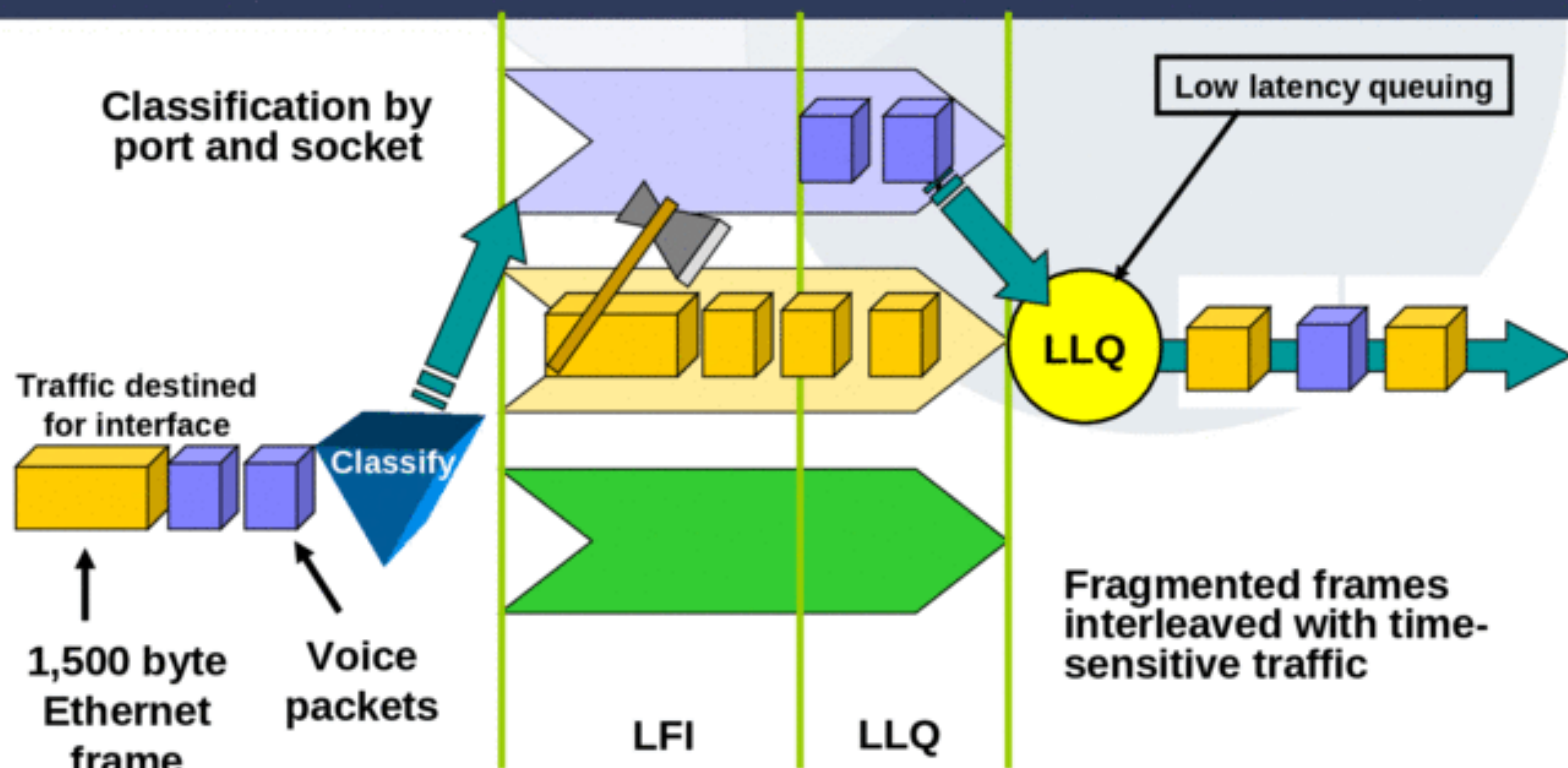
# RTP Header Compression



- RTP header compression saves bandwidth by compressing packet headers across WAN links

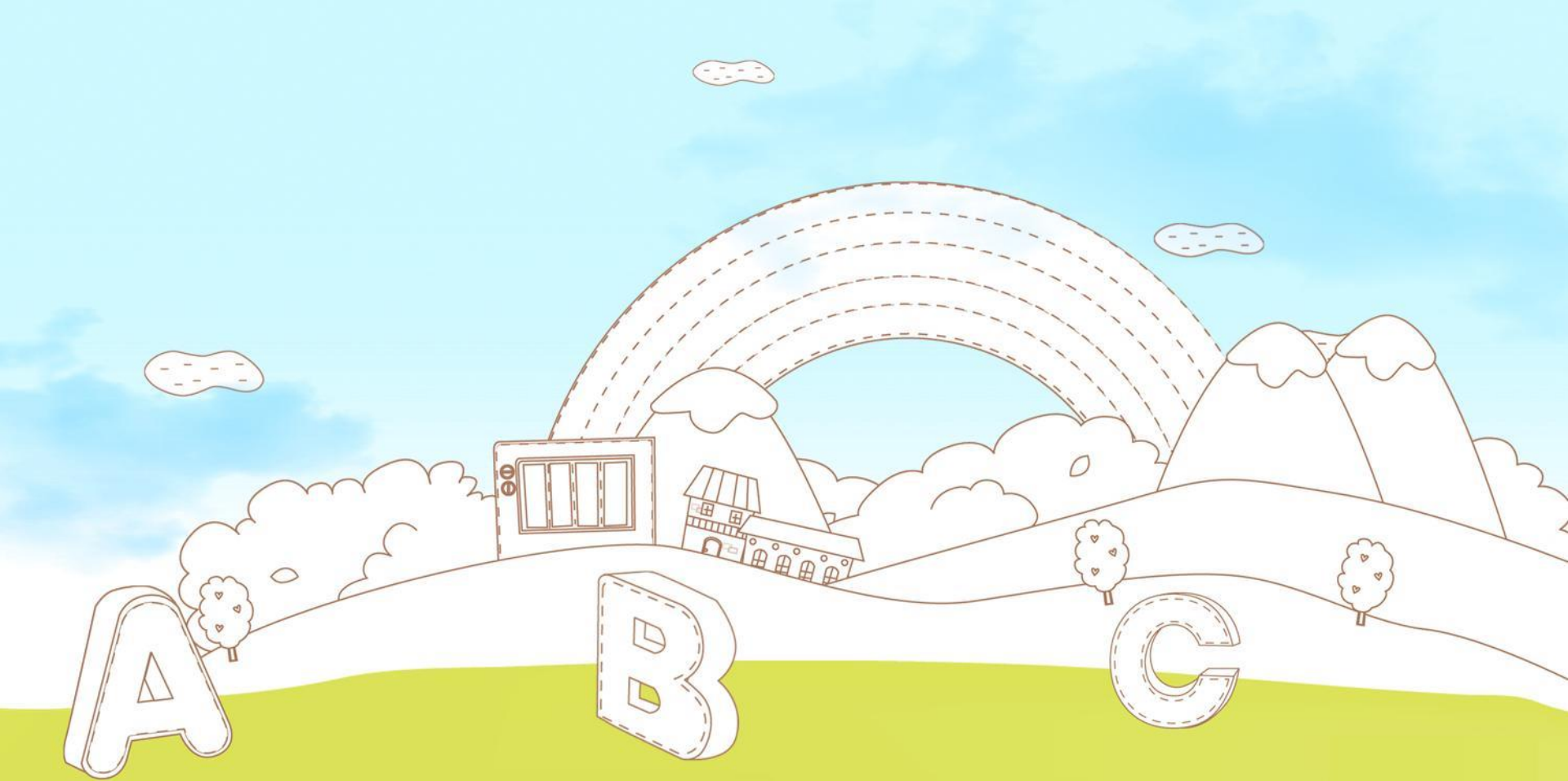
# LFI

## Fonctionnement du LFI

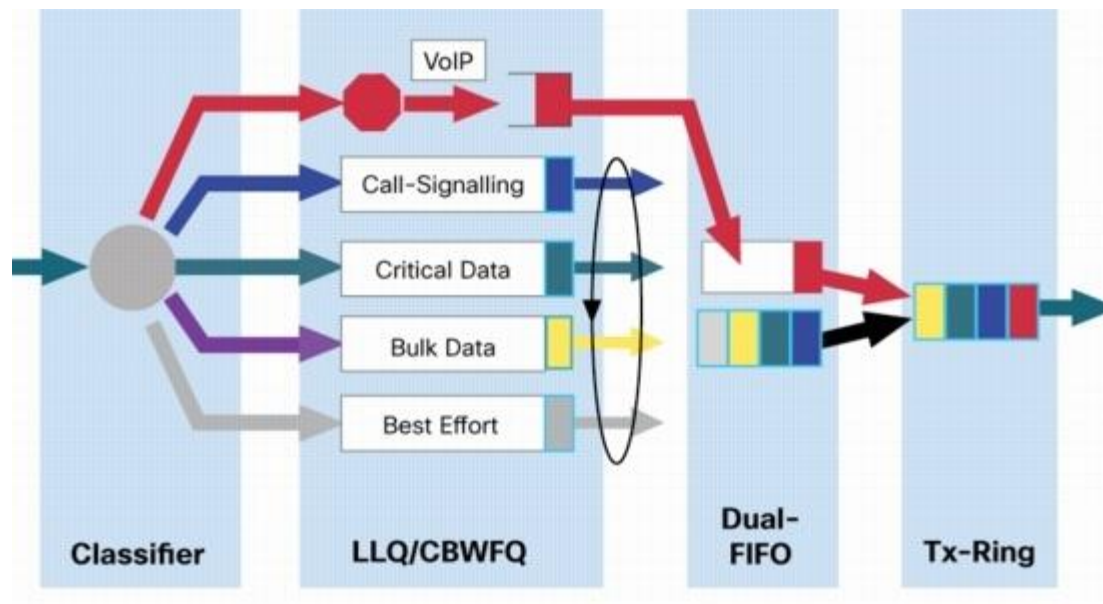


Fragment size based  
on serialization delay





**Thank You !**



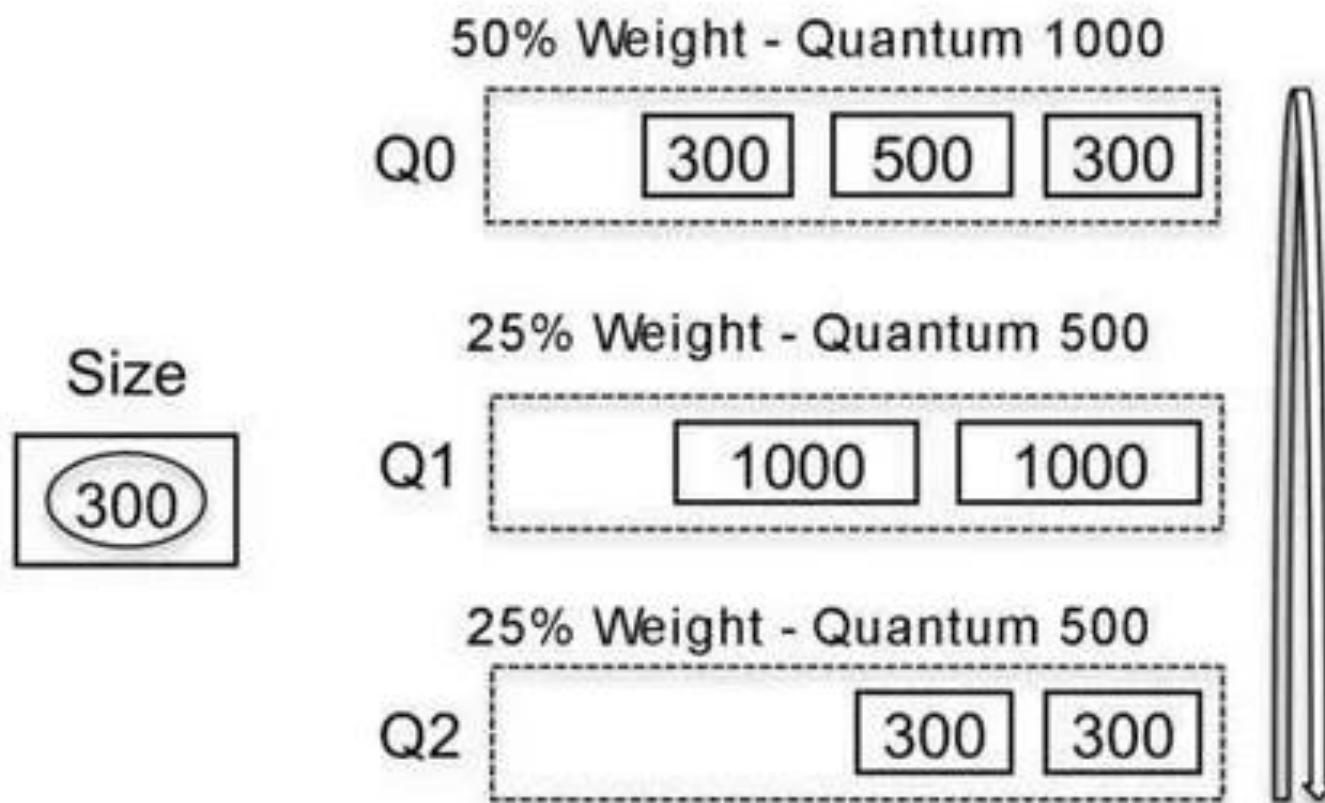
A

B

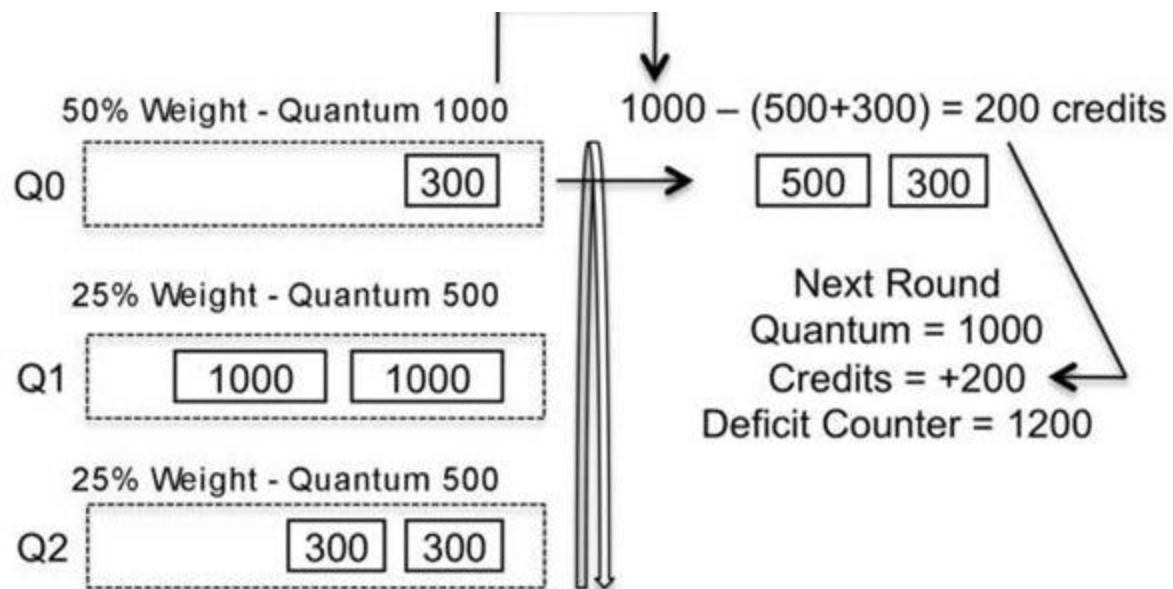
C

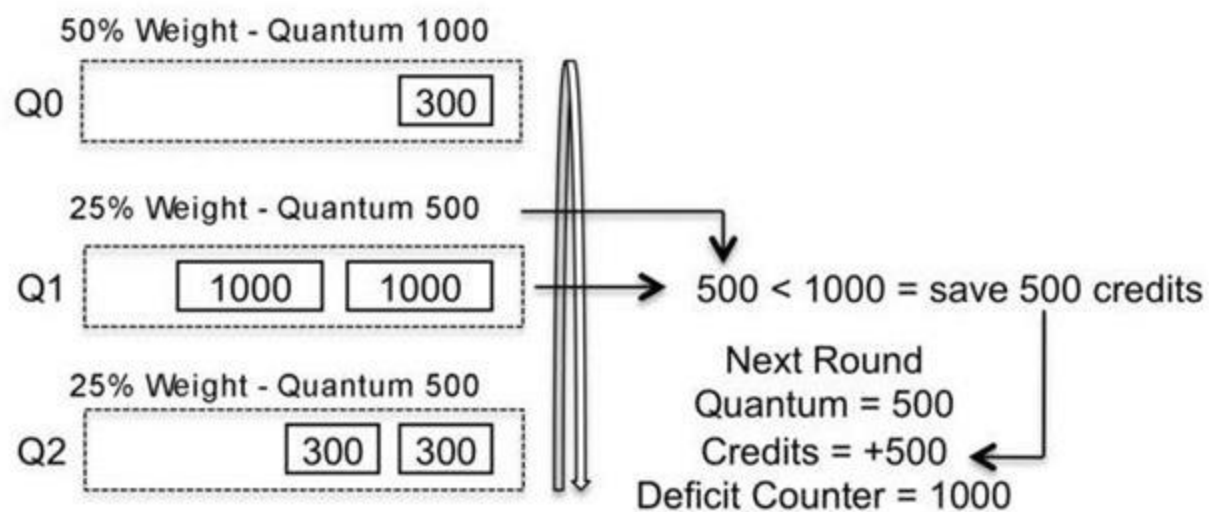


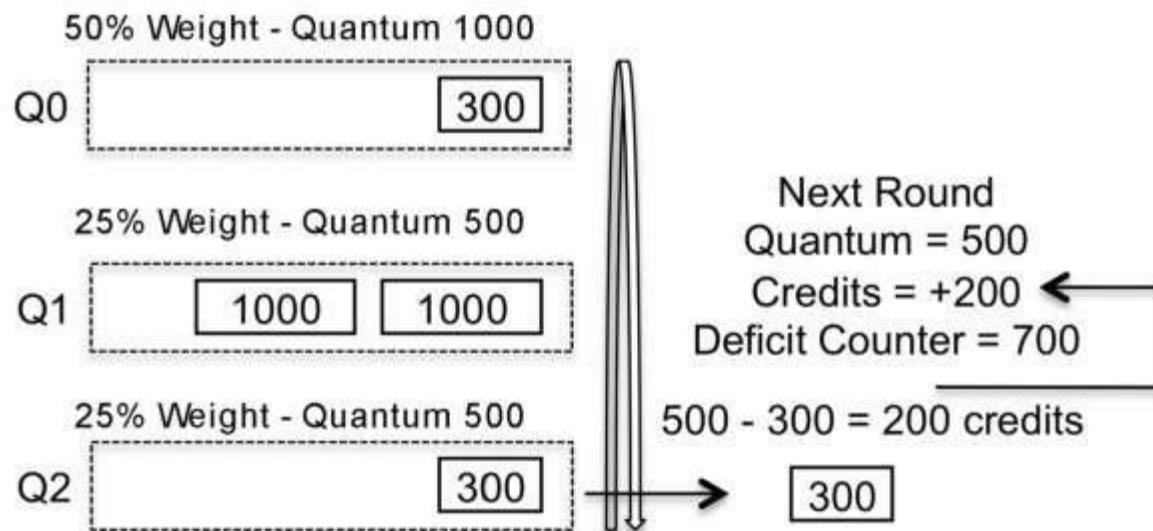
# Deficit Weighted Round Robin

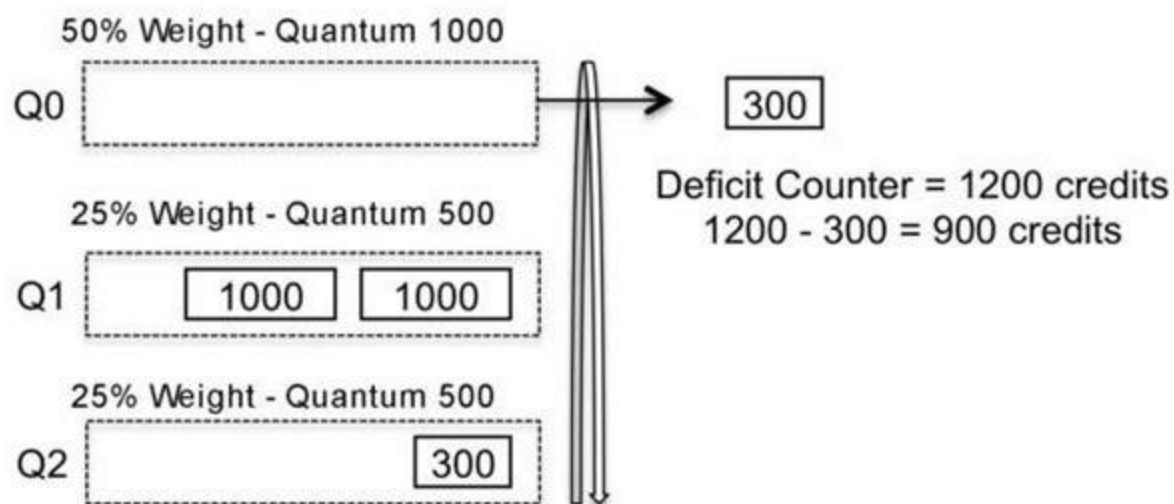


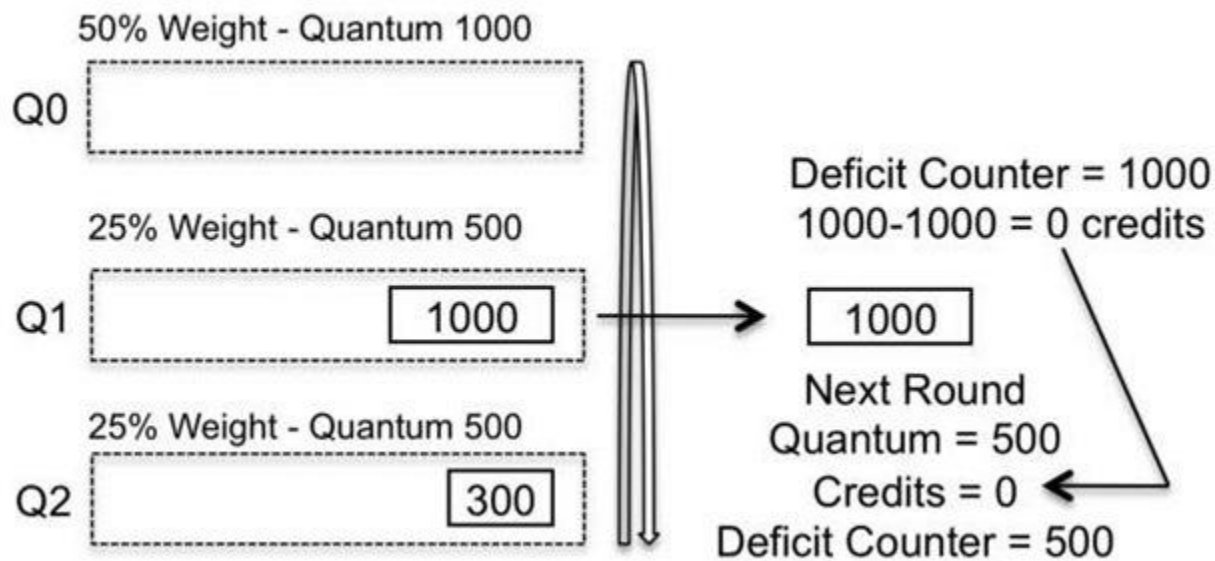
⑩加权差额循环调度

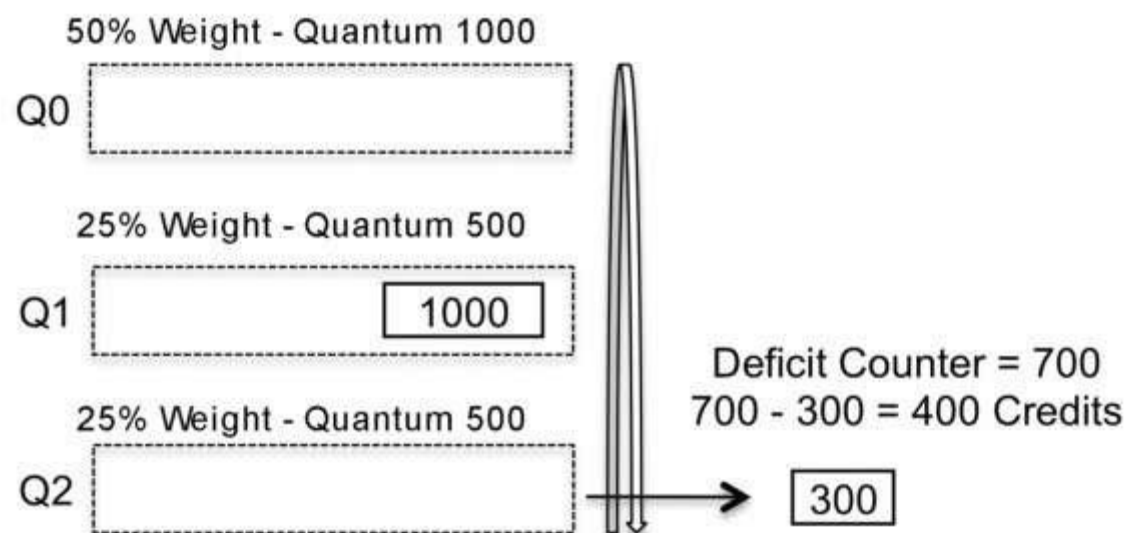


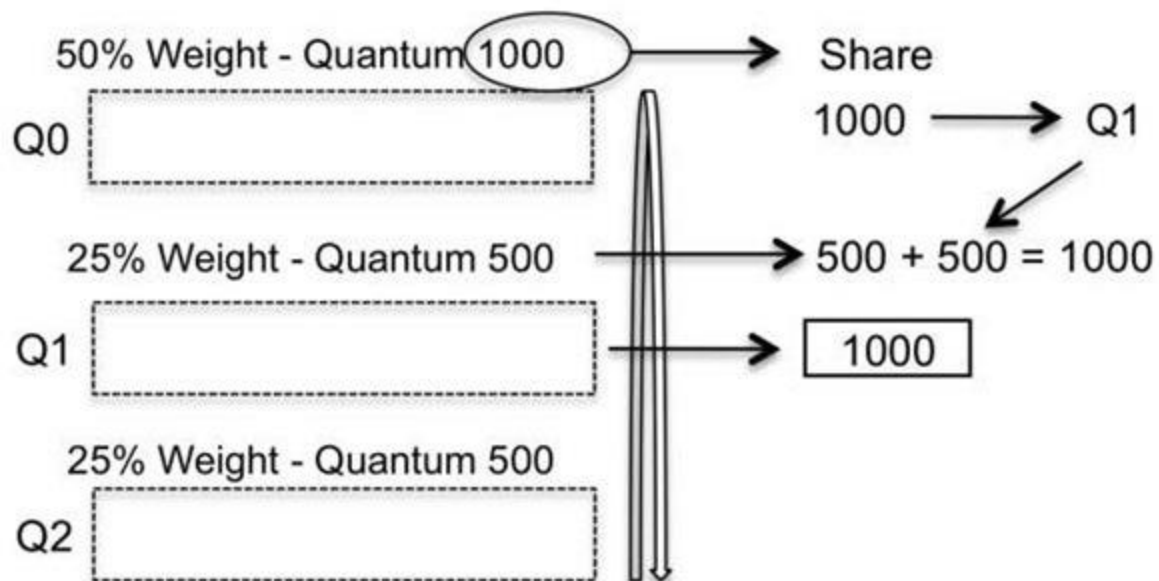














(no) shutdown QoS  
QoS enable / disable  
**Global**

