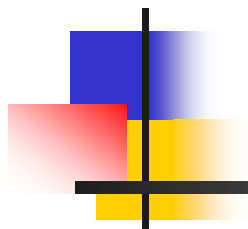


宽带通信网



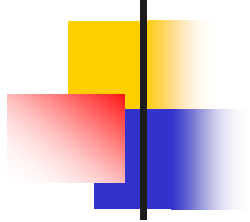
信息与通信工程学院

靳浩



宽带通信网

- 通信网技术的发展
- ATM 技术原理
- 宽带网交换技术
- IP网络体系结构与关键技术
- IP网络的QoS技术
- IP网络安全与管理技术
- MPLS技术及其发展
- 移动IP技术及其发展
- 下一代网络技术



IP网络的QoS技术

- QoS技术概述
- IP的QOS技术
- IP的业务模型
- IP的QOS技术实施



QoS的定义

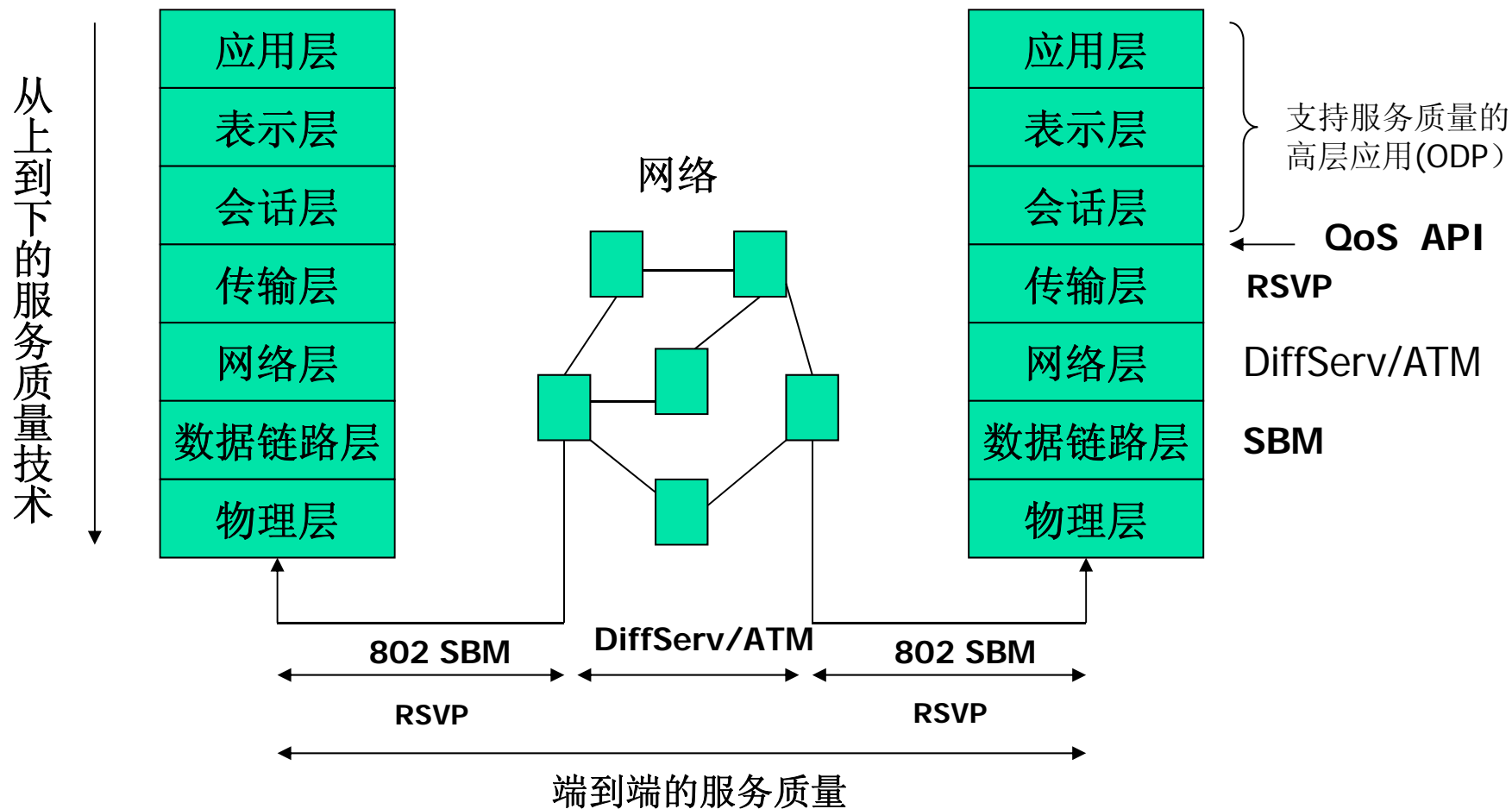
- 在通信和计算机网络中，服务质量简称QoS(Quality of Service)；
- 目前学术界普遍认为QoS有广义和狭义之分：
 - 狭义QoS：指技术指标，例如传输延迟、抖动、丢失率、带宽要求、吞吐量等；
 - 广义QoS：指资源调配与利用、层与层之间的协商，从而涉及不同层次的QoS。
- 目前主要研究狭义QoS。



QoS标准及应用层次

- 目前涉及QoS控制和管理的标准有：
 - 应用层：ISO/OSI提出了基于ODP分布式环境的QoS控制；
 - 网络层：ATM论坛提出了QoS控制的策略和实现；
 - IETF的综合业务模型和差分业务模型用于解决因特网的QoS控制和管理；
 - 数据链路层：以太网络中提出的802.1p、802.1q以及SBM (Subnet Bandwidth Management) 。

QoS标准及应用层次





QoS标准及应用层次

- 不同企业使用的服务质量参数不同，导致各层次上的参数的语法、语义和语用不一致，使得提出的种种QoS技术难以协同地为建立复杂的分布式系统而服务。
- 系统中高层应用程序与下层不同层次中单元的QoS参数含义、表示和度量不同，需要的QoS转换太多，导致系统复杂低效。



目前主要的服务质量等级及映射

- 以下标准化组织对服务质量等级进行了

定义：

- ITU-T
- ETSI的TIPHON
- 3GPP



目前主要的服务质量等级及映射

■ ITU-T:

- Y.1541只将IP业务QoS分为6类：主要根据IPTD（传输时延）、IPDV（时延变化）、IPLR（丢包率）、IPER（错误率）四个方面综合划分QoS类别；各级与延迟敏感性、丢包率之间有按照等级的映射；（语音最大400ms，抖动50ms，丢失率为 10^{-3} ）；
- H.323 Annex N定义的业务类别分为两大类，即GSC（对时延、抖动敏感）和CSC（无要求），各自又分别分为4类，从错误率（有/无）和带宽（CBR/VBR等）的角度进行映射。
- F.700基于业务流的方向将业务划分为会议业务、谈话业务、分发业务、检索业务、采集业务和消息业务等六类；



目前主要的服务质量等级及映射

■ ETSI的TIPHON

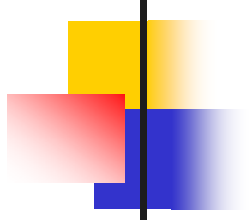
- 基于VoIP将QoS业务分为三大类，即宽带、窄带和BE，其依据是端到端延迟；
- 三类业务的时延限值依次递增，对应于用户感知的语音质量的满意度则是依次递减；
- 窄带业务又根据时延大小细分为三类：即高、中和可以接受，分别对应于不同等级的应用。



目前主要的服务质量等级及映射

■ 3GPP

- 主要针对移动网络，将QoS类别分为 conversational、streaming、interactive 和 background 四类，分类依据是业务对时延的敏感程度。
 - 前两类主要用于实时流量业务，区别在于对时延的容许程度；
 - 后两类主要用于传统的IP应用，都有误码率要求，前者多用于交互式场合，后者主要用于后台业务。



IP的QoS概述

- IP QoS的提出
- IP QoS等级
- IP QoS的演进
- IP QoS的定义和性能参数
- IP的QoS实施



IP的QoS概述

■ IP QoS的提出

- 未来的网络由原先不同网络支持不同业务趋向于单一的基于分组的IP网络；
- Internet需要支持包括VoIP、视频会议在内的各种业务需求；
- IP分组通常使用报头的5个流字段（源IP地址、目标IP地址、IP协议、源及目标端口）对流进行区分，尽力而为的服务方式对流不提供任何服务保证；
- IP网络无法运载对网络资源和服务有特定要求的通信。



IP的QoS概述

■ IP 网络中的QoS等级

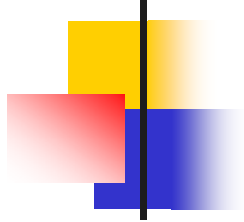
- 尽力而为的服务（只提供基本连接，路由器输入/输出队列用尽才丢弃分组），不属于QoS范畴；
- 区分服务（通常称为CoS，区分服务本身不提供QoS保证，只是优先处理某些通信，也称为软QoS，适合于带宽密集型应用）
- 保证服务（需要预留资源，满足特定的通信要求，也称为硬QoS，可以基于单个流预留/聚合预留，VoIP是一个典型应用）



IP的QoS概述

■ IP QoS的演进

- IPv4报头；
- 最初是针对主机的QoS，Nagle的拥塞崩溃（TCP）是IP QoS的开始；
- 1986年，Van Jacobson的TCP终端的拥塞避免机制（慢启动）；
- 快速重传（快速恢复）机制（1990年）
- 端到端的QoS（路由器上的FIFO/WFQ/WRED）
- IP QoS的业务模型（RSVP协议和差分服务）
- 移动环境中对IP QoS的支持（业务模型、移动IPv6）。



IP的QoS概述

■ IP QoS的定义

- IP网络的QoS是指IP数据包流过一个或多个网络时表现的性能属性，通常以一组可以测量的参数来表征。



IP的QoS概述

■ IP QoS的性能参数

- 带宽：用来描绘给定介质、协议或连接的额定吞吐量，实际上是指应用程序在网络上通信所需要的管道大小。
- 分组延迟：在两个参考点之间，发送和接收IP数据包的时间间隔，包括串行化延迟、传播延迟和交换延迟。
 - 串行化延迟：输出速率一定时，设备同步一个分组所需要的时间，它取决于链路的带宽以及分组的大小，也称为传输延迟。（3Mb/s同步64字节的分组需要171us传输。）
 - 传播延迟：一个数据从发送方到达接收方所需要的时间，通常取决于距离和介质，与带宽无关，以ms计；
 - 交换延迟：设备从收到分组到开始传播的时间，通常小于10us。



IP的QoS概述

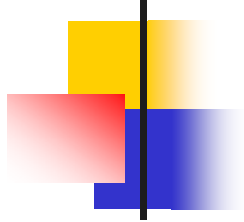
■ IP QoS的性能参数

- 抖动：分组延迟的变化程度，从抖动可以估算接收方分组的最大延迟，以便调整缓存区补偿延迟。
- 丢失率：通常是指在特定时间段内丢失的分组占传输的分组总数的比例；网络拥塞时丢失分组是分组丢失的主要原因。



IP网络的QoS技术实施

- IP的业务模型
- IP的QoS机制
 - 网络边界流量控制器（分组分类/标记、速率管理）
 - 每一跳行为（PHB）
 - 资源分配（分组调度机制）
 - 拥塞避免和分组丢弃（TCP慢启动和拥塞避免、随机早期侦测RED、加权随机早期侦测WRED、流WRED、显式拥塞通知ECN、选择性分组丢弃SPD等）
 - 第二层QoS和MPLS QoS与IP QoS的互联
 - ATM、FR、LAN等二层的QoS及其与IP的互联
 - MPLS的QoS
 - QoS信令
 - 交换
 - 基于策略的路由



IP网络的业务模型

- 综合业务模型 (IntServ)
- 差分业务模型 (DiffServ)



综合业务模型 (IntServ)

- 概述
- 服务类型
- RSVP操作
- RSVP的适用范围
- RSVP的扩展性
- 小结



概述

- IETF于1994年启动了综合业务模型的研究；
- RSVP定位为综合业务模型的信令协议，可以通过服务参数对应用程序的QoS进行需求的满足；
- RSVP最初针对的是多媒体通信（音频/视频会议和广播），但对网络文件系统、VPN等单播业务也可以使用；
- RSVP沿着网络中可用的路径传达资源预留要求，其数据通信和控制通信的路径是由底层的路由协议决定的；
- 模块化设计方式：当路由协议随着网络拓扑变化进行调整时，RSVP预留将在新的路径上传输；
- 资源预留基于流，从接收端向发送端方向进行。



服务类型

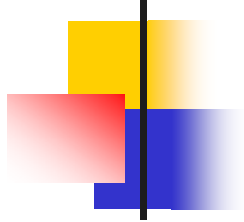
■ 控制负载服务：

- 网络保证预留的流到达目的地，受尽力而为通信的影响最小；
- 主要针对当前Internet上运行的许多对过载敏感的应用程序，它们在未满载时工作很好，在过载情形下，性能急剧降低（FTP）；

■ 保证比特率服务：

- 在网络部件无故障、路由没有变化时，能为顺从数据报提供没有排队丢失的限定延迟服务。网络可以确保尽力而为的干扰最小，预留通信之间的隔离和量化的最差状况延迟；
- 最差状况延迟是指最大排队累积延迟，而不是最大端到端延迟。（取决于通信路径）
- 适合于重放和实时应用程序，重放采用抖动缓冲区弥补分组的延迟变化，通过排队延迟，可以估算所需要的抖动缓冲区大小。

■ 尽力而为服务：无服务质量保证的服务。



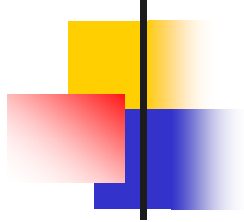
RSVP 消息

■ RSVP的消息集

RSVP在操作中使用了七个消息，其中2个是必需的消息，5个是可选的：

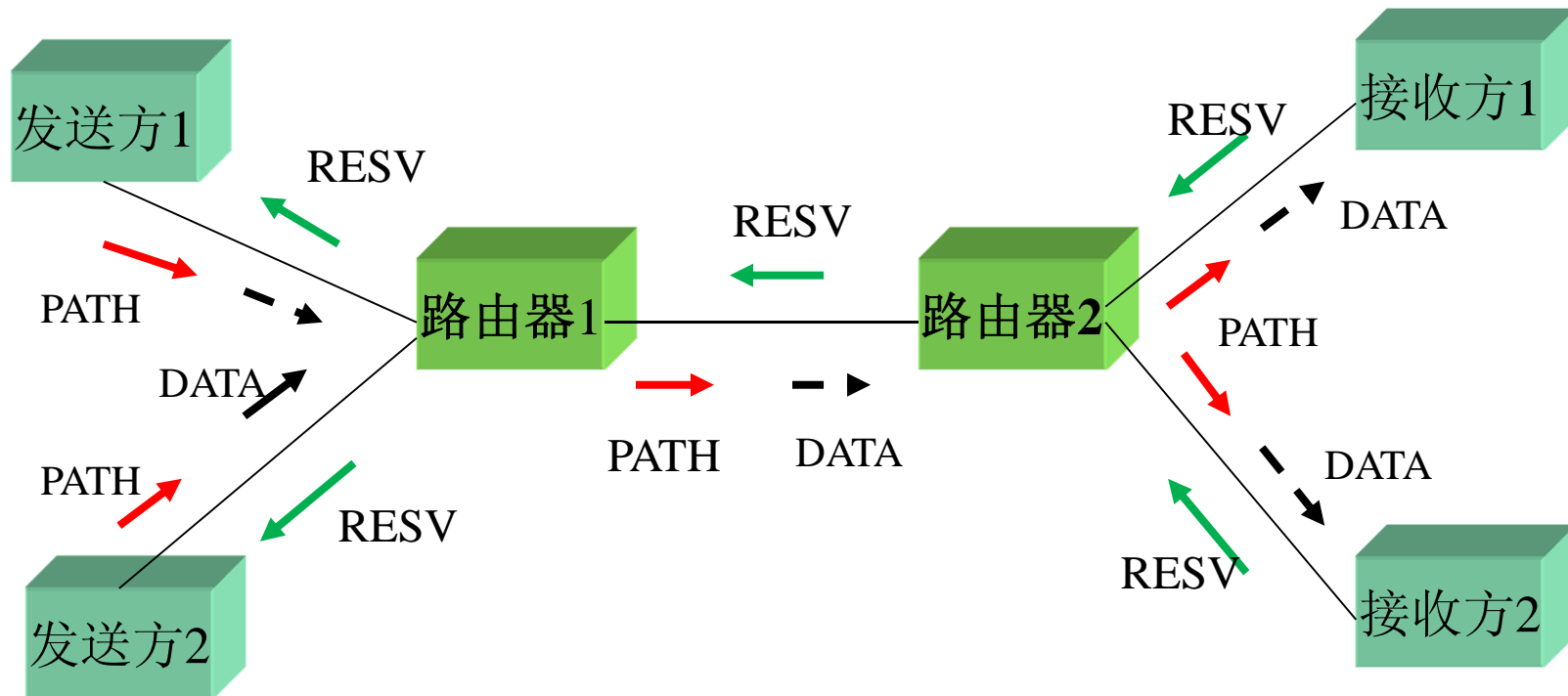
- 必选消息：PATH和RESV；
- 可选消息：PATH ERROR、PATH TEARDOWN、RESV ERROR、RESV CONFIRM和RESV TEARDOWN。

RSVP通常直接运行于IP之上，RSVP路由器和客户使用它们来创建和维护预留状态（软状态），需要定期刷新；



RSVP操作

■ RSVP的预留建立过程





RSVP的适用范围

- 适合于点对点介质（串行线路），不太适合共享介质环境。
 - 共享介质环境下，无法确保网段内其它人不发送过量负载，当网段上负载很小或源数量很少时，也可以很好地工作；为此开发了基于IEEE802网络上的扩展SBM；
 - 在ATM上支持RSVP，可以采用RSVP为ATM网络上的每个预留都创建一个可变比特率（VBR）的SVC，然后将预留通信都重定向到SVC。



RSVP的适用范围

- RSVP的状态信息量随着流预留数目的增加而增长，而流粒度的状态信息不易扩展。
 - 中等企业内部网络（小于或等于DS3）的RSVP扩展性能良好；
 - 大型内部网络或ISP主干网，可以通过在网络边界使用大型多播组、大型静态类或流聚集（不是单流预留）来扩展；RSVP的预留聚集是指将多个入口和出口路由器相同的端到端预留聚集成一个大型的端到端预留；
 - 在网络边界使用RSVP，在网络主干使用差分业务模型；
 - 一个路由器上既可以支持差分业务，也可以同时支持综合业务。
 - 充裕带宽解决扩展问题，不用为每个呼叫分配资源，只是占总容量的一部分。



移动环境下的IntServ和RSVP

- 目前的RSVP不适合移动IP网络，主要原因是它无法感知主机的移动，因此无法在移动主机即将访问的位置上提前预留资源，为了克服RSVP的缺陷，提出了以下四个方案：
 - MRSVP：本协议要求预测主机未来可能到达的位置，并在这些位置上提前预留资源，包括主动/被动两种资源预留方式。
 - 支持隧道的RSVP：本协议在隧道的两个端点之间为通过隧道的总业务量预留资源。
 - 基于组播的RSVP：在本协议中，每个移动主机用一个组播地址来标识，所有移动主机在发送、接收RSVP消息和IP数据包时都进行组播，主机的移动可视为组播组成员的变动，还采用了运动预测机制，另一方面，用剪枝的方法将主机已经离开的位置从树上删除。
 - DRSVP：是一种支持可变服务质量的动态资源预留协议，用户能根据网络资源的变化动态调整服务质量要求，在RESV消息中增加描述业务流的适应范围，报告上下游的资源瓶颈，引入新的带宽分配算法，允许预留的资源有一个波动范围，提高服务质量和资源利用率。
 - 基于RSVP的移动IPv6服务质量体系：将RSVP隧道和移动IPv6协议结合，实现移动主机在当前位置的路径上建立和维持预留资源。



综合业务模型的特点

- 要求端到端的信令支持，沿途的每个路由器为每个请求预留资源。
- 路由器的数据状态增大了路由器的负担，网络扩展性不好，如果业务流增加，路由器的负担会增大。
- 有些终端支持端到端的RSVP，有些则不支持RSVP。
- 必须在发送者和接收者之间的所有路由器都支持RSVP，中间有不支持的路由器，就会使网络失去真正意义的资源预留。
- 支持端到端行为，是复杂的全程端到端信令过程。
- 用户认证、优先权管理需要烦琐的上层策略机制来管理。



差分业务模型

- 概述
- 差分业务模型的体系结构
- DS编码与业务类别
- 差分业务模型(小结)



差分业务模型

■ 概述

- 在IP报头中定义区分服务字段和服务类型；
- 分组在每一个网络节点可以得到特定的转发处理或每一跳行为（PHB）；
- 合理定义小型基本部件；
- 不支持端到端的QoS。



差分业务模型的体系结构

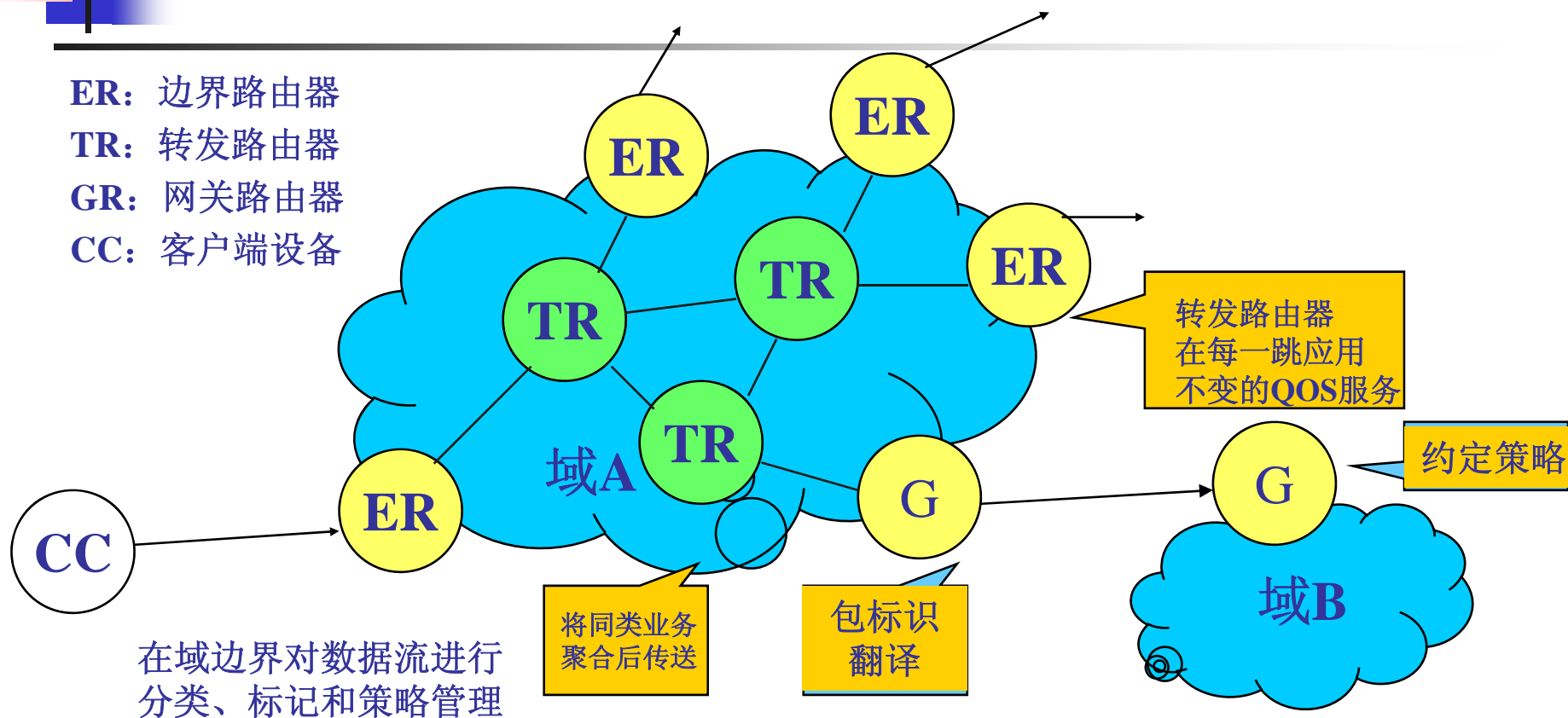
■ 域的概念:

差分域 (Differentiated Services

Domain) 是指Internet的一部分, 在本部分

中具有一致的业务分类策略和管理机制。

差分业务模型的体系结构





差分业务模型的体系结构

- 服务提供商可以为客户提供各种网络服务，客户只需将分组的区分编码点（DSCP）标记为特定值，即可以定义分组在服务提供商网络中的每一跳行为（PHB）；服务提供商和客户可以共同协商一个配置文件（Profile），用来描述各种服务等级通信的提交速率，如果提交速率超过了规定，则无法提供规定的服务等级。
- 业务网接口处对用户分级，分级基于数据包的不同标识，同一级别的业务在该网络域中汇聚统一传送，还负责调节进入域的通信量（输入接口实现分组分类和流量控制）
- 差分业务模型基于业务流分级，包括流的聚合以及适于不同业务类别的PHB，即定义一些重要的分组传输特征（一条路径某个方向上的吞吐量、延迟、抖动以及分组丢失率、访问网络资源的相对优先级等），网络上的所有节点都按照PHB提供网络性能。



差分业务模型的体系结构

- 在域之间，对不同类别业务，其QoS都有一定的约定和包标识的翻译机制。
- 差分业务模型不提供从发送者到接收者的端到端的QoS，只在域的范围内容保证与业务分类相对应的QoS。



DS编码与业务类别

- 差分业务模型的DS字段含义
 - 采用IPv4头中的服务类型 (ToS) 字段 (8bits)
 - 采用IPV6头中的业务类型 (Traffic Class) (8bits)

DS5	DS4	DS3	DS2	DS1	DS0	CU	CU
-----	-----	-----	-----	-----	-----	----	----

DSCP: 区分服务的编码点 (DS5-DS0)

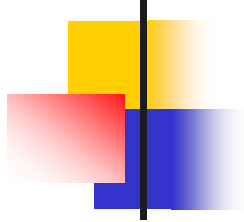
CU: 未用 (2位)



DS编码与业务类别

■ IETF工作组定义的DSCP编码

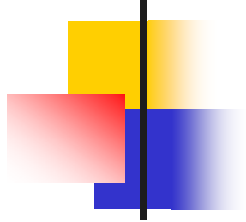
- 默认DSCP：定义为000 000；
- 类选择器：定义为向后与IP优先级兼容；
- 加速转发（EF）：定义了奖赏服务，推荐101110；
- 确定转发（AF）：定义了4个服务等级，每个服务等级有三个分组丢失优先级；



DS编码与业务类别

类选择器DSCP

类选择器	DSCP
优先级1	001000
优先级2	010000
优先级3	011000
优先级4	100000
优先级5	101000
优先级6	110000
优先级7	111000



DS编码与业务类别

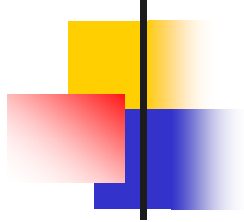
AF的编码点

丢失优先级	级别1	级别2	级别3	级别4
低	001010	010010	011010	100010
中	001100	010100	011100	100100
高	001110	010110	011110	100110



差分业务模型的特点

- 简单的QoS保证机制。
- 不提供端到端的QoS。
- 网络设备的负担小。
- 综合业务模型与差分业务模型的结合使用。



IP的QoS技术实施

- 网络边界流量控制器
- 每一跳行为 (PHB)



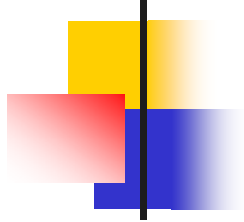
网络边界流量控制器

- 网络边界的流量整形可以提供以下功能：
 - 分组的分类
 - 分组标记
 - 通信速率管理



分组的分类

- 分组分类是一种根据分组的一个或多个字段标识分组所属的类的手段，标识功能可以很简单，也可以很复杂，可以支持以下分类类型：
 - 基于5个流参数标识IP流：即源IP地址、目标IP地址、IP协议字段和源端口、目标端口；
 - 基于分组的优先级或DS字段标识；
 - 基于其它TCP报头，例如分组长度标识；
 - 基于源和目的MAC地址标识；
 - 基于端口号、Web通用资源定位地址进行应用标识；
(即基于应用程序分类)
 - 根据路由器内部的信息分类，例如可以根据分组到达的输入端口或内部分组数据结构中的QoS字段标识。



分组标记

- 分组标记的方式很多，可以采用以下方式标记：
 - IP优先级
 - DSCP
 - QoS组

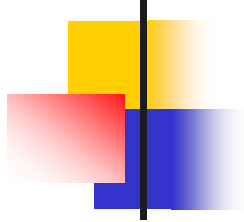
分组标记

■ IP优先级

- 分组的IP报头中的IP优先级字段用来指示分组处理的相对优先级，由分组报头中的ToS字段中的3位组成；
- 除了优先级字段之外，ToS字节还包含了ToS位，ToS用来包含指示在网络中应该如何处理分组的值，目前很少用。
- 默认情况下，所有路由控制通信均使用优先级6，优先级7也保留给网络控制通信使用，故不推荐使用6和7。

IP优先级 (P2-P0)
服务类型 (DTR)

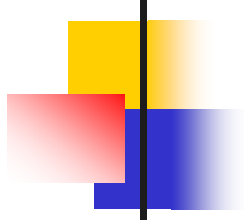
P2	P1	P0	D	T	R	CU	CU
----	----	----	---	---	---	----	----



分组标记

■ IP优先级

IP优先级	IP优先级位	IP优先级名称
0	000	一般(路由)
1	001	优先
2	010	立即
3	011	瞬间
4	100	最优先
5	101	关键
6	110	互联控制
7	111	网络控制



分组标记

■ IP优先级

- 设置IP优先级字段描述分组的工作可以由发起通信的应用程序完成，也可以由网络中的节点完成。
- 支持本功能的机制包括承诺的接入速率（CAR）、基于策略的路由（PBR）以及边界网关控制的QoS策略传播（QPPB）。



分组标记

■ DSCP

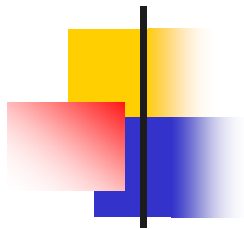
- 用来在网络中指示特定的PHB，由IP分组报头中的6位组成，IETF标准化后的DSCP是原先ToS字节的6位。
- DSCP字段是IP优先级字段的超集，其使用和设置与IP优先级字段类似；
- DSCP字段的定义向后与IP优先级字段兼容。



分组标记

■ QoS组

- QoS组是路由器内部分组数据结构中的一个字段，用于标记与用户定义的分类标准相匹配的分组，它是路由器内部的标签，不是IP分组报头的组成部分。
- 分组到达网络边界时，其IP优先级或DSCP字段已经设置好，即使分组进入网络之前已经被标记了，网络运营商还是想在分组进入网络之前，对其所属的类及其提供的服务等级，在网络边界强制进行适当的标记。



分组标记

■ 三种标记方式的比较

属性	IP优先级	DSCP	QOS组
分类范围	整个网络，在IP分组报头内运载	整个网络，在IP分组报头内运载	路由器内部，不在IP分组中运载
类数	8 (0-7)	64 (0-63)	100 (0-99)

使用IP优先级、DSCP和QOS标记通信



通信速率管理

- 即使网络边界的少数路由器发送的通信流量超过网络核心规定的流量，通信负载的增加都会导致网络拥塞，网络性能的下降导致无法为所有的通信提供QOS。
- 在网络中提供 QOS，必须在网络边界路由器上对进入服务提供商网络的通信流量进行控制，以确保通信速率在服务限制范围内。
- 通信速率管理可以采用**承诺的接入速率 (CAR)**和**流量整形 (TS)**功能进行。



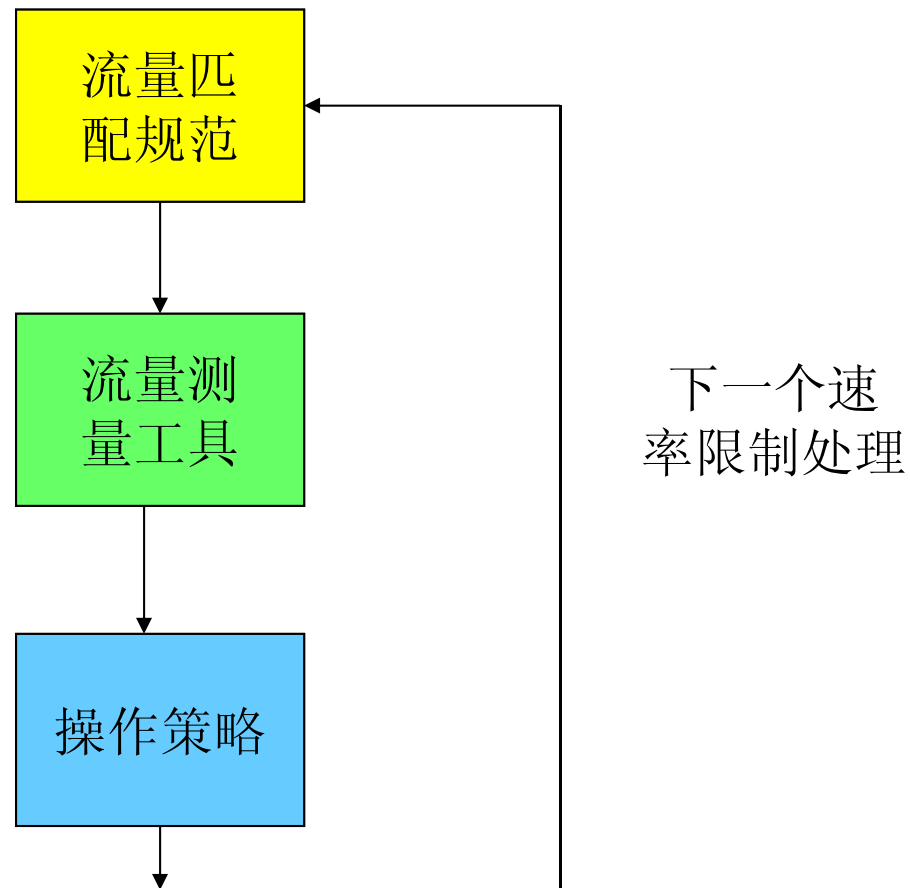
通信速率管理

■ 令牌桶方案

- 用来报告分组是否与为之配置的速率参数匹配;
- 四个关键参数:
 - 平均速率或承诺信息速率 (CIR)
 - 顺从突发量 (Bc): 瞬间可以超过令牌桶的流量, 也称为常规突发量;
 - 扩展突发量 (Be): 允许顺从突发量和扩展突发量之间有一定差别。
 - 时间间隔T: $T = Bc / CIR$
- 令牌桶能保存的最大令牌数等于Bc, 对于应用了令牌限制的分组, 将从令牌桶中删除与分组的字节数相同的令牌数, 如果令牌足够, 则分组将被传输, 否则:
 - 对于没有扩展功能的令牌桶: 丢弃分组;
 - 对于有扩展功能的令牌桶, 则允许流暂借更多的令牌, 直到超过扩展突发量为止; 超过Be的分组将被丢弃。

通信速率管理

■ 流量控制





通信速率管理

■ 通信匹配规范

- 用来定义哪些分组是CAR感兴趣的，每个速率限制语句将依次查找匹配，可以定义匹配规范来匹配每个分组。
- 匹配规范的4种方式：
 - 匹配所有通信；
 - 使用速率限制访问列表匹配某个IP优先级值；
 - 使用速率限制访问列表匹配某个MAC地址；
 - 使用IP标准或扩展访问列表进行匹配。

■ 流量测量工具

- 令牌桶是流量测量的手段。



通信速率管理

■ 操作策略

- 可以为每个速率限制语句单独定义顺从和超出通信的操作策略。
- 顺从和超出操作可以是以下任何一种操作：
 - 传输；
 - 丢弃；
 - 继续（转到列表的下一条速率限制语句）；
 - 设置优先级并传输；
 - 设置优先级并继续；
 - 设置QOS组并传输；
 - 设置QOS组并继续；



通信速率管理

■ 控制功能和整形功能的差别

控制功能 (CAR)	流量整形功能 (TS)
以高达线路速率的速度发送通信，允许突发	平缓通信，将数据以恒定速率发送出去
令牌耗尽时丢弃分组	令牌耗尽时，将分组存放在缓存区，等到有令牌时再发送
可以用于输入和输出通信	只能用于输出通信
TCP检测到线路速率，当分组被丢弃时，通过减小窗口尺寸调整到配置的速率。	TCP发现线路速率较低，从而调整重传计时器，使重传更少，对TCP友好。



每一跳行为 (PHB)

- 资源分配 (QoS调度机制)
- 拥塞避免和分组丢弃策略



每一跳行为 (PHB)

- 资源分配 (QoS调度机制)

- 支持QoS的调度技术
- 不同的分组调度/排队策略



支持QoS的调度技术

■ 对调度技术的要求

- 网络不拥塞时，所有的调度方案都有效，因为路由器上没有队列，当网络拥塞时，路由器上将建立队列，路由器给流分配资源的方式取决于排队系统中排队分组的调度规则，调度规则决定接下来传输队列中的哪个分组，流的分组得到服务的频率决定了流的带宽（资源分配）；
- 调度技术只涉及接下来哪个分组将得到服务的问题；
- 提供QoS的调度算法至少应该能区分队列中不同的分组，并指导每个分组的服务等级；
- 调度算法应能通过给流分配资源和指定优先级来保证性能；
- 完成的粒度为单个流或者来自不同流的分组组成的通信类；
- 需要给优先级相同的流提供公平和保护调度机制；
- 易于实现、对要求资源保证的流易于实现许可控制；
- 排队理论中的守恒定律：任何调度规则只能通过增加其他流的延迟来降低某个流的平均延迟。



支持QoS的调度技术

- FIFO排队
- 最大-最小公平-份额分配方案
- 广义处理器共享



支持QoS的调度技术

■ FIFO排队

- 目前路由器上最常见的调度机制；
- 没有区分流的机制，不能为具有相同优先级的流提供公平和保护；
- 使用端到端自适应流控制方案（TCP动态窗口控制）
- 流获得服务的频率与它向网络发送数据的速率成正比。



支持QoS的调度技术

■ 最大-最小公平-份额分配方案

- 根据不同用户对资源需求的不同，可以按照对资源的需求量从小到大的顺序对他们进行分类：
 - 按需求量从小到大的顺序分配资源；
 - 任何用户得到的资源份额都不会超过其需求的量；
 - 需求未被满足的用户的资源份额相同。
- 可以按照公式计算公平分配：
 - $\text{公平分配} = (\text{资源总量} - \text{已经分配给用户的资源总量}) / (\text{还需要资源的用户数})$
- 所有需求未被满足的用户将获得同样的分配份额；
- 扩展方案：每个用户指定一个权值，则用户的公平份额与其权值成正比。（加权最大-最小公平-份额分配方案）



支持QOS的调度技术

■ 广义处理器共享 (GPS)

- 对权值相同的流提供最大-最小公平份额的理想机制；
- 将每个流都放在各自的逻辑队列中，依次为每个非空队列传输无穷小的数据量，每一轮都只传输无穷小的数据量，因此，对所有非空队列都是公平的；
- 如果为每个流都指定权重，则GPS每一轮传输的数据量都与其权重成正比，从而实现了加权最大-最小公平份额；
- 是一种理想的最大-最小公平份额模型，却无法实现，实际的调度算法是接近GPS的模型。



不同的分组调度/排队策略

- FIFO
- FQ
- WFQ
 - 基于序列号计算的WFQ
 - 基于流的WFQ
 - 基于类的WFQ
 - 基于ToS/QoS组的DWFQ
- 优先级排队
- 定制排队



每一跳行为 (PHB)

■ 拥塞避免和分组丢弃策略

- TCP慢启动和拥塞避免
- 随机早期侦测 (RED)
- 加权随机早期侦测 (WRED)
- 流WRED
- 显式拥塞通知 (ECN)
- 选择性分组丢弃 (SPD)



拥塞避免和分组丢弃策略

■ TCP慢启动和拥塞避免

- TCP使用拥塞窗口避免拥塞，拥塞窗口指出了发送方在收到确认前可以通过TCP会话传输的最大数据量；
- TCP源的慢启动算法以收到对方确认的速率将分组发送到网络，因此TCP是自同步的；
- TCP的分组丢失率是拥塞的预示，重传计数器超时未受到确认时，就认为发生了拥塞；这时，将拥塞窗口重置为1，重新启动慢启动算法，同时将慢启动阈值减少到需要重传时的拥塞窗口大小的一半，会话建立后，阈值设为对方声明接收窗口值。
- 如果分组丢失不是网络拥塞引起的，则等待会影响性能，快速重传和恢复算法避免了这种影响。



拥塞避免和分组丢弃策略

■ TCP通信行为的尾丢弃方案

■ 尾丢弃方案

- 一般来说，当队列达到最大长度时，到达队列的分组将被丢弃，丢弃行为一直持续到队列由于分组被传输而减小，这种技术称为尾丢弃；
- 丢弃分组告诉TCP源发生了拥塞，只有队列充满时才会丢弃，启动TCP的缓慢启动算法，源的流量减少；
- 尾丢弃方案导致的大量TCP会话的分组丢失，引发网络流量瞬时下降，当源持续增大窗口时，再次导致拥塞和分组丢弃；这种队列长度呈波浪形变化的恶性循环称为全局同步现象；
- 全局同步导致通信的延迟抖动非常大，降低了网络的总吞吐量。



拥塞避免和分组丢弃策略

■ RED（随机早期侦测）

- 尾丢弃引起的TCP源通信行为需要在队列充满之前发出拥塞通知，并控制队列长度，迅速降低排队延迟；
- Sally Floyd和Van Jacobson提出了RED，它是一种积极的队列管理技术，旨在降低队列长度，缩短排队延迟，提供比传统尾丢弃方式更好的性能；
- 采用预防拥塞的方法，当平均队列长度超过最小门限值后，则开始以非零的丢弃几率丢弃分组，而不是等队列充满时。这样，避免了全局同步；
- 丢弃分组向TCP发出缓慢信号，使TCP进入慢启动模式；
- 如果随机丢弃时，平均队列长度继续增大，则丢弃几率线性增大，以便控制平均队列长度，在队列长度从最小门限值增大到最大门限值的過程中，丢弃率线性增长，当平均队列长度超过最大门限值时，将丢弃所有到来的分组。



拥塞避免和分组丢弃策略

■ WRED

- 在RED基础上改进，引入了根据分组丢弃几率为分组提供服务等级的功能，允许使用基于IP优先级的选择性RED参数，即WRED会更多地丢弃特定优先级的分组；
- 默认情形下，所有优先级的最大门限值都相同，最小门限值随着分组的优先级而不同，因此低优先级的分组比高优先级的分组丢弃的可能性更大。对优先级为0的通信，默认最小门限值是最大门限值的一半；
- 可以使用基于通信类的WRED，允许每个通信类具有不同的RED参数；
- 可以基于硬件或软件实现。



拥塞避免和分组丢弃策略

■ 显式拥塞通知 (ECN)

- ECN为路由器提供了向TCP源发出拥塞信号的功能，通过标记分组报头实现；
- 需要IP报头的CE位和支持ECN的传输协议支持：
 - IP报头中ToS字节的第6、7位组成，称为ECT (ECN功能传输) 和CE位 (经历拥塞)，TCP源设置ECT指示传输协议的端点支持ECN功能，路由器设置CE位，向端点指示发生了拥塞；
 - 差分业务模型中，第6、7位未使用；
- TCP需要支持以下ECN功能：
 - ECN功能协商：建立TCP连接时协商是否具备本功能；
 - ECN回应标记：TCP接收方使用本标记通知TCP源，它已经收到CE分组；
 - 拥塞窗口缩小标记：TCP源使用本标记通知TCP接收方，拥塞窗口已经被缩小。



拥塞避免和分组丢弃策略

■ SPD（选择性分组丢弃）

- 用来使路由器区分重要控制通信（如路由器协议分组）和常规数据通信，将路由器分组排到常规通信之前，确保路由协议分组顺利通过；
- 在IP处理队列中实现，只应用于处理交换通信；
- 可以运行于以下模式：
 - 禁用：禁止路由器上的SPD；
 - 常规：IP输入队列的长度小于最小队列门限值，则不丢弃分组。



拥塞避免和分组丢弃策略

■ SPD（选择性分组丢弃）

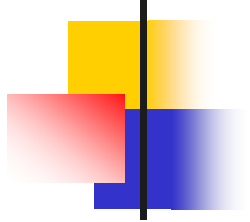
■ 可以运行于以下模式：

- 随机丢弃：也称为SPD刷新，重要的IP控制通信仍将被加入到队列中。
 - 当IP输入队列的长度大于最小门限值，但小于最大门限值，分组可能被丢弃，其丢弃几率 = $(\text{队列长度} - \text{最小门限值}) / (\text{最大门限值} - \text{最小门限值})$
- 完全丢弃：IP输入队列大于最大门限值，所有的常规IP分组都将被丢弃，重要的IP控制信息则仍然被接收到一个叫做优先队列的特殊处理等级队列中，该队列优先于常规队列获得服务；
- 主动丢弃：是一种针对没有通过健全检测的IP分组的特殊主动丢弃模式，当输入队列的长度大于最小门限值时，所有损坏的IP分组都将被丢弃。



IP的QoS技术及其发展（小结）

- QoS技术概述
- IP的QoS技术
 - IP的业务模型
 - 综合业务模型
 - 差分业务模型
 - IP的QoS技术实施
 - 网络边界流量控制器（分组分类/标记、速率管理）
 - 每一跳行为（PHB）
 - 资源分配（分组调度机制）
 - 拥塞避免和分组丢弃（TCP慢启动和拥塞避免、随机早期侦测（RED）、加权随机早期侦测（WRED）、流WRED、显式拥塞通知（ECN）、选择性分组丢弃（SPD））
 - MPLS的QoS技术
 - IP的业务模型在MPLS中的实现（综合业务模型在MPLS中的实现、差分业务模型在MPLS中的实现）



思考题

- 什么叫QoS?
- IP的QoS采用哪些衡量QoS的性能参数?
- IP网络中的业务模型有哪两个? 其各自的特点如何?
- 举例说明几种IP的QoS机制及其工作原理。