身份证 130425198708092057，210181199209046813，222302197312260017，350128197311050034 ，410611197403020058

联通手机号 18575414999 ，13026844666

电信手机号 18919628498 ， 18919628499

移动手机号 18756500265 ，18555918746 15665541872

邮箱 zhangsan@163.com ，lisi@qq.com

银行卡 6216610200016587010 ， 6221882600114166800

财务报表

内部数据

内部资料

保密

秘密

机密

密码口令

超级用户

家庭地址身份证 130425198708092057，210181199209046813，222302197312260017，350128197311050034 ，410611197403020058

联通手机号 18575414999 ，13026844666

电信手机号 18919628498 ， 18919628499

移动手机号 18756500265 ，18555918746 15665541872

邮箱 zhangsan@163.com ，lisi@qq.com

银行卡 6216610200016587010 ， 6221882600114166800

财务报表

内部数据

内部资料

保密

秘密

机密

密码口令

超级用户

家庭地址

Chapter 25 Qos简介

25.1 QoS简介

在传统的 IP 网络中,所有的报文都被无区别的等同对待,每个路由器对所有的报文均采用先入先

(FIFO) 的策略进行处理.它尽最大的努力(Best-Effort)将报文送到目的地,但对报文传送的可靠性,传

送延迟等性能不提供任何保证.

网络发展日新月异,随着 IP 网络上新应用的不断出现,对 IP 网络的服务质量也提出了新的要求.例

如 VoIP 等实时业务就对报文的传输延迟提出了较高要求.如果报文传送延时太长,将是用户所不能接受

的(相对而言 E-Mail 和 FTP 业务对时间延迟并不敏感).为了支持具有不同服务需求的语音,视频以及数

据等业务,要求网络能够区分出不同的通信,进而为之提供相应的服务.传统 IP 网络的尽力服务不可能识

别和区分出网络中的各种通信类别,而具备通信类别的区分能力正是为不同的通信提供不同服务的前提.

所以说传统网络的尽力服务模式已不能满足应用的需要 QoS(Quality of Service ,服务质量技术)的出

现便致力于解决这个问题

25.1.1 一个简单的 QoS案例

下面用简单的例子对比了在网络发生拥塞时报文在无 QoS 保证和有 QoS 保证网络中的不同处理过

程,下图所示为发生拥塞时网络设备的一个接口在不支持 QoS 的情况下报文的发送情况

所有要从该接口输出的报文,按照到达的先后顺序进入接口的 FIFO 队列尾部。而接口在发送报文

时从 FIFO 先入先出队列的头部开始依次发送报文。所有的报文在发送过程中没有任何区别也不对报文

传送的质量提供任何保证

当然 ,我们需要寻求更好的处理方式, 下图为一个优先级对列进行 Qos 处理的过程。

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 667

在报文到达接口后,首先对报文进行分类,然后按照报文所属的类别让报文进入所属队列的尾部。

在报文发送时,按照优先级总是在所有优先级较高的队列中的报文,发送完毕后再发送低优先级队列中

的报文。这样在每次发送报文时总是将优先级高的报文先发出去。保证了属于较高优先级队列的报文

有较低的时延报文的丢失率和时延。这两个性能指标在网络拥塞时也可以有一定的保障。

25.1.2 QoS的作用

QoS 旨在针对各种应用的不同需求为其提供不同的服务质量例如提供专用带宽减少报文丢失率降

低报文传送时延及时延抖动等为实现上述目的 QoS 提供了下述功能

报文分类和着色

网络拥塞管理

网络拥塞避免

流量监管和流量整形

QoS 信令协议

QoS 可以控制各种网络应用和满足多种网络应用要求如

控制资源：

如可以限制骨干网上 FTP 使用的带宽,也可以给数据库访问以较高优先级

可裁剪的服务：

对于 ISP 其用户可能传送语音视频或其他实时业务 QoS 使 ISP 能区分这些不同的报文

并提供不同服务

多种需求并存：

可以为时间敏感的多媒体业务提供带宽和低时延保证,而其他业务在使用网络时也不

会影响这些时间敏感的业务

在一个网络中需要以下的三个部分来完成端到端的 QoS

各网络元件路由器以太网交换机等支持 QoS 提供队列调度流量整形等功能

信令技术来协调端到端之间的网络元件,为报文提供 QoS

QoS 技术控制和管理端到端之间的报文在一个网络上的发送

而每个网络元件提供如下功能

报文分类,对不同类别的报文提供不同的处理

队列管理和调度来满足不同应用要求的不同服务质量

流量监管和流量整形限制和调整报文输出的速度

接入控制来确定是否允许用户信息流使用网络资源

25.2 QoS服务模式

网络应用是端到端的通讯结构,比如两个不同网络的主机进行通讯,中间可能跨越各种 router 和

核心 switch,那么想整体的实现所谓的 QOS,就必须全局考虑,QOS 的服务模型的概念就是采用通过什么

模式全局实现服务质量保证,一共分成三种。

Best-Effort service 尽力而为服务模型

Integrated service 综合服务模型 简称 Intserv

Differentiated service 区分服务模型 简称 Diffserv

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 668

25.2.1 尽力而为的服务模型

Best-Effort 是一个单一的服务模型,也是最简单的服务模型,应用程序可以在任何时候,发出任

意数量的报文,而且不需要事先获得批准,也不需要通知网络,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。其实 best-effort 并非是什么 QOS,就是互

联网的简单数据传输方式而已,有什么传什么,阻塞也就阻塞了,丢且也就丢弃了。

25.2.2 集成服务模型

Intserv：集成服务模型,它可以满足多种 QoS 需求。这种服务模型在发送报文前,需要向网络申

请特定的服务。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。

网络在收到应用程序的资源请求后,执行资源分配检查 Admission control 即基于应用程序的资

源申请和网络现有的资源情况,判断是否为应用程序分配资源,一旦网络确认为应用程序的报文分配了

资源,则只要应用程序的报文控制在流量参数描述的范围内,网络将承诺满足应用程序的 QoS 需求。而

网络将为每个流 flow 由两端的 IP 地址、端口号、协议号确定、维护一个状态,并基于这个状态执行

报文的分类、流量监管、policing、排队及其调度来实现对应用程序的承诺。

在 IntServ 服务模型中,负责传送 QoS 请求的信令是 RSVP(Resource Reservation Protocol)资源

预留协议,它通知路由器应用程序的 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。 Intserv 实际上是一种对服务的预定机制,通过申请来获取相应得服务,这里面主要依靠

的就是 RSVP——资源预留协议。

RSVP 是第一个标准 QoS 信令协议,它用来动态地建立端到端的 QoS,它允许应用程序动态地申请网

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 669

络带宽等。RSVP 协议不是一个路由协议,相反,它按照路由协议规定的报文流的路径为报文申请预留资

源,在路由发生变化后,它会按照新路由进行调整,并在新的路径上申请预留资源。RSVP 只是在网络节

点之间传递 QoS 请求,它本身不完成这些 QoS 的要求实现,而是通过其他技术来完成这些要求的实现。

(RSVP 只是一种用来预定的协议)

RSVP 的处理是接收方发出资源请求,按照报文发送的反向路径发送资源请求,所以它可以满足非

常大的多播组,多播组的成员也可以动态变化,RSVP 协议是针对多播设计的 单播可以看作是多播的一

个特例。

由于 RSVP 在 Internet 上还没有得到广泛的推广,在主机不支持 RSVP 的情况下,我们可以通过配

置 RSVP 代理,即代替不支持 RSVP 的主机发送 RSVP 报文来获得这种服务,对报文流路径上不支持 RSVP

的路由器,它只需要简单的转发 RSVP 报文 所以对 RSVP 协议不会有太大影响,但这些节点不会对报文提

供所要求的 QoS 。(这是 RSVP 的一个缺点)

RSVP 信令在网络节点之间传送资源请求,而网络节点在收到这些请求后,需要为这些请求分配资

源,这就是资源预留。网络节点比较资源请求和网络现有的资源,确定是否接受请求,在资源不够的情况

下,这个请求可以被拒绝,可以对每个资源请求设置不同的优先级。这样,优先级较高的资源请求可以在

网络资源不够的情况下,抢占较低优先级的预留资源,来优先满足高优先级的资源请求。

资源预留判断是否接受资源请求,并承诺对接受了的资源请求提供请求的服务,但资源预留本身

不实现承诺的服务,需要通过队列等其他技术来实现。

Intserv 有它的好处,但是也有严重缺点,首先就是 RSVP 协议数据太多,而且不断刷新,并且这种

给单一数据流的路径进行带宽预留的解决思路在浩瀚的 Internet 上实现简直是不可能的,而且 RSVP 的

部署,厂商之间设备的互联,业务管理方面 等存在着种种问题,所以这么模型在 1994 年推出之后就没

有获得任何规模的商业应用。

25.2.3 区分服务模型

DiffServ 是一个多服务模型,它可以满足不同的 QoS 需求,与 IntServ 不同,它不需要使用 RSVP

即应用程序在发出报文前,不需要通知路由器为其预留资源,对 DiffServ 服务模型,网络不需要为每个

流维护状态 ,它根据每个报文指定的 QoS 来提供特定的服务 可以用不同的方法来指定报文的 QoS,如

IP 报文的优先级位 IP Precedence) ,报文的源地址和

目的地址等,网络通过这些信息来进行报文的分类、流量整形、流量监管和队列调度。

DiffServ 一般用来为一些重要的应用提供端到端的 QoS 它通过下列技术来实现

CAR：

它根据报文的 ToS 或 CoS 值(对于 IP 报文是指 IP 优先级或者 DSCP 等等)IP 报文的五

元组(指源地址目的地址协议端口号)等信息进行报文分类,完成报文的标记和流量监

管。

队列技术：

WRED、 PQ、 CQ、 WFQ、 CBWFQ 等队列技术对拥塞的报文进行缓存和调度,实现拥塞管理。

通常在配置 DiffServ 时,边界路由器通过报文的源地址和目的地址等对报文进行分类,对不同的

报文设置不同的 CoS 值,而其他路由器只需要用 CoS 值来进行报文的分类

在 MPLS 上应用 DiffServ 用以下两种方法来解决：

在以太网等网络中,MPLS 报文在二层链路层和三层网络层之间有一个薄层(shim).我们扩展

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 670

薄层中未用的字段---EXP.由这几个位来决定报文的队列调度及丢弃的优先级.

在 ATM FR 等网络中,其 MPLS 报文没有薄层(shim), 可针对 FEC ForwardingEquivalance Class 转

发等价类和 QoS 请求的组合来分配标签,而不同于以前仅针对 FEC 分配标签.这样在收到一个 MPLS 报文

后,根据收到报文的标签就可以确定发出报文的标签及报文所要求的服务.

25.2.4 Intserv与Diffserv之间的互通

一般来讲,在提供 IP 网络的 QoS 时,为了实现规模适应性,在 IP 骨干网往往需要采用 Diffserv 体

系结构,在 IP 边缘网可以有两种选择:采用 Diffserv 体系结构或采用 Intserv 体系结构.目前在 IP 边

缘网络采用哪一种 QoS 体系结构还没有定论,也许这两种会同时并存于 IP 边缘网中.在 IP 边缘网采用

Diffserv 体系结构的情况下,IP 骨干网与 IP 边缘网之间的互通没有问题.在 IP 边缘网采用 Intserv 体

系结构的情况下,需要解决 Intserv 与 Diffserv 之间的互通问题,包括 RSVP 在 Diffserv 域的处理方

式,Intserv 支持的业务与 Diffserv 支持的 PHB(Per-Hop Behavior, 单中继段行为)之间的映射.

RSVP 在 Diffserv 域的处理可以有多种可选择的方式.例如

1. 一种方式为RSVP对Diffserv域透明,RSVP在Intserv域边界路由器终结,Diffserv域对

Intserv域采用静态资源提供方式:

2. 一种方式为Diffserv域参与RSVP协议处理,Diffserv域对Intserv域采用动态资源提供

方式.

前一种互通方式实现相对简单,可能造成 Diffserv 域资源的浪费.后一种互通方式实现相对复杂,

可以优化 Diffserv 域资源的使用.

除此以外,还需要解决 Intserv 支持的业务与 Diffserv 支持的 PHB 之间的映射问题.映射标准为

两者支持的应用是否相同或相近. 为了说明这个问题我们首先回顾一下 Intserv 支持的业务,它支持

的业务包括保证服务(Guaranteed Service)负载控制服务(Controlled-Load Service).前者可以为用

户应用提供严格的端到端时延及带宽保证,适用于实时应用.后者在网络负荷较重的情况下为用户应用

提供与网络轻负荷情况下相近似的性能,不能保证端到端的时延.

Diffserv 提供的 PHB 包括 EF (Expedited Forwarding 加速转发)AF(Assured Forwarding 确保

转发)

EF 用于支持低丢失率,低时延,确保带宽的应用 AF 可以保证在应用向网络发送的业务流量没有超

过约定值的情况下,该应用的报文丢失概率非常低 AF 有 4 类每一类可以设置 3 个不同的丢弃优先级.

从上面的叙述易于获得 Diffserv 与 Intserv 之间的映射关系

将Intserv中的保证服务映射为Diffserv中的EF

将Intserv中的负载控制服务映射为Diffserv中的AF

25.3 Diffserv服务模式

25.3.1 Diffserv体系结构

用来在网络中提供 QoS 的 Diffserv 方法应用了一组合理定义的小型基本部件，使用这些部件可

以构筑一系列的服务。其目标是在 IPv4 报头中定义区分服务(DS)字节和服务类型(ToS)字节，在 IPv6

中定义通信类(TrafficClass)字节，并标记分组中的标准化 DS 字节，使分组在每一个网络节点得到特

定的转发处理或单中继段行为(PHB)。

区分服务体系结构提供了—个框架，在这个框架下，服务提供商们可以为客户提供各种网络服务，

并根据性能来区分每一种服务。客户只需将分组的区分服务码点(DSCP)标记为特定值，便可以选择每

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 671

个分组的性能等级，这个值指定了分组在服务提供商网络中的单中继段行为(PHB)。通常，服务提供商

和客户可以共同协商一个配置文件(profile)，它描述了各种服务等级通信的提交速率。如梁提交的分

组速率超过了规定，则可能无法获得规定的服务等级。

Diffserv 体系结构只规定处理分组的基本机制，将这些机制作为基本部件可以构建各种服务。服

务定义了一些重要的分组传输特征，例如，网络小一条路径某个方向上的吞吐量、延迟、抖动以及分

组丢失率等。另外，还可以使用访问网络资源的相对优先级米表征服务。服务被定义后，提供这种服

务的网络上的所有节点都将被指定一个 PHB，并给 PHB 分配一个 DSCP。PHB 是网络节点提供给所有带

特定 DSCP 值的分绍的转发行为，需要特定服务等级的通信，其分组的 DSCP 字段将被指定为相应的值。

25.3.2 DSCP值

Diffserv 域(domain)中所有的节点都将根据分组的 DSCP 字段来遵守 PHB。另外，Diffserv 域边

界的节点还负责调节进入域的通信量。流量调节涉及诸如分组分类和流量控制(policing)等功能，通

常在通信进入域的输入接口中实现。流量调节在 Diffserv 域的流量工程中担任重要角色，对于所有进

入该域的通信，网络都能监测其单中继段行为

IETF Diffserv 工作组 正在进行使用户可以使用 IP 报头中的 ToS 字节中的 6bit 用来标记 DSCP.

TOS 整个字段一共 8 位，下图中是 ip precedence 的标识方法，使用前三位，也就是 P0、P1、

P2 一共三位，共 8 个 bit 位，也就是 8 个优先级，分别是 0-7，其中 6 和 7 一般保留，常用的是 0-5,

图中下面部分介绍的事 IP 优先级的含义，提供这个表格的意义在于 进行更改数据包优先级等配置时，

我们既可以使用数字，也可以使用名称。

分类是没有范围限制的，也就是说我们可以对数据进行灵活的分类，比如说把某一个源 ip 到目

的 ip 的流量优先级进行更改 IP 优先级或者 DSCP 的操作。也可以定义去更改某一个特定的流量（扩

展控制列表）的优先级。

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 672

通常于网络边界处对报文进行分类时，同时标记 IP 优先级或 DSCP。这样，在网络的内部就可以

简单的使用 IP 优先级或 DSCP 作为分类的标准，而队列技术如 WFQ 等 CBWFQ 就可以使用这个优先

级来对报文进行不同的处理。

什么是标识，实际上就是通过更改这些优先级字段将这些数据分出种类来，即便是上面的图标中

有所谓的什么 0-7 优先级，好似 7 要比 0 就会优先级大一些，但是一定要清楚这只 是区分，执行

的策略要靠后面的队列机制来解决，实际上之所以这样定义优先级我认为只是 为了制定一个共同遵守

的类别优先标准，没有实际的意义，真正的操作是在配置上针对于不 同优先级采用的措施——例如在

队列里面使用什么标识的数据包属于什么队列等等。

上表是 DSCP 与 WRR 中应用的队列匹配。

25.3.3 Cos值

Cos 值是一种基于 多层的优先级表示方式: 如下图所示

如上图所示，正常的以太网 frame中是不存在标记的，但是 ISL和 dot1q的 frame 中有

三个 bit 定义服务级别，一共有6个服务级别可以使用

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 673

下表是 COS 和 DSCP 彼此匹配时的对应关系。

25.3.4 PHB

支持 Diffserv 的网络节点使用 IP 报头中的 DSCP 字段来为分组选择特定的 PHB，PHB 描述了

Diffserv 节点对具有相同 DSCP 的分组采用的外部可见的转发行为。可以用资源优先级来定义 PHB，也

可以用一些可见的通信服务特征如分组延迟、分组丢失率或分组抖动来定义。可以将 PHB 看做一个黑

盒子，因为它定义了一些外部可见的转发行为，但没有指定特定的实现方式。

在 Diffserv 网络中，可以认为其默认的 PHB 是尽力而为的行为。Diffserv 为每个 PHB 推荐了特

定的 DSCP 值，但网络提供商可以在自己的网络中选择使用不同的 DSCP。推荐的尽力而为行为的 DSCP

是 000000。

特定通信类的PHB取决于下列因素：

通信类的到达速率或负载：这是通过网络边界的流量调节控制的。

通信类的资源分配：这是由Diffserv域中节点的资源分配控制的

通信丢失率：取决于Diffserv域中节点的分组丢弃策略。

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 674

EF PHB

加速转发PHB(Expedited Forwarding PHB)可以用来实现的丢失率,低延迟,低抖动,确定

带宽,端到端服务.EF PHB针对VoIP和视频会议这样的应用. 导致分组延迟和分组抖动 的主要

原因是排队延迟, 排队延迟通常是在拥塞时,队列过长引起的,EF PHB确保通信不需要排队或

队列足够短.

路由器可以通过使用不同的EF功能实现方法, 为某个端口上特定的离开速率分配资源.

当EF通信可以在高度加权的队列上传送,例如CBWFQ,WRR,DRR等分组调度技术提供了这个功能.

对于网络中的 EF通讯,推荐使用DSCP 101110.

AF PHB

AF PHB可保证业务流量在没有超过阈值的情况下，确保报文保持较低的丢包率。

AF PHB对应4种AF Class

Class 1 Class 2 Class 3 Class 4

+-------------+------------+-------------+-----------------------+

Low Drop Prec | 001010 | 010010 | 011010 | 100010 |

Medium Drop Prec | 001100 | 010100 | 011100 | 100100 |

High Drop Prec | 001110 | 010110 | 011110 | 100110 |

-----------+------------+------------+-------------+-------------+

缺省的PHB编码 000000 对应 Best-effort traffic.

PHB编码的解释: 以AF11 (0010100)为例， 前3个bit表示IP报文的优先级，接下来的2

个bit表示丢包优先级，最后一个bit总是为0

AF11 == 001-01-0

AF Class级别越高，其对应的IP报文的优先级越高，相对来说提高的QoS质量也越较好。

Class 1 Class 2 Class 3 Class 4

------------------+------------+-------------+--------------------------+

Low Drop Prec | 10 | 18 | 26 | 34 | (decimal)

Medium Drop Prec | 12 | 20 | 28 | 36 |

High Drop Prec | 14 | 22 | 30 | 38 |

------------------+------------+-------------+--------------------------+

25.3.5 数据报标记

数据报标记就是修改 IP 优先级或者 DSCP，但是由于 IP 优先级和 DSCP 都 是占用 TOS 字段，后者

相当于前者的扩展,所以不能同时设置这两种值,如果同时设置了这两 种值,那么只有 IP DSCP 的值生效。

标记是后续很多 QOS 策略应用的基本，使用的是 policy map。

配置方法

1.定义class map

Class map是一个匹配表，类似于ACL。

所有的policy map实质上是对class map进行操作的

nimokaka(config)#class-map [match-all|match-any] {map - name}

参数中match-all 表示匹配所有条件，match-any表示至少符合一个条件

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 675

2.class map的匹配

nimokaka(config-cmap)#

match access-group {ACL} 匹配IP ACL （主要就是对应数据包了）

match protocol {protocol} 匹配协议（这个在NBAR中使用）

match input-interface {interface} 匹配进站接口

match qos-group {Group ID} 匹配组ID（不知道干啥的）

match destination-address {mac mac-address} 匹配目标MAC 地址

match source-address {mac mac-address} 匹配源MAC 地址

match ip {dscp dscp} 匹配IP DSCP 值

match ip {precedence precedence} 匹配IP 优先级

match class-map {map-name} 匹配class map（class map嵌套）

match vlan {vlan-id} 匹配VLAN

Class-map可以嵌套：在创建class map的时候去调用一个已有的class map

1、管理方便,在已有的基础上增加一个修改进行平滑的过度。

2 、允许用户在同一个class map 里分别使用匹配所有(match-all) 和匹配任何

(match-any)。

比如4个匹配标准：A、B、C和D。现在想让class map 匹配A，或匹配B，或同时匹

配C和D，就可以使用class map 的嵌套：创建一个新的class map ，定义为匹配所有

(match-all)新标 准为匹配E即同时匹配C和D；然后定义另一个匹配任何(match-any)的

class map，去匹配A， 或B，或E(即同时匹配C和D)。

3.设置policy map

nimokaka(config)#policy-map {policy-name}

nimokaka(config-pmap)#class {class-map}

4.配置优先级和DSCP值

nimokaka(config-pmap-c)#

一些用于标记的动作选项:

set ip {precedence precedence} 设置IP优先级

set ip {dscp dscp} 设置IP DSCP 值

set qos-group {Group ID} 设置组ID

set cos {cos} 设置CoS

priority {kbps|percent percent} [Bc] 定义优先级流量保留带宽以及突发流量

bandwidth {kbps|percent percent} 定义保留的带宽

police {CIR Bc Be} conform-action {action} exceed-action {action}

violate-action{action}]

5.将配置挂接到接口上

nimokaka（config-if）service-policy [input|output] policy-name

6.检查配置

CCIE R/S & Service Provider Exam Certification Guide

P a g e | 676

nimokaka#show policy-map [policy-name]

查看接口的policy map 信息:

nimokaka#show policy-map interface [interface]