目　录

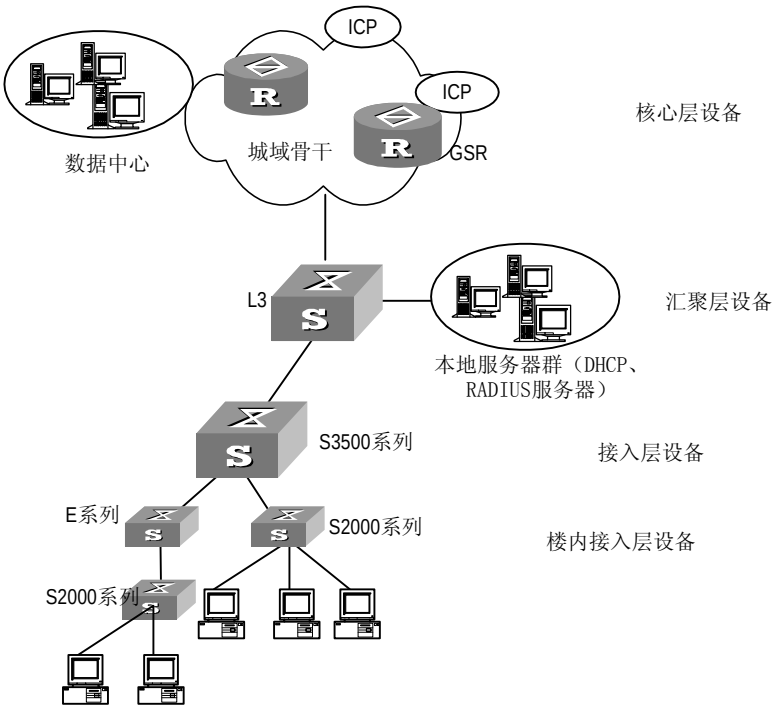
附录 A 交换网络数据转发流程

A.1 引言

随着因特网的高速发展，人们对通信的需求已从传统的电话、传真、电报等低速业务逐渐向高速的因特网接入、可视电话、视频点播等宽带业务领域延伸。用户对上网速率的需求也越来越高，以太网接入因其成本低、使用简单、速度高而倍受市场的关注。面对迅猛发展的宽带网络建设需求，华为公司根据不同的客户类型需求，推出了 Quidway 系列以太网交换机及其它网络设备。使用华为公司的网络设备，可以构建可运营、可管理的网络。那么在网络中，数据是怎样转发的呢？本文将简要讲述数据在交换网络中的转发流程。

A.2 简单的转发流程

下面给出一个简单的组网示意图，以便说明。



图A-1 小区组网示意图

在上图中，L3 表示的是三层交换机，GSR 为 Gigabit Switch Router 的缩写，即 G 比特交换路由器，ICP 为 Internet Content Provider 的缩写，即因特网内容提供商。

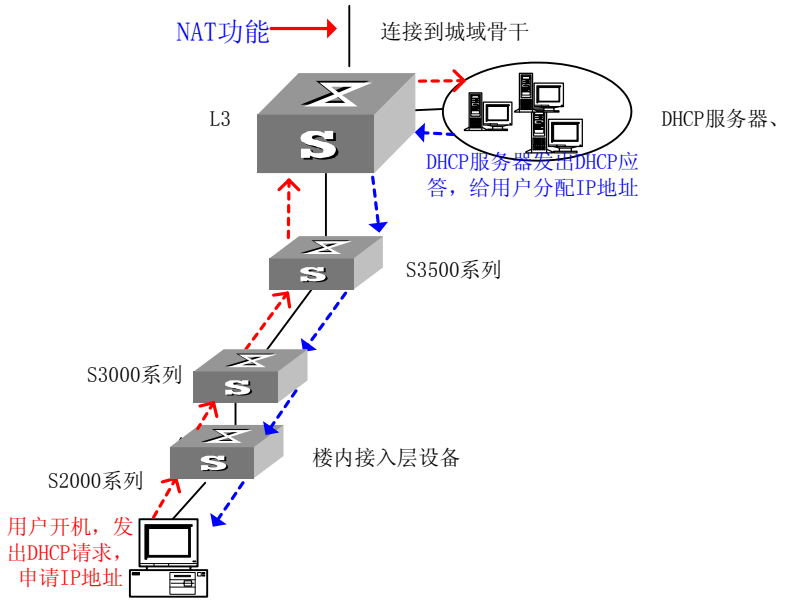
在组网中，接入层设备为华为 S2000 系列和 S3000 系列以太网交换机。汇聚层设备为 S3500 系列以太网交换机或者 S5516、S6506 等三层以太网交换机。小区内部有 AAA（Authentication, Authorization and Accounting）服务器、DHCP（Dynamic Host Configuration Protocol）服务器、DNS（Domain Name Server）服务器。

📖 **说明：**

实际组网中，用户的计算机可能采用的是固定 IP。用户通过固定 IP 上网与动态申请 IP 上网相比，仅仅是缺少了申请 IP 地址这一过程。所以在下面的例子中，以用户的计算机通过 DHCP 协议动态申请 IP 地址为例进行介绍。
用户计算机配置为自动获取 IP 地址。

下面介绍小区内部数据的转发过程，对数据到达城域网后的过程不作介绍。

用户计算机开机后会通过 DHCP 报文申请 IP 地址。接入层交换机、汇聚层交换机会将此请求报文转发给 DHCP 服务器。DHCP 服务器通过应答报文给用户 PC 分配 IP 地址。有了 IP 地址后，用户就可以上网了。



图A-2 IP 地址请求过程

整个申请 IP 地址的过程如下：

(1) 客户机通过 DHCP Discover 广播提出请求。如果客户机有一个永久性的租用地址，它可以直接请求那个地址。

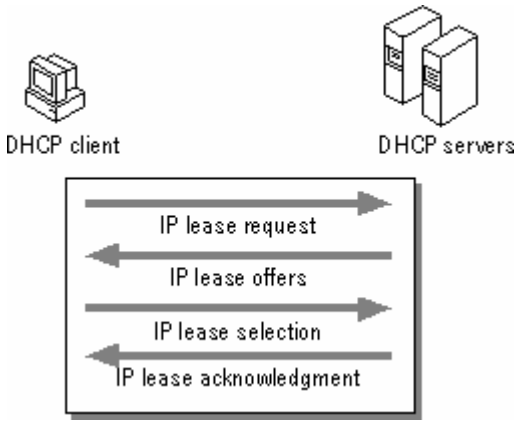
(2) 服务器一旦收到 IP 请求，会从地址池中取出一个地址并返回一个附有可用 IP 地址的 DHCP Offer 报文。

(3) 如果客户机收到多个 IP，它会选择第一个或其所请求的那一个。

(4) 客户机广播标识服务器的 DHCP Request 报文并等待。

(5) 每一个服务器检查报文，若发现不是它的标识，它会丢弃报文。当被标识的服务器接收了报文后，它会发回一个 DHCP Ack 报文，如果所请求的 IP 被分配也就是说租用已中止，会发回 DHCP Nak 报文。

(6) 如果客户机收到 DHCP Ack 报文，它可以开始使用 IP 地址。如果它收到 DHCP Nak，它会重新开始整个过程。假如 IP 有问题，客户机会发送一个 DHCP Decline 报文给服务器并重新开始。



图A-3 IP 地址请求过程示意图

📖 **说明：**

这里有一个问题：如果给每个用户分配的都是公网 IP，会需要大量 IP 地址，这非常浪费，也不现实。所以一般情况下，小区内用户分配的是私网 IP，而在图中的 L3 上实现 NAT 功能。

A.2.1 同一 VLAN 内的通信

用户计算机通过 ARP 在本 PC 上建立一个 IP 与 MAC 地址的对应表格（通往 VLAN 外部的 IP 地址将对应为 VLAN 网关的 IP 地址）。这个表称为 ARP 表。同时 ARP 在存储器中维护一个 cache，这个 cache 称为 ARP cache。通常用户计算机要向外发送报文的时候（进行通信的应用程序不知道对端的物理地址），有如下步骤：

(1) 首先搜索 ARP cache 对目的 IP 进行匹配，如果匹配成功，ARP 就反馈 IP 地址对应的 MAC 地址给进行通信的应用程序；假如没有匹配成功就检查 ARP 表，如果匹配成功，ARP 就反馈 IP 地址对应的 MAC 地址给进行通信的应用程序。

(2) 假如 ARP 没找到一个匹配的 IP 地址，它就会向网络上发布消息，这个消息被称为 ARP 请求。ARP 请求被广播到局域网上的每一个设备。

(3) 如果局域内存在目的 IP 对应的设备，则此设备会向发起 ARP 请求的计算机反馈应答，将自己的 MAC 地址反馈给用户计算机。如果局域网内不存在目的 IP 对应的设备，则网关会将自己的 MAC 地址反馈给用户计算机。

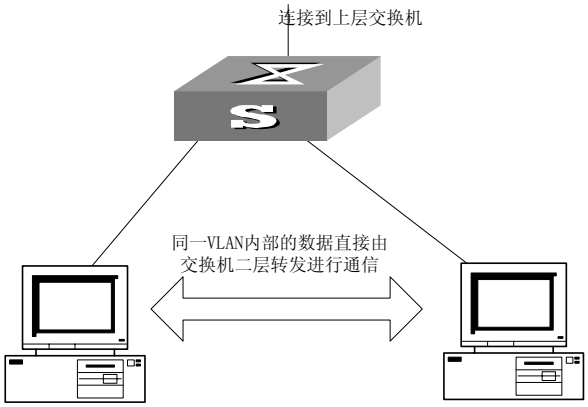
(4) 用户计算机上进行通信的应用程序根据找到的 MAC 地址封装报文，并发送出去。同时用户计算机上的 ARP 会将新找到的 MAC 地址和其对应的 IP 地址作为一个表项添加到 ARP 表和 ARP cache 中。

(5) 交换机收到用户计算机的报文，进行判断，如果通信的源端和目的端在同一个 VLAN 内部，进行二层转发；如果二者不在同一个 VLAN 内，就交给网关进行三层转发。

如果用户在同一个 VLAN 内部通信，只需要进行二层的点到点通信。

📖 **说明：**

VLAN 是虚拟局域网，是将有相同需求的网络设备从逻辑上划分在一个局域网内，而不是按照物理位置划分局域网。VLAN 的详细描述可以参见 IEEE 802.1Q 协议和配置指导中 VLAN 配置模块。



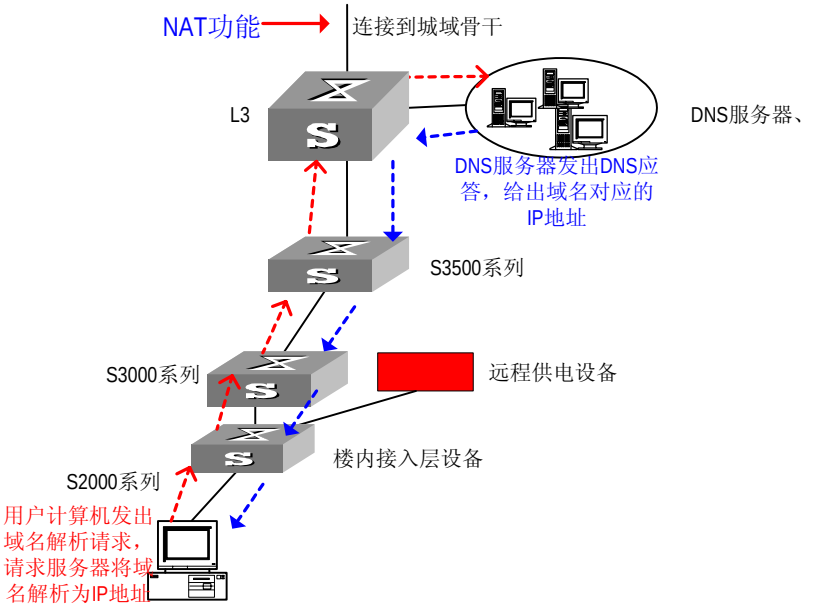
图A-4 同一 VLAN 内的通信示意图

A.2.2 不同 VLAN 间的通信

交换网络中如果没有配置 PVLAN，则不同 VLAN 间的计算机进行通信需要经过路由来实现，这里就不再详细介绍。

A.2.3 用户登录因特网的数据流程

如果用户想要上网，则需要将报文发送给网关；网关再进行三层的路由转发。一般用户会使用域名来访问因特网，这时在登录到指定的网站之前有一个域名解析的过程：用户键入域名后，计算机会向小区内的域名服务器发送一个 DNS 请求报文，小区内的域名服务器会向用户返回域名对应的 IP 地址（用户的计算机上需要正确配置域名服务器的 IP 地址）。



图A-5 域名解析的过程

说明：

当小区网络内无 DNS 服务器，而使用小区网络外面的 DNS 服务器时，就会出现内部用户不能通过域名访问 DNS 服务器的情况。原因是：内部 PC 通过域名访问网络时，会到外部的 DNS 上请求 IP 地址，由于 DNS 是在外部，所以它会返回一个公网的地址或找不到地址。这样导致内部 PC 通过域名访问时，得到是外部的地址或者得不到地址，导致小区内部用户不能正常访问小区内部的服务器。

📖 **说明：**

Quidway S2008B、S2016B 以太网交换机支持远程馈电。对这些设备进行远程馈电需要专门的供电设备，同时必须通过指定的交换机端口才能实现远程馈电。一般情况下，供电设备设置在小区的机房中，对小区内所有的需要远程馈电的交换机进行供电。

用户计算机在取得了域名对应的 IP 地址后，就可以访问因特网：

(1) 用户计算机以域名对应的 IP 地址为目的地址，自己的 IP 地址为源 IP 地址封装用户的 TCP 或者 UDP 数据，向网关发送此 IP 数据包；

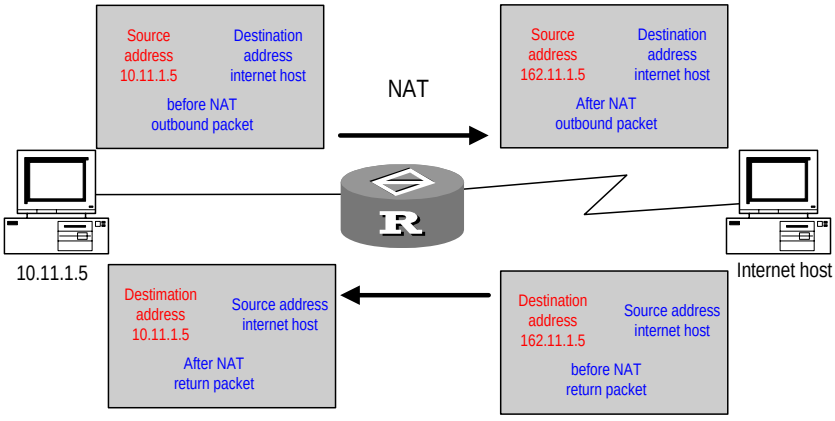
(2) 网关交换机收到此数据后根据路由表将此数据交给图中的 L3 设备；

(3) L3 设备对此数据进行一次 NAT 转换，将源地址改为 L3 的地址池中的一个公网 IP 地址，然后将此数据发送到城域网上；

(4) 当 L3 收到从城域网返回的数据后，进行一次 NAT 操作，将目的 IP 地址转换为相应的私网 IP 地址，然后转发给下挂的相应的交换机；数据沿交换机一层层下发，直到发给用户计算机。

📖 **说明：**

NAT（Network Address Translation），实现私网 IP 地址和公网 IP 地址之间的转换。详细内容可以参见 L3 设备配置指导手册的相关描述或其它技术文档。目前华为的 S8016 交换机实现了 NAT 功能。



图A-6 NAT 的地址转换过程

通过对交换机作简单的配置，小区内的用户就可以上网了。但是，怎样对用户进行计费、认证、授权等操作呢？这就需要构造一个可运营、可管理的交换网络。下面就讲述一下可运营、可管理的网络中的数据转发流程。

A.3 可运营、可管理网络中的数据转发流程

华为 Quidway 系列以太网交换机提供多种特性，可以为运营商构造可运营、可管理的网络。

在用户启动计算机准备上网时，需要首先通过认证，然后才能获取 IP 地址，进行后续的上网过程。

1. 用户的 802.1x 认证过程

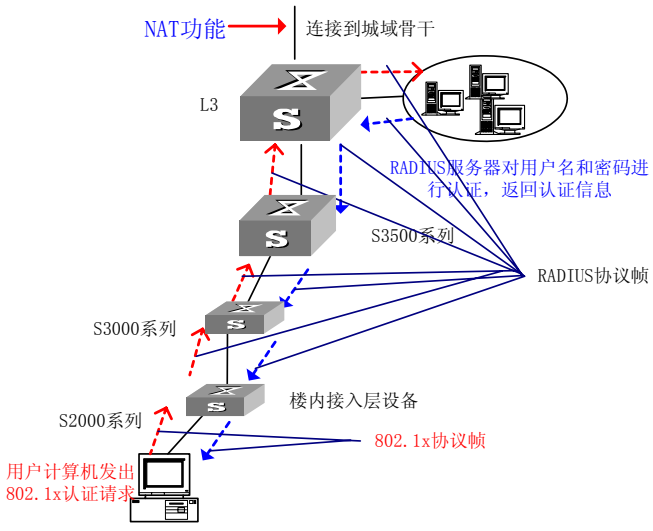
华为 Quidway 系列以太网交换机提供 802.1x 特性，可以对用户进行 802.1x 认证。

计算机上网必须先进行认证：在本计算机上启动 802.1x 终端软件，输入用户名和密码；交换机在收到用户名和密码后有两种认证方式，可以在本地进行认证，也可以通过 RADIUS 服务器进行远端认证。

本地认证需要在交换机上配置相应的用户名和密码，这种方法数据交互流程比较简单，但管理起来比较麻烦，需要在交换机上作很多配置，在组网中一般不会用到。下面就介绍通过 RADIUS 服务器进行认证的数据交互过程。

交换机上首先需要启动 802.1x 认证，并作了相应的配置。配置过程可以参见交换机用户手册的“AAA 及安全协议配置”模块。

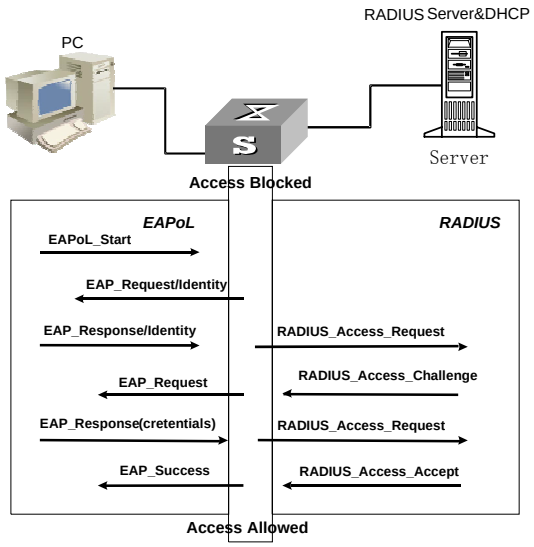
一般情况下交换机和 RADIUS 认证服务器之间传输的是标准的 RADIUS 报文，下面介绍在这种情况下 802.1x 认证的数据交互过程。



图A-7 802.1x 认证数据转发过程

用户计算机使用 802.1x 协议帧封装认证信息，和与之相连的交换机进行交互。交换机则使用标准的 RADIUS 协议封装用户认证信息，这样该认证数据就可

以穿越复杂的网络到达 RADIUS 服务器进行认证。下图为认证的数据交互示意图。其中 EAPoL（EAPOL 是 Extensible Authentication Protocol over LAN 的缩写）数据构成 802.1x 协议帧，RADIUS 数据构成 RADIUS 协议帧。



图A-8 802.1x 认证数据交互过程

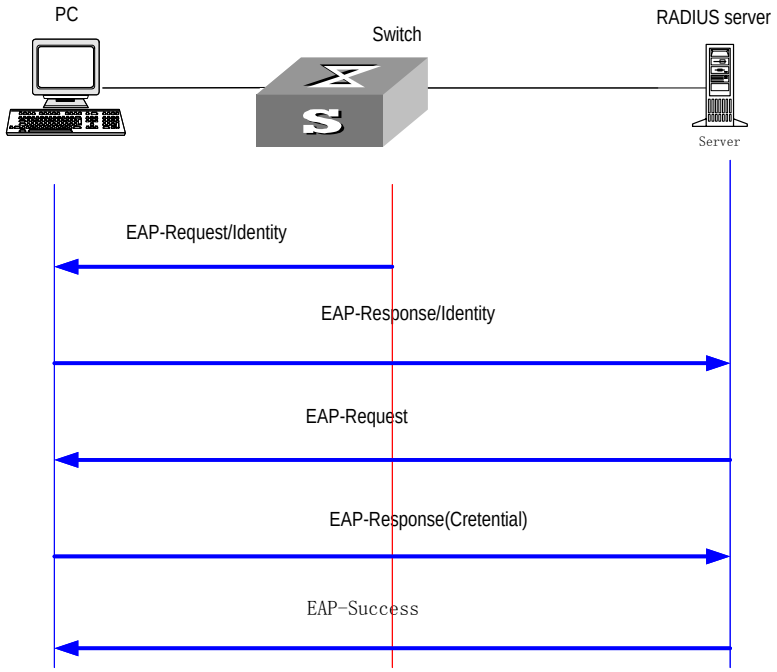
交换机也可以透明传输 802.1x 的 EAP 报文给 RADIUS 服务器。下面介绍这种情况下的 802.1x 认证的数据交互过程。

在这种情况下，从认证客户端到认证服务器直接传递的都是 802.1x 的 EAP 报文。首先交换机和认证客户端之间会进行 EAP 的协商，在协商完成后，会进行 802.1x 认证过程。

📖 **说明：**

在交换机上作了相应的配置之后，交换机就可以透明传输 802.1x 的 EAP 报文。相应的配置信息请参见交换机用户手册配置指导分册的“AAA 及安全协议配置”模块。

802.1x 认证的数据交互过程如下图所示。



图A-9 交换机透明传输 802.1x 认证报文

关于 802.1x 认证更详细的描述，可以参见 IEEE 802.1x 标准文档、RFC2869 和支持 802.1x 的华为网络设备的用户手册。

在经过了 802.1x 认证之后，用户就被允许访问网络资源，进行后续的上网过程，同时运营商也可以对该用户进行计费操作。

📖 **说明：**

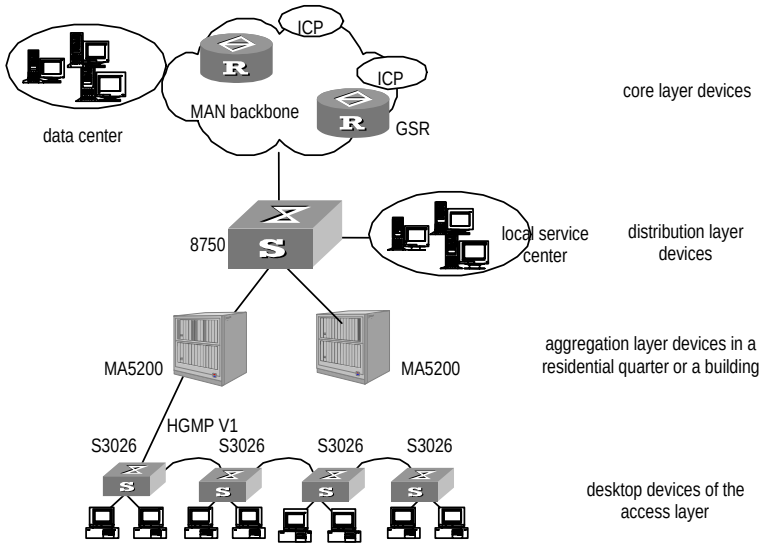
除了提供 802.1x 认证，华为公司还提供 portal 认证、Web 认证等方式对用户进行认证。对于这些认证的数据交换过程，可以参见相关的技术文档和支持这些功能的华为网络设备的用户手册。

下面引入一个问题：如果需要限制某些用户的上网时段；限制某些用户的带宽，例如某些用户虽然使用 10Mbit/s 的带宽和交换机相连，但是他只付了 2Kbit/s 带宽的费用。对于运营商来说，该怎样去控制网络设备适应不同的需求呢？可以通过人工关闭某个交换机的端口、到每个交换机上进行限制带宽的配置等满足这些需求，但是这样费时费力，不能满足可运营、可管理网络的需求。华为公司充分考虑了这种需求，为运营商提供了电信级的用户管理和运营能力。下面简单介绍在这种网络中的数据转发流程。

2. 在可运营、可管理网络中对用户进行集中管理

(1)　MA5200＋S3026 的组网方案

第一种方案：对前面小区的组网重新设计，利用“MA5200”加“S3026”进行组网。S3026 和 MA5200 之间的通信可通过 HGMP V1 来承载。通过 HGMP V1，MA5200 可以实现对 S3026 的集中管理；MA5200 可以通过上行高速口直接连接到 8750 或者直接连接到城域网或骨干网，多台 S3026 直接挂在 MA5200 的 100Mbit/s 以太网光接口上，然后 S3026 连接到小区用户桌面计算机上。



图A-10 通过 MA5200 实现电信级的用户管理和运营

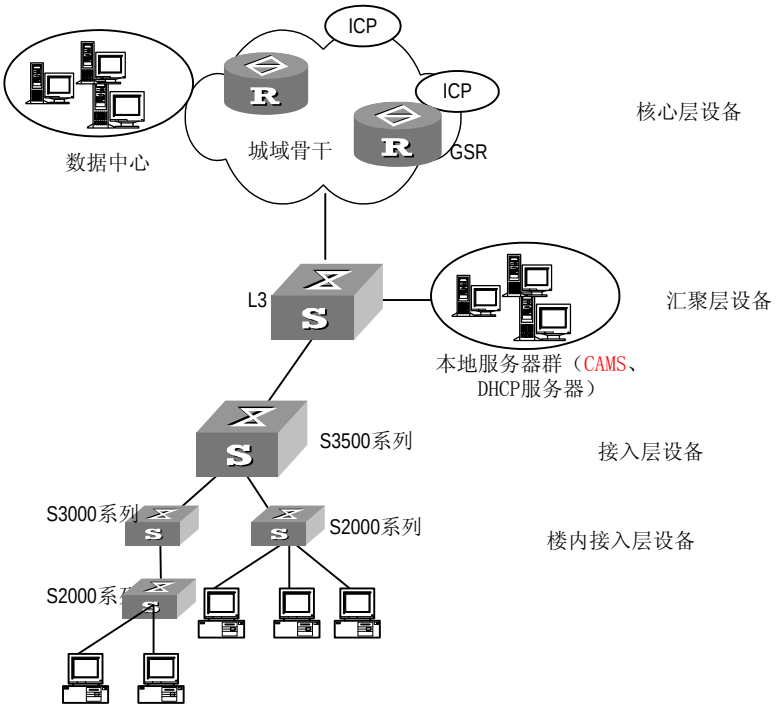
MA5200 可以对上网用户进行认证、授权、计费的功能，同时可以对用户使用的带宽、上网的时段进行控制，对用户进行按时按量计费，使之既有以太网接入经济、成熟的特色，又有电信级的用户管理和运营能力。

MA5200 可以在本地实现对上网用户进行认证、授权、计费，也可以把认证、授权、计费交给 RADIUS 服务器来做，这由 MA5200 上的配置所决定。

MA5200 将对用户的上网时段限制等配置信息通过 HGMP 报文直接下发到用户相连的 S3026 交换机上，并可以实现基于用户的带宽限制，从而实现对所有用户的集中管理。MA5200 的详细配置可以参见 MA5200 的用户手册。

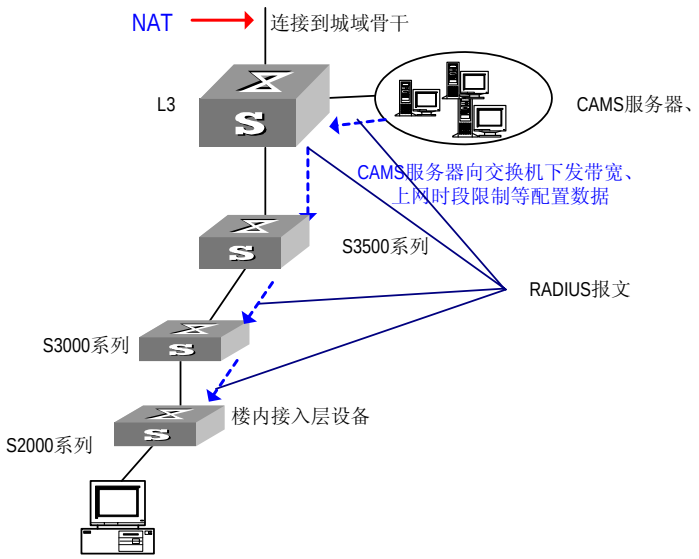
(2)　使用 CAMS 构造电信级交换网络

第二种方案，利用华为公司的 CAMS 服务器，实现电信级的可运营、可管理的交换网络。



图A-11 利用 CAMS 构造电信级交换网络

在小区的数据中心增加 CAMS 服务器。CAMS 服务器可以把对用户的上网时段限制等配置直接下发到各个交换机上，并且可以实现基于用户的带宽限制，从而实现对所有用户的集中管理。



图A-12 CAMS 下发配置数据的过程

📖 **说明：**

CAMS 下发的配置数据在缺省情况下封装在标准的 RADIUS 协议报文中。如果用户在交换机上作了如下配置之后：
Quidway(config-radius-huawei)#server-type huawei
CAMS 下发的配置数据将封装在华为扩展的 RADIUS 协议报文中。

同时 CAMS 服务器可以实现对用户的认证、授权、计费等功能，使网络具备电信级的可运营、可管理能力。对用户进行 802.1x 认证的过程和前面讲述的认证过程是一样的，只不过 RADIUS 服务器被 CAMS 服务器代替。

📖 **说明：**

CAMS 支持多设备，多业务，多协议，实现宽窄带的一体化的管理。它不会仅对一个小区网络提供服务，它在网络中的位置应该位于城域网的数据中心。此时它向交换机下发配置数据和 802.1x 认证的过程基本上没有改变，只是数据在小区和城域网之间需要进行一次 NAT 转换。

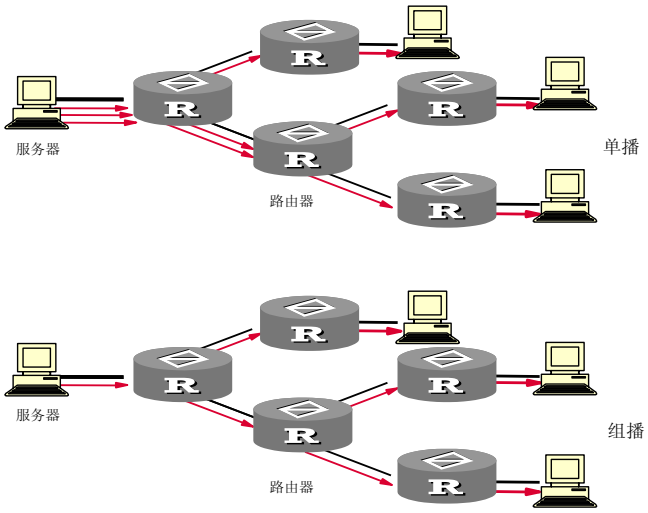
A.4 组播业务

宽带网络支持视频点播等业务。如果仅在网络层实现组播，则组播数据将在二层交换网络中进行广播，既浪费了大量的带宽，而且给网络设备带来巨大冲击，很可能会引起网络设备瘫痪。华为 Quidway 系列路由器和以太网交换机都支持组播特性，在网络层和数据链路层都实现了组播，解决了上述问题。

A.4.1 IP 组播

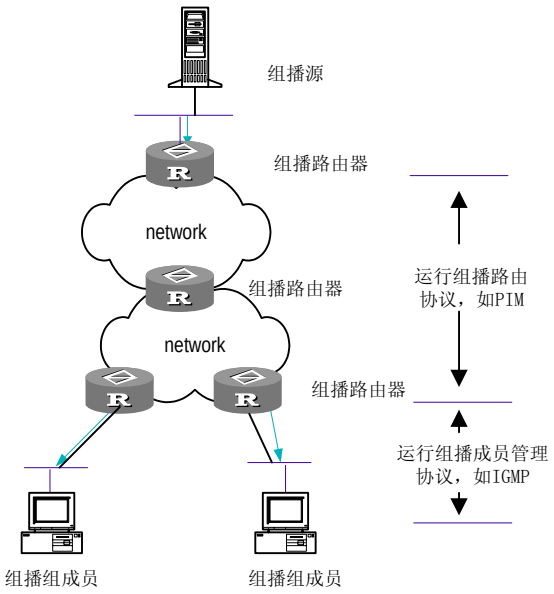
IP 组播是一种节省带宽的技术，它把一个数据流同时传送给许多接收者，组播源将需要传播的数据包发送一次，被传递的数据包在网络关键节点处不断地进行复制和分发。通过组播方式，数据包能被准确高效地传送到每个数据包接收者。IP 组播可以减少大量网络的流量。

IP 组播数据的转发和单播数据的转发如下图所示。



图A-13 单播与组播传送数据的对比

成功的进行组播需要组播路由协议和组播成员管理协议相互配合。组播路由协议负责维护组播路由信息，它跟踪、了解哪些组播报文需要在路由器之间转发，进而将组播报文发送到与组播路由器之间相连的局域网。组播成员管理协议是组播成员管理的协议，用来在 IP 主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。所有参与组播的主机必须实现组播成员管理协议，组播成员管理协议是 IP 组播路由协议的直接支持协议。



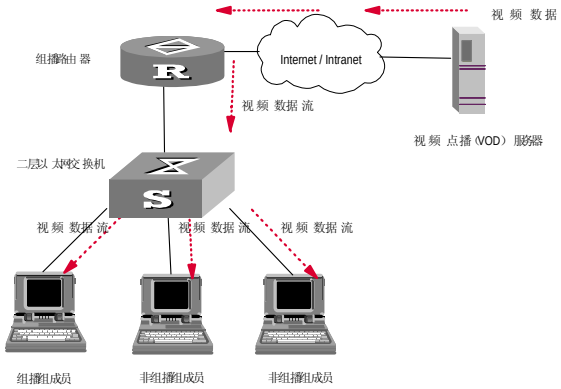
图A-14 IP 组播示意图

如上图所示，组播数据在组播路由器之间根据组播路由协议维护的路由表进行转发，在组播数据接收者和其相邻的组播路由器之间则根据组播成员管理协议维护的成员关系进行转发。

要了解组播协议更详细的描述，请参见 RFC 文档（如 RFC1112、RFC2236 等）和相关产品的用户手册。

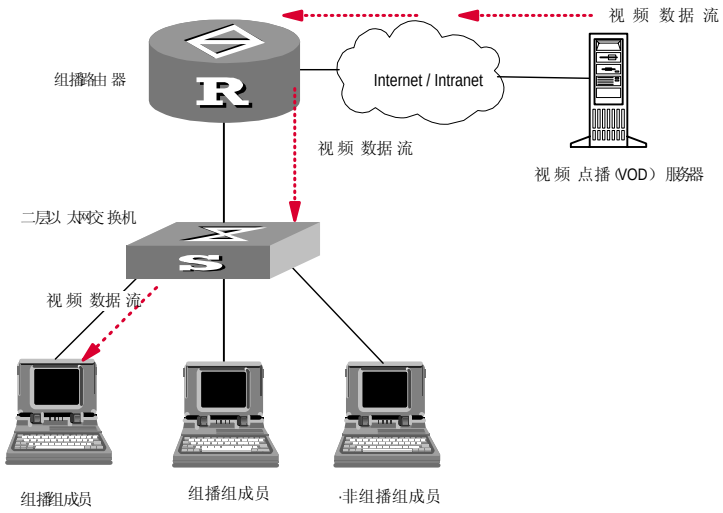
A.4.2 二层组播

在实现 IP 组播后，组播数据进入局域网后将在局域网内进行广播。这将造成网络带宽的极大浪费，同时也对网络设备带来很大冲击。华为 Quidway 系列以太网交换机实现了二层组播协议，使报文在二层也实现组播，从而解决了上述问题。



图A-15 没有实现二层组播时组播报文的传播过程

实现了二层组播以后，组播数据的转发流程如下图所示：



图A-16 实现二层组播后组播报文传播过程

二层组播协议是在交换机上维护二层组播转发表，表项由组播组、该组播组的转发端口列表等构成，这样，组播数据只会发送到连接有组播组成员的端口上。

关于二层组播的详细描述可以参见相应路由器和交换机的用户手册和相关技术文档。

附录 B 缩略语表

A		
AAA	Authentication, Authorization and Accounting	认证、授权和计费
ACL	Access Control List	访问控制列表
ARP	Address Resolution Protocol	地址解析协议
C		
CLI	Command Line Interface	命令行接口
F		
FTP	File Transfer Protocol	文件传输协议
G		
GARP	Generic Attribute Registration Protocol	通用属性注册协议
GE	Gigabit Ethernet	千兆以太网
GVRP	GARP VLAN Registration Protocol	GARP VLAN 注册协议
GMRP	GARP Multicast Registration Protocol	GARP 多播注册协议
H		
HGMP	Huawei Group Management Protocol	华为组管理协议
I		
ICMP	Internet Control Message Protocol	因特网控制消息协议
IGMP	Internet Group Management Protocol	因特网组管理协议
IP	Internet Protocol	因特网协议
M		
MAC	Medium Access Control	介质访问控制
MIB	Management Information Base	管理信息库
N		
NMS	Network Management System	网络管理系统
NVRAM	Nonvolatile RAM	非易失随机存储器
Q		
QoS	Quality of Service	服务质量
R		
RMON	Remote Network Monitoring	远程网络监视
RSTP	Rapid Spanning Tree Protocol	快速生成树协议

S		
SNMP	Simple Network Management Protocol	简单网管协议
STP	Spanning Tree Protocol	生成树协议
T		
TCP/IP	Transmission Control Protocol/ Internet Protocol	传输控制协议/互联网协议
TFTP	Trivial File Transfer Protocol	简单文件传输协议
TTL	Time To Live	生存时间
U		
UDP	User Datagram Protocol	用户数据报协议
V		
VLAN	Virtual LAN	虚拟局域网
VOD	Video On Demand	视频点播
VT	Virtual Terminal	虚拟终端
VTY	Virtual Type Terminal	虚拟类型终端，用于虚拟线路