

## H3C SR8800 QoS技术介绍 (V3R3)

ISSUE 2.0



日期：2007-11-2

密级：公开

杭州华三通信技术有限公司

# 课程目标

学习完本课程，您应该能够：

- 了解QoS的基本概念
- 掌握SR8800 QoS基本实现方法
- 掌握SR8800 QoS技术特点与配置思路

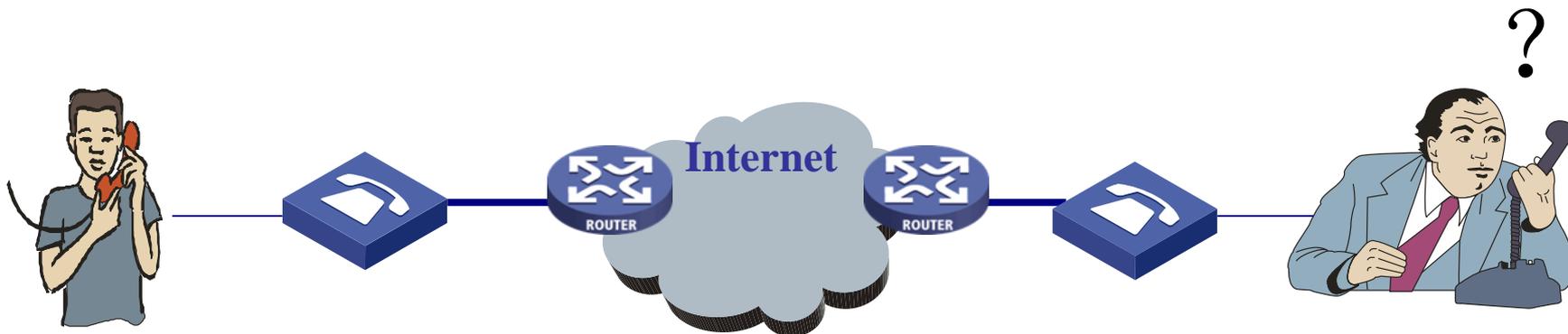




# 目录

- **QoS基本概念与技术发展**
  - **SR8800 QoS的实现方法**
  - **SR8800 QoS技术基本特性与配置**
  - **SR8800 QoS的典型应用**
- 

- **QoS: Quality of Service**（服务质量）是指网络通信过程中，允许用户业务在丢包率、延迟、抖动和带宽等方面获得可预期的服务水平。更简单地说：**QoS**就是针对各种不同需求，提供不同服务质量的网络服务。

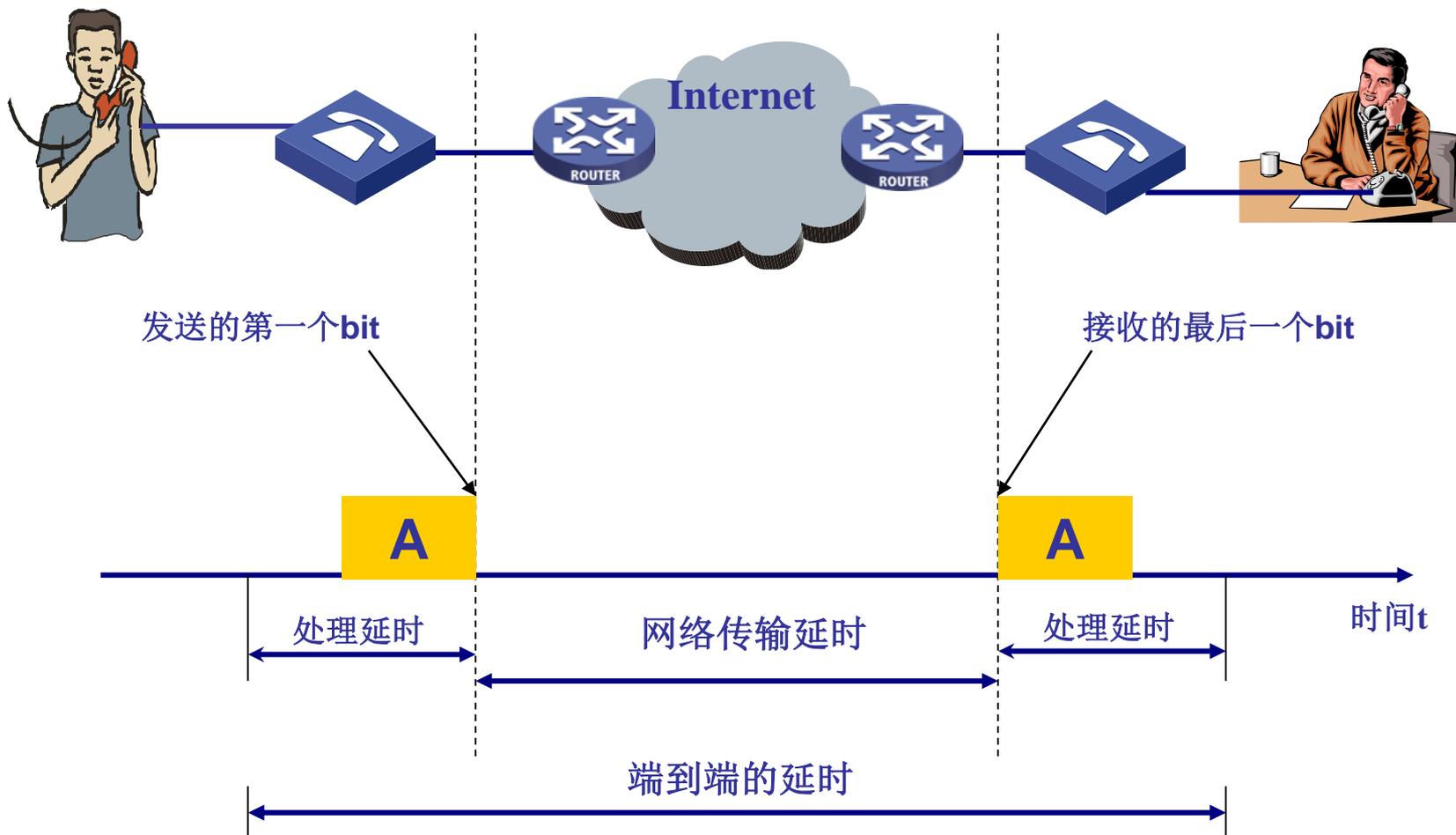


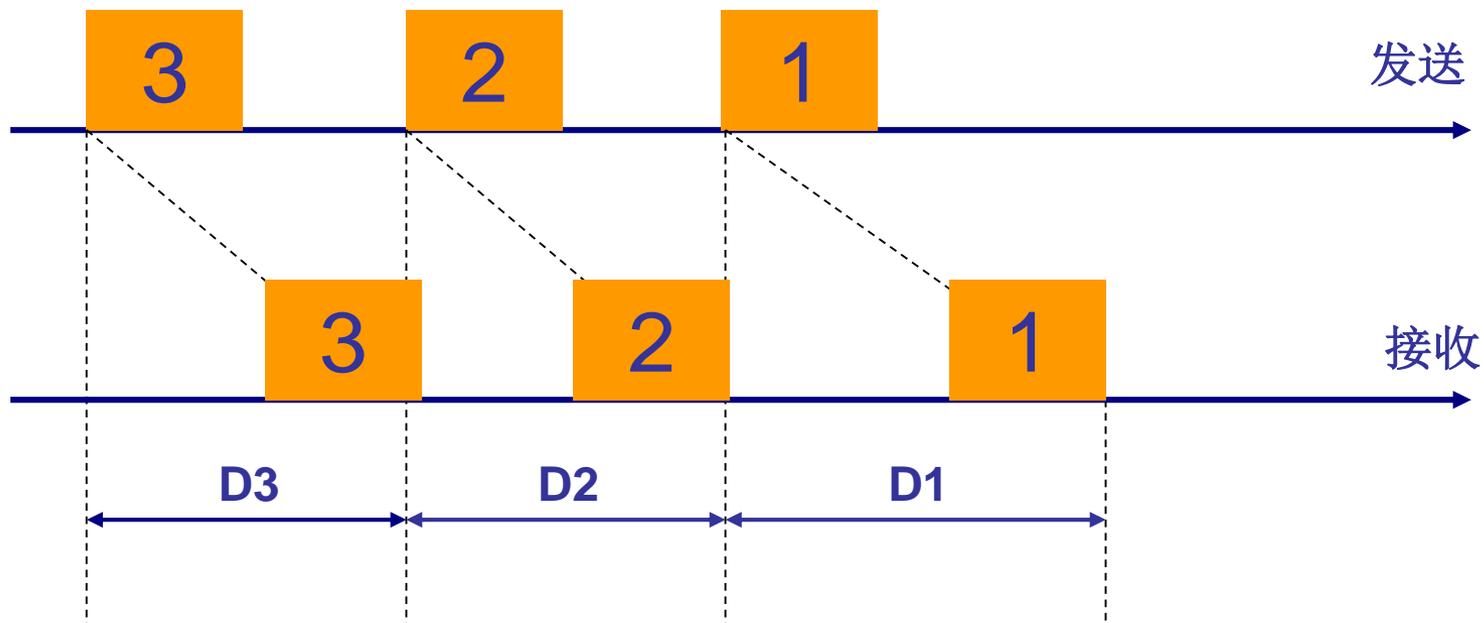
我 是 张 三 啊 .....

本地这么说 .....

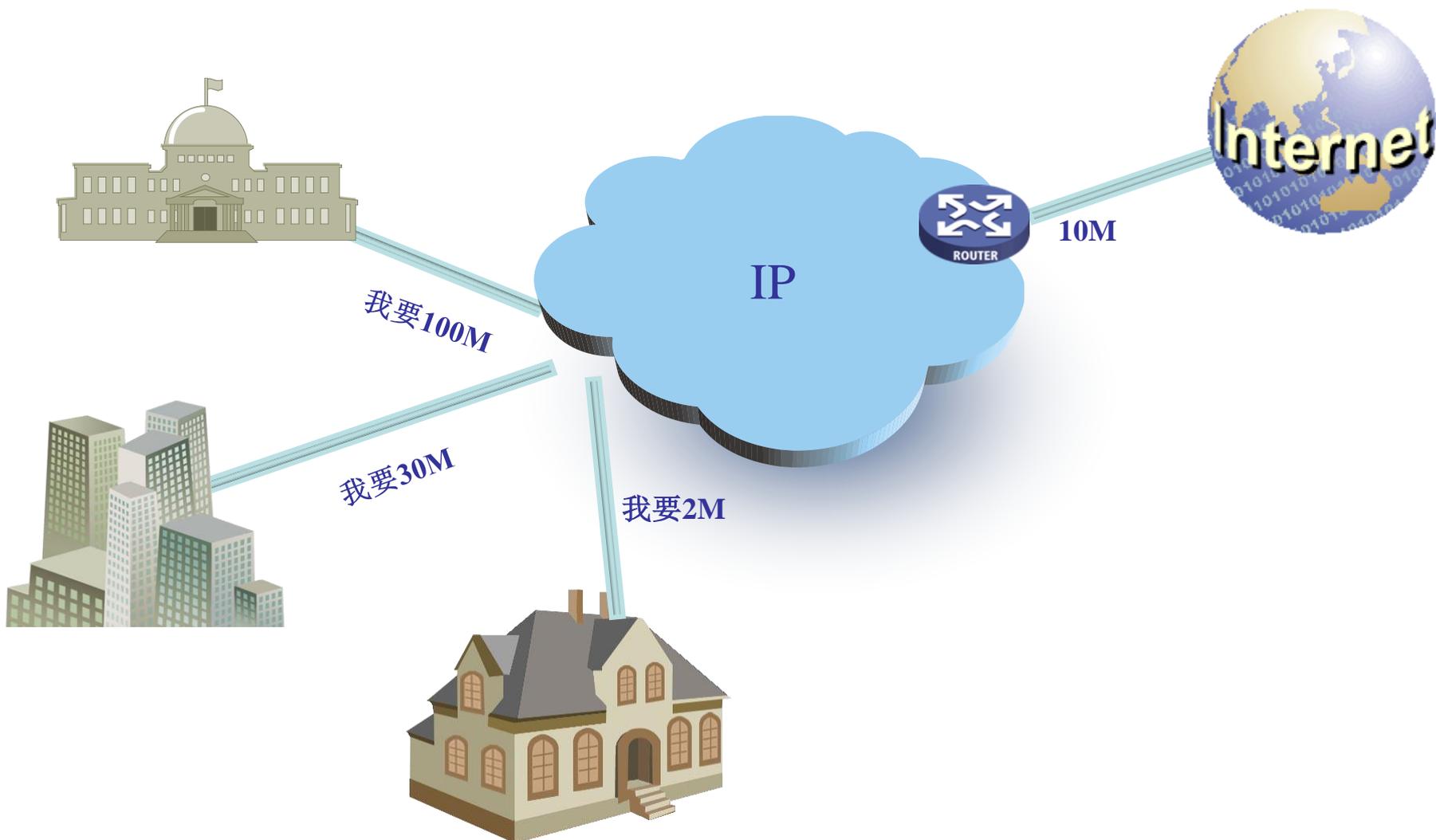
我 是 三 啊 .....

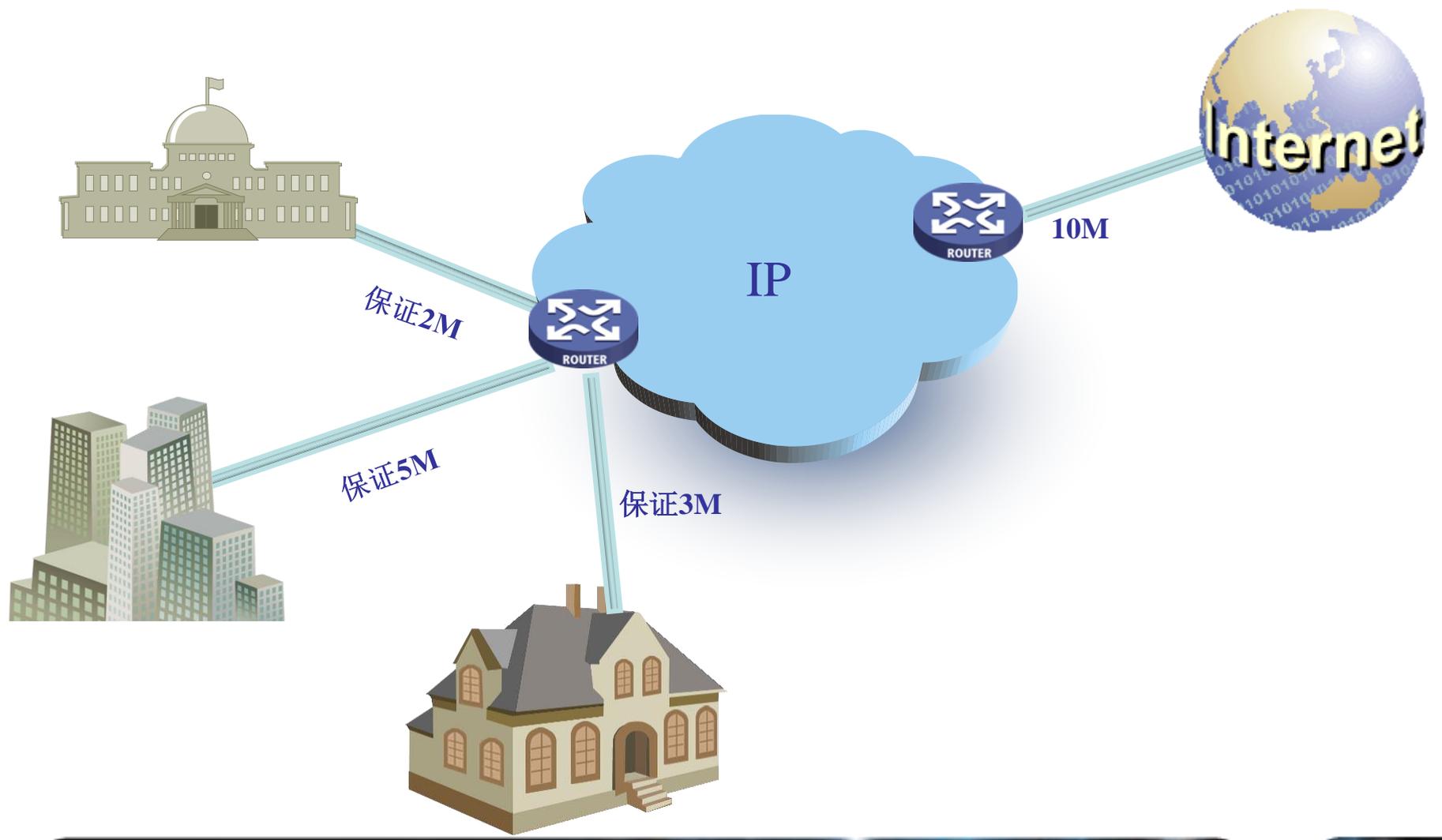
对方听到的是 .....





$$D_3 = D_2 \neq D_1$$

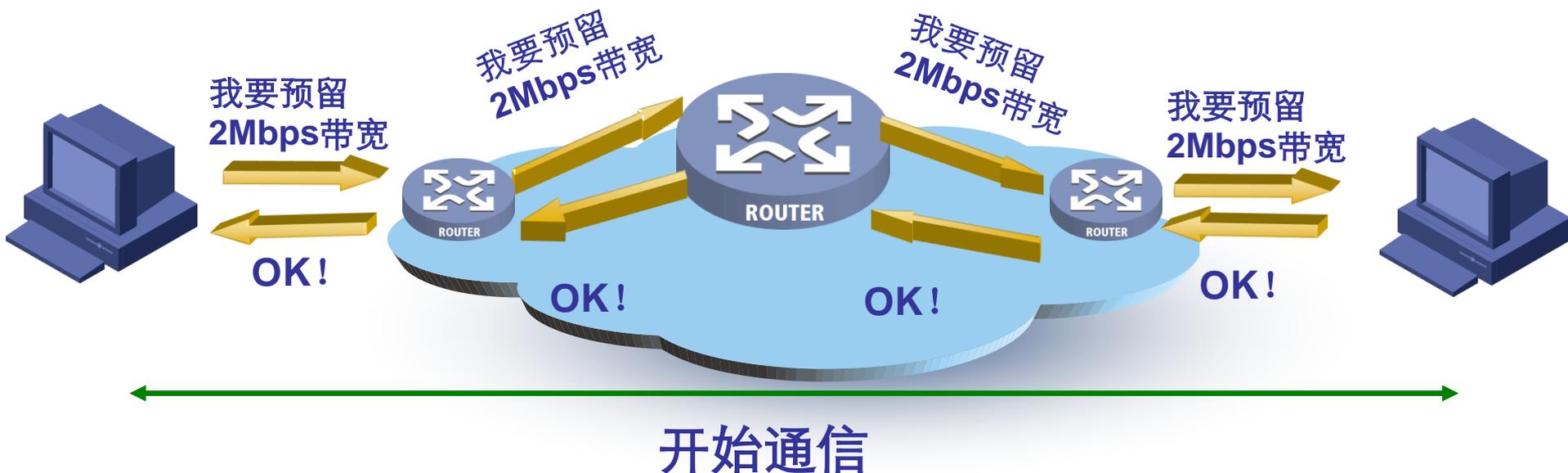




- **Best-Effort 模型**：是目前Internet的缺省服务模式，主要实现技术是先进先出队列(FIFO)；
- **IntServ模型**：业务通过信令向网络申请特定的QoS服务，网络在流量参数描述的范围内，预留资源以承诺满足该请求；
- **DiffServ模型**：当网络出现拥塞时，根据业务的不同服务等级约定，有差别地进行流量控制和转发来解决拥塞问题。

- **Best-Effort**是单一的服务模型，也是最简单的服务模型；
- 应用程序可任意发送任意报文，不需要事先得到批准或通知网络；
- 网络尽最大可能发送这些报文，但对时延、可靠性等性能不提供任何保障。

- 为应用提供可控制的、端到端的服务
- 网络单元支持QoS的控制机制
- 应用程序向网络申请特定的QoS服务
- 信令协议在网络中部署QoS请求
- **RSVP**是主要使用的信令协议



- 目的

- 当网络出现拥塞时，根据不同的服务等级要求，有差别地进行流量控制和转发来解决拥塞。

- 方法

- 采用相对优先级机制，有区别地控制不同信息流的分组聚类 and 转发行为，从而在保证服务质量和解决拥塞之间取得一个良好的折衷。

## ● BA: Behavior Aggregate

→业务在进入DS (DiffServ) 网络时, 基于优先级进行业务流聚合, 汇聚成的行为集合;

## ● PHB: Per-hop Behavior

→DS节点对特定的BA使用的转发行为, DS节点通过判别分组头中的DSCP来执行PHB。

→PHB有三种类型:

- EF (Expedited Forwarding): 快速转发
- AF (Assured Forwarding): 确保转发, 分为四个类
- BE (Best Effort): 尽力传送。

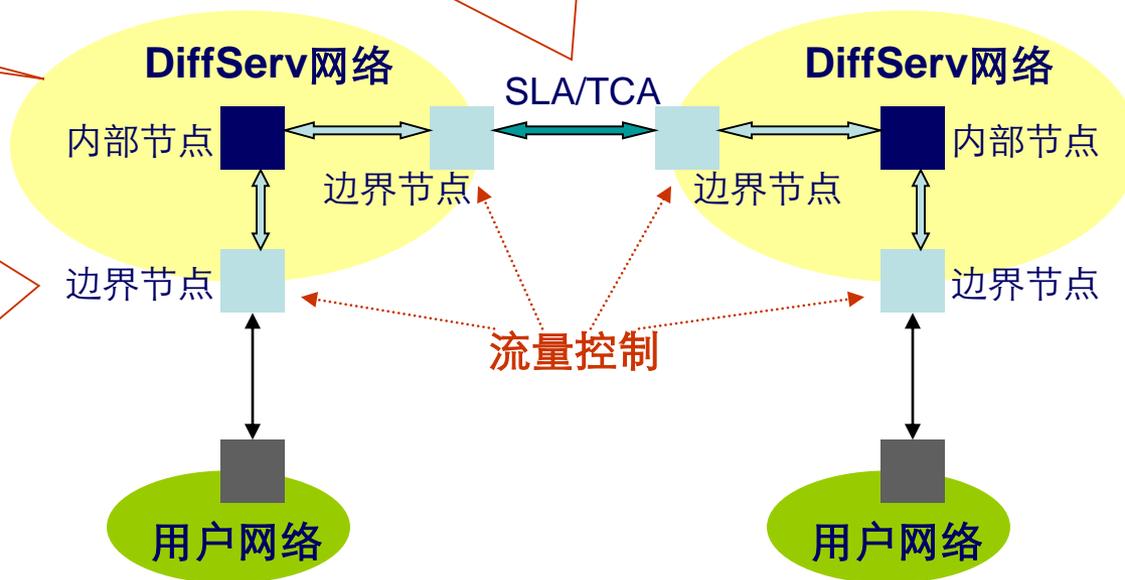
不同DS区域可有不同的PHB，以实现不同的服务提供策略，它们之间通过SLA与TCA协调提供跨区域服务：

- . SLA：服务等级协定，关于业务流在网络中传递时应当获得的待遇。
- . TCA：流量调整协定，关于业务分类准则、业务模型及相应处理的协定。

DS区域的服务提供策略由PHB决定。  
DS节点根据PHB属性转发。

在网络边缘进行业务分类和流量调整。

- 业务分类
  - . 基于DS域
  - . 基于其他特征
- 流量调整
  - . 测量
  - . 标记
  - . 丢弃
  - . 整形



- 多个业务流聚合成一个行为集合（BA），在各网络单元上使用相同的PHB进行转发处理，简化了业务的处理和存储过程；
- QoS保证是基于每个报文的，不需要额外处理信令；
- DiffServ避免了IntServ所需的per-flow状态和每一跳的RSVP协议，进而获得了良好的可扩展性。

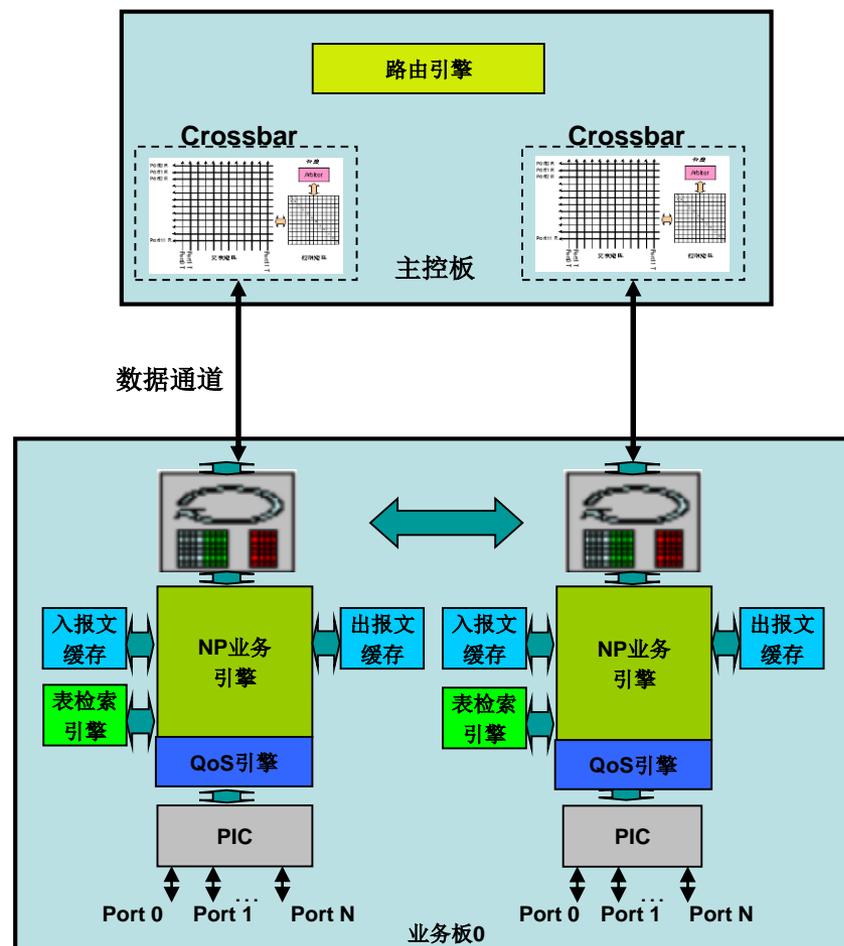


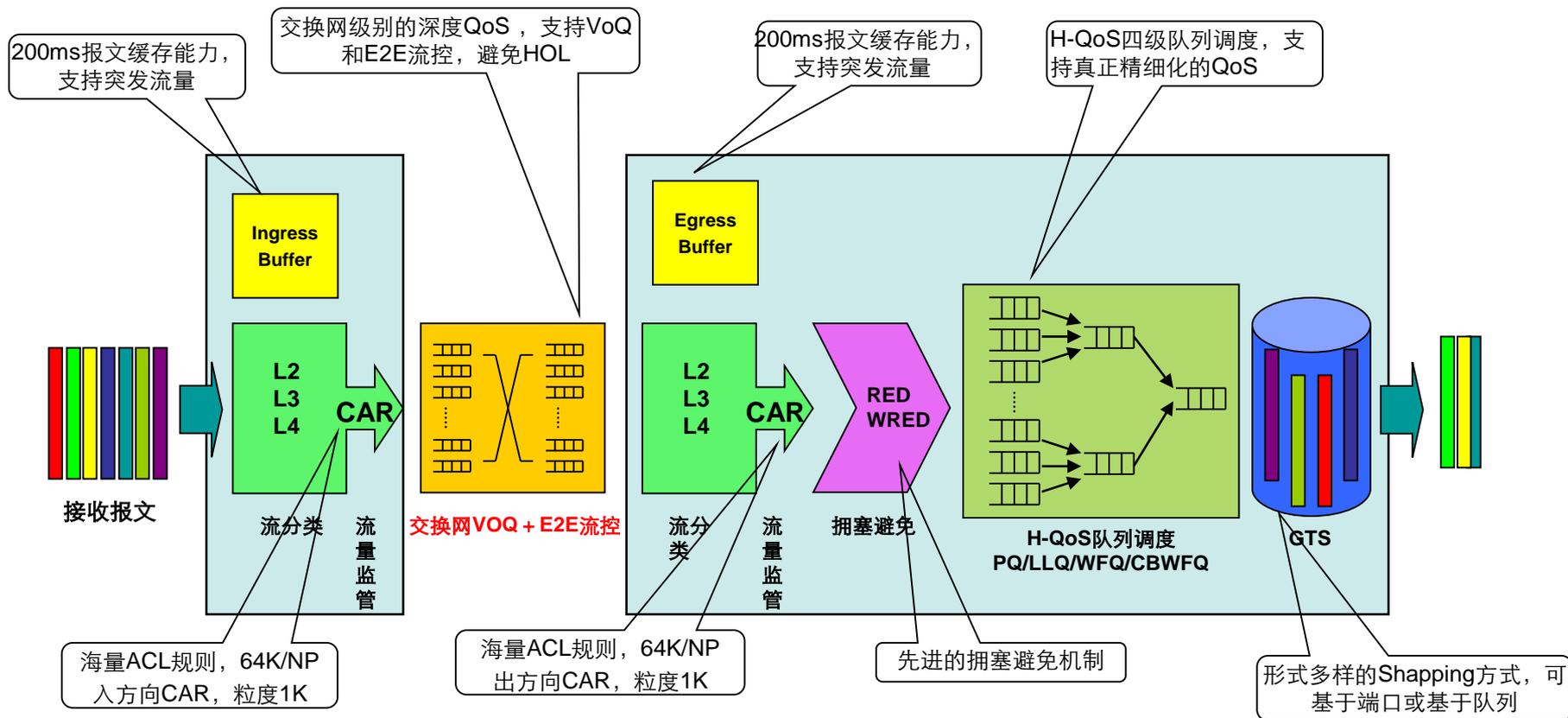
# 目录

- QoS基本概念与技术发展
  - **SR8800 QoS的实现方法**
  - SR8800 QoS技术基本特性与配置
  - SR8800 QoS的典型应用
- 

- **SR8800全面支持Diff-Serv域边缘路由器、核心路由器的QoS要求**
  - 流分类 (复杂流分类)
  - 优先级映射 (简单流分类)
  - 流量监管
  - 流量整形
  - **接口限速(V3R3B03版本新增)**
  - 拥塞管理
  - 拥塞避免
- **支持IntServ模型**
  - MPLS TE
  - TE带宽预留

- SR8800采用专用引擎固化表项查找和QoS调度功能，大幅提高处理性能。
- NP业务引擎：实现了灵活的流分类与流量统计、控制；
- QoS引擎：专用流量管理(TM)芯片；
- ACL表检索引擎：实现大规格下ACL的线速查找；
- 大容量缓存：能缓存200MS报文，应对网上突发流量，增强流量整形效果；





- SR8800支持双向ACL，报文从入接口进入NP完成入方向ACL，实施CAR、流过滤等QoS动作，经过交换网抵达出接口，完成出方向ACL处理，实施CAR等流动作，在接口发送时根据配置进行拥塞避免（WRED）、队列调度、流量整形、带宽保证处理，在需要精细化业务管理时，可以配置分层QoS。
- 入接口缓存能够处理网络中的Burst流量，减少对NP的冲击；出接口大容量缓存还可以完成极佳流量的效果。



# 目录

- QoS基本概念与技术发展
  - SR8800 QoS的实现方法
  - **SR8800 QoS技术基本特性与配置**
  - SR8800 QoS的典型应用
- 

- 简单流分类是实现报文外部优先级和内部优先级之间的映射；
- SR8800**缺省不信任报文所携带的外部优先级，可以通过在接口下发如下配置，信任报文的优先级；

```
[interface GigabitEthernet1/1/1] qos trust auto
```

- 信任报文所携带的外部优先级可以是**DSCP**、**IP**优先级、**dot1p**、**EXP**等，根据接收到的报文类型来决定，如收到一个**MPLS**报文，信任的就是报文的**EXP**值；
- 不配置端口信任模式时，**SR8800**不识别报文所携带的外部优先级，以报文进入端口的端口优先级作为报文的外部优先级，端口优先级也可以通过命令行调整，缺省为**0**：

```
[interface GigabitEthernet1/1/1] qos priority priority-value
```

- SR8800产品内部实现有一张映射表，该表将识别出来的用户优先级与设备内部处理时的本地优先级进行映射关联，其中用户优先级称为up，设备本地优先级称为lp，缺省映射关系如下：

UP	LP
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7

- 通过命令行可以修改up-lp的映射表：

```
[SR8800] qos map-table outbound up-lp
```

- 复杂流分类是指根据五元组（源IP地址、源端口号、协议号码、目的IP地址、目的端口号）、TCP SYN等报文信息对报文进行分类；
- SR8800通过MQC（Modular QoS Commandline）的配置方法实现复杂流分类；
- MQC包含了三个要素：流分类、流行为、流策略，MQC配置不仅仅用在QOS复杂流分类中，在QOS的流量监管、拥塞管理等实现中也采用MQC的配置方法。在一些非QOS技术，如重定向、镜像、流量统计等等技术实现中也采用MQC的配置方法。
- SR8800支持在入接口或者出接口对报文进行复杂流分类。

- 流模板主要用来对硬件下发的ACL规则中所能包含的信息进行限制。
- 设备支持的流模板包括：
  - 缺省流模板
  - 用户自定义流模板
- 流模板在全局配置，在接口应用
  - 配置用户自定义流模板

```
[SR8800] flow-template flow-template-name basic { sip | dip |  
smac | dmac | sport | dport | ethernet-protocol | ip-protocol |  
customer-vlan-id | service-cos | icmp-type | icmp-code |  
fragments | tcp-flag | tos | dscp | ip-precedence }
```

→在接口上应用用户自定义流模板

```
[SR8800] flow-template flow-template-name
```

- 定义类，即定义 **traffic classifier**
- 进入类视图：

```
[SR8800] traffic classifier tcl-name [ operator { and | or } ]
```

- 定义各种规则，例如：

1	定义ACL匹配规则	<b>if-match acl</b> <i>acl-number</i>
2	定义IPv6 ACL匹配规则	<b>if-match ipv6 acl</b> <i>acl-number</i>
3	定义目的MAC匹配规则	<b>if-match destination-mac</b> <i>mac-address</i>
4	定义DSCP匹配规则	<b>if-match dscp</b> <i>dscp</i>
5	定义8021p匹配规则	<b>if-match 8021p</b> <i>8021p-value</i>
...	.....	

- 定义流行为，即定义 **traffic behavior**
- 进入流行为视图：

```
[SR8800] traffic behavior behavior-name
```

- 定义复杂流分类规则，例如：

配置标记报文的DSCP值	<code>remark dscp dscp-value</code>
配置标记报文的IP优先级值	<code>remark ip-precedence ip-precedence-value</code>
配置标记报文的Dot-1p值	<code>remark dot1p 8021p</code>
配置标记报文的exp值	<code>remark mpls-exp exp-value</code>
配置标记报文的本地优先级	<code>remark local-precedence local-precedence</code>
.....	

- 其中标记本地优先级，只能下发在端口出方向。

- 定义类，即定义 **traffic policy**
- 进入 **Policy** 视图：

```
[SR8800] traffic policy policy-name
```

- 定义 **classifier-behavior** 对，可定义多对，当前产品支持256对。

```
[SR8800- traffic policy- policy-name] classifier classifier-name  
behavior behavior-name [ precedence classifier-precedence ]
```

## ●traffic-policy 应用在接口上

```
[SR8800-interface GigabitEthernet1/1/1] qos apply policy policy-name { inbound | outbound }
```

## ●traffic-policy应用在VLAN上

```
[SR8800] qos vlan-policy policy-name vlan vlan-id-list { inbound | outbound }
```

## ●支持流分类与流行为以及**ACL规则**的动态修改，而不需要先从接口删除**Policy**。

→包括ACL规则、C-B对、if-match 增加和删除、if-match AND/OR 修改；

→其中动态修改ACL规则 V3R3版本才可以支持。

- SR88产品支持普通CAR和聚合CAR，其中聚合CAR实现对多个端口上的业务流使用同一个CAR进行流量监管，使得多个端口的流量之和在聚合CAR设定的流量监管范围之内。
- 配置普通CAR，采用MQC的配置方法，在流行为的视图下配置CAR：

```
[SR8800- traffic behavior- behavior-name] car cir committed-information-rate
```

- 配置聚合CAR，在系统视图定义聚合CAR，并在流行为视图下引用聚合CAR

```
[SR8800] qos car car-name aggregative cir committed-information-rate
```

```
[SR8800- traffic behavior- behavior-name] car name car-name
```

- 根据Policy的配置，SR8800支持出、入方向的CAR；其中，CIR取值范围为64~10000000 kbps，粒度1Kbps。

- **TS (Traffic Shaping, 流量整形)** 是一种主动调整流量输出速率的措施。一个典型应用是基于下游网络节点的**TP**指标来控制本地流量的输出。
- **SR88支持基于端口和队列的流量整形，其中V3R3B03版本开始支持以太网子接口GTS (SPC板不支持)。**
- **基于端口的GTS配置：**

```
[SR8800-interface GigabitEthernet1/1/1] qos gts any cir  
committed-information-rate
```

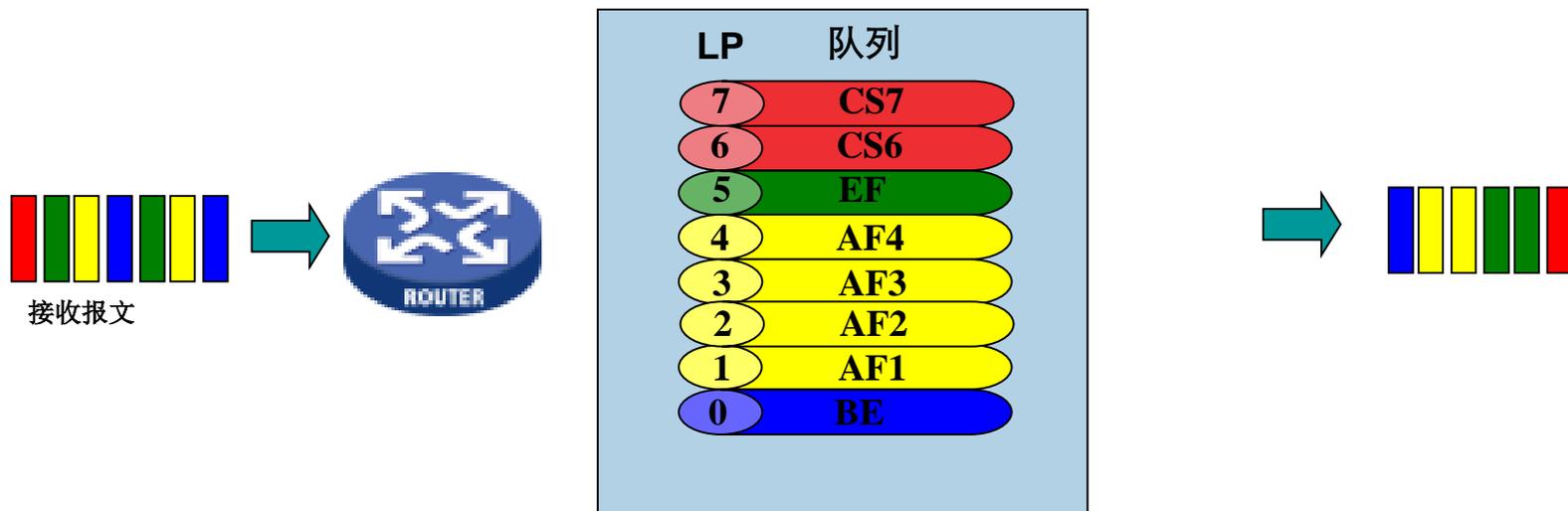
- **基于队列的GTS配置：**

```
[SR8800-interface GigabitEthernet1/1/1] qos gts queue queue-  
number cir committed-information-rate
```

- LR（Line Rate，物理接口限速）可以在一个物理接口上限制发送报文（包括紧急报文）的总速率。
- SR88支持对各种端口进行接口限速，包括以太网子接口。
- 端口限速的配置：

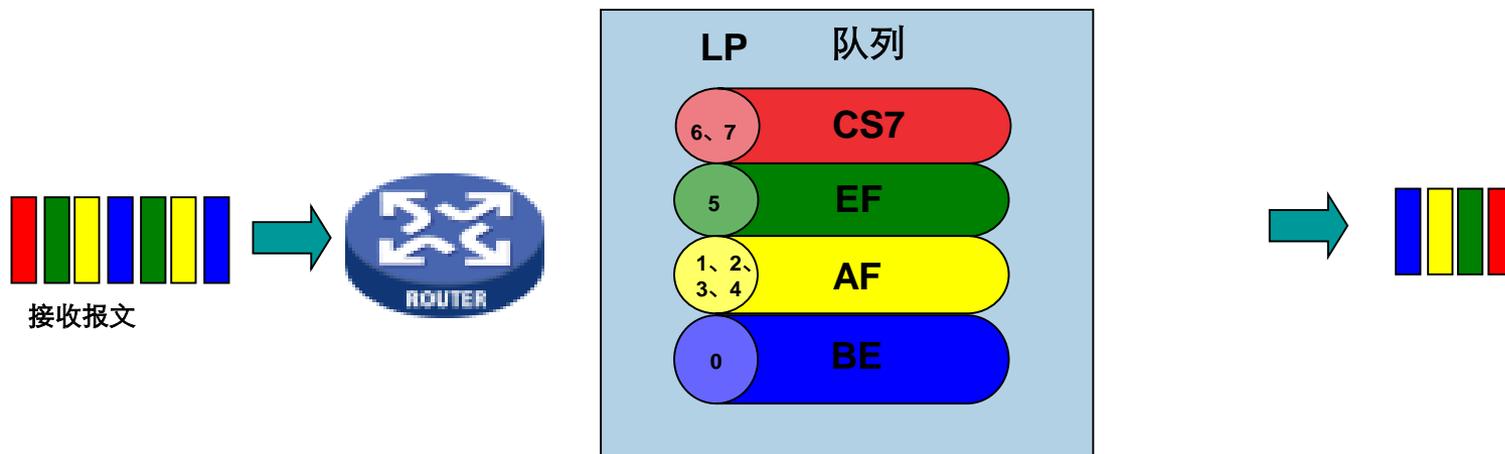
```
[SR8800-interface GigabitEthernet1/1/1] qos lr outbound cir  
committed-information-rate [ cbs committed-burst-size ]
```

- SPC单板接口暂不支持该功能



- SR8800每端口（包括以太口、POS、Serial、MP等）默认8个队列。
- 队列属性默认为：0队列为BE、1/2/3/4为AFx队列、5为EF队列、6/7为CS队列。
- 报文在端口的出方向根据报文的LP（本地优先级）值，选择进入不同的队列，LP值与队列的一一映射如上图所示。通过配置流分类，可根据用户需求，使得不同的用户报文进入不同的队列进行转发，从而实现拥塞管理。

- SR8800不同的队列类型之间默认采用PQ调度方式。其中优先级CS>EF>AF>BE，绝对保证高优先级流量；
- 配置PQ只需要配置流分类，将用户不同报文映射到不同的优先级。
- SR88支持四种类型的队列，所以最多可实现四种类型报文PQ转发。



- **SR8800高优先级CS、EF队列均为LLQ队列，满足低延时、抖动小的需求；一般用于视频和生产业务流；**
- **SR8800的CS7是协议队列，一些多跳的协议报文将采用该队列进行转发，不建议给用户业务使用，以免影响协议。**

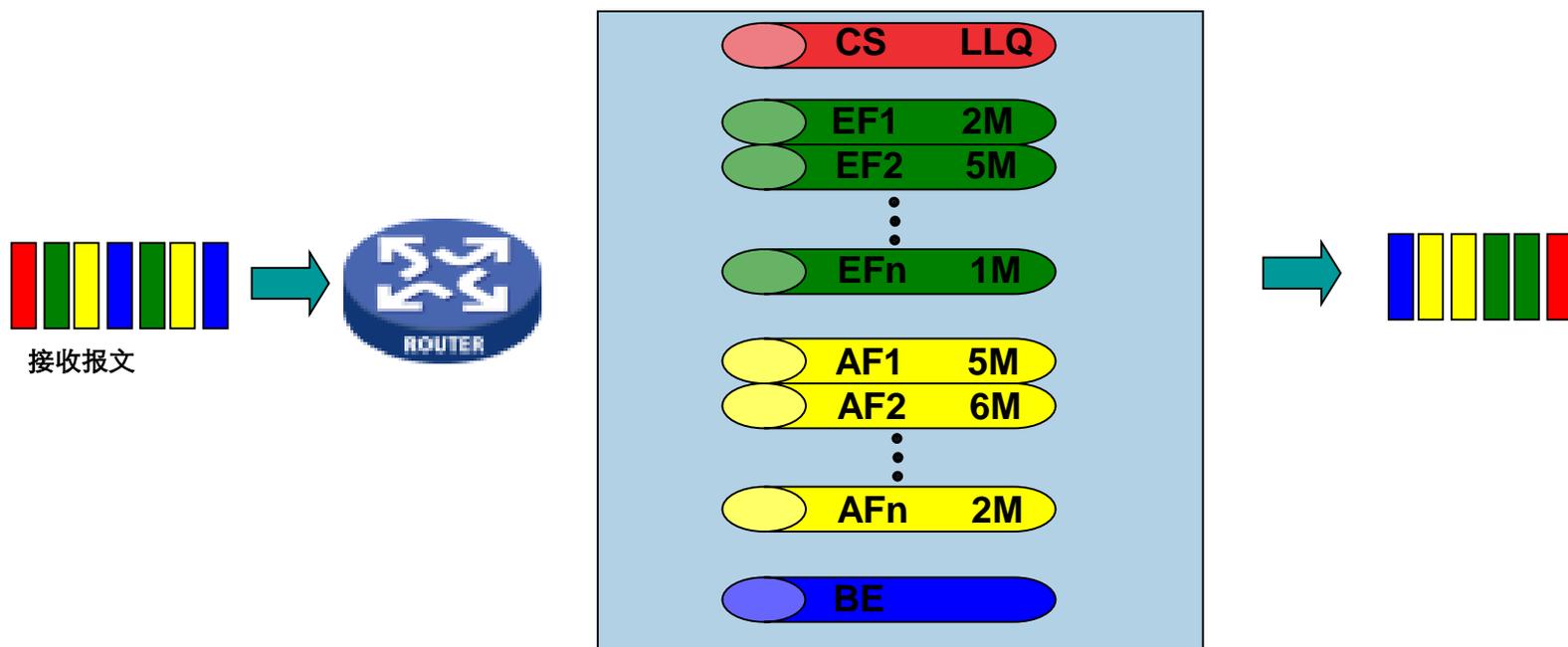
- **SR8800**同种类型的队列之间采用**WFQ**的方式进行调度，如**AFx** ( $x = 1, 2, 3, 4$ ) 之间采用**WFQ**调度；所有队列**WFQ**默认权重为**1**；根据队列带宽要求配置相应的权重。
- 调整**WFQ**权重配置：

```
[SR8800-interface GigabitEthernet1/1/1] qos wfq queue weight  
schedule-value
```

- V3R1及V3R2版本，SR8800实现CBQ的时候只能采用传统的CBQ配置方法，V3R3版本可以继承，升级后可不做调整；
- 传统CBQ配置SR8800支持对BE以外的队列配置带宽保证，对于高速接口（如GE、POS等），配置带宽保证以后，超过保证带宽的流量降级到BE队列，与原BE流一起WFQ调度；
- 对于低速接口如（CPOS、E1等），配置带宽保证后，只有EF队列支持降级BE，其他队列不降级。
- 配置举例：

```
[SR8800-interface GigabitEthernet1/1/1] qos bandwidth queue  
queue-number min bandwidth-value
```

- 传统CBQ最多可以对用户的七种业务进行带宽保证；
- 传统CBQ对于低速接口来讲，对CS队列的带宽保证没有意义，因为超过保证带宽的流量不降级，仍然优先于EF、AF、BE队列，CS队列的所有流量仍然将得到优先转发；
- 传统CBQ对于低速接口来讲，对AF队列的配置带宽保证时，通过对AF队列权重的调整，可以保证4个AF队列在保证带宽以内的流量得到正常转发，但是超过保证带宽的流量因为不降级，优先级高于BE，可能造成BE饿死的情况。



- SR8800从V3R3版本开始，支持基于MQC配置的CBQ功能。
- 通过这种配置方法，EF、AF队列数量没有限制，由ACL规格和QOS策略中C、B对规格来决定。

- **SR8800基于MQC的CBQ配置实现效果如下：**

- EF实现：EF超出保证带宽的业务流量降低至比BE还低的级别；

- AF实现：AF超出保证带宽的业务流量与BE形成WFQ；

- **配置时需要注意：必须配置缺省流进BE队列。**

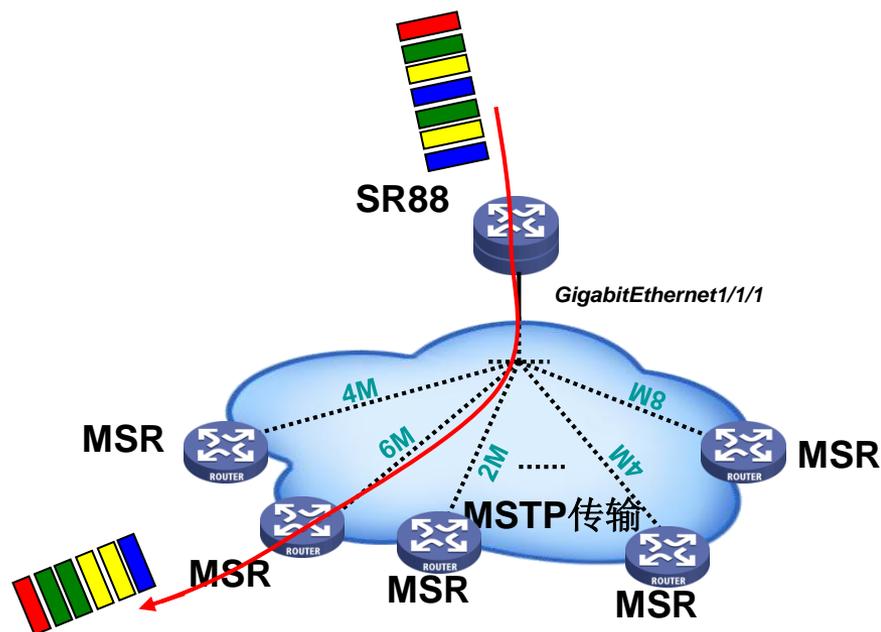
- **配置采用MQC的配置方法，其中流行为部分配置保证带宽，例如：**

```
[SR8800] traffic behavior behavior-name
```

```
[SR8800- behavior- behavior-name] queue ef bandwidth  
bandwidth [ cbs burst ]
```

```
[SR8800- behavior- behavior-name] queue af bandwidth  
bandwidth [ cbs burst ]
```

- 传统QOS可以配置带宽保证的流量数量有限，基于MQC配置的CBQ可以配置带宽保证的流量数量大幅增多，可以满足用户业务更多的划分；
- 传统QOS在低速接口只有EF队列支持超过保证带宽的流量降级，基于MQC配置的CBQ不存在这样的问题，在低速接口与高速接口EF和AF队列都能降级；
- 传统QOS超过带宽保证的流量与BE同一级别，无法满足用户优先丢弃超带宽保证流量的业务部署需求，基于MQC配置的CBQ，EF流超过带宽保证的流量降级至比BE还低的级别，满足该需求；
- 传统QOS直接对接口上的队列进行配置，以太网子接口无法部署。基于MQC配置的CBQ，QOS策略可以在以太网子接口下发，可以满足新一代的MSTP组网的需求。



- 以太网子接口下首先部署LR或GTS匹配传输分配的带宽（V3R3B03版本开始支持）；
- 再在子接口下部署基于MQC配置的CBQ实现拥塞管理，区分用户业务进行带宽保证；
- V3R3B03版本支持子接口的流量统计功能（可通过命令行或网管查看子接口流量），但是该功能与基于MQC配置的CBQ功能冲突，不能同时部署。

- **SR8800缺省采用WRED的拥塞避免方法，建议正常情况下，直接选用设备的缺省WRED参数取值，不用做任何WRED的相关配置。**
- **调整WRED的相关参数，采用表配置方式：在系统视图下配置WRED表，然后在接口上应用WRED表。**

```
[SR8800] qos wred queue table table-name
```

```
[SR8800-wred-table-queue-table ] queue queue-value  
[ drop-level drop-level ] low-limit low-limit high-limit  
high-limit[discard-probability discard-prob ]
```

```
[SR8800-interface GigabitEthernet1/1/1]qos wred apply  
table-name
```

## ● V3R3版本变更：

- 支持基于MQC配置的CBQ功能（V3R3B02）
- 以太网子接口支持GTS（V3R3B03）
- 支持端口限速 LR，包括以太网子接口（V3R3B03）

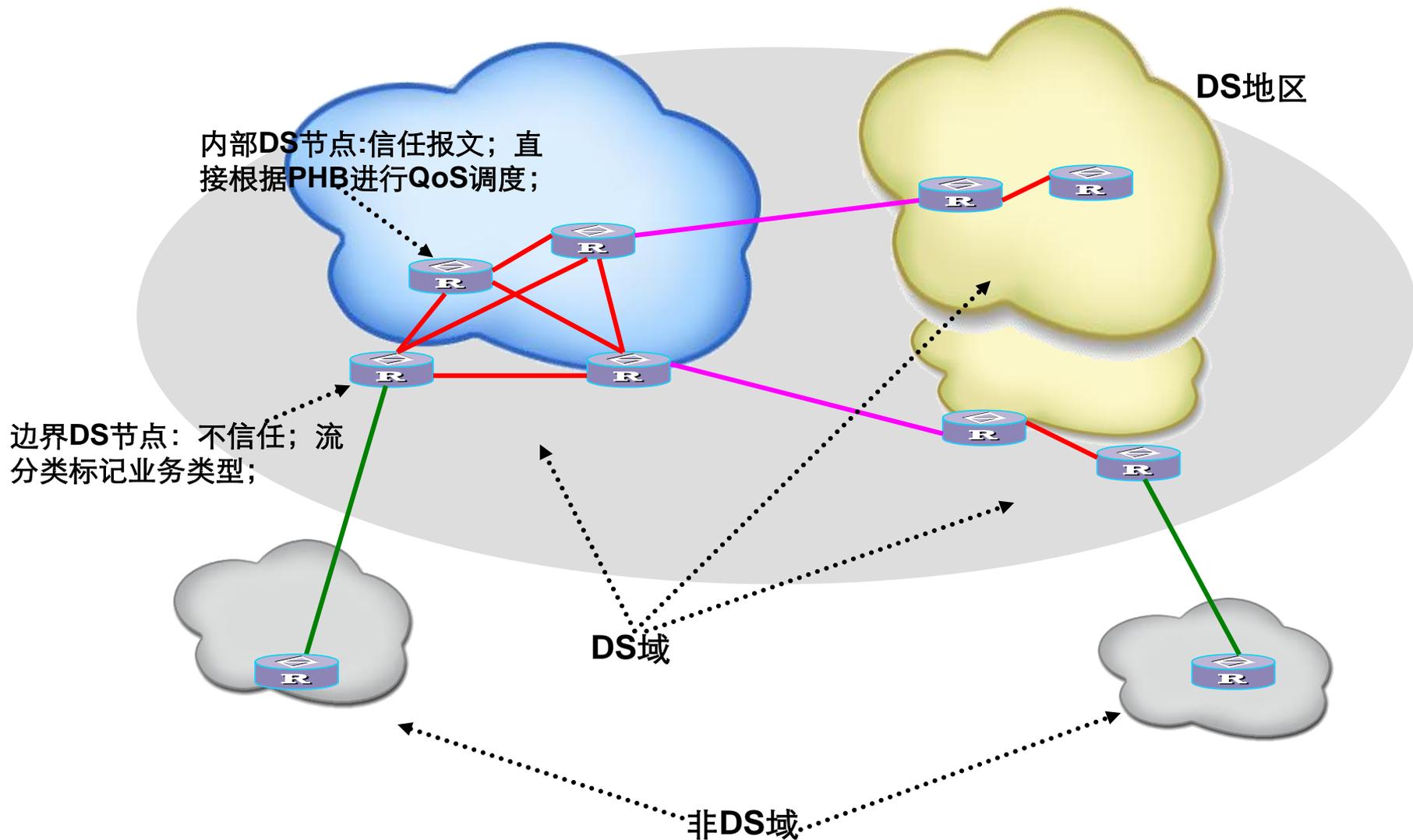
## ● SPE板与SPC板差异：

- 拥塞管理部分，SPC板实现与95E类似，支持SP和WRR，不支持CBQ功能。
- SPC板上的GE口，可以配置子接口，但不支持GTS或LR功能，更不能再不是CBQ，不适合MSTP组网应用。



# 目录

- QoS基本概念与技术发展
  - SR8800 QoS的实现方法
  - SR8800 QoS技术基本特性与配置
  - **SR8800 QoS的典型应用**
- 





## 本章总结

- 介绍了QoS的基本概念
- 介绍了SR8800 普通QoS的实现方式
- 介绍SR8800分层QOS的实现与典型配置
- 详细介绍了SR8800 QoS的应用和配置

# H3C

IToIP 解决方案专家

杭州华三通信技术有限公司

[www.h3c.com.cn](http://www.h3c.com.cn)