内 容 提 要

本书是全国信息化计算机应用技术资格认证（CCAT）项目的指定教材，属于工程师级认证体系。CCAT资格认证项目设立的目的除了培养学生掌握相应专业的理论知识，注重学员动手能力、创新能力的训练外，还注重培养和提高学员的企业管理能力，为社会和企业培养既懂技术、又懂管理的复合型人才，以改变人才培养中存在的重理论轻实践、重文凭轻能力的缺陷。

本书共分为8章，包括园区网设计、多层交换机连接和基本配置、VLAN和VTP技术、生成树协议、多层交换技术、VLAN间路由、多层交换机的高可用性、组播技术等。随书配有多媒体教学光盘，方便读者实际操作，让读者在最短时间内掌握最多的知识和技能。

本书可作为计算机科学、计算机网络专业的教材，也可作为网络专业人员了解和学习网络交换知识的参考书。

图书在版编目（CIP）数据

多层交换网络设计标准教程 / 曾勋炜，鄢志辉主编；全国信息化计算机应用技术资格认证管理中心组编 . —北京：北京理工大学出版社，2007. 2

全国信息化计算机应用技术资格认证指定教材

ISBN 978 - 7 - 5640 - 0927 - 4

Ⅰ. 多…　Ⅱ. ①曾…②鄢…③全…　Ⅲ. 计算机网络 - 设计 - 资格考核 - 教材　Ⅳ. TP393. 02

中国版本图书馆 CIP 数据核字（2007）第 008187 号

出版发行/ 北京理工大学出版社

社　　址/ 北京市海淀区中关村南大街 5 号

邮　　编/ 100081

电　　话/（010）68914775（办公室）68944990（批销中心）68911084（读者服务部）

网　　址/ http://www. bitpress. com. cn

经　　销/ 全国各地新华书店

印　　刷/ 北京圣瑞伦印刷厂

开　　本/ 787 毫米 ×1092 毫米　1/16

印　　张/ 15. 5

字　　数/ 354 千字

版　　次/ 2007 年 2 月第 1 版　　2007 年 2 月第 1 次印刷

印　　数/ 1 ~ 4000 册　　责任校对/ 张　宏

定　　价/ 29. 00 元　　责任印制/ 吴皓云

目　录

前　言

为贯彻中共中央、国务院《关于进一步加强人才工作的决定》，培养高层次、高技能和复合型的社会急需人才，全国信息化计算机应用技术资格认证管理中心受人事部中国高级公务员培训中心和教育部全国高等学校计算机教育研究会的委托，组织编写了全国信息化计算机应用技术资格认证（简称“CCAT 资格认证”）项目的指定教材。CCAT 资格认证项目是全国性的 IT 培训认证项目，其主要特色是为社会培养动手能力和管理能力兼备的人才。该培训认证与在国际上享有盛誉的瑞士管理论坛（Swiss Management Forum ，简称“SMF”）已实现了国际互认。本书属于 CCAT 资格认证项目中工程师级认证体系。

随着硬件技术的发展，多层交换技术在园区网络中得到普及应用。本书适于想学习多层交换技术的读者，适用于作为培训教程，对于负责网络规划、配置和故障排除的技术人员也有参考价值。本书的配置命令以思科公司产品的 IOS 系统为例。

本书共分为 8 章，包括园区网设计、多层交换机连接和基本配置、VLAN 和 VTP 技术、生成树协议、多层交换技术、VLAN 间路由、多层交换机的高可用性、组播技术等。

本书在编写过程中力求体现下列特点：

（1）本书有大量的插图、范例和表格来帮助读者能更直观地来了解、掌握知识。

（2）内容阐述循序渐进，图文并茂、条理清楚，便于自学。

（3）配有多媒体教学光盘，使读者能在最短的时间内掌握最多的知识和技能。

（4）配有一套标准题库，该题库中的每个例子都对不同知识点进行了练习，对于读者掌握这些知识点及使用技巧都有很大的帮助。

本书是 CCAT 资格认证指定教材，适用于社会各界人士以及在校学生参加“全国信息化计算机应用技术资格认证”考试的需求，尤其适用于高等院校、大中专学校等进行课程置换，作为相关课程的教材，亦可作为计算机职业技能考试及继续教育的培训教材或自学教材。

由于作者的水平有限，加之时间仓促，书中难免存在疏漏之处，恳请各相关单位和读者在使用本书的过程中给予关注，不吝指正。

编　者

第 1 章　交换园区网和设计模型

园区网（Campus Network）主要由通过以太网连接起来的交换机和路由器组合而成。随着园区网的不断发展及技术的不断成熟，组建园区网除了传统意义上的集线器、以太网交换机和路由器之外，网络设计者还可以有更多的选择。如果仅通过使用交换机来代替现有的共享网络是不够的，这样只能在设计和层次设置正确的情况下减轻阻塞并且增大可用带宽。本章将介绍一种被称为多层交换的园区网解决方案，通过对网络的逻辑设计，可以建立具有高可靠性、高性能、易于扩展，能满足语音、视频和数据等新兴需求的网络。多层交换是将第 2 层交换功能和第 3 层路由相结合。本章还将讨论第 2 层、第 3 层和多层交换设备在园区网中的正确模型和设计。

本章涉及下列几个主题目：

- 园区网概述。
- 新兴的园区网。
- 多层交换网络模型及功能定义。
- 多层交换技术。
- 层次化设计的概念。
- Cisco 交换产品线。
- 构建区块解决方案。

通过本章的学习，对于所给定的一系列交换功能，读者将能够辨别出开放系统互联（OSI）参考模型中与它们相关联的各层。对于给定的一系列特性，读者将能够正确辨别出分级模型中的各层。对于给定的一组用户需求，读者将能够恰当地选择 Cisco 产品解决方案。

1.1　园区网概述

本节包含了对传统园区网的综合介绍，并描述了导致园区网络运行方式改变的一些主要问题和解决方案，在这一节中还讨论了园区网络流量模式的演变。

1.1.1　传统园区网络

“园区”是指一栋建筑物或一组建筑物，它（们）连接到一个由多个局域网（LAN）组成的企业网上。“园区”也可以被进一步定义为在一个固定地理区域的一个公司或一个公司的一部分。

园区网环境的一个明显特征是：拥有该园区网的公司通常也拥有该园区网内所有的物理线路。园区网的拓扑结构主要是一种将建筑物或建筑群中的末端系统连接起来的局域网结构。园区网一般采用局域网技术，例如以太网、令牌环网和光纤分布式数据接口（FDDI）。图 1–1 所示为园区网的一个示例。

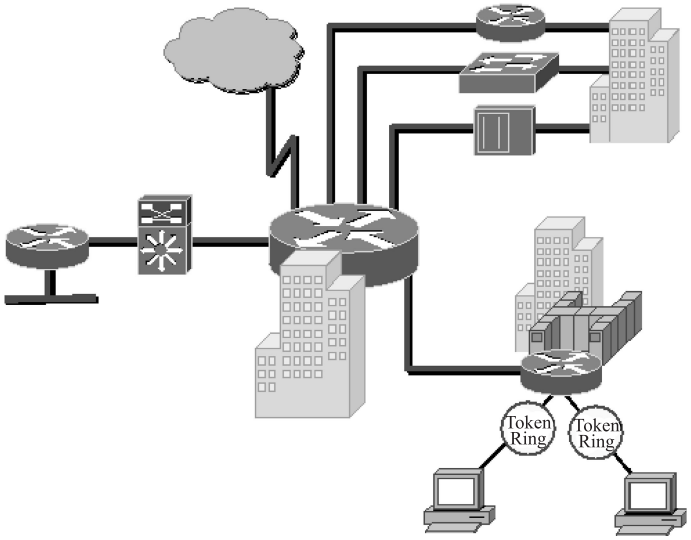

图 1–1 传统园区网

网络设计人员一般都想应用这样一种园区网设计方案，即能够获得平滑的、运行于现有物理线路上的、运行速度最快的功能构架。当今网络设计人员和管理人员所面临的主要挑战是如何让他们的园区 LAN 保持高效运行。要实现这一目标，就必须要了解、实施并管理整个网络的数据流量。

最初，园区网仅由一个局域网组成，新用户就被加到这个局域网上。由于受距离所限，园区网通常被限制在一栋建筑物或相隔很近的几栋建筑物中。LAN 是将设备连接起来的一个物理网络。在以太网的情形中，所有设备共享 10 Mb/s 的半双工数据通道。出于以太网所用载波检测多路访问/碰撞检测（CSMA/CD）机制的原因，这整个局域网被认为是一个碰撞域。

在过去几乎不考虑为用户提供访问骨干网的要求。由于受以太网的限制，物理上邻近的用户被连接到单个接入设备上，以减少与骨干网的接头数量。尽管集线器可以满足这一要求并变成了网络多路介入的标准设备，但不断增加的用户需求很快影响到网络的运行性能。

1.1.2 传统园区网问题

传统的园区网存在的两个主要问题是可用性和性能。这两个问题都受有效网络带宽的影响。在单一的冲突域中，帧对 LAN 中所有的设备来说都是可见的，并且也是自由碰撞的。

使用多口的第 2 层设备（例如网桥和交换机）把 LAN 分为若干独立的冲突域，同时只把第 2 层的数据帧转发到目的地址所在的段上。例如，一个 24 口的交换机有 24 个冲突域。因为第 2 层设备的端口将 LAN 划分为相互独立的物理网段，所以它也能够帮助解决以太网的长度限制问题。

但是，因为网桥只看数据帧中所含的 MAC，具有广播媒体访问控制（MAC）地址的帧仍然能够在全广播中传播。同时单个网络的第 2 层设备可能会出现故障，产生超长数据帧、错误帧使整个网络充斥着“噪声”，这都可能让网络不可用。这就是引入路由器的原因，如图 1–2 所示。很多应用使用广播来通信，如 IP 地址解析协议（ARP）、NetBios 名字请求、IPX

协议、DHCP 协议等。

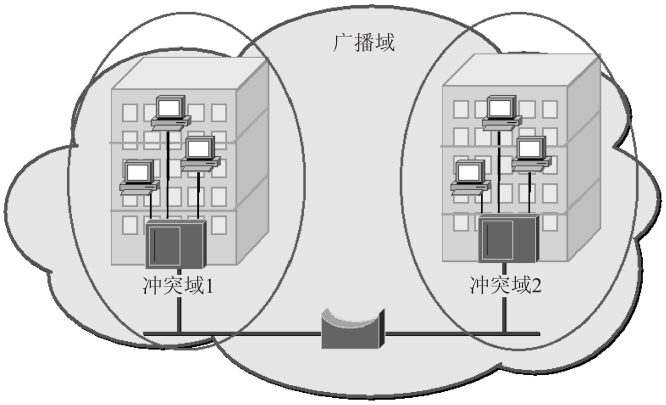

图 1–2　桥划分冲突域，无法划分广播域

因为路由器工作在开放系统互联（OSI）模型的第 3 层，所以能够判断数据流到的网段源和来自哪个网段。路由器不转发广播信息，所以如果由于大量的广播信息形成的噪声使得一个逻辑网段遭到破坏的话，路由器不会把这样的广播信息转发到其他的逻辑网段上去。

轮询网络（Poll Network）上设备的状态或可用性的流量以及广播网上设备的状态或可用性的流量可能会影响到网络的性能。轮询网络的两种常见的广播类型是 IP 地址解析协议（ARP）请求和 NetBios 名字请求。这些广播信息通常是在整个子网内部传播，并且希望目的设备直接应答到广播。

除了广播之外，多播流量可能会消耗掉大量的带宽。多播流量流向一些特定的用户组，如果部署正确，它实际的消耗等同于广播流量的带宽。依据多播组内用户的数量或包含在多播分组中的应用数据类型，多播流量可能会消耗更多（即使不是全部）的网络资源。多播应用的一个例子是 Cisco 的 IP/TV 解决方案，它利用多播分组传输如视频和音频的多媒体信息。

随着网络的发展，网上的广播流量也随之增长。过多的广播信息会减少终端用户获得的网络带宽，同时迫使终端用户节点在不必要的过程上浪费 CPU 周期。在最坏的情况下，广播风暴会独占带宽而使网络宕掉。

有两种方法可以解决大型交换 LAN 站点的广播问题。

- 使用路由器创建多个子网，对流量进行逻辑分段。
- 在交换网络中使用虚拟 LAN（VLAN）。

如果使用路由器的话，如图 1–3 所示，广播信息不会跨越路由器。尽管这种方法能够抑制广播流量，但是传统路由器的 CPU 不得不处理每一个分组。这样的方案可能会造成网络的瓶颈。如果一台路由器的一个端口连接的子网产生了广播风暴，这个路由器的 CPU 就可能受到干扰，以至于路由器不能及时处理分组路由到其他网络。

在本书中，VLAN（Virtual LAN，虚拟局域网）基本上与广播域定义一致。一个 VLAN 包括多个位于物理 LAN 网段和交换机端口的一组终端设备。位于同一 VLAN 中的设备的通信就像它们位于同一个 LAN 网段一样。带有 VLAN 功能的 LAN 交换机的一个重要的优点就是它们可以有效地控制广播流量和管理数据流量。在 VLAN 之间的流量传输需要路由器。尽管路由器几乎总是 VLAN 所在地方的一部分，但交换机也能阻止广播流量由一个 VLAN 影响到另一个 VLAN。

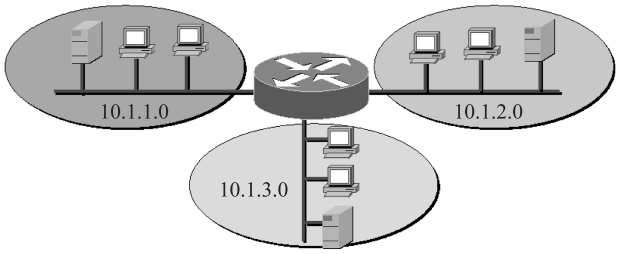

图 1–3　路由器建立独立的广播域

1.1.3　网络流量 80/20 规则

要设计和组建一个成功的园区网，必须对处于使用中的应用所产生的流量以及用户群体的流量有一个详细的了解。网络上的所有设备将产生需要在网络中进行传输的数据。每一台设备都将涉及到以不同的模式和输入产生数据的应用。不同的应用，如电子邮件、打印、文件传输和 Web 应用，带来了从源到目的可预测的流量模式。然而，新的应用如视频点播、网络电视和 IP 电话，有一个动态的用户群体，从而使得流量模式难以预测和建模。

理想情况下，有共同兴趣或工作方式相似的最终用户应被放在同一个逻辑网络中，他们经常访问的服务器也应放在这同一个逻辑网络中。出于逻辑网络的定义，这些工作组中大多数的流量被限制在这个本地网段。这样可以减少网络主干的负载。

80/20 规则是指：在一个设计恰当的网络环境中，一个给定的网段上 80%的流量是本地的，不超过 20%的网络流量需要通过主干。网络主干发生阻塞则说明流量模式没有符合 80/20 规则。在这种情况下，网络管理员不需添加交换机或者对集线器进行升级，可以通过下面几种方法之一来改善网络性能。

- 将资源如应用、软件程序和文件从一台服务器转移到另一台，将流量限制在工作组本地。
- 如果不是物理地转移用户，就逻辑地转移用户，以使工作组能更准确地反映实际的流量模式。
- 添加服务器以使用户可以在本地进行访问而不必通过网络主干。

如图 1–4 所示，80/20 规则说明在一个设计适当的网络中，给定网段的 80%流量都是本地的，不超过 20%的网络流量需要在主干网上传输。

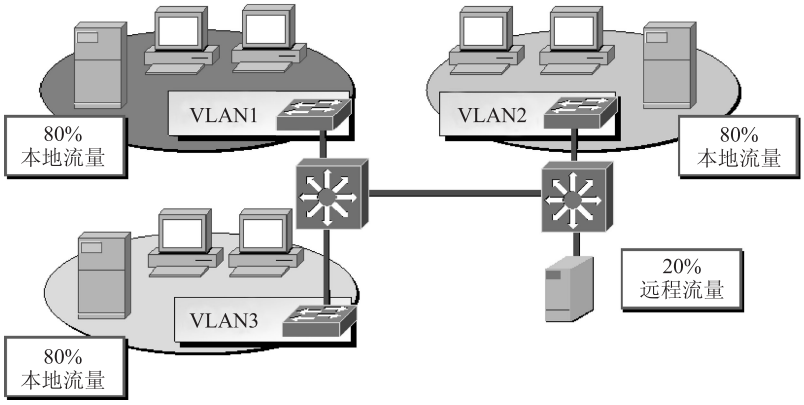

图 1–4　80/20 流量模型

1.1.4　当前的园区网络

现在大多数的园区网由两部分组成：局域网交换机和路由器。通过生成较小的第 2 层广播域并用第 3 层功能将它们链接起来，网络管理员可以过滤广播数据流、互联多协议工作组，并提供一定安全级别的数据传输。

网络上的设备和运行的相关应用软件都会产生数据流量。用户网络上可能至少有几种典型的应用，例如文字处理、文件传输和电子邮件等。这些应用不需要太多带宽，而且它们的流量模式也比较直观。

但是，新兴的园区局域网需要的应用比这些基本应用多得多。多界面应用，比如桌面出版、电视会议和 WebTV 多点广播节目，正在变得日益普及。这些应用的特征并不总是能容易地进行预测。

1.1.5　网络流量 20/80 规则

流量模式正向着现在被称为 20/80 的模型转换。在 20/80 模型中，只有 20%的流量是在本地工作组局域网的，而 80%的流量需要流出本地网络。

有两个因素导致了流量模式的改变。

- 通过基于 Web 的计算，比如 Internet 的应用，一台 PC 可以既是信息的接收者又是信息的发布者。其结果是信息可以来自网上的任何地方，这样就会生成很多必须穿过子网边界的流量。用户通过使用超链接可以透明地在整个企业网络的各服务器间跳来跳去，而不需要知道数据所在的具体位置。
- 导致本地通信减少的第二个因素是向服务器整合的转变。企业正在部署集中式的服务器机群，因为这样可以降低持有成本，提高安全性，同时也易于管理。所有从客户子网到这些服务器的数据流都必须通过园区网主干。

流量模式的变化意味着现在流量的 80%必须通过第 3 层设备。因为路由是 CPU 密集型进程，所以进行第 3 层处理的点可能产生网络瓶颈。流量模式的这种变化要求网络的第 3 层运行能力要与第 2 层运行能力相匹配。

新的 20/80 规则使网络管理员管理 VLAN 发生困难。网络管理员不希望花费时间跟踪流量模式并重新设计网络。因为 VLAN 的创建是基于大多数流量在工作组内的前提，末端站点需要在同一广播域中才能充分利用交换架构。

对于新的 20/80 规则，末端设备需要访问多个 VLAN，然而，这些末端设备仍然需要在它们当前的 VLAN 内运行。

随着当前和将来的流量模式不断远离传统的 80/20 规则，更多的数据流必须要在子网和 VLAN 间进行传输。同时，对特定的设备的访问也需要进行控制。要实现这些功能，网络中就需要路由技术。

这要求人们考虑按新的园区网模型，对传统网络进行重新设计。如图 1–5 所示，在 20/80 流量模式中，只有 20%的流量保留在工作组所在的 VLAN 中，而 80%流量都要离开本地网络去其他的 VLAN。

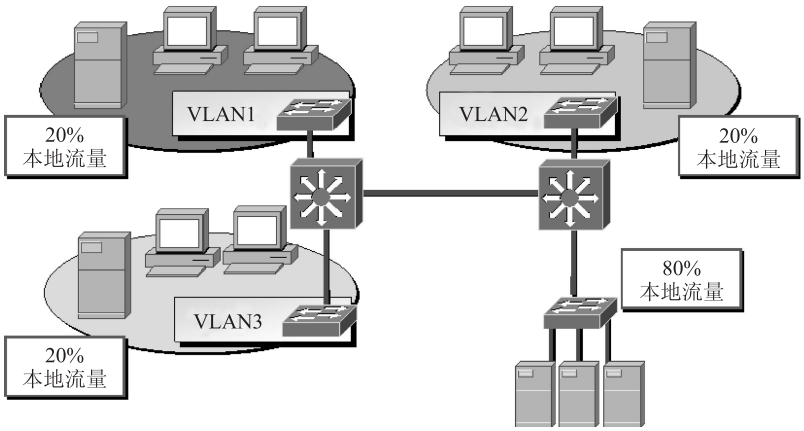

图 1–5　20/80 流量模型

1.1.6　园区网结构发展的需求

现代园区网有很多共同需求。在设计任何园区网时都必须考虑以下这些需求：

- 快速收敛。
- 确定的路径。
- 规模和吞吐量具有可扩展性。
- 集中存储。
- 20/80 原则。
- 多协议支持。
- 组播支持。

这些需求指导用户在设计网络时做出决定。下面逐一阐述各个需求。

1. 快速收敛

收敛是第 2 层和第 3 层交换机利用第 2 层和第 3 层协议来适应网络变化的机制。网络变化可能是链路断开、路由器崩溃、网桥崩溃或加入了新的路由器、链路和网桥。随着网络的扩展，链路、主机和路由器的数目也逐渐增长，从而使网络中随机的改变更有可能发生。

无论如何，如果网络按计划发生改变或突然发生改变，网络本身要有能力迅速适应这种变化。这种能力使网络具有可扩展性并使网络在发生意外期间瘫痪时间最短。

2. 确定的路径

确定性是网络的优点。不管是在流量模式、响应时间、抖动（Jitter）上，还是在正常运行时间上，用户和网络管理员都要求具有确定性。保证确定性的最大模块就是逻辑拓扑结构，

强迫流量按事先预定的链路集流动。这种事先预定性使路径选择具有确定性。

如果两台主机能通过多条路径进行通信，路径的选择就是随机的，两条路径之间的响应时间也会相差很大。在随机选择的路径不正常时，管理人员就要解决这个问题。这一征兆不会很快使管理人员面临导致不必要的瘫痪时间及网络管理压力等问题。确定的路径使网络性能有确定性并有助于减少排除疑难问题的时间。

3. 规模和吞吐量具有可扩展性

网络通常与用户数量和个人用户需求同步扩展。这意味着用户的网络设计必须能处理连接数目的增长，也要能处理更多的网络内部链路带宽。

如果网络设计时没考虑这一点，当用户碰到意外的情况时，不改变基本设计就无法处理新链路的加入。这样不仅花费大量的时间，而且花费大量的金钱。

4. 集中存储

在多用户计算的起步阶段，主机是核心设备，容纳了大部分存储和处理能力。由于 PC 的出现和快速发展，每张桌子上一台 PC 成为可能，这种完全分布式的计算模型所遇到的问题是信息系统的不一致性和管理的不一致性。这种新的需求类似于主机的应用管理。与用户从远程终端访问主机不同，用户现在从桌面访问服务器或服务器群。这种对文件和应用程序管理的集中化方法要求所有用户对该资源有足够的访问权限。

5. 20/80 原则

过去，网络中发出的流量有 80%保持在网络内部，这种估计是恰当的。这意味着网络或工作组局域网之间的带宽可以宽一点，而连接其他网络或工作组的带宽可以窄一点。

由于现代园区网的灵活性和集中存储，这条原则已经彻底改变了。现在，恰当的估计应该是，网络中发出的流量 20%保留在网络内部，其余的就到达其他网络，这意味着网络之间的带宽必须很宽且具有可扩展性。

6. 多协议支持

虽然现在流行 IP 协议，但也必须支持其他老协议，包括 Novell IPX 和苹果计算机的 Apple Talk。这些协议都很少使用，因为新的网络产品大部分都是以 IP 为中心的。即使考虑纯 IP 网络，IP 上面也会有许多新的协议，这些协议包括多媒体协议、组播协议和路由协议。要使用户网络设计的生命周期尽可能长，必须支持这些协议。

7. 组播支持

由于电信远程会议、视频点播和新网络协议等新应用的出现，组播支持是必需的。点到部分的概念提供了一种方法可以大量减少网络中不必要的流量，并增加了可靠性和对新的多媒体电信远程会议应用的响应。不考虑组播的新网络设计很可能效率不高并且将来需要进行代价昂贵的升级。这里所讨论的需求列表也因为对园区网的新期望、新的应用功能和新的网络功能的出现而不断变化。该列表肯定会增长，但现在这些需求是设计园区网的基本目标。

1.1.7 园区网结构

不断增长的用户要求和复杂的应用迫使网络设计人员把精力集中在网络的流量模式上。网络不再仅仅基于用户的数量来划分子网。运行全局应用的服务器的出现也直接影响了网络的负载。整个网络上的更高流量负载导致了对高效路由和交换技术的需求。

在新的园区网模型中，流量模式决定最终用户所需服务的定位，这些服务可以分为3类：

- 本地服务。
- 远程服务。
- 企业级服务。

1. 本地服务

本地服务是指提供服务的实体与用户处于同一个子网，也即同一个 VLAN 中。本地服务保持在网络的特定区域内。来往本地服务的数据流被限制在服务器、交换机和最终用户之间。本地数据流不进入网络主干或通过路由器。

为了服务本地化数据流，第 2 层交换机走向网络的边缘并进入接线柜。这些交换机将最终用户设备和服务器连接成为共同工作组。

2. 远程服务

远程服务是指提供服务的实体在地理位置上可能靠近最终用户，但不与该用户处于同一子网或 VLAN 中，来往远程服务的数据流可能通过也可能不通过网络主干。因为这些服务与发出请求的最终用户不在同一个子网或 VLAN，所以远程服务的数据流将穿过广播域边界。因此，交换机要连接到第 3 层设备以容许数据流通过广播域边界。路由器也控制通过网络主干的数据流类型。

3. 企业级服务

企业级服务是指对所有用户通用的服务，比如电子邮件、Internet 访问和电视电话会议等。因为所有用户都需要访问企业级服务，所以一般将这些服务器和服务都放在距网络主干很近的一个独立的子网上。因为企业级服务存在于最终用户的广播域之外，所以需要通过第 3 层设备来访问这些服务。企业级服务可能通过也可能不通过第 2 层交换机进行分组。

将企业级服务器放在靠近主干的地方可以确保每个用户的距离都相同，但是，这也意味着所有到企业级服务器的数据流都要经过主干。

图 1–6 展示了上面所述的三种服务，以及流量模式怎样决定这些服务的位置。

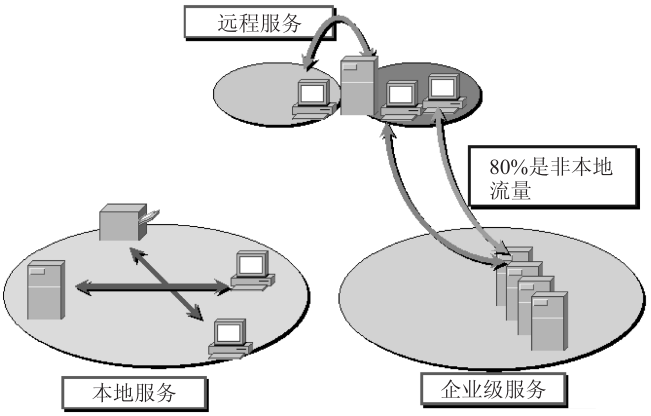

图 1–6　新兴园区网结构

1.2　交换技术

由于新兴的 20/80 规则，网络管理员想利用交换技术的高吞吐率，同时还要在网络中保留第 3 层路由功能。因此，需要一种新模型来支持这些需求，这种模型采用了为第 2 层、第 3 层和第 4 层功能提供交换技术的概念。

1.2.1　分层术语

大多数通信环境所用的模型是将通信功能和应用处理程序划分到不同的层，每一层都执行某项具体的功能。本教材集中在这一模型的第 2，3 和 4 层。图 1–7 给出了基本的分层术语。

各层通过它自己层的协议与另一系统中的同级分层进行通信。每一层的协议在同级分层之间交换被称为协议数据单元（Protocal Data Unit，PDU）的信息。各层可以对它的 PDU 使用一个更具体的名称。表 1–1 给出了第 2，3 和 4 层具体 PDU 的示例，以及处理这些 PDU 的设备类型。

表 1–1　OSI 分层模型对应的 PDU 和设备类型

模型层	PDU 类型	设备类型
数据链路层（第 2 层）	数据帧	交换机/网桥
网络层（第 3 层）	数据包	路由器
传输层（第 4 层）	TCP 数据分段	TCP 端口

各同级分层协议使用其下层的服务。这样传输控制协议（TCP）数据分段就被封装在第 3 层数据包中，第 3 层数据包又被封装在第 2 层数据帧中。各层设备只负责处理其所负责 PDU 的头部，如图 1–7 所示。

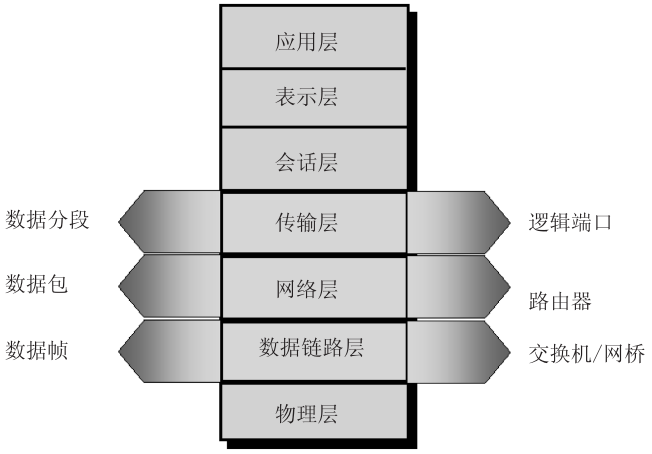

图 1–7 ISO 分层术语

1.2.2 硬件交换与软件交换

业界术语“硬件交换”表示通过专门的硬件组件（ASIC）在任何层（第 2～7 层）处理数据包的行为。ASIC（Application-Specific Integrated Circuit，应用专用集成电路）通常能够达到线速的吞吐量，而不会因为某些高级特性，例如 QoS（Quality of Service，服务质量）标记、ACL（Access Control List，访问控制列表）处理或 IP 重写等，降低性能等级。

术语“硬件方式”、“使用 ASIC”和“基于硬件”也用于描述“硬件交换”。此外，本书贯穿全文交替使用这些术语。MLS（Multilayer Switching，多层交换）是用于描述硬件交换的另外一个常用术语。术语 MLS 可能会带来一定程度的混淆。在如今的术语中，MLS 描述使用 ASIC 的高级特性，例如 NAT（Network Address Translation，网络地址转换）、QoS 和访问控制等，能够达到以线速（所有端口同时发送信息流，全双工，接口的最高速率）对数据帧进行路由或交换的能力。第 5 章将更加详细地讨论 MLS 术语。对于接下来的几章，MLS 和硬件交换都只表示以硬件方式对数据包和数据帧进行高速的交换与路由选择。

与通过 CPU 实现的传统的 “软件交换”相比较，硬件（或硬件交换）中交换和路由选择通信流量的速度更快。很多 ASIC（特别是第 3 层路由选择所使用的 ASIC）需要使用被称为 TCAM（Ternary Content Addressable Memory，三重内容寻址内存）的专用内存，并且还需要结合使用数据包匹配算法来获得高性能，然而 CPU 只是用更高的处理速度来达到更高程度的性能。

一般而言，ASIC 能够比 CPU 获得更高的性能和可用性。此外，ASIC 易于在交换体系结构中扩展，而 CPU 则不能。不仅 Supervisor Engine 上的 ASIC 能够以分布方式集成到硬件交换数据包中，而且 Catalyst 交换机线路模块的 ASIC 也能够以分布方式集成到硬件交换数据包中。

1.2.3 第 2 层交换

第 2 层交换严格集中在数据链路层，也就意味着第 2 层交换机只能根据 MAC 地址对数

据包进行交换。第 2 层交换机能够方便地增加网络带宽和端口密度。术语第 2 层交换暗指交换机所转发的帧不会被采用任何方式而修改。然而，某些第 2 层交换机的确能够修改帧，例如支持第 4 层 QoS 标记和访问控制的 Catalyst 2950 和 2970 交换机。一个第 4 层 QoS 标记的例子是，根据 TCP 报头中的 TCP 端口号在 IP 报头中标记 DSCP（Differeniated Services Code Point，差异化服务编码点）比特。

由于众多因素的作用，早期的第 2 层交换机受到网络扩展性的限制。基于上述原因，老式的第 2 层交换机中全部网络设备都必须属于相同的子网，并且为进行地址解析而交换广播数据包。对于被分组为共同交换广播数据包的网络设备，它们组成单个广播域。第 2 层交换机在整个广播域中是扩散未知的单播、多播和广播的通信流量。基于上述原因，广播域内的全部网络设备需要处理所有的扩散通信流量。伴随着广播域规模的扩大，由于需要执行任务来处理很多不必要的超过少数的早期的第 2 层交换机。

然而，当前流行的和大部分老式的 Cisco Catalyst 交换要都支持 VLAN（Virtual LAN，虚拟局域网）。VLAN 将通信流量分段为单独的广播域，也就是划分为单独的子网。VLAN 克服了基本第 2 层网络的若干限制。本书将在第 3 章中讨论 VLAN 的相关内容。

第 2 层交换是基于硬件的桥接。在第 2 层交换机中，ASIC 负责处理帧转发。此外，第 2 层交换机将增加带宽的能力给予配架，而无需给网络增加不必要的复杂性。在第 2 层，当帧在第 1 层接口（例如快速以太网到 10 Gb 以太网）之间传输的时候帧内容不需要进行任何修改。

简单地讲，目前的第 2 层交换机具有的网络设计属性如下：

- 设计接近线速的性能。
- 使用高速专用 ASIC。
- 低延迟。
- 可扩散第 3 层功能，例如 IGMP（Internet Group Management Protocol，Internet 网组管理协议）监听和 QoS 标记。
- 大型网络中有限的可扩展性，没有第 3 层边界。

某些第 2 层交换机能够为 QoS 标记而进行数据包重写。例如，Cisco Catalyst 2950 交换机能够使用第 2 层 CoS（Class of Service，服务类别）值或第 3 层 DSCP 值标记入口帧。

1.2.4　第 3 层交换

路由器和第 3 层交换机对数据包交换操作的主要区别在于物理的实施。在通过的路由器上，一般由基于微处理器的引擎执行数据包交换的操作，而第 3 层交换机通过硬件执行数据包交换的操作。

第 3 层设备，比如路由器或第 3 层交换机，执行两个基本的功能。第一个功能是根据在第 3 层地址中找到的信息来决定路径。这是通过用路由选择协议建立路由表而完成的。路由表用来决定数据包怎样通过网络。必须执行的第二个功能是被称为“包交换”的工作。包交换过程包括重写第 2 层信息，减少生存时间（TTL）值，以及将数据帧传送到下一个接口。该包交换过程不应该与基本的第 2 层交换混淆。

第 3 层交换包括第 3 层路由选择功能。目前的很多第 3 层 Catalyst 交换机都能够使用选择协议来制定最优的转发决策，例如 BGP，RIP，OSPF（Open Shortest Path First，开放式最

短路径优先）和 EIGRP（Enhanced Interior Gateway Routing Protocol，增强型内部网关路由选择协议）等。第 3 层 Catalyst 交换机不仅能够执行 PIM 多播，而且还能够使用 HSRP（Hot Standby Routing Protocol，热备份路由选择协议）或 VRRP（Virtual Router Redundancy Protocol，虚拟路由器冗余协议）获得冗余。本书后续章节将讨论这些第 3 层特性。

对于 Catalyst 交换机产品，Cisco 当前有两种主要的第 3 层交换实施方案：多层交换和 Cisco 快速转发（CEF）。本书会较深入地介绍多层交换和 Cisco 快速转发（CEF）。

因为是被设计来高性能地处理局域网流量的，所以第 3 层交换机可以放在网络中的任何地方，经济有效地取代传统的路由器。

第 3 层交换是基于硬件的路由选择。通过提供路由选择域，第 3 层交换克服了第 2 层可扩展性不足的缺点。ASIC 和其他专用电路负责处理第 3 层交换机中的数据包转发。第 3 层交换机对数据包的处理程序就如同传统路由器所为，其中包括如下工作：

- 根据第 3 层信息确定转发路径。
- 通过第 3 层校验和验证第 3 层数据包的完整性。
- 验证数据包的 TTL（Time To Live，生存时间）是否到期并递减。
- 在 IP 重写的过程中，更新第 2 层 CRC。
- 处理并响应数据包中的任何选项信息，例如 ICMP（Internet Control Message Protocol，Internet 控制消息协议）记录。
- 更新 MIB（Management Information Base，管理信息库）中的转发统计数据。
- 如果需要的话，实施安全控制和 QoS。

第 3 层路由选择要求具有数据包重写的能力。数据包重写发生在任何的路由边界。

1.2.5 第 4 层交换

第 4 层交换是指考虑了应用基于硬件的第 3 层路由。数据包包头中的信息通常包括第 2 层和第 3 层地址以及第 3 层协议类型，加上与第 3 层设备有关的更多的字段，例如 TTL 和校验和等。数据包中也包含通信主机中有关高层的信息，比如协议类型和端口号。

第 4 层交换的简单定义是：不仅基于 MAC 地址或源/目的 IP 地址，同时也基于这些第 4 层参数来作出转发决定的能力。在 TCP 或用户数据报协议（UDP）流中，应用类型被作为端口号编码在数据分段的头中。第 4 层交换对于设备厂商来说是中性的，甚至当被添加到原先已经存在的网络环境中也是有益的。

Cisco 路由器可以根据第 4 层信息来控制流量。控制第 4 层流量的一种方法是采用标准的或扩展的访问控制列表。另一种方法是通过 NetFlow 交换来提供流的第 4 层统计。

最后，当执行第 4 层功能时，交换机/路由器读取 TCP 和 UDP 字段以决定数据包所承载信息的类型。网络管理员可以设置交换机根据应用来对数据流划分优先级。这项功能使网络管理员能够为最终用户定义服务质量（QoS）。当用于 QoS 目的时，第 4 层交换意味着电视会议应用比电子邮件可能会得到更多的带宽。

如果策略规定根据应用对流量进行细化控制，或者需要按应用进行流量统计的话，第 4 层交换是必需的。