

计算机网络服务质量QoS研究

1023040806侯震

May 2024

摘要

计算机网络的服务质量（QoS）是当今国际网络研究领域最重要、最富有魅力的研究方向之一，是计算机网络研究与开发的热点，被称为新一代计算机网络的核心问题之一。本文为5部分第一部分是QoS的产生背景及定义；第二部分是QoS的应用，阐述了QoS当代的主要应用场景；第三部分是QoS的模型，对QoS三种主要的模型进行了简介；第四部分是QoS的原理，对QoS流量分类，流量监管，拥塞管理等方面做了简要介绍；第五部分是其他QoS重要技术的简要说明。

关键词

QoS，应用场景，流量管理，拥塞控制，增强技术等

1 QoS的产生背景及定义

1.1 产生背景

网络的普及和业务的多样化使得互联网流量激增，从而产生网络拥塞，增加转发时延，严重时还会产生丢包，导致业务质量下降甚至不可用。

- **丢失数据包**- 当数据包到达一个缓冲器（buffer）已满的路由器时，则代表此次的发送失败，路由器会依网络的状况决定要丢弃、不丢弃一部份或者是所有的数据包，而且这不可能在预先就知道，接收端的应用程序在这时必须请求重新传送，而这同时可能造成总体传输严重的延迟。

- **延迟**- 或许需要很长时间才能将数据包传送到终点，因为它会被漫长的队列迟滞，或需要运用间接路由以避免阻塞；也许能找到快速、直接的路由。总之，延迟非常难以预料。

所以，要在网络上开展这些实时性业务，就必须解决网络拥塞问题。解决网络拥塞的最好的办法是增加网络的带宽，但从运营、维护的成本考虑，这是不现实的，最有效的解决方案就是应用一个“有保证”的策略对网络流量进行管理。

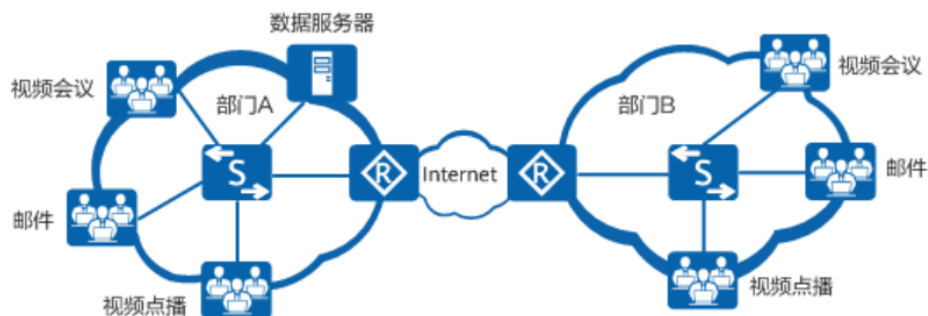


图2 QoS企业办公例图

1.2 定义及概述

QoS技术就是在前面这种背景下发展起来的。QoS (Quality of Service) 即服务质量，顾名思义，服务质量往往体现了消费者对服务者所提供服务的满意程度。在计算机系统，特别是计算机网络系统，作为计算和信息等服务的提供者，同样存在服务质量 (QoS) 优劣的问题。从计算机系统诞生伊始，人们就一直孜孜不倦地致力于提高系统的服务性能和服务质量。目前，关于QoS一词的定义有多种形式。RFC2386中，将QoS定义描述为：网络在传输数据流时要求满足的一系列服务请求，具体可以量化为带宽、延迟、延迟抖动、丢失率、吞吐量等性能指标。QoS反映了网络在保证信息传输和满足服务要求方面的能力。

简单的说，服务质量(QoS)是一组服务要求，网络必须满足这些要求才能确保数据传输达到适当服务的级别，数据在网络中传输的过程中，QoS机制通过制包括流量分类，访问控制，流量整形，拥塞控制与管理，调度机制等。QoS的提供增加网络的性能可预知性，网络资源能够有效地分配，提高网络的吞吐量，减少网络丢弃数据包尤其是较重要的数据包的概率。

2 QoS的应用场景

以企业办公为例，除了基本的网页浏览、工作邮件外，在较集中的工作时间内还需要保证Telnet登录设备、异地的视频会议、实时语音通话、FTP文件的上传和下载，以及视频播放等业务的网络质量。对于不同网络质量要求的业务，可以配置不同的QoS子功能，或者不部署QoS。如下图所示：

2.1 网络协议和管理协议（如OSPF、Telnet）

这类业务要求低时延和低丢包率，但对带宽的要求不高。因此可以通过QoS的优先级映射功能，为此类报文标记较高的服务等级，使网络设备优先转发此类报文。

2.2 实时业务（如视频会议、VoIP）

视频会议要求高带宽、低时延和低抖动。因此可以通过QoS的流量监管功能，为视频报文提供高带宽；通过QoS的优先级映射功能，适当调高视频报文的优先级。

VoIP是指通过IP网络进行实时语音通话，它要求网络做到低丢包、低时延和低抖动，否则通话双方可以明显感知质量受损。因此一方面可以调整语音报文的优先级，使其高于视频报文；另一方面通过流量监管功能，为语音报文提供最大带宽，在网络产生拥塞时，可以保证语音报文优先通过。

2.3 大数据量业务（如FTP、数据库备份、文件转储）

大数据量业务是指存在长时间大量数据传输行为的网络业务，这类业务需要尽可能低的网络丢包率。因此可以为这类报文配置流量整形功能，通过数据缓冲区缓存从接口发送的报文，减少由于突发流量导致拥塞而产生的丢包现象。

2.4 流媒体（如在线音频播放、视频点播）

由于这些音视频节目都是提前制作好的，观看者的终端通常可以先进行缓存再进行播放，因此降低了对网络时延、丢包和抖动的要求。如果需要降低这类业务的丢包和时延，可以通过QoS的优先级映射功能，适当提高对应报文的优先级。

2.5 普通业务（如HTML网页浏览、邮件）

这类业务对网络无特殊要求、重要性也不高。管理员可以对其保持默认设置，不需要额外部署QoS功能。

3 QoS模型

最初的Internet只是面向非实时的、单数据类型的通信（如FTP、Email设计的，所有资源共享，不关注QoS的控制和管理，不保证网络按照用户所需的质量要求传递发送数据，这种类型的服务无法控制分组的传输时延、吞吐量等QoS参数，也不会给数据提供QoS保证。随着服务质量QoS概念的提出，不同国际标准化组织相继提出不同的控制机制和策略，较为典型的有因特网工程任务组(IETF)提出的集成服务体系模型(IntServ)、服务体系模型(DiffServ) 尽力而为（Best-Effort）服务模型。

3.1 IntServ模型

IntServ（Integrated Service，综合服务）模型由RFC1633定义，在这种模型中，节点在发送报文前，需要向网络申请资源预留，确保网络能够满足数据流的特定服务要求。

IntServ可以提供保证服务和负载控制服务两种服务，保证服务提供保证的延迟和带宽来满足应用程序的要求；负载控制服务保证即使在网络过载的情况下，也能对报文提供与网络未过载时类似的服务。

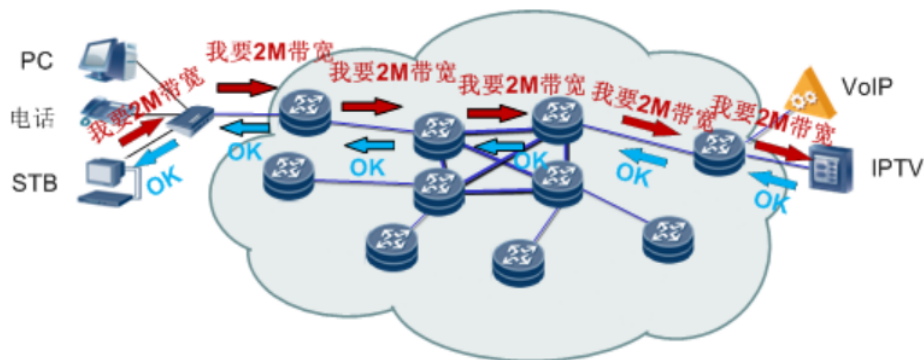


图3.1 IntServ模型

在IntServ模型中，网络资源的申请是通过信令来完成的，应用程序首先通知网络它自己的流量参数和需要的特定服务质量请求，包括带宽、时延等，应用程序一般在收到网络的确认信息，即确认网络已经为这个应用程序的报文预留了资源后，才开始发送报文。同时应用程序发出的报文应该控制在流量参数描述的范围以内。负责完成保证服务的信令为RSVP（Resource Reservation Protocol，资源预留协议），它通知网络设备应用程序的QoS需求。RSVP是在应用程序开始发送报文之前来为该应用申请网络资源的，所以是带外信令。

保证服务要求为单个流预先保留所有连接路径上的网络资源，而当前在Internet主干网络上有成千上万条应用流，保证服务如果要为每一条流提供QoS服务就变得不可想象了。因此，IntServ模型很难独立应用于大规模的网络，目前主要与MPLS TE（Traffic Engineering，流量工程）结合使用。IntServ模型如下图所示：

3.2 DiffServ模型

IntServ模型是指应用程序在发送报文前，需要通过信令（signaling）向网络描述它的流量参数，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。在收到确认信息，确定网络已经为这个应用程序的报文预留了资源后，应用程序才开始发送报文。应用程序发送的报文应该控制在流量参数描述的范围内。网络节点需要为每个流维护一个状态，并基于这个状态执行相应的QoS动作，来满足对应用程序的承诺。

IntServ模型使用RSVP（Resource Reservation Protocol）作为信令，类似于ATM SVC的方式，是面向连接的。RSVP工作在传输层，但不参与应用数据的传送，是一种网络上的控制协议（类似于ICMP），用于在网络节点间传递资源预留信息以及建立资源预留。

在通过RSVP建立端到端通信的过程中，沿途的各路由器（包括核心路由器）必须为每个要求服务质量保证的数据流保存状态信息，称为“软状态”。

“软状态”是一种临时性状态，被定期的RSVP信息更新。通过RSVP信息的预留，各路由器可以判断是否有足够的资源可以预留。只有所有的路由器都给RSVP提供了足够的资源，“路径”方可建立。DiffServ模型如下图所示：

与Intserv模型相比，DiffServ模型不需要信令。在DiffServ模型中，应用程序

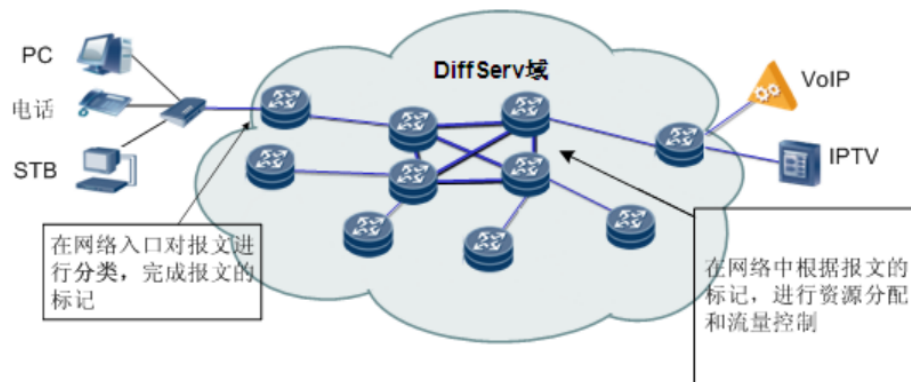


图3.2 DiffServ模型

发出报文前，不需要预先向网络提出资源申请，而是通过设置IP报文头部的QoS参数信息，来告知网络节点它的QoS需求。网络不需要为每个流维护状态，而是根据每个报文流指定的QoS参数信息来提供服务，对报文的服务等级划分，有差别地进行流量控制和转发，提供端到端的QoS保证。

DiffServ模型充分考虑了IP网络本身灵活性、可扩展性强的特点，将复杂的服务质量保证通过报文自身携带的信息转换为单跳行为，从而大大减少了信令的工作。因此，DiffServ模型不但适合运营商环境使用，而且也大大加快了IP QoS在实际网络中应用的进程。

3.3 Best-Effort服务模型

Best-Effort是最简单的QoS服务模型，应用程序可以在任何时候，发出任意数量的报文，而且不需要通知网络。对Best-Effort服务，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服务模型适用于对时延、可靠性等性能要求不高的业务质量保证。Best-Effort是现在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。

4 QoS基本原理

4.1 流量分类与标记

流量分类，就是将流量划分为多个优先级或多个服务类，如使用以太网帧中802.1Q头保留的User Priority(用户优先级)字段标记服务级别，可以将以太网帧最多分成 $2^3 = 8$ 类；使用IP报文头的ToS (Type of service, 服务类型) 字段的前三位 (即IP优先级) 来标记报文，可以将报文最多分成 $2^3 = 8$ 类；使用DSCP (Differentiated Services Codepoint, 区分服务编码点, ToS域的前6位)，则最多可分成 $2^6 = 64$ 类。在报文分类后，就可以将其它的QoS特性应用到不同的分类，实现基于类的拥塞管理、流量整形等。

对于MPLS网络报文，则一般是根据MPLS报文中的EXP域进行处理。EXP域

包括3位，虽然RFC 3032把它叫做实验域，但它通常作为MPLS报文的CoS域，与IP网络的ToS或DSCP域等效。

对于流量的分类，上面提到的关于以太网帧的Cos域、IP报文的ToS域等与MPLS报文的EXP域等仅是分类的一种情况，其实几乎可以对报文的任何信息段进行分类，比如也可以根据源IP地址、目的IP地址、源端口号、目的端口号、协议ID等进行流量的分类。

虽然流量分类几乎可以根据报文的任何字段进行，但是流量分类标记则一般只对802.1Q 以太网帧的CoS域、IP报文的ToS域、MPLS报文的EXP域进行标记。

流量的标记主要的目的就是让其他处理此报文的应用系统或设备知道该报文的类别，并根据这种类别对报文进行一些事先约定了的处理。

例如，在网络的边界做如下分类和标记：1) 所有VoIP数据报文聚合为EF业务类，将报文的IP优先级标记为5，或者将DSCP值标记为EF；

2) 所有VoIP控制报文聚合AF业务类，将报文的IP优先级标记为4，或者将DSCP值标记为AF31。

当报文在网络边界被标记分类之后，在网络的中间节点，就可以根据标记，对不同类别的流量给予差别服务了。例如：对上述例子中的EF类业务保证时延和减少抖动，同时进行流量监管；对AF业务类在网络拥塞时仍然保证一定的带宽，等等。

4.2 流量监管和流量整形

4.2.1 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。通过监控进入网络的某一流量的规格，限制它在一个允许的范围之内，若某个连接的报文流量过大，就丢弃报文，或重新设置该报文的优先级（比如限制HTTP报文不能占用超过50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) 等流量参数，对超出SLA约定的流量报文可指定给予pass (通过)、drop (直接丢弃) 或markdown (降级) 等处理，此处降级是指降低服务等级 (Service Class)，或者是提高丢弃等级 (Color)，即报文在网络拥塞时将被优先丢弃，从而保证在SLA约定范围之内的报文享受到SLA预定的服务。

转发 (pass)：对测量结果不超过CIR的报文通常处理为继续正常转发。

丢弃 (discard)：对测量结果超过PIR的报文通常进行丢弃。

重标记 (remark)：对处于承诺速率 (CIR) 与峰值速率 (PIR) 之间的流量通常执行Remark动作，此时的报文不丢弃，而是通过Remark降低优先级进行尽力而为转发。

流量监管采用承诺访问速率CAR (Committed Access Rate) 来对流量进行控制。CAR使用令牌桶算法进行流量速率的评估，依据评估结果，实施预先设定好的监管动作。

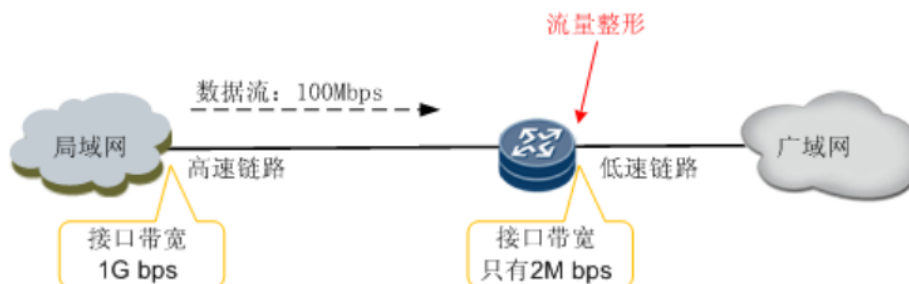


图4.2.2.1流量整形

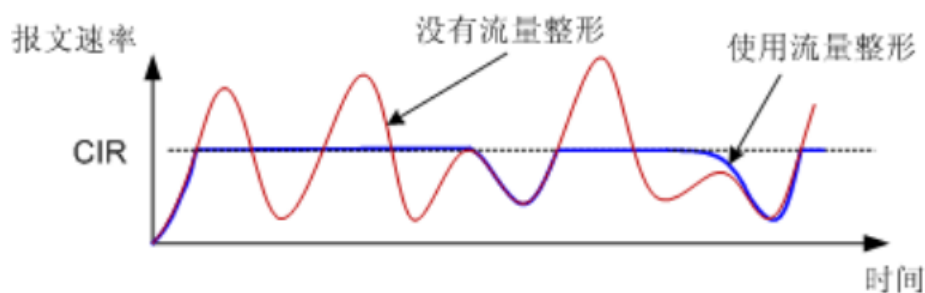


图4.2.2.2 流量整形的效果

4.2.2 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。流量整形通常是为了使报文速率与下游设备相匹配。当从高速链路向低速链路传输数据，或发生突发流量时，带宽会在低速链路出口处出现瓶颈，导致数据丢失严重。这种情况下，需要在进入高速链路的设备出口处进行流量整形。如下图所示：通过在上游设备的接口出方向配置流量整形，将上游不规整的流量进行削峰填谷，输出一条比较平整的流量（如下图所示）解决下游设备的瞬时拥塞问题。

4.2.3 两者比较

相同点：

流量监管和流量整形的相同点主要有：

- 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。

差异点：（如下表所示）

表4.2.3.1 流量监管和流量整形的比较

流量监管	流量整形
丢弃超额流量或将超额流量重标记为低优先级量	缓存超出策略/协定SLA规定的超额流量
不需要额外的内存资源，不会带来延迟和抖动	需要内存去缓存超额流量，可能会带来延迟和抖动
丢包可能引发重传	较少的丢包，因而较少导致重传
可以重标记流	不能重标记流

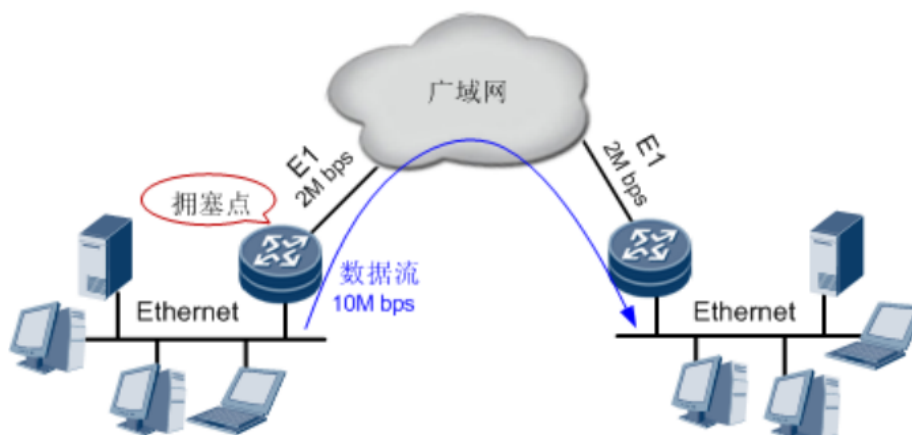


图4.3.1 网络拥塞

4.3 拥塞管理

4.3.1 拥塞管理的概念

拥塞是在共享网络上多个用户竞争相同的资源（带宽、缓冲区等）时发生的问题。例如，由于广域网的带宽通常要比局域网的带宽小，当一个局域网的用户向另一个局域网的用户发送数据时，由于广域网的带宽小于局域网的带宽，数据将不可能按局域网发送的速度在广域网上传输。此时，处在局域网和广域网之间的路由器将发生拥塞，如图4-3-1所示。

4.3.2 队列

拥塞管理的中心内容是当拥塞发生时如何制定一个策略，用于决定报文转发的处理次序和丢弃原则，一般采用队列技术。

队列指的是在缓存中对报文进行排序的逻辑。当流量的速率超过接口带宽或超过为该流量设置的带宽时，报文就以队列的形式暂存在缓存中。报文离开队列的时间、顺序，以及各个队列之间报文离开的相互关系则由队列调度算法决定。

队列调度算法：

本文介绍几种常见队列调度算法：

先进先出FIFO（First In First Out）

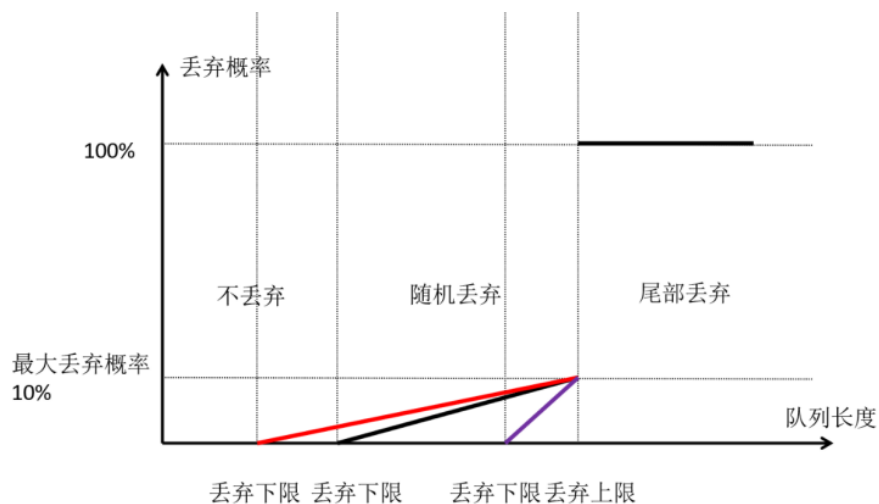


图4.4 WRED的策略

严格优先级SP (Strict Priority)

轮询RR (Round Robin)

加权轮询WRR (Weighted Round Robin)

差分轮询DRR (Deficit Round Robin)

加权差分轮询WDRR (Weighted Deficit Round Robin)

队列WFQ (Weighted Fair Queuing)

4.4 拥塞避免

受限于设备的内存资源，按照传统的处理方法，当队列的长度达到规定的最大长度时，所有到来的报文都被丢弃。对于TCP报文，如果大量的报文被丢弃，将造成TCP超时，从而引发TCP的慢启动和拥塞避免机制，使TCP减少报文的发送。当队列同时丢弃多个TCP连接的报文时，将造成多个TCP连接同时进入慢启动和拥塞避免，称之为：TCP全局同步。这样多个TCP连接发向队列的报文将同时减少，使得发向队列的报文的量不及线路发送的速度，减少了线路带宽的利用。并且，发向队列的报文的流量总是忽大忽小，使线路上的流量总在极少和饱满之间波动。

为了避免这种情况的发生，队列可以采用加权随机早期检测WRED (Weighted Random Early Detection) 的报文丢弃策略 (WRED与RED的区别在于前者引入IP优先级，DSCP值，和MPLS EXP来区别丢弃策略)。采用WRED时，用户可以设定队列的阈值 (threshold)。当队列的长度小于低阈值时，不丢弃报文；当队列的长度在低阈值和高阈值之间时，WRED开始随机丢弃报文 (队列的长度越长，丢弃的概率越高)；当队列的长度大于高阈值时，丢弃所有的报文。WRED的策略如下图所示：

5 其他QoS技术简介

链路效率机制 链路效率机制，用于改善链路的性能，间接提高网络的QoS，如降低链路发包的时延（针对特定业务）、调整有效带宽。链路效率机制有很多种，下面介绍两种比较典型的链路效率机制及其基本原理。

链路分片与交叉（Link Fragment Interleave, LFI）

对于低速链路，即使为语音等实时业务报文配置了高优先级队列（如RTP优先队列或LLQ），也不能够保证其时延与抖动，原因在于接口在发送其他数据报文的瞬间，语音业务报文只能等待，而对于低速接口发送较大的数据报文要花费相当长的时间。采用LFI以后，数据报文（非RTP实时队列和LLQ中的报文）在发送前被分片、逐一发送，而此时如果有语音报文到达则被优先发送，从而保证了语音等实时业务的时延与抖动。

RTP报文头压缩（RTP Header Compression, CRTP）

CRTP用于RTP（Real-time Transport Protocol）协议，对IP头、UDP头和RTP头进行压缩，通常在低速链路上使用。可将40字节的IP/UDP/RTP头压缩到24个字节（不使用UDP校验和可到2字节），提高链路带宽的利用率。CRTP主要得益于同一会话的语音分组头和语音分组头之间的差别往往是不变的，因此只需传递增量。

链路层QoS技术

链路层QoS技术主要针对ATM（Asynchronous Transfer Mode，异步传输模式）、帧中继等链路层协议支持QoS。ATM作为一种面向连接的技术，提供对QoS最强有力的支持，而且可以基于每个连接提供特定的QoS保证；帧中继网络确保连接的CIR（Committed Information Rate，承诺信息速率）最小，即在网络拥塞时，传输速度不能小于这个值。

ATM QoS

ATM是异步传输模式（Asynchronous Transfer Mode）的简称，以信元为基本单位进行信息传输、复接和交换。ATM信元具有53字节的固定长度，其中5个字节构成信元头部，主要用来标识虚连接，另外也完成了一些功能有限的流量控制，拥塞控制，差错控制等功能，其余48个字节是有效载荷。ATM是面向连接的交换，其连接是逻辑连接，即虚电路。每条虚电路（Virtual Circuit, VC）用虚路径标识符（Virtual Path Identifier, VPI）和虚通道标识符（Virtual Channel Identifier, VCI）来标识。一个VPI/VCI值对只具有本地意义，不具有全局有效性。它在ATM节点上被翻译。当一个连接被释放时，与此相关的VPI/VCI值对也被释放，它被放回资源表，供其它连接使用。

FR QoS

FR（Frame Relay，帧中继）是一种统计复用的协议，它能够在单一物理传输线路上提供多条虚电路。每条虚电路用DLCI（Data Link Connection Identifier，数据链路连接标识）来标识。每条虚电路通过LMI（Local Management Interface，本地管理接口）协议检测和维护虚电路的状态。

参考文献

1. 崔恒，计算机网络服务质量（QoS）的理论研究
2. 华为文档，端到端QoS服务模型
3. CSDN，深入浅出QoS详解
4. 华为论坛，交换机在江湖，QoS专题-第1期-QoS理论篇
5. 腾讯云，2万字带你学习Qos原理