摘 要

本学位论文首先简要介绍了MPLS、VPN和MPLS VPN技术的发展情况，对MPLS VPN技术的原理及用户需求进行了详细分析，针对用户需求给出各种不同的组网解决方案；接着对ZXB10多业务路由交换机的软硬件实现进行了讨论，结合上述分析的用户需求，对于在ZXB10设备上如何实现MPLS VPN，包括系统方案、实现原理、所需设备情况、开发环境以及ZXB10 MPLS VPN具有哪些特色等，进行了详细的阐述。

论文结合总参骨干网的建设，对其组网问题进行了分析，阐述了如何用ZXB10 MPLS VPN来满足用户需求，讨论了在开局过程中遇到的困难及解决办法，总结了一些用来指导我们研发设计及工程组网的经验教训。实践表明，所研制的ZXB10 MPLS VPN能够经受实际商用系统的考验，具有较强的市场竞争力。论文作者参与了这一课题的研究，不仅提高了自己的技术水平，同时也为公司创造了良好的经济效益。

本学位论文共分八章，第一章介绍了VPN的演进，分析了其发展现状和优缺点；第二章介绍了数字通信网络技术的发展、MPLS技术的兴起及其优点；第三章主要分析了MPLS VPN现状、市场需求和各厂商的解决方案；第四章阐述了MPLS VPN原理和典型的网络结构；第五章对课题研究的物理平台ZXB10的软硬件系统进行了简要说明；第六章是本论文之重点，详细描述了如何在ZXB10上实现MPLS VPN，从分析用户需求开始，然后给出设计的依据和原则，最后阐述该分系统的详细框图和实现功能；第七章结合总参ATM骨干网建设的实践，分析了MPLS VPN如何满足实际工程需要，总结了一些行之有效的经验教训；最后是研究成果应用情况介绍和本人的收获总结。

关键词： 多协议标记交换，虚拟专用网，异步传输模式，边界网管协议，多业务路由交换机

Abstract

This thesis reviews the development of MPLS, VPN and MPLS VPN, analyzes user requirements and gives out corresponding networking solutions, especially the technical principles of MPLS VPN. After discussing software and hardware of the ZXB10 multi-service routing switch, it shows how to implement MPLS VPN using the ZXB10 devices, which includes system scheme, principles, required equipment, development environment and system features.

Analyzing the network solution of the backbone network for Headquarters of the General Staff as an example, this thesis first elaborates how to satisfy user requirements using the ZXB10 MPLS VPN, and then discusses possible difficulties and solutions during the network implementation, summarizes finally some experiences for later system design and project implementation. The ZXB10 MPLS VPN has been put to commercial use for a long time and has been proved now very competitive. Joining efforts in this topic, the author not only wishes to enhance his own technology capability, but also creates favorable economic benefits for his company.

There are eight chapters in the thesis. Chapter 1 discusses VPN evolution, current status, VPN pros and cons. Chapter 2 presents the development of digital communication network technologies, emerging of MPLS and MPLS advantages. Chapter 3 analyzes the present MPLS VPN status, market requirements and solutions of manufacturers. Chapter 4 elaborates MPLS VPN principles and typical network structures. Chapter 5 introduces software and hardware systems of the physical platform of the ZXB10. Chapter 6, the highlight of this thesis, describes in detail how to implement MPLS VPN over the ZXB10. Analysis of user requirements, system design considerations and guidelines, subsystem design block diagram, and system functions are given. Chapter 7 shows how to satisfy actual engineering demands using MPLS VPN through the example of the backbone network for Headquarters of the General Staff and summarizes some tried and true experiences. Lastly, the application of the research findings and the author's personal gains are presented.

Keywords: Multiprotocol Label Switching (MPLS), Virtual Private Networks (VPN), Asynchronous transfer mode (ATM)、Border Gateway Protocol (BGP)、VPN routing/forwarding instance (VRF)、Multi-service Routing Switch (ZXB10)

东南大学学位论文独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得东南大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名：顾曙明 日　期：2004.7.1

东南大学学位论文使用授权声明

东南大学、中国科学技术信息研究所、国家图书馆有权保留本人所送交学位论文的复印件和电子文档，可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外，允许论文被查阅和借阅，可以公布（包括刊登）论文的全部或部分内容。论文的公布（包括刊登）授权东南大学研究生院办理。

研究生签名：顾曙明 导师签名：沈兰荪 日　期：2004.7.1

第一章 VPN概述

虚拟专用网VPN（Virtual Private Networks）指的是依靠服务提供商ISP（Internet Service Provide）和网络服务提供商NSP（Network Service Provide），在公用网络中建立专用的数据通信网络的技术。在VPN中，任意两个节点之间的连接并没有传统专网所需的端到端的物理链路，而是利用某种公众网的资源动态组成的。

VPN被不太精确地定义为这样一种网络，即多个站点（site）之间的客户连接性部署在相同的底层结构之上，其访问及安全策略与专用网络相同。

VPN是运营商通过其公网向用户提供的虚拟专有网络，即在用户的角度VPN是用户的一个专有网络。对于运营商来说公网包括公共的骨干网和公共的运营商边界设备。地理上彼此分离的VPN成员站点通过用户网络边缘设备CE（Customer Edge）连接到对应的提供商边缘设备PE（Provider Edge），通过运营商的公网组成客户的VPN网络。

1.1 VPN的演进过程

最初的计算机网络是使用两种主要的技术实现的：

1. 对于要求永久连接的情况，使用租用线。
2. 对于只需偶尔连接的情况，则使用拨号。

但由于下述两个原因，实现的成本效益不高：

1. 网络中两个站点之间的数据流量随一天的不同时间、一个月的不同天、一年的不同季节而异（例如：在元旦、春节期间，超市的流量将增加）。
2. 终端用户总是要求快速响应，因此要求站点之间有很大的带宽，但租用线的专用带宽只有部分时间被使用（用户处于活动状态时）。

这两个因素促使数据通信行业和服务提供商开发并实现大量的统计复用方案，这些方案为客户提供的服务几乎与租用线相同，但由于服务提供商可以从大量的客户群中获得统计性益处，因此这些服务的价格更便宜。最初的虚拟专网是诸如X.25、帧中继FR（Frame Relay）等技术的，后来则是基于交换式多兆位数据服务SMDS（Switched Multimegabit Data Service）和异步传输模式ATM（Asynchronous Transfer Mode）等技术。

随着新技术被引入到服务提供商网络中以及新的客户需求的出现，VPN的概念变得越来越复杂。我们可以使用4种标准对VPN进行分类：

1. 从“VPN要解决的业务问题”可分：企业内部网、企业外部网和虚拟拨号专网。
2. 从“服务提供商在哪一个ISO层与客户交换拓扑信息”可分：覆盖模型和对等模型。
3. 从“在网络中实现VPN服务的第二层或第三层技术”可分：X.25、FR、SMDS、ATM或IP。
4. 从“网络的拓扑结构”可分：涵盖了从简单的集线器和机架拓扑到大型网络中的全网网络和多级层次式拓扑。

下面分别对它们进行简要描述。

1.2 基于业务问题的VPN分类

企业需要使用虚拟专网解决的问题有3个：

1. 企业内部通信（企业内部网）：通常没有通过终端主机或防火墙进行很好的保护，故用于实现企业内部通信的VPN服务需要较高的隔离性和安全性，它通常使用诸如X.25、帧中继或ATM等传统技术实现的。
2. 与其他企业的通信（企业外部网）：企业之间的通信常常是在企业的中央站点之间

进行的——通常使用专用的安全设备，如防火墙与加密装置，这些通信对服务质量的要求不那么苛刻。

3. 远程用户通常从非固定位置的地方接入到企业网络，这通常会引起安全性问题，必须使用诸如加密或一次性口令等技术在端到端的基础上解决这种问题——网络接入服务器、本地网关和 VPDN 隧道（L2F 或 L2TP）。

1.3 覆盖 VPN 模型和对等 VPN 模型

有两种 VPN 实现模型得到了广泛的应用：

1. 覆盖模型（overlay model）：在这种模型中，服务提供商给客户提供仿真租用线路。

   覆盖 VPN 模型最容易理解，因为这种模型中客户和服务提供商的职责非常清晰，可以使用大量的交换式第二层技术（包括 X.25、FR、ATM 或 SMDS）来实现：

   (1) 服务提供商给客户提供一组仿真租用线路（VC，Virtual Circuit），VC 可以是始终可用的（PVC，Permanent Virtual Circuit）或根据需要建立（SVC，Switched Virtual Circuit）。

   (2) 客户通过服务提供商提供的 VC 在 CE 之间建立路由器到路由器的通信。

   虽然理解和实现覆盖 VPN 模型相对比较容易，但它有一些缺点：

   (1) 这种模型非常适合包含的中央站点不多，而远程站点非常多的非冗余配置，但对于连接性更为复杂的配置，则管理的难度非常高。

   (2) 要正确底提供 VC 容量，则必须详细了解站点间的流量情况，而这没有现成的数据可用。

2. 对等模型：服务提供商和客户交换第三层路由信息，提供商以最优的路径中继客户站点之间的数据，而不需要客户的参与。在该模型中，PE 是一台路由器，它直接与 CE 路由器交换路由信息，它有共享和专用路由器两种模式。

   对等模型有许多优于传统覆盖模型的地方：

   (1) 路由工作（从客户角度来说）变得非常简单，因为客户路由器只与一台（或几台）PE－路由器交换路由信息，而在覆盖 VPN 网络中，邻接路由器的数量可能增加到非常多。

   (2) 客户站点之间的路由通常是最优的，因为提供商的路由器知道客户的网络拓扑，因此可以实现站点间的最优路由。

   (3) 带宽供应工作更为简单，因为客户只需规定每个站点的出站带宽和入站带宽和承诺的传送速度，而不用规定简单之间的流量情况。

   (4) 加入新站点更为简单，因为服务提供商只需提供一个站点，并修改与之相连的 PE－路由器的配置。在覆盖 VPN 模型中，服务提供商必须提供从新加入的站点到客户 VPN 中其他站的一整套 VC。

   但对等模型还是存在以下缺点，使得其难以得到广泛的应用：

   (1) 所有的客户共享相同的 IP 地址空间，使得客户无法按照 RFC1918 采用其专用的 IP 地址。

   (2) 客户无法将缺省路由插入到其 VPN 中。这种限制使得无法实现某些路由优化功能，并使客户无法通过另一个服务提供商接入到 Internet。

1.4 典型的 VPN 网络拓扑

不同的 VPN 拓扑主要是试图解决不同的业务问题，它主要可以分为以下几种类型：

1. 受覆盖 VPN 模型影响的拓扑：包括中心和辐条拓扑、部分和全网式拓扑和混合拓扑。

2. 企业外部网拓扑：包括两两相连的企业外部网和集中服务式外部网络。
3. 特殊用途的拓扑：如VPDN主干和管理网络拓扑。

1.5 VPN发展现状及优点

随着Internet在企业领域应用的不断深化，VPN作为一种廉价安全的组网方案越来越受到人们的青睐。在北美和欧洲，VPN已经是一项相当普遍的业务；在亚太地区，该项服务也迅速开展起来。

从面世到成熟，VPN经历了技术不断完善的过程。目前，市场上的VPN解决方案有四种：

1. 第一种为基于拨号的VPN——VPDN（Virtual Private Dial Networks）

目前来看，用的最多的是VPDN。VPDN指用户利用拨号网络访问企业数据中心，用户从企业数据中心获得一个私有地址，但用户数据可跨公共数据网络进行传送，可利用PPTP（Point-to-Point Tunneling Protocol）、L2F（Layer 2 Forwarding）、L2TP（Layer 2 Tunneling Protocol）实现。VPDN主要适合一些企业用户，如企业的员工在出差、旅行情况下，需要与企业建立联系，可以借助VPN，通过某一地方的拨号进入公网提供的VPN的服务器，通过认证后，确认其为该VPN的客户，然后把这个客户认证转到企业自己的认证服务器上，进行二次认证，然后通过VPN实现与企业的联系。这种方式的缺点是难以保证在公网上的服务质量。

2. 第二种为基于路由的VPN——VPRN（Virtual Private Router Networks）

VPRN对网络服务商的硬件要求很高，但由于功能很强大，大型的企业客户应用较多。企业可以利用公共数据网络建立自己的私有企业网络。用户可自由规划企业各分支机构之间的地址、路由策略、安全机制等。实现协议包括GRE（Generic Routing Encapsulation）、L2TP、VTP（VPN Tunneling Protocol）、IPSec（Internet Protocol Security）、MPLS等。例如证券公司，在全国各地有许多分支机构的企业，并且分支机构和总部间需要频繁、长期的联系，需要网络带宽保证，所以VPRN在这类用户中很受欢迎。尤其是一些应用，比如IPTV，一定要采用基于路由的有质量保证的VPRN。

3. 第三种为基于虚拟专线的VPN——VLL（Virtual Leased Line）

VLL是基于虚拟专线的一种VPN，在公网上开出各种隧道，模拟专线来建立VPN。与实际的专线相比，VLL的好处是可以节省了不少费用，但是这种专线是在公网上运用第三层协议打出来的遂道专线，其质量和稳定性都不是很好。而且在企业的分支点或联系人比较多的时候，隧道入户比较多，管理起来很困难。

4. 最后一种为基于局域网仿真的VPN——VPLS（Virtual Private LAN Service）

VPLS是在公网上用隧道协议仿真出来一个局域网，从成本上、管理上VPLS都不是一个好方案，这种基于隧道局域网仿真的方案基本上没有人用。

由于VPN提供了安全、可靠的Internet访问通道，为企业进一步发展提供了可靠的技术保障，从而受到了众多企业的关注。与实际专线比较，VPN有很多好处，可以取代传统的客户租用的专线。总的来看，VPN的优点可以归纳为以下几点：

1. 一是成本节约，行业调查公司的研究报告显示拥有VPN的企业相比起采用传统的远程接入服务器或Modem池和拨号线路的企业能够节省30%到70%的开销。当使用Internet时，实际上只需付本地电话费，却实现了长途通信。因此，借助VPN，就可以节省大量的通信费用。此外，VPN还使企业不必投入大量的人力和物力去安装和维护广域网设备和远程访问设备，这些工作都可以交给服务商。
2. 其次，接受网络服务商的VPN服务，用户不需要直接管理，通过服务商去代管，从而可以集中精力从事企业的核心业务。

3. 第三，容易扩展，并且扩展时速度非常快。比如用户租一条 2M 线路，它可以买服务商两条 VPN，当它再想从 2M 扩展到 10M，建专线需要大概两个月的周期，但如果采用服务商的服务，从 2M VPN 扩展到 10M VPN 可能只需要几分钟的时间。
4. 最后，支持新兴应用。许多专用网对许多新兴应用准备不足，如那些要求高带宽的多媒体和协作交互式应用。VPN 则可以支持各种高级的应用，如 IP 语音，IP 传真，还有各种协议，如 RSVP、IPv6、MPLS、SNMPv3 等。

1.6 VPN 在国内应用现状及缺点

VPN 在国内的应用已经起步。广东电信在这方面处于国内的领先地位，其 VPDN 业务已经覆盖省内全部地区，用户通过拨 96366 进入企业虚拟专用网。不久前，中国电信在北京开了个 VPDN 产品推介及客户咨询会，将 VPDN 业务命名为“V 信通”，向社会开放，特服号为 17979。但总的来说，VPN 在国内的应用还不普及，这主要是因为 VPN 现在还存在若干缺点：

1. 一是 QoS 的保证能否实现。隧道协议不足以保证这一点，能比较好地解决这一问题的 MPLS 方式又因为标准不统一受到影响。虽然主要的网络设备厂家各有相应设备问世，但不同厂家产品之间的互联互通经常难以实现。
2. 另一个关键问题是安全性。虽然 VPN 为了实现专网的功能在设计时就考虑了这个因素，采取了加密、认证、防火墙等手段，但目前的 VPN 最不让人放心的就是这一点。按理说，银行这样需要各地连起专用网的企业是天生就需要 VPN 的，但据说目前国内的银行还没有敢吃这只螃蟹的。与银行相比对安全性要求低一点的证券企业，也只有三峡证券等少数先行者用上了 VPN。

第二章 MPLS 概述

随着 Internet 的迅猛发展，传统路由器因其固有的局限，已成为发展的瓶颈。ATM 作为宽带综合业务数字网 B-ISDN（Broadband Integrated Services Digital Network）的最终解决方案，已被国际电信联盟（ITU-T）所接受。各发达国家都已实施了试验网计划和商用业务计划。90 年代中期以来，Internet 的骨干网和高速局域网大都采用 ATM 来实现的。IP over ATM 已成为跨电信产业和计算机产业的多年持久的热点。先后有重叠模式的 CIPOA（Classical IP over ATM）、局域网仿真 LANE（Local Area Network Emulation）和基于 ATM 技术的多协议信息传送 MPOA（Multiple Protocol Over ATM），集成模式的 IP 交换机和标记交换机等多项技术出现。多协议标记交换 MPLS（Multi-Protocol Label Switching）在综合前述技术的基础上，提出了更好的 IP over ATM 解决方案。因特网工程组 IETF 先后发布一系列关于 MPLS 的建议草案，并于 1999 年 3 月的会议正式颁布其中几个主要的草案，申请 RFC 号码，并已获得批准。

97 年吉位线速路由交换机的商品化，给 ATM 带来了冲击性的影响。在局域网上，因其明显的性能价格比优势已在取代 ATM。在 Internet 骨干网上也在逐步取代 ATM。Everything over IP 和 IP over ererything 已作为通信体制革命的重要论点出现在各种刊物上。

吉位线速路由交换机对 ATM 发展的影响不能低估，应给予充分重视，这是市场需求驱动引起的冲击，主要表现在性能价格比上。吉位线速路由交换机和传统路由器相比有以下三点改进和提高：

1. 依靠微电子技术和光技术的进步，极大地提高了线速，达到 Gbit/S，甚至可达到 Tbit/S。从而成百倍地提高了处理数据分组的能力，由传统路由器最高为几个 MP/S 提高到 GP/S 的水平。缓解了传统路由存在的瓶颈问题。
2. 改变了控制结构，将路由选择功能和数据分组地转发功能分开。前者由 CPU 完成，后者由专用集成电路 ASIC 实现，简化和提高了转发能力。
3. 引入空分交换器（Cross bar switch）。将输入数据分组经过交换机构，直通到输出队列，消除了因存储转发引起地交换时延。

吉位线速路由交换机与 ATM 相比，其优势表现在以下方面：

1. 性能价格比优势，在传送相同性能的 IP 分组时，吉位线速路由交换机单个端口的价格远低于 ATM。这是由于 ATM 要满足不同业务不同网络连接成无缝大网的要求，因而技术复杂，且长期以来这些标准未能及时制定，未能大规模得到应用，成本居高不下，许多性能、功能的优势未能体现。
2. ATM 需将 IP 分组重新封装为 ATM 信元，增加了内部开销。约占整个数据的 24%左右，相比之下，IP 分组的内部开销约占 2.3%，有明显的效率优势。

纵上所述，吉位线速路由交换机是标志 IP 网有重大进展的一项技术。已经形成对 ATM 技术的严重冲击。但吉位线速路由交换机没有从根本上解决传统路由网上存在的任何一个问题，只是用速率的提高来掩盖和缓解这些问题。随着终端数量的增加和速率的提高，这些问题仍将呈现出来。

MPLS 源自 IP over ATM 的需要。早期工作在网络层，集中于 IP（Ipv4 、Ipv6）协议，但其核心技术同样适用于其他网络层协议（如：IPX、Appletalk、Dcnet、CLWP 等）。在链路层 MPLS 没有限制于某一特定的链路层，但主要工作仍集中在 ATM 上。随着 IP 网的发展，尤其在吉位线速路由交换机上，希望由 IP/SDH/OPTICS 模式直接发展成为 IP/OPTICS（DWDM）时，MPLS 是必须应用和发展的技术。因为从 IP 到光的密集波分复用 DWDM，从层次的概念看，中间有一链路层，即用于传输、交换和转发的一层。现有适用于 IP 分组在链路层传送的技

术，只有同步传递模式 STM 的 SDH，和异步传递模式 ATM 的信元两类。MPLS 是同时适用于 SDH 和 ATM 并可适用于未来发展的任一特定的链路层制式的技术。MPLS 还蕴含着支持网络管理、流量工程、QOS（Quality of Service）和 COS（Community of Service）等各项功能。IP 必须通过 MPLS（当然也可以采用其他相应的方式）才能直接在 OPTICS 上传送的。

事实上，MPLS 已不是仅在 IP over ATM 上的一项应用技术，而是作为 L3 层和 L2 层之间的“垫层”的网络技术，作为一种体系结构在研究和发展。当前可直接应用于 ATM 网和 FR 网。进一步已成为正在研究和发展的 IP over OPTICS 的必不可少的首选技术被重视。其至有人提出 MPLS 是 ATM 的终结者。

无论如何，MPLS 和 ATM 相互是不能替代的。这是由于两者的功能定位是不可覆盖的。ATM 执行 B-ISDN 网四层参考模型的 ATM 信元层和 ALL 适配层，相当于 ISO-OSI 七层模型中的第二层链路层功能。而 MPLS 执行七层模型中第三层网络层第二层链路层之间的一个相对独立的垫层或夹心层的功能，不具备完整的链路层功能。即 MPLS 要实现链路层实际转发功能时，必须依靠特定的链路层来完成，如通过 ATM 的信元层或帧中继的 FR-SDH 层。MPLS 与 IP 网、ATM 网分层功能定位可参考图 2-1：

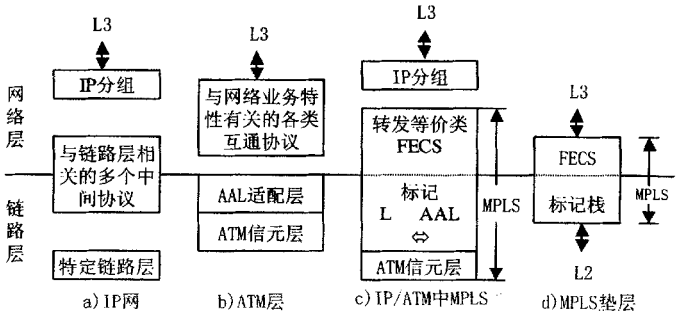


图 2-1 MPLS 与 IP 网、ATM 网分层功能定位

从图 2-1 可知，MPLS 垫层极大地简化和规约了 L3 和 L2 之间的转换协议。相比之下，IP 网的 IP 分组到特定的链路层之间，每一种都要有相应的多个中间协议；ATM 网中在链路层中只有 ATM 的一种协议系列，但在网络层对每一种业务都要有相应的多个互通协议。MPLS 垫层在 IP over OPTICS 上也是必不可少的，从网络层的 IP 到物理层的 OPTICS，要有一个链路层，经过 MPLS 垫层，当前可满足现有的 FR-SDH 和 ATM 链路层要求，将来也可适应任何一种新的链路层协议。

2.1 数字通信网络技术发展概况

迄今为止，在数字通信领域发展了三种网络技术，即 STM、IP 和 ATM。

STM 是基于连接的电路网络技术，源于保证实时话音的传输交换需要而发展起来的。从 60 年代初第一套 PCM 到 90 年代的 SDH，经历了出生、成长到成熟的全过程。IP 是源于计算机网络互联需要的一种异构网互联通信协议。能容纳和掩蔽不同网络硬件细节，使计算机独立于物理网络的具体连接进行工作；ATM 是经 ITU-T 认可的未来 B-ISDN 网的最终解决方案，是一种独立于终端和业务的通用信息传输和交换的体制。

三种体制各有优缺点和重点服务对象，在当前都有不同程度的发展。概括来讲，STM 是一种成熟的标准化体制，能满足电话网和低速数据网的各项业务性能要求，能提供可靠的有保证的服务；同时在现阶段还给 ATM 和 IP 提供低层服务（链路层低层和物理层）。IP 是当

前发展最为迅速的体制，本身尚在演变过程中。是源于计算机异构网络互连要求发展起来的一种体制，能最有效地满足互连网络的传送需求，在其上已传送无保证的语音和视频通信。ATM 是一项新技术，用于综合 STM 和 IP 的各项业务的统一网络技术平台，作为宽带综合业务数字网 B-ISDN 的最终解决方案。但 ATM 发展不快，在 90 年代作为链路层技术平台提供给 IP 网构建高速骨干网和局域网，速率为 155.52/622.08Mbps。随着吉位线速路由交换机的出现，ATM 在 IP 中的应用受阻，至今为止 ATM 仍无法替代 STM 中端到端话音业务，ATM 在当今最有应用前途的是在公众多媒体通信网上，即有多业务接入的传输的网络上。

目前通信产业界的发展重点主要为：

1、 发展宽带/窄带混合系统。综合应用 STM、IP、ATM 和 Mobile 等方面的成熟经验和技术，逐步演进。

2、 重点发展 ATM。加快 VOA 的实现，优选 Ip over ATM 技术方案（MPLS）挖掘纯 ATM 的潜力，大力降低成本。

3、 直接发展 IP over OPTICS，实现 everything over IP。

2.2 多协议标记交换 MPLS 结构及标记调换转发过程

MPLS 的基本目标是集成标记调换转发技术和网络层选路技术。其核心是标记的语义、基于标记的转发方法和标记的分配方法。

为了实现这一新技术的应用，IETF 工作组负责有关的标准化。先后提出了多个标准草案，其中主要标准草案已申请 RFC 号码，并获得批准，于 1999 年三月在正式会议上颁布。

MPLS 使用的主要术语：

1. 转发等价类 FEC：一组可用同一方式处理的 IP 分组，即同一路径、同一转发处理。
2. 标记 L Lable：一个短而定长的标识符，用以标记传输段落上的 FEC，局部有效。
3. 标记交换路径 LSP：在对等层面上，通过一个或多个 LSRS 的一条路径，和对应一个已有一组 IP 分组所映射的特定 FEC。
4. 标记交换路由器 LSR：具有 MPLS 节点功能的处理设备，并具有转发纯 L3 层 IP 分组的能力。
5. MPLS 域：一个运行 MPLS 协议的节点的邻接集合，相应于一个自治系统或一个 LSR 管理域。
6. MPLS 节点 MPLS node：一个运行 MPLS 协议的节点，节点能被 MPLS 控制协议发现、邻接和对话。执行一个或多个路由协议，具有标记调换转发功能。建议具有纯 L3 层 IP 分组的处理能力。
7. MPLS 边缘节点 MPLS edge node：MPLS 节点，连接 MPLS 域和一个域外节点，域外节点可不运行 MPLS。
8. MPLS 输入节点 MPLS Ingress：MPLS 边缘节点，用以处理输入到 MPLS 域的 IP 分组流量。
9. MPLS 输出节点 MPLS Egress：MPLS 边缘节点，用以处理 MPLS 域输出的 IP 分组流量。

在常规路由器网，一个 IP 分组是沿路由器逐跳传送的。每一个路由器都要独立读出 IP 分组组头，分析目的地址，运行路由算法，选择下一跳路由器。事实上 IP 分组组头所包含的信息远多于简单选择下一跳所需的信息。选择下一跳可概括为二种功能的组合：第一种功能是将所有需转发的 IP 分组，按转发方向规约为转发等价物的集合 FECs；第二种功能将每一个 FEC 映射到下一跳。MPLS 规定每个特定的 IP 分组映射到特定的 FEC，只在 IP 分组进入 MPLS 域时分配一次；FEC 是基于 IP 分组的目的地址划分的，并在进入 MPLS 域输入节点 Ingress 到 MPLS 域输出节点 Egress 之间建立一条与特定 FEC 相映射的标记交换路径 LSP；

沿 LSP 的两个相邻的标记交换路由器 LSR，及其连接的链路上，FEC 被编码为一个短而定长的标记 L；标记与 IP 分组一起传送，携带标记的 IP 分组，称标记分组；在后续的每一跳路由器上，不再需要对 IP 分组组头进行读出分析处理，只使用标记分组的标记作为指针，指向下一个新的标记和到达下一跳的一个输出端口，标记分组用新标记替代旧标记成为新标记分组，由指定输出端口传送到下一跳。上述标记调换转发过程同 FR 网中按 DLCI 和 ATM 中按 VPI/VCI 的转发过程，在实质上是一致的。其区别是 FR 网中 DLCI 只是链路的标志在 ATM 中 VPI/VCI 只是信元的标志，而在 MPLS 中的 FEC 是远比链路和信元要复杂得多的是概念：FEC 是对数据流、链路、端口等各种独立的对象进行集中提升抽象了的概念。MPLS 转发是按标记实现的，因而可以用交换机来进行转发；通常情况下交换机不能直接用来转发 IP 分组，因为交换机不能或不具有合适的速度来读出分析处理 IP 分组的组头。

2.3 MPLS 使用标记调换转发的优点

MPLS 与常规路由器网相比具有的优点：

1. 简化转发：MPLS 可按标记直接转发，而 IP 分组则需应用最长地址最长匹配算法进行转发。
2. 高效的明确路由：明确路由是由源主机指定的一条通过互联网到达目的地址的路径。明确路由也称为源路由，是功能非常强大的能用于多种目的的技术。在常规路由器网上，源路由用于网络测试，在纯数据报传送时，IP 分组是禁止携带完全的明确路由信息的，在 MPLS 中允许只在建立 LSP 时携带完全的明确路由信息，而不需要由每个 IP 分组来携带；这意味着在 MPLS 上能实际应用明确路由，能充分利用明确路由上的许多先进特性。
3. 流量工程：流量工程是指由数据流量选路的一种选择过程。用于按规则均衡网络中各种链路、路由器和交换机上的流量负荷。今天的 IP over ATM 是由 PVC 来实现的，PVC 通常是人工配置的，因而在 IP over ATM 的网上，流量过程的典型方式是人工调配。MPLS 允许数据流从特定的输入节点到特定的输出节点分别标识，即 MPLS 提供了对每对输入输出节点，进行测量的直接机制。
4. 服务质量 QOS：QOS 路由是指一种选路方法。这个方法是为特定的数据流选择路由，选出的路由应满足特定数据流的 QOS 要求。在许多情况下，QOS 路由要使用明确的路由，因 QOS 路由中最重要的一项是带宽保证，这同流量工程的要求是相同的。
5. 复杂的业务类别：特定用户在 Internet 上的特定业务要求日益增加。MPLS 能提供有效的方法去设别一个与 COS 和 QOS 相关的任何一个特定的 IP 分组。MPLS 是在 MPLS 域的输入节点 Ingress 上，一次性地完成 IP 分组到特定 FEC 地映射的；使得 IP 分组到合适的 COS 和 QOS 等价复杂映射变得容易，其他方式是不易做到。
6. 功能划分：MPLS 必须支持数据流的聚集转发：标记就具有粒度性质，最细可标识一个原始的用户数据流，最粗可标识由全部通过交换机或路由器的数据流聚集成的一个数据流。这就可能将路由处理功能分级划分给不同的网络单元。
7. 不同的业务类型采用单一的转发方式：MPLS 能用单一的转发方式在同一网络上提供给多种业务类型。如 IP 业务、帧中继业务、ATM 业务、IP 隧道和 VPNs 等。

MPLS 与 ATM 网 FR 网相比的优点：

1. 路由协议的伸缩性：在 IP over ATM 的核心网上，对等层路由器相互连接时要建立 n2 个逻辑链路。而在 MPLS 中对等层的每个路由器需要的通信减少到与其直接连接的路由器个数；在整个对等层上所需的处理传输交换的最高能力按 O（n）流要求。
2. 能在数据分组和信元媒质上通用操作：MPLS 对分组和信元媒体的路由和转发采用通用方法。允许对流量工程、QOS、COS 和其他性能功能要求采用通用方法。这就

意味着同一的标记可用于 ATM、帧中继和其他的链路层媒质。

3. 容易管理：对多种类型的媒质使用通用的路由协议，通用的标记分配方法，可以期望简化 MPLS 网的网络管理。
4. 路由风暴问题的消除：MPLS 消除了 ATM 网上用下一旅程简化协议 NHRP（Next Hop Resolution Protocol）和按需直接建立 SVC 需要，这就消除更新路由引起的争抢 SVC 问题，同时也消除了直接建立 SVC 有关的时延问题。

2.4 MPLS 发展背景

多协议标签交换 MPLS（Multiprotocol Label Switching）是一种二层半的面向连接的转发技术，采用并拓展了 IP 路由协议。支持多种网络协议，如 IPv4、IPv6、APX、AppleTalk 等。支持多种链路层技术：如 ATM、帧中继、Ethernet、RPR 等。工作方式为边缘路由、核心交换。

MPLS 发展简介：

1. 1996 年 Ipsilon 提出 IP 路由交换技术，使具有 ATM 交换机性能的设备执行路由器的功能。
2. 1996 年 Cisco 提出 TagSwitching 技术，采用拓扑驱动交换路径的建立，是 MPLS 主要的技术来源。
3. 1996 年 IBM 提出 ARIS（Aggregate Route-Based IP Switching）技术。
4. 1997 年 4 月，在 Cisco 的大力促成下，成立了 IETF MPLS 工作组。
5. MPLS 正在变成万能灵丹：
   (1) IETF TEWG 工作组：流量工程。
   (2) IETF PPVPN 工作组：网络服务商提供的 VPN。
   (3) IETF PWE3 工作组：二层线路仿真 VPN。
   (4) IETF CCAMP 工作组：GMPLS 光交换的核心协议。
   (5) MPLS 论坛：VoMPLS。
   (6) ITU NGN：IP/MPLS 是其传输骨干网。

MPLS 最新发展趋势：

1. MPLS 技术已成熟了 80%。
2. 国外大型电信运营商从 1999 年开始部署 MPLS 流量工程和 MPLS VPN。
3. 国外大型电信运营商基本上都提供 MPLS VPN（IP VPN）服务。
4. 已有大型电信运营商提供基于服务质量的金银铜牌服务。

第三章 MPLS VPN现状及市场需求

MPLS VPN是在网络路由和交换设备上应用MPLS技术，简化核心路由器的路由选择方式，利用标记交换，并且结合传统的路由技术实现的虚拟专网。

3.1 MPLS VPN概述

MPLS技术提供了类似于虚电路的标签交换业务，这种基于标签的交换可以提供类似于帧中继、ATM的网络安全性。同时相对于传统的VPN技术来说，MPLS VPN可以实现底层标签自动的分配，在业务的提供上比传统的VPN技术更廉价，更快速。同时MPLS VPN可以充分的利用MPLS技术的一些先进的特性，比如说MPLS流量工程能力，MPLS的服务质量保证，结合这些能力，MPLS VPN可以向客户提供不同服务质量等级的服务，也更容易实现跨运营商骨干网服务质量的保证。同时MPLS VPN还可以向客户提供传统基于路由技术VPN无法提供的业务种类，比如像支持VPN地址空间复用。对于MPLS的客户来说，运营商的MPLS网络可以提供客户需要的安全机制，以及组网的能力，VPN底层连接的建立、管理和维护主要由运营商负责，客户运营其VPN的维护和管理都将比传统的VPN解决方案简单，也减低了企业在人员和设备维护上的投资和成本。基于MPLS的VPN可以作为传统的基于二层专线的VPN、纯三层的IP VPN和隧道方式的VPN的替代技术，在现阶段可以作为传统VPN技术的有效补充。

具体到MPLS VPN的实现方式，根据运营商边界设备PE是否参与客户的路由，运营商在建立基于IP/MPLS的VPN时有两种选择：

1. 第三层的解决方案，通常称作是Layer3 MPLS VPNs。
2. 第二层的解决方案，通常称作是Layer2 MPLS VPNs。

衡量一个VPN解决方案的优劣主要基于以下几点的考虑：

1. 支持的业务种类。
2. 可以向用户提供的连接的种类。
3. 扩展性。
4. 部署的复杂度。
5. 业务开展的复杂度。
6. 管理和维护的复杂度。
7. 部署的成本。
8. 管理和维护的成本。

当然这些因素并不是绝对的，实际的应用中很难简单的说这两个方案谁优谁劣。两个方案都有其优缺点，有其特定的业务模式，也都还处在不断完善发展的阶段，选择一个方案的关键是运营商实际的网络运营环境，和运营商自身的业务定位，要向客户提供什么样的服务模式。

3.2 MPLS VPN现状

MPLS VPN可提供个人Internet业务、IP电话以及为企业团体用户提供电视会议、远程教学、信息公告、各种应用服务和网络商务等。MPLS VPN的产品已陆续进入国内市场。

VPN作为一种新业务和增值业务，是在20世纪90年代才发展起来的。目前，在国内提供VPN业务的主要是几大运营商，如中国电信、中国联通、中国网通、中国移动等；还有一些运营服务提供商，如中企通信、首创网络、Unihub等；当然，也有一些企业在自建VPN和外包ISP服务。

一项电信研究院集团客户市场调查显示，到 2003 年，有 35%的集团客户开始应用了虚拟专用网业务。另外，对于最重要性在前三位的业务中，VPN 占 0.1%，对于未来三年内需要迫切程度排在前三位的业务，VPN 占 30.7%。可见 VPN 服务在短期内已经取得了较大的发展，而且有巨大的市场潜力。它之所以这么快获得市场认可，与其相对传统专网的优越性分不开。

首先，企业在考虑 VPN 服务时，最为突出的关注是能否节省费用。当一个企业自行组建 VPN 时，客户的费用主要由三个部分组成：一次性投资（如开户费、设备费等）、接入费（即按用户的端口数、带宽等来收费）和通话使用费。从一次性投资和接入费来说，相比单独建设 PBX 或 CENTREX、移动集团电话等都具有优越性。有关的机构估算，如果企业放弃租用专线而采用 VPN，其整个网络的成本可节约 21%～45%，至于那些以电话拨号方式连网存取数据的公司，采用 VPN 则可以节约通讯成本 50%～80%。

其次，企业选择 VPN 服务还在于其简单便捷的特性。在传统专网上，如果客户要实现新业务应用的话，一般要求客户自己来设计应用新的技术和业务，而在 VPN 上，则一切都可由专业的电信部门来提供，这在节省了大笔费用的同时，也使企业更集中于主营业务，提高竞争力。同时，这还能实现专业化程度更强的网络运行维护管理。

再次，对于企业来说，采用 VPN 服务能实现灵活的成本管理。如果采用传统专网，客户的成本投入较大；而且，开发新业务的高成本又会提高客户在网络应用方面的边际成本。但是，如果采用 VPN 服务的话，则客户不需要在业务应用开发方面投入，它的成本主要来自通信使用费；而在使用业务方面，只要向 VPN 提供商申请就可以了，这些都是可以灵活控制和管理的。

有了这些相对优势，企业用户越来越多地采用了 VPN 服务。据预测，在北美将会出现近百亿美元的市场规模。但是，在国内，由于通信市场发展程度相对落后，加上国内对 VPN 的一些消费偏见，如网络的安全性、服务质量、可直接控制和管理性等，VPN 的成功应用还需要提供商加强服务力度，提供更丰富的客户服务类型。

据调查显示，在世界排名 1000 家的大公司中，有超过 84%的公司正在开发基于内部 IP 专用网的应用。在被调查的 501 家大公司中，已有 130 家采用了 VPN，175 家计划在未来的 18 个月内组建 VPN。通过对最终企业用户的调查表明，企业用户对运营商的需求已经从提供简单的 Internet 接入过渡到可管理 VPN 方案和业务的需求。专家们预见，到 2003 年，将有 70 亿到 290 亿美元的巨大市场。

3.3 MPLS VPN 市场需求

3.3.1 MPLS VPN 在运营商网中的应用

随着 MPLS VPN 技术的日益成熟，MPLS VPN 在各种网络中的应用也日益广泛，包括各种运营商网络、运营支撑网及各种企业网络。

在 MPLS 技术出现以前，运营商曾经选择不同的技术，建立了多种骨干网，例如，通过 PSTN 骨干网承载语音业务，用 FR 骨干网承载 FR 数据，用 ATM 骨干网来承载 ATM 数据，随着 IP 业务的爆炸性增长，又建设了 IP 骨干网。与各种骨干网技术相对应的是多种多样的接入网，这些不同类型的网络很难互联互通。如何实现这些不同类型网络的融合，提高现有网络的利用率，为用户提供更丰富的服务，是一个迫待解决的问题。MPLS 技术的成熟为骨干网的融合提供了可能。当前数据网络的发展趋势是骨干网逐渐统一为 MPLS 网络。而由于现有接入网络在很长一段时间仍然继续存在，接入网络呈现多样化的趋势，如 ATM / FR PVC、E1、LAN、xDSL、Cable、Wireless 等丰富的接入方式，并提供 IP、ATM、IPv6、CLNP（ISO Connectionless Network Protocol）、IPX 等多样化的服务。采用 MPLS L2 VPN 技术，IP 骨干网可以连接多样化的接入网络，实现对原有数据网络的改造及增强。在建成 MPLS 骨干

网之后，传统的数据通信网 ATM / FR 等可以下移为接入网，但对 ATM / FR 用户而言，感觉不到网络结构的变化。通过 Kom pella 方式的 MPLS L2 VPN 技术，各种不同的接入网协议还可以实现互操作，如 ATM 用户与 FR 用户之间也可以实现互通。通过 MPLS VPN 技术，可以在一个统一的物理网络上划分出多个逻辑上相互独立，相互隔离的业务专网，比如：为语音业务构建语音 VPN，为视频业务构建视频 VPN 等等。

3.3.2 MPLS VPN 在城域网上的应用

宽带城域网直接面向个人用户及企业用户，因此一方面宽带城域网要提供丰富的接入手段，满足各种用户的接入需求，另一方面，宽带城域网必须具有良好的可扩展性和可运营性，以满足用户不断增长的网络需求。

作为运营商的基础网络，宽带城域网需同时服务多种不同的用户，承载多种不同的业务，且宽带城域具有多种接入方式，这一特点决定宽带城域网需同时支持 MPLS L3 VPN、MPLS L2 VPN 及 VPLS，网络运营商可以根据网络的实际情况及用户的需求开通相应的 VPN 业务，例如，可以为以下类型的企业开通 MPLS L2 VPN 服务：

1. 有较强的网络维护能力的用户，如一些大型企业。同时这些企业不想将自己的路由管理及 Qos 管理等外包给运营商，希望自己管理自己的 VPN 网络。
2. 只想租用带宽的用户，如专线用户迁移。

同时，由于城域网中需求的多样性，还可能需要 L2TP，GRE 等多种 IP VPN 技术。在电信运营网络上开展 MPLS VPN，运营商最关心的往往是跨域解决方案及运营计费问题。

3.3.3 MPLS VPN 在运营支撑网上的应用

原有的运营支撑网络采用建专网的思路，及为不同的业务系统分别建立各自专网，一方面网络建设投资巨大，同时由于各个专网需独自进行管理，业务系统之间工作协调存在问题。通过采用 MPLS VPN 技术，将原来分散管理的多个专网统一在一个物理网络上实现，从而实现了统一平台，统一管理，可以提高网络的利用率及劳动生产力。

3.4 各厂商 VPN 解决方案

目前大型网络厂商都把虚拟专用网络作为重要市场目标。诸如 3Com、Cisco、Nortel Networks、Lucent 和 Shiva 公司等公司纷纷出击，提供各具特色的 VPN 解决方案。

3.4.1 3Com VPN 解决方案

3Com 为企业和 NSP 提供多种端对端的 VPN 解决方案。所有的 3Com VPN 产品彼此兼容，还配备了 TranscendWare 软件，使用户可以对常规网络和 VPN 执行统一的策略。TranscendWare 软件使边界设备可以与终端设备通信，进而强制执行网络策略。这些设备通过监视 VPN 隧道，可以更好地管理拨号端口、带宽分配、网络负载等，对 VPN 环境进行有效的控制。已配备 NETBuilder II 或 SuperStack II 桥接器/路由器的企业，可以在这些设备上增加 VPN 服务器能力(隧道终止器功能)。这样的一个设备，可以通过利用线路、帧中继、ISDN、SMDS 或 Switched 56 连接到 NSP，又可以通过 Ethernet、Token Ring 和 ATM 提供 LAN 连接。

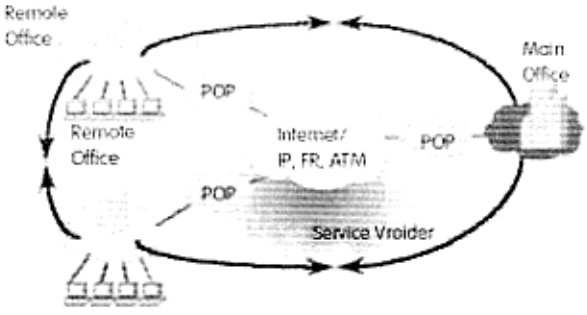


图 3-1 Intranet VPN 应用示意图

3Com 公司的 VPN 解决方案的优点在于能够很容易地集成到多厂商产品的网络中，且支持多种环境：移动用户和居家工作者、通过虚拟租用线路连接分支机构办公室和通过安全的 Extranet 连接商业伙伴。3Com 公司推出的基于 L2TP(Layer2 Tunneling Protocol)和扩展型 IPsec(Internet Protocol Security)的最新强化软件，从而进一步为用户扩展了 VPN 解决方案。

3Com 的隧道交换机提高了 VPN 的安全性和灵活性，可以带来以下好处：

1. 很容易区别对待远程职工的信息流和合作伙伴与客户的信息流。
2. 各部门可以共用一个 Internet 接口。
3. 可以最终利用 IP 地址空间。
4. 使远程用户很容易成为 VLAN(Virtual LAN)的成员。

3.4.2 Lucent VPNWorX

Lucent VPN 解决方案将电路交换式语音 VPN 和智能网络功能的优良传统，与原 Ascend 通讯公司的 VPN 体系——以数据为中心的 MultiVPN 创新技术结合起来，创建了语音和数据相融合的新型 VPN 服务。

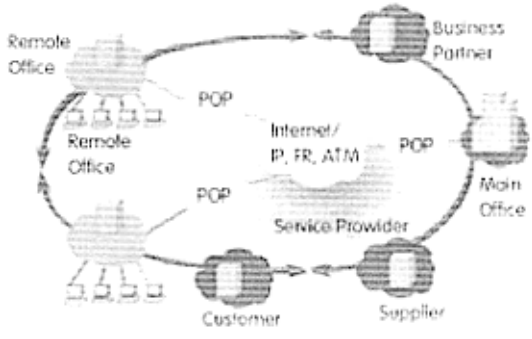


图 3-2 extrant VPN 应用示意图

Lucent VPNWorX，对企业来讲具有无需牺牲类似专用网的控制能力，可以获得 VPN 的所有好处。VPNWorX 解决方案中的产品，涵盖了供应商/用户环境，包括 Lucent 完善的产品方案，如电信级接入平台 MAX 和 MAX TNT，多业务交换机 GBX500、GX550、B-STDX8000/9000 和 Nexabit NX64000。IP Navigator MPLS 使服务供应商能够建立一个可以提供全新 IP 服务—Quality IP 的基础设施。Internet 安全产品包括 VPN Gateway 和 SecureConnect。Pipeline 和远程接入产品提供了集成化 VPN 路由和防火墙保护功能。并且，该方案具有通过 QoS 和 SLA 保障，提供关键型性能、策略化联网能力、端到端服务和网络管理及智能网络等增值功能。

3.4.3 Cisco EVPN

Cisco推出的企业级虚拟专用网EVPN全面解决方案，在可扩展的平台、安全性、服务、应用和管理五大VPN实施要素方面具有基于标准的开放式体系结构、可扩充的和端到端的网络互联能力。

和专用网络一样，Cisco为VPN提供了有保障的带宽和差别控制的服务与应用。在Cisco IOS软件12.0版中，丰富了QoS功能和机制，甚至支持IP QoS，使企业用户能够在不同应用和用户之间强制实现优先级和带宽分配，并内置了一个专门用来检测服务水平的响应时间报告器（RTR）。

Cisco提供了一系列的VPN安全功能来实现安全保障。基于硬件的集成式IPsec加密，不仅可以实现高速加密，还提高了WAN的可用带宽。PIX防火墙、NetRanger入侵检测系统和NetSonar安全扫描程序三者配合，保证企业VPN的可靠性和安全性。而这些功能都是被集成进Cisco IOS软件中的，用户可以通过简单的软件升级，透明地保护网络安全。

目前，所有的Cisco路由器平台都可以方便地实现VPN，经过优化的Cisco路由器集成了VPN功能、高速加密、安全、带宽管理和与WAN连接的能力，降低了VPN的复杂度和成本。该产品系列包括用于企业和地区办公环境的Cisco7500、7200VXR和7200高端路由器，以及用于小型地区、分支机构及远程个人的Cisco3600、2600、1720和800路由器。所有的Cisco VPN路由器都完全可以互操作，为从园区到广域网ISP，以窄带或宽带速率进行多媒体通讯提供了一条可扩展的端到端链路。

3.4.4 华为VPN方案

华为提供了各种有效的VPN解决方案，包括：Access VPN、Intranet VPN、Extranet VPN及结合防火墙的VPN解决方案。各方案中，Quidway系列路由器具有VPN网关功能，是组建VPN的重要设备之一。Quidway系列路由器支持各种VPN技术，包括隧道技术、IPsec、密钥交换技术、防火墙技术、QoS与配置管理等，并可继续发展，支持越来越多的先进技术。

其中，VPN与防火墙的结合使用，可以利用防火墙的许多安全特性在VPN上建立更加安全的网络环境，增强抵御黑客攻击、禁止非法访问的能力。公司内部特定的用户可以访问Internet网络，但是Internet网络内的主机不能通过防火墙访问公司的内部网络，只允许访问位于DMZ（非军事化区）的内部服务器主机，公司的外地办事处机构可以利用VPN和总部建立安全的私有隧道，以访问总部的资源。

第四章 MPLS VPN 原理及解决方案

目前涉及 MPLS VPN 原理及规范的有：

1. RFC2547 BGP MPLS VPNs。
2. BGP MPLS VPN 组网技术应用规范。

4.1 MPLS VPN 原理

要建立 MPLS VPN，涉及以下术语：

1. 站点（Site）：站点是指虚拟专用网络中具有路由邻接关系的用户设备（包括主机和其他网络设备）的集合，一个 Site 的主机和网络设备可以由位于相同地理位置组成，也可以是一个地理分布的网络。Site 中的设备采用二层链路或者 GRE 等隧道机制的逻辑链路建立路由邻接。多个 Site 通过公共基础设施（运营商网络）连接构成虚拟专用网。
2. 用户边缘设备 CE（Customer Edge）：用户网络中和运营商网络直接相连的设备，可以是主机、路由器等三层设备，在 BGP MPLS VPN 对 CE 设备没有特殊的要求（如 VRF 要求等）。
3. 运营商边缘设备 PE（Provider Edge）：运营商网络和用户网络直接相连的设备，PE 设备需要支持 MPLS 功能。PE 设备负责运营商网络同 VPN 客户网络的交互。PE 设备处理来自 CE 设备的数据，并转发到相同 VPN 中的其他 PE 设备，PE 设备也要能够处理来自网络的到所属 VPN 的站点的数据。PE 设备能够理解和处理 VPN 用户的私网地址，并提供不同 VPN 用户网络之间的隔离性。
4. 运营商核心设备 P（Provider）：运营商骨干网络中不和 CE 相连的路由器称为 P 设备，它是转发从 PE 设备发过来的 VPN 数据的设备，如果 PE 设备之间使用 MPLS 隧道，需要相应的 P 设备支持 MPLS 和 LDP。如果 PE 设备之间使用其他隧道机制，P 设备可以不支持 MPLS 和 LDP，只需要按照该隧道机制的要求支持相应技术。
5. RD（Route Distinguisher）：这是一个全局唯一的 8 字节数值，可以为每个 VRF 配置一个唯一的 RD，也可以为每个 VPN 配置一个唯一的 RD，在运营商骨干网络中用 BGP 来发布路由信息时，这些值附加到 VPN 的 IPV4 地址前缀之前，形成一个新的地址族：VPN-IPV4 地址族，这样，即使两个 VPN 具备相同的重叠 IPV4 地址空间，他们的 VPN-IPV4 地址空间是唯一的。
6. 路由目标（Route Target）：Route Target 属性标识了可以使用某路由的站点的集合，即该路由可以被哪些 Site 所接收，PE 路由器可以接收哪些 Site 传送来的路由。与 Route Target 中指明的 Site 相连的 PE 路由器，都会接收到具有这种属性的路由。PE 路由器接收到包含此属性的路由后，将其加入到相应的路由表中。PE 路由器存在两个 Route Target 属性的集合：一个集合用于附加到从某个 Site 发送的路由上，称为 Export Targets；另一个集合用于决定哪些路由可以引入此 Site 的路由表中，称为 Import Targets。通过匹配路由所携带的 routetarget 属性，可以获得 VPN 的成员关系。匹配 Route Target 属性可以用来过滤 PE 路由器接收的路由信息。
7. VPN 路由/转发实例 VRF（VPN routing/forwarding instance）：PE 设备中为不同的 VPN 站点分别维护的路由转发表，它主要包括：IP 路由表、标签转发表、使用标签转发表的一系列接口以及管理信息（包括 RD、路由过滤策略、成员接口列表等）。VPN 是由多个 Site(站点)组成的。在 PE 上，每个 Site 对应一个 VRF。

用户 Site 和 VPN 不存在一对一的关系；一个 Site 可以同时属于多个 VPN。在实现中，每一个 Site 在 PE 上关联一个单独的 VRF。VPN 中 Site 的 VRF 实际上综合了该 Site 的 VPN 成员关系和路由规则。报文转发信息存储在每个 VRF 的 IP 路由表和标签转发表中。系统为每个 VRF 维护一套独立的路由表和标签转发表，从而防止了数据泄漏出 VPN 之外，同时防止了 VPN 之外的数据进入。如果一个特定的 VPN 站点连接到特定的 PE 路由器，运营商必须在该 PE 路由器上配置一个 VRF，PE 上与 VPN 的每一个站点连接的接口或子接口都应该配置成与相应的 VRF 关联，这些接口可以是借用地址 unnumbered，或者在 VPN 的地址空间中分配一个地址。一般而言，在每条链路上应该运行一个路由协议（也可以使用静态路由），可以使用 EBGP 或 IGP，如 RIP 等，如果在链路上运行 IGP，则应该在 PE 这些链路上分别运行一个 IGP 的多实例，以和在运营商骨干网中运行的 IGP 以及相同 PE 上其他链路上运行的 IGP 区分开来。

VRF 中将包含从与 PE 相连的相应 CE 中通过在 PE/CE 链路上运行的路由协议分发的路由，并记忆从远端相同 VPN 其他 VRF 中通过 MP-BGP 分发的路由，然后通过标准的路由选择过程自动选择最佳路由，在 VRF 中静态配置是可选方案。在每个 VRF 中必须配置如下三个参数：

1. Route Distinguisher(RD)。
2. 一个或多个输出路由目标(Export Route Targets，Export RT)，它作为一种 MP-BGP 扩展团体属性，随 VRF 中向外发布的路由一起由 MP-BGP 携带。
3. 一个或多个输入路由目标 Import RT（Import Route Targets），在从远端 VRF 学习 MP-BGP 路由时用来和相应路由的输出路由目标 Export RT（Export Route Targets）相匹配以确定是否将该路由学习到本 VRF 中。

在最简单的情况下，一个 VPN 中的 Export RT，Import RT 和 RD 可以完全相同，这样，相同 VPN 中的 VRF 相互分发路由组成一个典型的 intranet，在更加复杂的情况下，他们可以设置成不同的情况，允许在 VRF 之间精确控制路由的发布，这样可以用来生成 extranet，或者在 VPN 中增加各种用户的策略，在复杂的情况下，特定的 Export RT 可以通过路由器的管理机制分配给特定的路由。在 VPN 中增加一个新的站点是将该站点的 CE 连接到 PE 路由器，配置接口，如果在 PE 路由器上已经存在和该 VPN 对应的 VRF。这些配置的改动自动通过 MP-BGP 通告到其他 PE。

可以利用将运营商的标识后面跟随所标识的运营商分配的号码来构造 RT 和 RD 中，以便他们在整个网络中唯一，运营商可以通过他们的 AS 号码，或者该运营商拥有的注册 IP 地址来标识。虽然 RT 被当作一种 BGP 扩展团体，但它的编码方式和其他类型的 BGP 扩展团体不同，它需要占用 8 个字节，编码方式可以和 RD 完全相同。

4.2、典型 BGP MPLS VPN 网络结构

IETF RFC 2547-bis 系列协议和草案提供并规定了 BGP MPLS VPN 的框架，下面给出基于 IETF RFC 2547-bis 系列协议和草案的 BGP MPLS VPN 的典型结构。

1. 基本网络模型

如图 4-1 所示，一个基本的 BGP MPLS VPN 网络由 CE 路由器、PE 路由器和 P 路由器构成。CE 作为客户边缘设备，是客户站点中连接运营商网络的路由器或者交换机；VPN 功能由 PE 路由器提供；P 和 CE 路由器没有特别的 VPN 配置需求。

为了将一个 VPN 的路由与公共的 Internet 路由和其它 VPN 的路由进行分离，PE 路由器为每个 VPN 生成了一个分离的路由 / 转发实例（VRF）。PE 路由器为每个有 CE 路由器连

接的 VPN 生成一个 VRF 表。任何属于 VPN 的客户和站点只能访问这个 VPN 的 VRF 表。

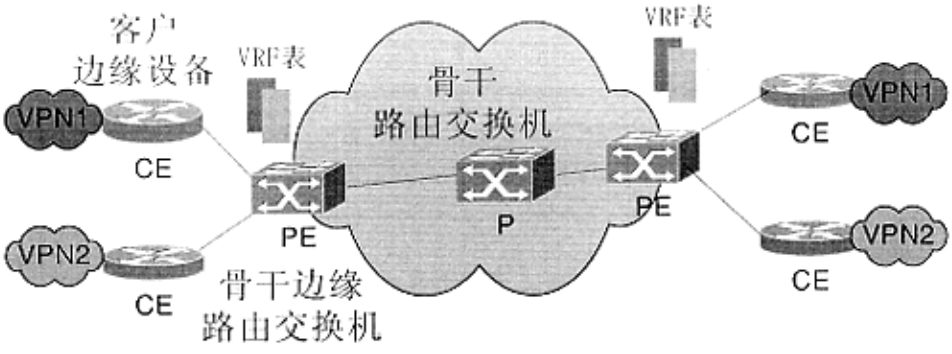


图 4-1 基本 BGP MPLS VPN 网络模型

2. 具有路由反射器的 BGP MPLS VPN 网络模型

在组建 BGP MPLS VPN 网络时，在每个 PE 路由器上必须运行 MP-BGP（Multiprotocol Border Gateway Protocol），BGP MPLS VPN 中 PE 之间使用 MP-BGP（RFC 2858）在 PE 之间进行 VPN 路由的学习和通告，MP-BGP 继承了 BGP 协议要求在同一个路由域中运行 IBGP（Internal Border Gateway Protocol）的对等体之间进行全连接以在路由域内通告 BGP 路由，当 VPN 中的 PE 数量很大时，这种 IBGP 全连接的数量会很大，有严重的 N 平方问题和可扩展性问题，为了改善这种状况，可以使用路由反射器来解决。

路由反射器允许保存不直接相连接的 VPN 的信息的设备，但是没有必要要求网络中的任何一个路由反射器都有全网的所有 VPN 的所有 VPN-IPv4 路由。下面列出了两种在几个路由反射器间划分 VPN-IPv4 路由职责的办法：

(1) 每个路由反射器都预先配置一系列的 RT。

(2) 每个 PE 路由器都是网络上路由反射器的某个子集的客户。

3. 跨 AS 的 BGP MPLS VPN 网络模型

如果同一个 VPN 的两个 Site 位于不同的自治系统内，相应的 PE 路由器就无法使用 IBGP 连接来转发 VPN-IPv4 路由，这时必须使用 EBGP 来在自治系统间传播 VPN-IPv4 路由，目前主要有几种可行方法：

(1) 背靠背的 VRF 方法。

(2) 使用 EBGP 从一个自治系统到另一个自治系统分发有标记的 VPN-IPv4 路由。

(3) 利用 Multi-hop EBGP 分发 VPN-IPv4 路由。

4.3、小结

本章简明阐述了 MPLS VPN 原理，并探讨了其解决方案。

第五章 ZXB10 产品软硬件平台

作者在选题过程中，结合了本人实际工作情况，对 VPN、MPLS 原理进行深入研究，MPLS VPN 市场需求进行充分调查，以及竞争对手产品进行充分剖析。在学校导师和本单位导师的指导下选择了 MPLS VPN 在多业务路由交换系统上的实现及应用作为课题，而 ZXB10 多业务路由交换机是 MPLS VPN 实现的软硬件平台。

5.1 系统原理

ZXB10-BX/AX/M1000 是基于 ATM 信元交换的多业务路由交换系统。ATM 技术本质上是一种高带宽、低时延的交换和复用技术，它同时支持语音、视频和数据等多种业务。ZXB10-BX/AX/M1000 利用 ATM 技术构建高速数据通信网络，可以向用户提供高速数据、LAN 互联、图象处理、视频和多媒体服务。同时 ZXB10-BX/AX/M1000 充分顺应因特网的迅猛发展的潮流，利用集成方式的 MPLS 技术将 ATM 与 IP 技术完美融合在一起，向用户提供具有 QoS 保证的高速路由转发系统。

1. ZXB10-BX/AX/M1000 在网络中的位置

作为多业务路由系统，ZXB10-BX/AX/M1000 在网络中的位置可以用图 5-1 描述.

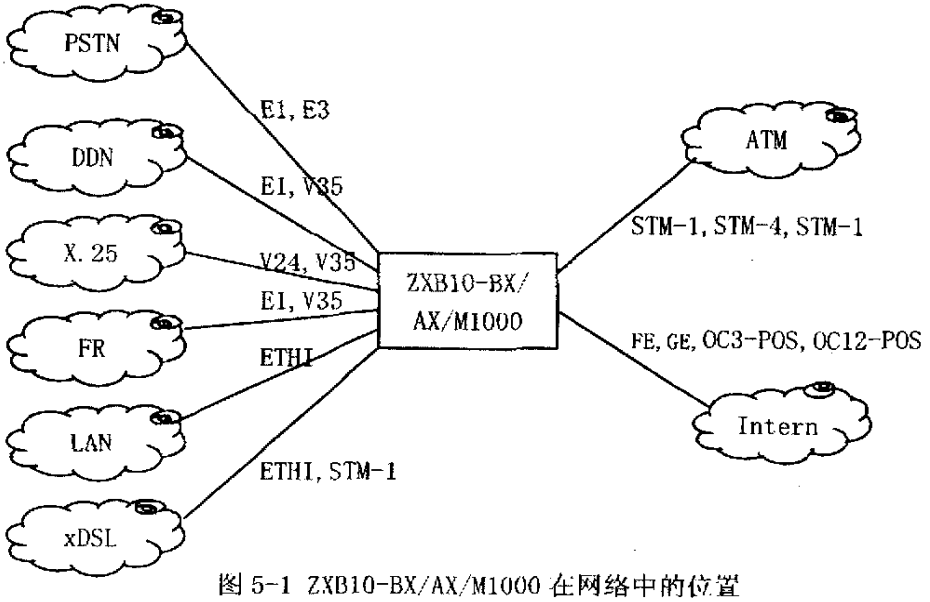


图 5-1 ZXB10-BX/AX/M1000 在网络中的位置

由于 ZXB10-BX、AX、M1000 各自的交换容量和端口数目不同，它们在网络中的具体位置也各不相同。作为全系列的宽带多业务路由交换机，ZXB10-BX/AX/M1000 应用范围从网络的核心/骨干到边缘/接入，覆盖宽带网络的各个层次。

ZXB10-BX 用于组建网络的骨干节点（MSCP，Multi-Service Converge Point），提供高速端口、大的交换容量，具有优异的性能，特别是在包转发速率、网络的处理能力、路由选择、可靠性、安全性上利用了先进的技术（ASIC、高效的算法等），形成网络的核心高速通道和交换平台。同时 ZXB10-BX 也提供少量的业务接入，便于核心节点接入 PBX、IP 或 FR 等业务。

ZXB10-AX 是骨干、接入一体化的路由交换机，提供丰富的业务接入能力，具有强大的适配能力和协议处理能力，可以作为城域网的骨干（MSCP）、边缘节点和广域网的边缘节点

（MSAP，Multi-Service Access Point）。

ZXB10-M1000 用于非 ATM 业务的接入，如 LAN、E1、ADSL 等，也包括常规的低速数据接入如 X.25、FR、RS232 等。M1000 适用于多业务、大流量的业务接入（MSAP）。

ZXB10-BX、AX、M1000 各自在网络中的位置如图 5-2 所示。

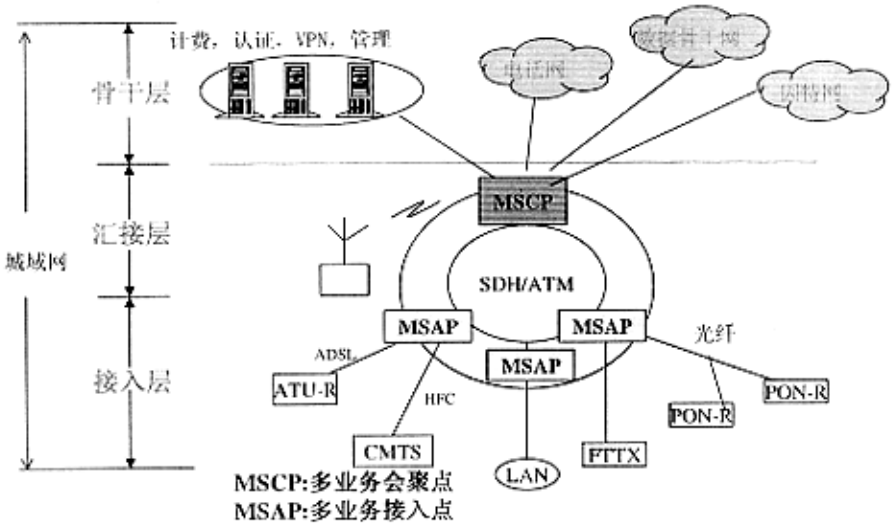


图5-2 ZXB10-BX、AX、M1000各自在网络中的位置

2. ZXB10-BX/AX/M1000 信号处理流程概述

ZXB10-BX/AX/M1000 可完成现有的各种网络的互联互通。总的来说，ZXB10-BX/AX/M1000 中的点到点通信的信号可以在任意两块接口卡、任意两块业务卡、任意接口卡和业务卡之间传送，也可以在某块单板（业务板或线路板）流入流出。所有信号的流动都必须经过交换网。对点到多点通信，信号可以在多块业务板或线路板之间流动，但每个分支与点到点通信一致。因此，ZXB10-BX/AX/M1000 中的典型信号流向可以归纳成如图 5-3 所示的八种：

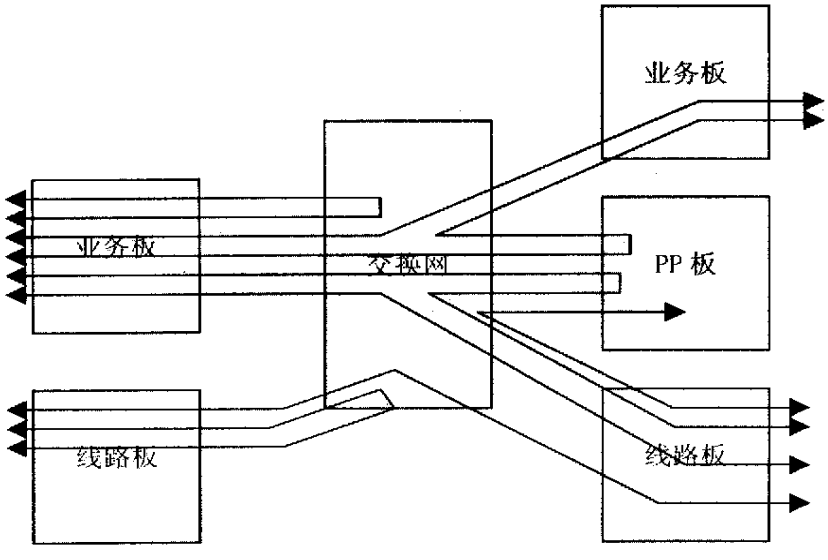


图 5-3 ZXB10-BX/AX/M1000 系统内的信号流向

(1) 业务板——交换网——业务板。
(2) 业务板——交换网——线路板。
(3) 业务板——交换网——PP 板——交换网——业务板。
(4) 业务板——交换网——PP 板——交换网——线路板。
(5) 业务板——交换网。
(6) 线路板——交换网。
(7) 线路板——交换网——线路板。
(8) 线路板——交换网——PP 板。

5.2 硬件总体架构

ZXB10 系统中 BX 、AX 和 M1000 都属于插箱式结构，支持业务的本地交换。业务单板在硬件上与机型无关，同一块业务单板可以分别配置在 BX、AX 和 M1000 中。BX、AX 和 M1000 硬件上的主要区别在于多功能板和交换网板的不同。

BX 的端口容量为 12.5G，AX 的端口容量为 5G，M1000 的端口容量为 622M。

5.2.1 ZXB10-BX 硬件总体

图 5-4 为 BX 的硬件总体框图，图中可以看出，硬件主要包括控制和业务接入单元、交换单元、时钟单元和电源系统四部分：

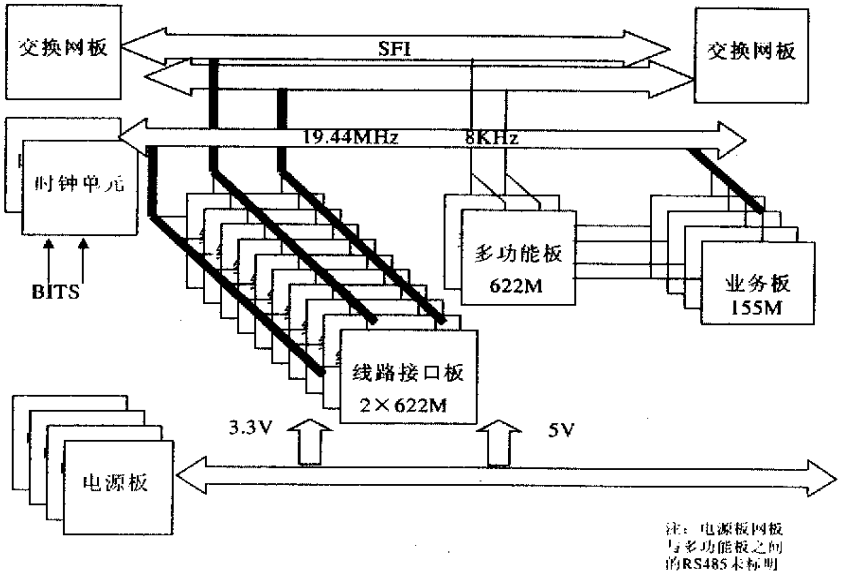


图 5-4 ZXB10-BX 硬件总体框图

5.2.1.1 控制和业务系统

BX 的控制系统是通过带内方式来实现的。多功能板（MFUC）占用了交换网板的一个端口，用来支持业务板和管理通道，如图 5-5 所示。控制信息是通过 AAL5 适配成 ATM 信元在物理通道上进行传输的。总的来说，系统控制部分包括管理模块和 ATM 适配模块，这两个模块分布在控制板（多功能板）、业务板和线路接口板上。ATM 适配模块根据情况可以是控制

系统独享的或与业务系统共享。系统的控制通道分三种：控制板－线路接口板、控制板－业务板以及控制板－网板/电源。

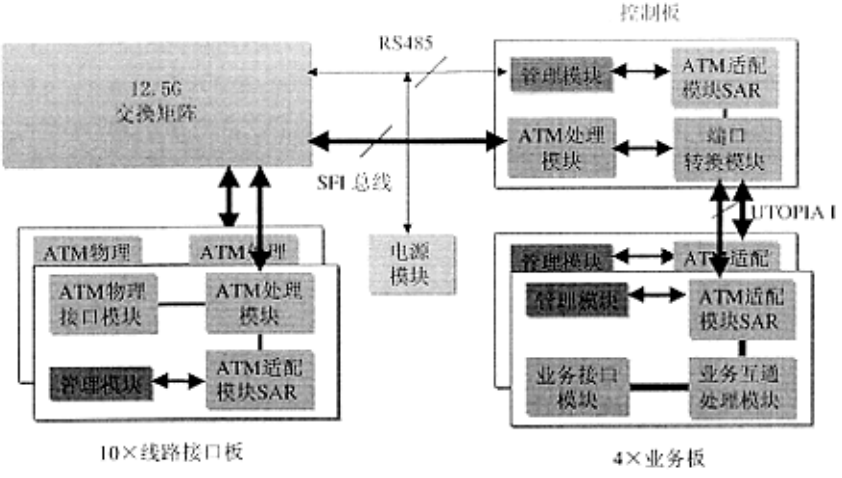


图 5-5 多功能板（MFUC）框图

控制板－线路接口板的控制通道需通过交换网板来实现，线路接口板根据物理端口号来分离信元中继数据流与管理控制数据流，其管理数据流向为：控制板－交换网板－线路接口板。

控制板－业务板的控制通道的管理数据流向为：控制板－交换网板－控制板－业务板，由于业务数据流和管理控制数据流共用一个物理端口，因此必须通过 ATM 连接来分离管理控制信息。

控制板－网板/电源的控制通道是通过 RS485 总线来实现的，主要是用来采集单板或模块的状态信息。

控制系统的主要载体控制板（多功能板）在系统中是主备冗余备份 。单板主备倒换电路包括主备通信通道和主备倒换逻辑。主备通信通道由 10M 以太网来实现，主备倒换逻辑包括主用占用信号、Watchdog 溢出信号、复位对机信号和本机状态指示信号等。

为充分利用端口资源，主备两块控制板（多功能板）只占用交换网的一个端口。如图 5-6 所示，通道主备切换通过高阻来实现。

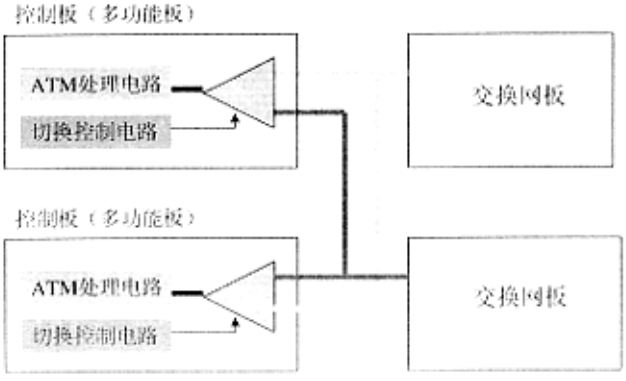


图 5-6 控制板与交换网接口示意图

ZXB10-BX 的业务板定义为背板速率低于 155M 的单板。如图 5-7 所示，业务板的流量由

多功能板 MFU 集中，多功能板利用交换网络的一个 SFI 端口汇聚各业务板和单板内管理通道的数据流，业务单板和单板管理通道通过物理端口来进行区分。

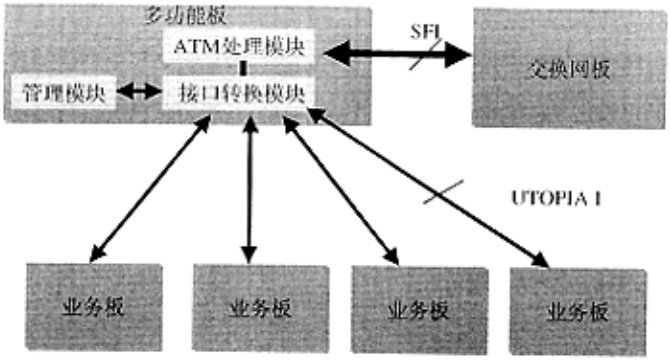


图 5-7 多功能板数据流管理示意图

5.2.1.2 交换结构

在 ZXB10-BX 中，仅用一个交换单元不能达到交换容量。而是采用多交换单元构成三级交换网，图 5-8 是 12.5G 的交换结构（20X20）：

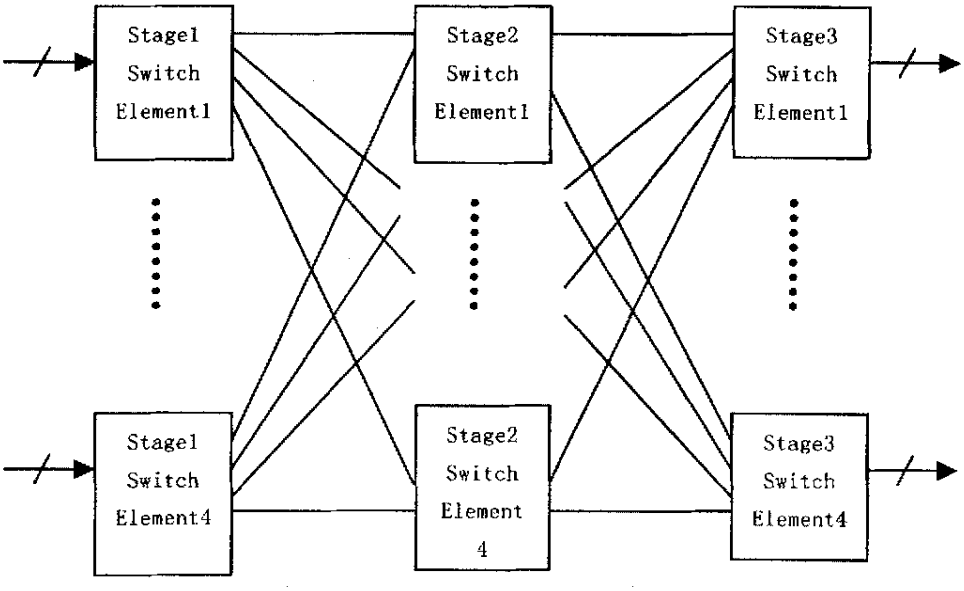


图 5-8 BX 交换结构示意图

20*20 的交换网由 12 个 8*8 的交换单元组成，其中第一、二级采用 5:8 和 8:5 的输入输出端口比例，第二级的交换单元类似于一、三级，但使用全部 8 个端口，没有信元缓存。这样构成一个 20*20 的交换网。从上面的描述中可以看出，在多级交换网的实现中，我们采取了如下几种方法来防止阻塞：

1. 从第一级到第二级时，输出处理器总是将数据平均的分配到各条出线上。
2. 增加内部通道数，这里采用 5:8 的比例。
3. 内部速率高于物理端口的速率。
4. 利用反压机制和缓存，完全消除阻塞。

5.2.1.3 定时系统

ATM 的异步传输特性决定了 ATM 在组网时不需要全网同步，但这并不意味着 ATM 设备本身不需要性能较好的时钟以及同步定时功能。因为在 ATM 所承载的某些业务中是需要定时同步的，以保证这些业务的信息可以实时地、正确无误地被传输和交换，另一方面为了最大限度地减少信元的抖动和漂移，所以 ATM 设备的时钟以及定时同步功能也必须按 ITU-T 的有关规定进行设计。

BX 的定时系统集成在多功能板上。其时钟定时系统可以向全系统提供高精度的 19.44MHz 时钟，定时系统的时钟基准可以来源于线路接口板或业务板，时钟分配及基准来源如图 5-9 所示。

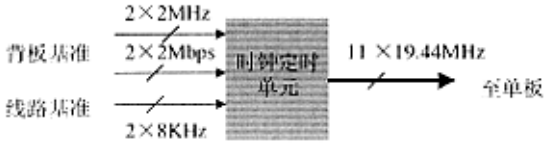


图 5-9 定时系统时钟分配及基准来源

详细设计原来及各项指标数据略。

5.2.1.4 电源系统

ZXB10-BX 的电源系统位于插箱的下方，高度为 4U。结构上分 5 个槽位，其中左边 4 个为电源板槽位，右边为电源监控板槽位。

ZXB10-BX 的电源系统包括电源板、电源控制板两部分。电源板输入电源电压为负 48V，输出 5V 和 3.3V。电源控制板对每一路电源电压实现监控检测，并通过 RS485 总线上报系统控制板。

电源板主要完成电源滤波、过压保护、电压转换等功能。每个电源板提供 200W 功率。由于单个电源板不能满足 BX 的供电要求，因此在电源板的电源开关上采用了并联式设计，即只要一个电源板的开关开启，所有在线的电源板都同时开启供电，只有当所有的电源板的开关都关闭时，电源板才同时关闭。这样可以保证 BX 在上电时不会由于供电不足导致异常状态。为方便维护，BX 的电源板都带有软启动电路，支持带电插拔。

电源监控板主要完成风扇电源输出、风扇状态指示、电源板电压检测、与控制板之间通信等功能。

5.2.2 ZXB10-AX 硬件总体

AX 与 BX 的硬件总体基本相似，相对来说，AX 比 BX 交换容量要小，只支持 7 个 622M 的线路接口槽位，而业务板的槽位则可以到 8 个，比 BX 多。

5.2.2.1 控制和业务系统

ZXB10-AX 与 BX 一样都是通过带内通道实现内部通信的。

5.2.2.2 交换结构

AX 的系统交换网络为 8×8 的交换结构，每个交换口为 622Mbps 的带宽，故总交换容量为 5G。AX 的交换结构比较简单，利用一片 ASX 即可以实现一个 8×8 交换网络。考虑到交换网络在系统中所处位置的重要性，采用双机双总线热备份的形式，通过故障切换、手动切换和软件切换可实现主备倒换。主控板与交换网板之间采用 RS485 进行通信，以实现单板状态信息的采集。

5.2.2.3 定时系统

AX 的时钟定时单元集成在交换网板上，其工作原理与 BX 一致，提供两种形式的 BITS 接口外时钟同步输入：2 个 2MHz 和 2 个 2Mbits/S，接口方式为 75 欧姆非平衡 E1 接口。系统内的同步输入可以来源于任意接口板的任意端口。时钟定时单元为系统内的 13 个接口板槽位各提供一组 19.44MHz 的同步时钟。

AX 的时钟系统的指标为：

时钟稳定度：$\pm 3\times 10^{-8}$ /天。

时钟精确度：$\pm 1\times 10^{-6}$。

5.2.2.4 电源系统

AX 与 BX 一样采用同样的电源系统。但由于 AX 的功耗比 BX 小，满配置为 200W 左右，因此 AX 的电源系统可以做到 1＋1 备份。另外 AX 的电源板电源开关单独工作，不需要联动开启。

5.2.3 ZXB10-M1000 硬件总体

M1000 的控制、交换、时钟、电源都不提供冗余备份，主要是为用户提供一种价格较低的汇聚层设备，其硬件原理框图如图 5-10 所示。其硬件模块也参照 AX、BX 分为控制接入单元、交换单元、时钟单元和电源系统。

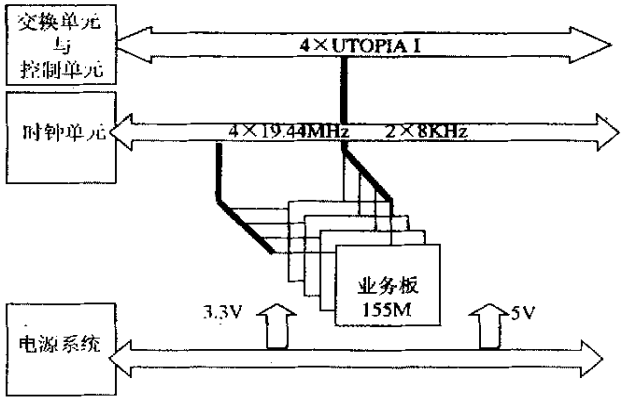


图 5-10 M1000 硬件原理框图

5.2.3.1 控制与接入系统

M1000 的控制系统与 AX/BX 有一些区别，首先，M1000 的多功能板不提供冗余备份，这主要是考虑到 M1000 的网络位置是位于用户端或网络的汇聚点，需要价格相对低廉。第二，M1000 不提供线路接口板和交换网板，因此控制模块主要是对业务板进行控制。第三，M1000 不对电源模块进行监控。

M1000 支持 4 块业务板，多功能板与业务板之间仍然通过 UTOPIA I 总线连接，基本结构与 BX 的业务接入方式非常相似。

5.2.3.2 交换结构

M1000 没有单独的交换网板，其交换能力由多功能板上的一ALM、ABM 提供，在 AX、BX 的多功能板上，ABM 芯片只是提供队列管理、流量整型等功能，必须与交换芯片 ASX、ACE 一起配合完成交换功能。而在 M1000 上 ABM 工作于 Standalone 模式，不需要与 ASX 连接自己内部可以实现 622M 端口交换能力。

5.2.3.3 定时系统

M1000 不提供外部 BITS 同步时钟，只提供线路恢复时钟同步机制。其时钟同步电路采用模拟锁相环，在基准时钟丢失的情况下，模拟锁相环输出的时钟精度较低为 30ppm 左右，因此此时采用外置高精度温补晶振提供同步时钟。温补晶振的主要参数指标为：

时钟稳定度：$\pm 1\times10^{-6}$ /年；时钟精确度：$\pm 1\times10^{-6}$

M1000 的时钟单元的工作原理框图如图 5-11 所示。

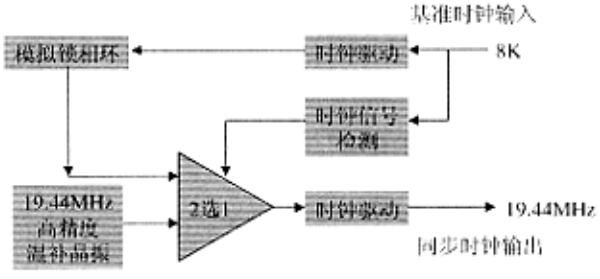


图 5-11 M1000 时钟单元工作原来框图

5.2.3.4 电源系统

M1000 的电源系统提供 2 种供电方式：220V 交流供电和负 48V 直流供电。电源系统为单板提供＋5V 和＋3.3V 两种电源电压，为风扇提供＋12V 电源电压。220V 交流供电方式采用外购电源模块，其功率为 70W，向系统提供 5V、3.3V 电源。其参数为输入电源电压范围：90～264V，输入频率：47Hz～63Hz。输出电源电压参数为：＋5V 标称电流为 13A，＋3.3V 标称电流为 13A，＋12V 标称电流为 0.8A。在标称范围内，三种输出电源电压功率可以平衡输出。

负 48V 供电方式下，电源电压输入范围为：36～75V，输出电源电压为＋5V 最大电流为 15A，最小电流为 1A；＋3.3V 的最大电流为 15A，最小电流为 1A；＋12V 的最大功率为 5W。＋5V 和＋3.3V 功率可以平衡输出。

5.3 软件总体架构

5.3.1 软件总体设计描述

ZXB10-BX/AX/M1000 软件总体设计采用分散式的模块结构，结构如图 5-12 所示：

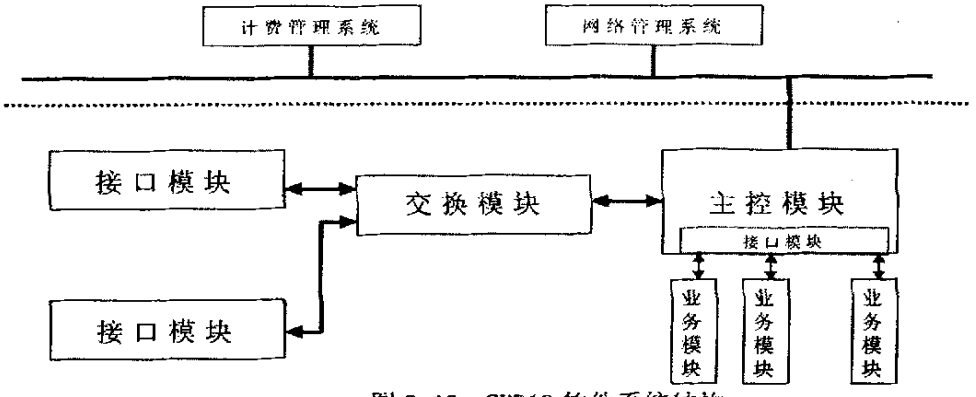


图 5-12　ZXB10 软件系统结构

每个模块的功能如下：

1. 网络管理系统（后台）：采用 client/server 结构，完成网管的四大功能（计费功能由专门的计费管理系统完成）。为了方便管理，网络管理模块与前台系统间采用的标准的网管协议 SNMP（一些功能如版本升级不采用 SNMP 协议）。
2. 计费管理系统（后台）：完成用户档案管理、计费信息收集、结算等计费功能。
3. 接口模块：
   (1) 接口模块包括：ATM 线路接口卡（LIC），千兆以太网接口卡（GEI），POS 接口板和多协议处理模块板（SMPP）。
   (2) LIC 负责提供 155M 或 622M 接口上的本地接口管理协议（ILMI）、连接控制、流量管理、拥塞控制、B-ISDN 信令协议、ATM 路由选择、连接计费、性能管理以及其它所有相关的操作维护及网络管理功能。
   (3) GEI 和 POS 负责提供二层和三层转发。
   (4) SMPP 负责处理路由协议和三层转发。
4. 主控模块（MFU）：该模块主要负责完成系统控制功能并提供操作维护和网络管理接口。同时提供 LIC 的功能。
5. 交换模块：该模块负责交换网的管理以及信元路由的实现。该模块硬件上对应交换网板。
6. 业务模块（SB）：该模块负责提供 FR、ELAN、CES、纯信元接入功能、性能管理以及其它所有相关的操作维护管理功能。

根据 ATM 交换机系统结构，整个系统由以下九个软件模块建筑在硬件基础上组成：系统承载模块、数据支撑模块、接入控制模块、信令处理模块、PVC 管理模块、业务处理模块、OAM 及网络管理模块、计费统计模块(实际上计费统计功能属于网管功能的一部分，单列以强调其重要性)、硬件驱动模块组成，图 5-12 所示的各个系统模块上的软件就是由以上软件模块组成，每个系统模块上的软件模块关系如图 5-13～5-15 所示。

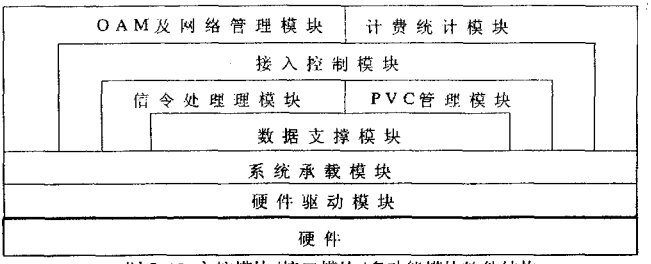


图 5-13 主控模块/接口模块/多功能模块软件结构

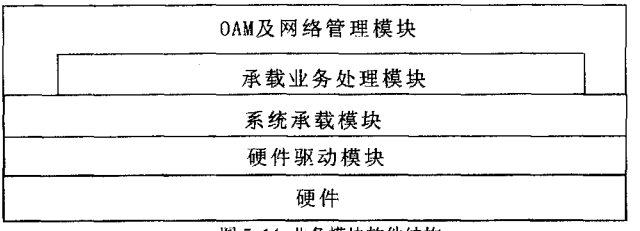


图 5-14 业务模块软件结构

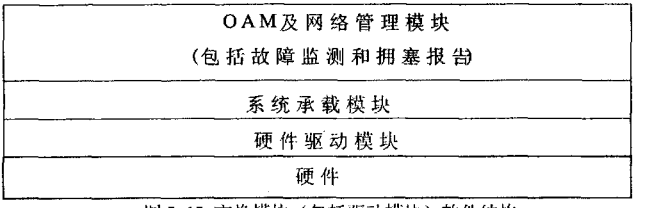


图 5-15 交换模块（包括驱动模块）软件结构

5.3.2 各层结构设计描述

5.3.2.1 系统承载模块

系统承载模块是整个软件系统的支撑模块，它负责完成如下功能：

1. 进程调度、定时器管理、内存管理。
2. 版本升级。
3. 文件管理。
4. 模块内通信。
5. 前后台通信、主备 MPU 通信。
6. 系统运行控制。
7. 告警。
8. 诊断测试。

5.3.2.2 硬件驱动模块

该模块直接对硬件操作，为上层应用程序提供对底层的操作接口。该模块的设计目的是要尽量为上层应用软件屏蔽底层硬件的复杂性，同时为各应用程序提供统一接口。该部分由

以下几个功能部分组成：

1. 硬件初始化：负责整个硬件环境的所有芯片的初始化工作。
2. 时钟级处理部分：负责扫描硬件芯片上提取的有用的信元，如 OAM 信元，以及其它一些信息，如：来自信令通道的 CPCS-PDU，帧中继上的数据包等。
3. 事件级处理部分：为应用程序提供回调函数，完成应用程序或网管的指定操作。中断级处理部分：有些消息非常重要，可以通过产生中断，然后再将消息传给应用程序。

5.3.2.3 数据支撑模块

数据支撑模块和系统支撑模块一样是整个 ZXB10 交换机系统的支撑模块，为了满足安全性、可靠性以及互通性要求，在设计时，要求系统必须满足以下要求：

1. 采用自设计的实时分布式数据库。
2. 采用面向对象的关系数据模式组织管理数据。
3. 数据维护管理系统采用实时修改内存方式。后台用户终端采用交互式图形接口，操作直观简便；同时提供命令行接口，使熟练操作员能方便快捷地维护系统。为交换机软件的各应用层提供数据库存取接口。

5.3.2.4 信令处理模块

本模块为 UNI 和 NNI 接口上提供建立 SVC 通路的信令要求并且提供信令 AP 接口。

5.3.2.5 接入控制模块

接入控制模块是整个 ZXB10 系统的又一核心模块，该模块体现了 ATM 交换系统与以往各种交换系统所不同的地方。该模块主要实现以下功能：

1. 呼叫接纳连接控制。
2. UPC 控制，它包括流量控制、流量整形、丢失信元优先率等。
3. 拥塞控制。
4. 快速资源管理。
5. 号码分析、路由选择。
6. 接续控制。
7. 计费信息收集。

5.3.2.6 承载业务处理模块

承载业务处理模块主要负责实现所有业务功能，目前可考虑实现的业务功能有：

1. 局域网仿真业务 LANE。
2. IPOA 业务。
3. MPLS 业务。
4. 电路仿真业务 CES。
5. 帧中继业务 FR。
6. 反向复用业务 IMA。

5.3.2.7 OAM 及网络管理模块

操作维护和网络管理系统将提供对 ZXB10 系列产品的综合管理。同时要提供标准的简单网络管理协议（SNMP，Simple Network Management Protocol）和公共管理信息协议（CMIP）接口与其他管理实体进行互操作。主要包括：

1. 配置管理。
2. 故障管理。
3. 性能管理。
4. 安全管理。
5. 计费管理。

5.4、运行环境

硬件环境：ZXB10-BX/AX/M1000。

操作系统：前台开发采用 PSOS 操作系统。
后台开发采用 Windows NT（PC 平台）或 UNIX（工作站平台）。

数据库系统：前台开发采用自己开发的数据库系统。
后台开发采用 Sybase 数据库系统。

开发工具：前台开发采用 PSOS 集成调试系统。
后台开发采用 Delphi（PC 平台）或 Motif（工作站平台）。

5.5、小结

本章主要对 MPLS VPN 模块的实现平台 ZXB10 AX/BX/M1000 的软硬件平台进行了简单阐述，该部分内容是作者进行 MPLS VPN 模块设计的前提与基础。

第六章 MPLS VPN 的实现

6.1 系统需求

在分析 MPLS VPN 的市场需求，解剖相关厂家设备的功能和组网方案，跟踪研究 MPLS VPN 最新协议标准后，我们得出必须尽快实现 MPLS VPN 功能，以增强我司数据产品的市场竞争力。首先我们决定将 MPLS 和 MPLS VPN 功能在 ZXB10 设备中加以实现，然后再在其他设备中加以推广，我参加的主要工作是在研究解剖竞争对手设备情况，积极跟踪 MPLS VPN 最新技术基础上，将 MPLS VPN 设计进 ZXB10 产品中，使其成为 ZXB10 系统的一个分系统；

在设计开发过程中，恰逢总参骨干网招标、建设和开通，该网目前为全球最大的 ATM 骨干网，在该网上开通 MPLS VPN 对于我们设计、研发、工程、售后维护方面的经验积累具有非常重要的现实指导意义。

6.1.1 MPLS VPN 在系统中位置

该 MPLS VPN 子系统主要包含：

1. PE-CE 路由学习模块。
2. PE-PE 路由信息广播模块。
3. 内层标签分配模块。
4. 外层标签分配模块。

其中 PE-CE 路由学习模块，PE-PE 路由信息广播模块是 MPP 板的 BGP 模块的扩展。内层标签/外层标签模块是 LIC 板的 LDP 模块的扩展，它与网管、PVC 建立模块都有接口。它根据 PE-CE 路由信息预先计算好内层/外层标记映射表，为 VPN 数据包的转发提供标记映射，使得在骨干 MPLS 网中转发 VPN 数据不需重新计算路由，提高了数据转发的效率。

6.1.2 实现功能

PE-CE 路由学习模块主要完成 PE 与所连接的 CE 之间的路由信息学习，通过扩展的 BGP-4 协议，采用 RD+IPv4 地址完成。

PE-PE 路由信息广播模块主要完成同一 VPN 内的不同 PE 之间相互广播自己所连接的 CE 路由信息，最终在每个 PE 端都保存一张完整的 VPN 路由信息表，表明该 VPN 包括哪些 CE，以及 CE 之间的路由信息。

内层标签分配模块主要完成功能：在源端 PE，从 CE 接收来的数据包，根据接收端口、目的 CE 地址，为其分配内层标签；在目的端 PE，根据接收到 VPN 数据包的内层标签，即可判断目的 CE 地址或出端口，从而将数据发送到目的 CE。

外层标签分配模块功能：为源端 PE 至目的端 PE 建立 LSP，为其分配 PVC VPI/VCI，SP 的 P 路由器根据外层标签进行数据包的转发。

本分系统设计的目的主要是实现和完善 LDP/CR_LDP 协议，使其具备支持 Diff_Ser、Loop 等功能；同时实现 RSVP_TE、MPLS/BGP/VPN 功能，并为 MPLS 对组播的支持作一先期的研究，为全网流量管理打下一个坚实的基础。

6.1.3 性能指标

ZXB10-AX 支持 7*8 个 VPN 用户(基于端口)。

ZXB10-M1000 支持 4*8 个 VPN 用户(基于端口)。

ZXB10-AX ETJI 支持>130M 转发带宽。

ZXB10-M1000 ETH 支持>130M 转发带宽。

支持 MPLS 提供的 Cos,Qos 能力。

6.2 设计依据

Requirements for Traffic Engineering Over MPLS(RFC2702)

Multiprotocol Label Switching Architecture(RFC3031)

MPLS Label Stack Encoding(RFC3032)

Use of Label Switching on Frame Relay Networks Specification(RFC3034)

MPLS using LDP and ATM VC Switching(RFC3035)

LDP Specification(RFC3036)

LDP Applicability(RFC3037)

VCID Notification over ATM link for LDP(RFC3038)

MPLS Loop Prevention Mechanism(RFC3063)

Carrying Label Information in BGP-4(RFC3107)

draft-ietf-mpls-ldp-mib-07.txt

draft-ietf-mpls-ldp-state-04.txt

draft-ietf-mpls-rsvp-lsp-tunnel-08.txt

draft-ietf-mpls-cr-ldp-05.txt

draft-ietf-mpls-te-mib-06.txt

draft-ietf-mpls-diff-ext-09.txt

draft-ietf-mpls-multicast-05.txt

draft-ietf-mpls-lsr-mib-07.txt

draft-ietf-mpls-crldp-applic-01.txt

draft-ietf-mpls-rsvp-tunnel-applicability-02.txt

LSP Modification Using CR-LDP

draft-ietf-mpls-lsp-hierarchy-02.txt

draft-ietf-mpls-lmp-02.txt

draft-ietf-mpls-recovery-frmwrk-02.txt

draft-ietf-mpls-ftn-mib-01.txt

Fault Tolerance for LDP and CR-LDP

draft-ietf-mpls-generalized-signaling-04.txt

draft-ietf-mpls-lsp-query-02.txt

draft-ietf-mpls-diff-te-reqts-00.txt

draft-ietf-mpls-diff-te-ext-01.txt

draft-ietf-mpls-generalized-cr-ldp-03.txt

draft-ietf-mpls-generalized-rsvp-te-03.txt

draft-nadeau-mpls-tc-mib-00.txt

6.3 设计原则

系统设计遵循全分散，独立于硬件体系和操作系统，数据管理、消息驱动独立于实际应用系统。系统采用模块化的设计尽量减少各协议间的偶合。标记转发在各单板（LIC、POS、ETH、FR）上做到与单板类型无关，对应用系统提供标准的 API 调用接口。人机命令采用完全仿 CISCO MPLS 命令行方式，并提供人机命令解释接口为人机命令在不同的应用环境下提供统一的界面。

6.4 基本原理

本分系统的设计主要遵循MPLS VPN工作组要提供的协议完成相应的功能。

6.5 分系统构成

本系统的V2.1可以参考以前设计的相应文档。在V2.2中主要增加以下几个功能：

1. 仿CISCO MPLS命令行方式的人机界面。
2. DIFF_SERR、LOOP检测支持。
3. 业务板分散支持。
4. RSVP_TE支持。
5. MPLS/BGP/VPN支持。

以上各功能的详细设计都可以看相应的详细设计文档。

6.5.1 环境框图

本分系统主要运行于ZXB10-BX/AX/M1000上，可运行于Psos和VxWorks操作系统。

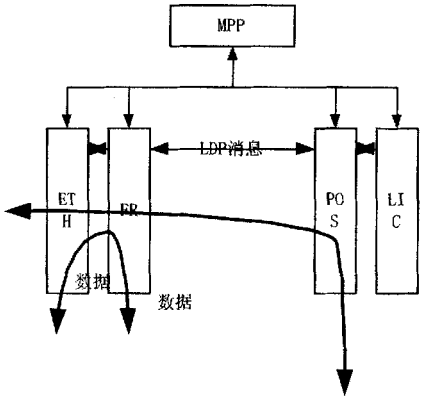


图6-1 MPLS VPN子系统环境框图

由图6-1可知MPP和各业务板交换路由信息，并通过命令行方式发送消息给单板，各单板通过交换LDP消息建立相应的LSP通路，通过LSP通路完成高速的数据交换。

6.5.2 总体框图

图6-2是整个MPLS VPN子系统所有模块之间的信息传递总体框图：

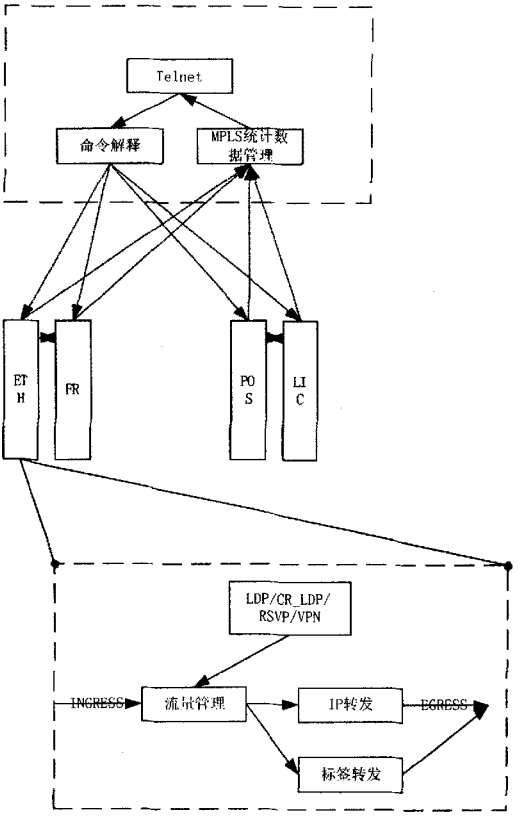
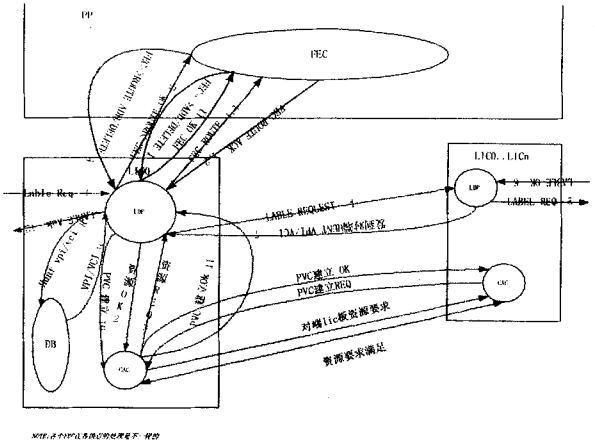


图 6-2 MPLS VPN 子系统总体框图

由上图知 PP 板完成人机命令解释，业务板完成 LSP 通路建立和数据转发。

6.5.3 功能框图

下面给出最主要的“BackBone 的 FEC 和 PVC 过程”、“Engress 的 FEC 和 PVC 过程”以及“Ingress 的 FEC 和 PVC 过程”功能框图。



图一：BackBone的FEC过程

图 6-3 BackBone 的 FEC 过程

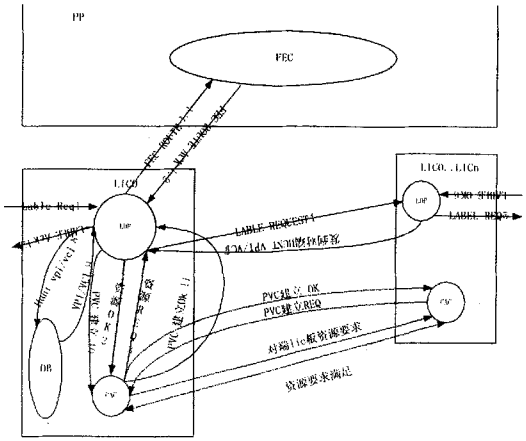
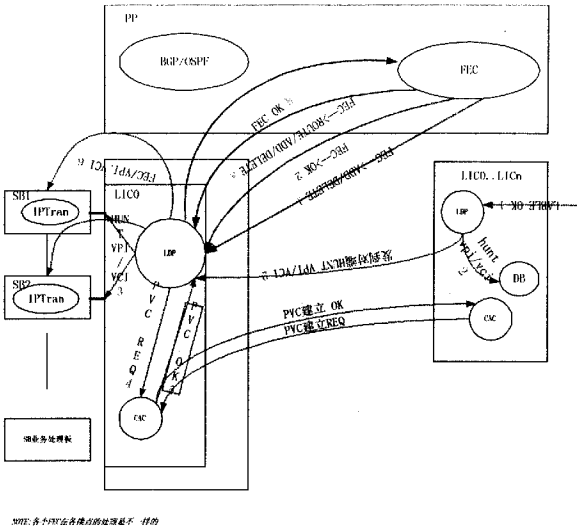


图 ：BackBone的PVC过程

图 6-4 BackBone 的 PVC 过程



图一：Engress的FEC过程

图 6-5 Engress 的 FEC 过程

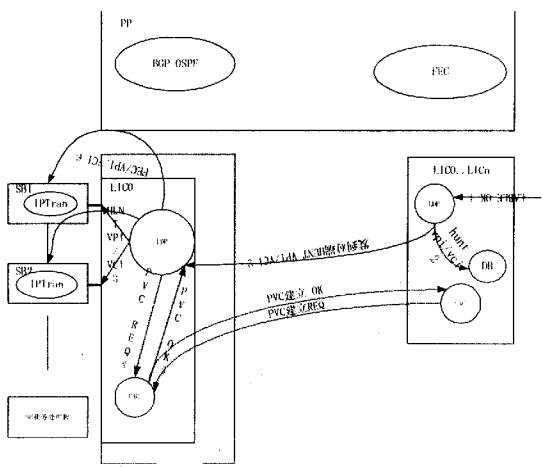
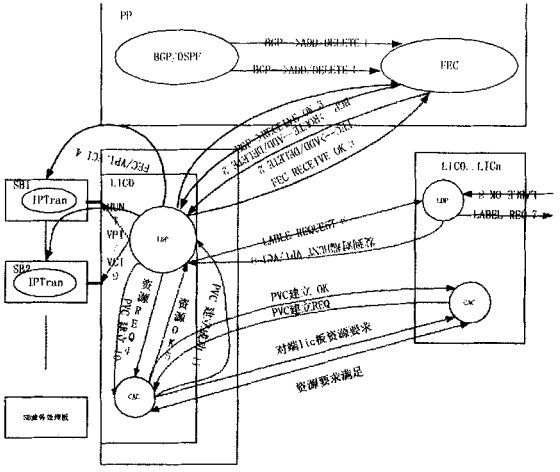


图 二：Engress的PVC过程

图 6-6 Engress 的 PVC 过程



图一：Ingress的FEC分配过程

图 6-7 Ingress 的 FEC 过程

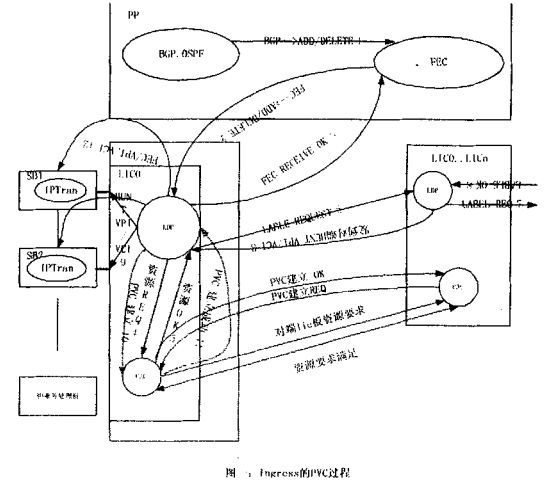


图 ：Ingress的PVC过程

图 6-8 Ingress 的 PVC 过程

6.5.4 模块框图

按照模块化要求，MPLS VPN子系统内包含众多模块，各模块之间数据传输关系如图 6-9 所示：

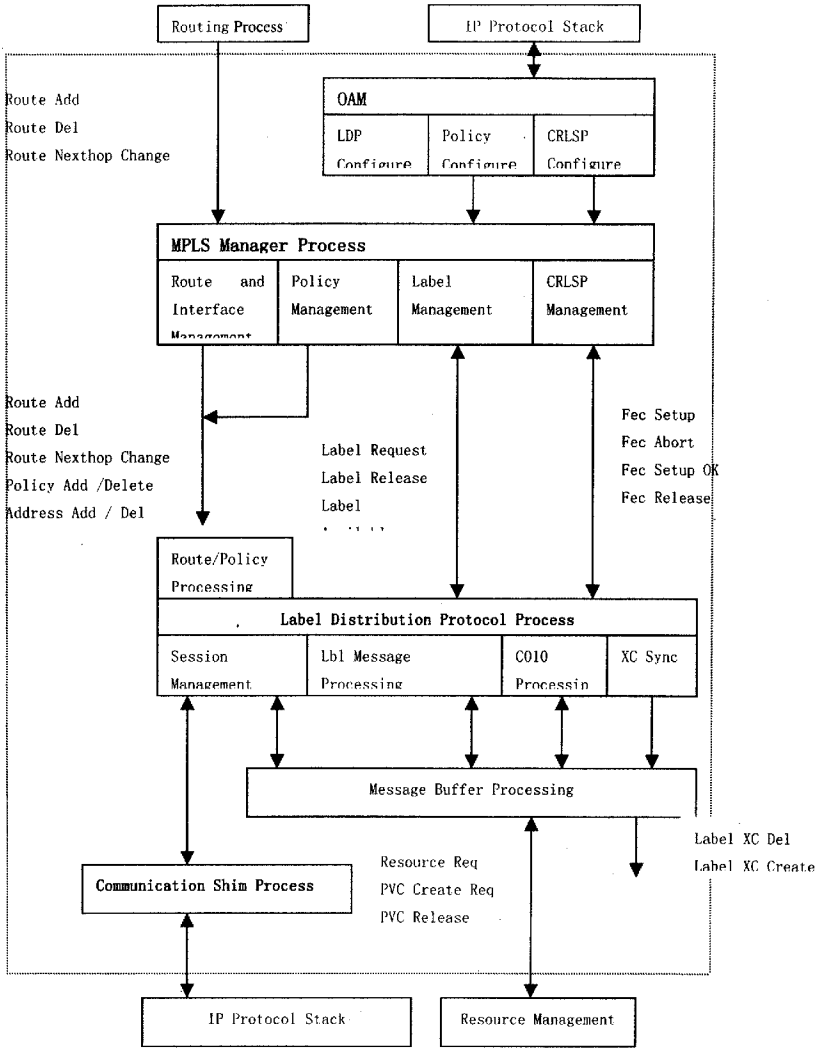


图 6-9 MPLS VPN 子系统模块关系框图

1. 图 6-9 中示例图 6-10：

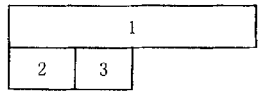


图 6-10 示例说明图

其中

(1) 1 表示：MPLS 模块功能划分，可能在系统中为一个进程（如 MPLS Manager Process、Label Distribution Process、Communication Shim Process），也可能不作为一个进程（如 OAM），只提供函数入口。

(2) 2、3 表示：该部分子功能划分。

2. 在 ZXB10 上，PVC 的操作需要和 ZXB10 C010 进程（Resource Management）交互，图中所示 Resource Management 表示 ZXB10 C010 进程。

3. MIB 管理，系统实现独立的资源 MIB 管理，分别在两个进程中实现，MPLS manager 和 LDP。

4. MPLS manager 标记管理（Label Management），系统提供两种标记管理形式：

(1) MPLS 模块完全负责标记管理（申请、释放），平台实现自身标记管理。如果采用 MPLS 模块完全负责标记管理，MPLS 模块提供对全局标记资源（供 LDP、MP-BGP 使用）、端口标记资源（供 LDP 使用）的管理。

(2) 如果采用平台实现自身标记管理，如 ZXB10 独立管理自身端口 ATM 标记资源，MPLS 模块只提供对全局标记资源（供 LDP、MP-BGP 使用）。

6.5.5 实现功能

MPLS VPN 子系统主要实现功能为：

1. 支持协议：LDP、CRLDP。

2. 支持接口类型：ATM 接口、GEI、FEI。

3. 支持标记类型：ATM 标记（vpi / vci 同时有效）、通用标记。

4. LDP 协议工作模式：

(1) 发现机制：基本发现机制，不支持扩展发现机制；链路 Hello 处理，不处理目标 Hello。

(2) 隧道机制：在输入节点建立 CRLSP Tunnel；在设备本身具备 GRE 等隧道机制的情况下，提供 MPLS over GRE 等机制。

(3) 标记分发模式：下游按需（DoD），下游自主（DU）。

(4) 标记分发控制模式：有序（Ordered），独立（Independent）。

(5) 标记保持模式：保守标记保持模式(Conserve)，自由标记保持模式(liberal)。

(6) 在 ATM 接口上支持 DoD、Ordered、Conserve 模式组合。

(7) 在 GEI 端口上支持 DU、Independent、liberal 模式；在实际组网中，网上设备同时存在工作于 DoD 模式或者 DU、Ordered、Conserve 模式组合的设备时，允许互操作。

5. 独立的标记管理能力：子系统维护一个标记库（包括 LDP 协议、BGP4），LDP、BGP4 在同一库里操作（详细描述见标记管理功能）。包括以下功能：

(1) 路由与端口管理功能。

(2) LDP 协议标记分发功能。

(3) LDP 协议标记分发控制功能。

(4) LDP 协议标记保持功能。

(5) 不同标记编码互操作功能。
(6) 转发层面 Implicit NULL Label / Explicit NULL Label 处理功能。
(7) 策略 Proxy Egress 功能。
(8) FEC 聚合功能。
(9) Egress Targeted Label Assign 功能。
(10) 中继标记映射消息（全网策略一致）功能。
(11) 转发层面映射流量到 LSP 功能。
(12) LDP 处理标记分发消息如下：
① Label Request 消息。
② Label Request Abort 消息。
③ Label Mapping 消息。
④ Label Release 消息。
⑤ Label Withdraw 消息。
⑥ 处理错误通告消息：Notify 消息。
⑦ LDP Address 消息生成与处理。
(13) 适应底层通信差异的消息缓冲功能；

6.5.6 功能模块设计说明

MPLS VPN 网络初始化逻辑流程如图 6-11 所示。

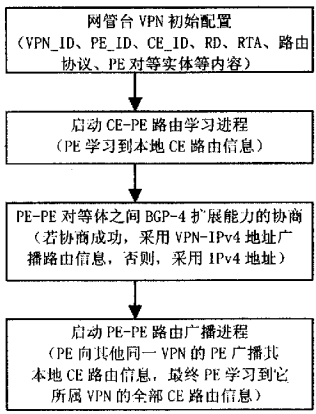


图 6-11 MPLS VPN 初始化逻辑流程图

VPN 数据包传送过程中，所涉及的处理流程如图 6-12 所示。

注：若在 IP GSR 路由器上实现 MPLS VPN，采用上述全部流程；若在 ATM 交换机上实现 MPLS VPN，则不需要内层标签的封装，不需要标签栈的压栈、弹栈操作，骨干网络上外层标签即源端 PE 至目的端 PE 的 PVC VPI/VCI，在目的端 PE，将 ATM 信元还原为 IP 数据包后，再根据目的 IPv4 地址，在目的端 PE 的 VPN 路由信息表中确定出接口至相应目的 CE。

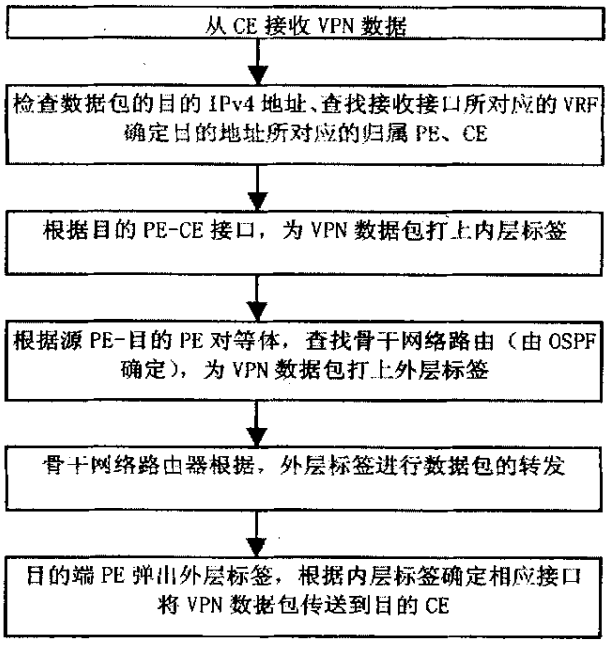


图 6-12 VPN 数据传输流程图

6.5.7 采用的关键技术极其功能

RFC2547bis 定义了 BGP/MPLS VPNS, BGP 用来分发经过骨干网的 VPN 路由并利用 MPLS 将属于一个 VPN 节点的数据转发到该 VPN 的另一节点。

基于 MPLS VPN 的技术优点：

1. MPLS VPN 技术在提供现有 VPN 网络所有能力的同时提供强有力的 QoS 能力。
2. MPLS VPN 用户享受到的保密性与帧中继或 ATM PVC 所提供的相同，这是因为 MPLS 把路由信息的传播限于归属某个特定 VPN 的路由器的范围内。
3. 对 VPN 用户不需要限定 IP 地址空间，从服务提供商的角度不同的用户可以使用重叠的地址空间。
4. 方便用户管理。这基于以下两个方面：不同用户 CE 路由器的路由交换由服务提供商来完成；VPN 用户不需要管理骨干网或虚拟骨干。
5. 不同的 VPN 能够构建于同一个服务提供商的骨干网，利于 VPN 的扩展。

参照图 6-13 MPLS VPN 示意图，将涉及 MPLS VPN 的相关概念加以说明：

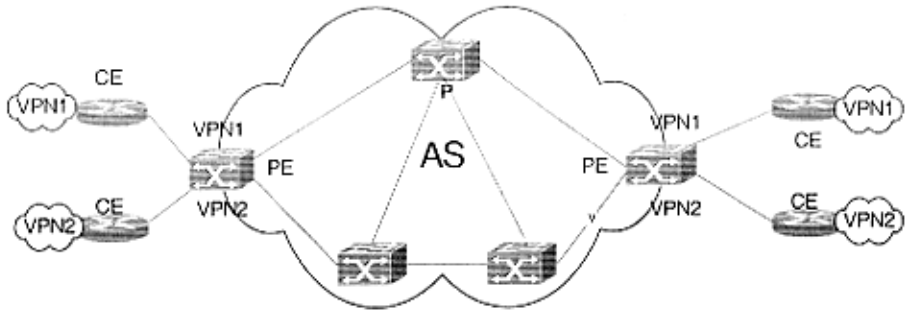


图 6-13 MPLS VPN 示意图

SITE

属于同一个 VPN 的一个 IP 系统的集合可以构成一个“Site”，条件是这些系统间有相互的 IP 连接，并且它们之间的通信不基于骨干网。通常在地理位置上相近的系统构成一个 Site。

CE 设备

CE 设备通过一条或多条数据链路连接 PE 以访问服务提供商的网络。CE 设备可以是一个主机或 L2 交换机，典型的是一个与 PE 直接相连的 IP 路由器。CE 路由器负责将本地 VPN SITE 的路由向 PE 通告并从 PE 获得远端 VPN 路由。

PE 路由器

PE 负责维护 VPN 路由信息。每一用户连接（如 FR VPC，ATM PVC 和 VLAN）都被映射到一特定的 VRF，这样使 PE 的端口与一个 VRF 转发表关联。

PE 从 CE 学习了本地路由之后通过 IBGP 与其它 PE 交换 VPN 路由信息，最终利用 MPLS 将 VPN 数据在提供商的网络中转发，而此时可将 ingress PE/egress PE 的作用看作相应的 ingress LSR/egress LSR。

P 路由器

骨干网路由转发路由器。P 路由器用来维护通达 PE 的路由并用作 MPLS LSR 转发 PE 间的 VPN 数据。

要在 ATM 网络中实现 MPLS VPN，需要实现两大功能：

1. 路由功能：主要包括 PE 与 CE 间的路由学习，PE 向同一 VPN 的其他 PE 路由器广播它所学到的本地 CE 路由信息，另外还需要对 BGP-4 协议进行扩展，以支持 VPN-IPv4 地址（RD+IPv4）。
2. VPN 数据包转发功能：主要为内层标签的分配和外层标签的分配，并且需要建立 VPN Routing and Forwarding （VRF）表。

MPLS VPN 的基本工作方式是采用第三层技术，每一个 VPN 具有独自的 VPN-ID，每一个 VPN 的用户只能与自己 VPN 网络中的成员进行通信，而也只有 VPN 的成员才能有权进入该 VPN。

ATM-MPLS VPN 的工作过程：

1. 用户端的路由器（CE）首先通过静态路由或 BGP 将用户网络中的路由信息通知提供商路由器（PE），同时在 PE 之间采用扩展的 BGP-4 协议传送 VPN-IPv4 的信息以及相应的标记（VPN 的标记，以下简称为内层标记），而在 PE 与 P 路由器之间则采用传统的 IGP-OSPF 协议相互学习路由信息，采用 LDP 协议进行路由信息与标记（骨干网络中的标记，以下称为外层标记）的绑定。到此时，CE，PE 以及 P 路由器中基本的网络拓扑以及路由信息已经形成。PE 路由器拥有了骨干网络的路由信息以及每一个 VPN 的路由信息。
2. 当属于某一 VPN 的 CE 用户数据进入网络时，在 CE 与入口 PE 连接的接口上（根据

网络层 IP 地址）可以识别出该 CE 属于哪一个 VPN，进而查看该数据包的目的 IPv4 地址，在该 VPN 的路由表中查找该地址所归属的目的端 PE 地址信息，并将它作为下一跳地址。同时，在前传的数据包中打上 VPN 标记（内层标记，在 ATM-MPLS VPN 中，可认为内层标记即数据包的目的 IPv4 地址；在 IP-MPLS VPN 中，内层标记是“出口 PE 与目的 CE 连接端口地址”映射成的内层标签）。这时得到的下一跳地址为与该 PE 作 Peer 的 PE 的地址，为了达到这个目的端的 PE，此时在 ATM 骨干网络中通过 BGP 获得骨干路由信息，建立一条源端 PE 至目的端 PE 的 LSP 通路，同时采用 LDP 在用户前传数据包中打上骨干网络中的标记（外层标记，即 ATM LSP VPI/VCI）。

3. 在骨干网络中，初始 PE 之后的 ATM 交换机均只读取外层标记的信息来决定下一跳，因此骨干网络中只是简单的标记交换。
4. 在达到目的端 PE 之前的最后一个 ATM 交换机时，将外层标记去掉，读取内层标记，找到 VPN，并送到相关的接口上，进而将数据传送到 VPN 的目的地址处。

本模块实现特点：

1. ATM 交换机上实现 MPLS 支持 VPN。
2. PE、CE 间的路由学习通过扩展 BGP-4 协议或静态路由协议实现。
3. MPLS 域内同一 VPN 的不同 PE 之间路由广播通过 EBGP 协议实现，其连接通过 ATM PVC 实现。
4. MPLS 域内 PE、P 之间路由学习通过传统的 IGP（如 OSPF）协议实现。
5. VPN 用户可以延用原有的私有地址 IPv4，不需要作任何修改，在骨干网络采用 VPN-IPv4（RD+IPv4），可以保持全网地址的唯一性。
6. 当新的 CE 加入到 VPN 网络时，只需在 PE 上作简单配置，其余的改动信息由 IGP/BGP 自动通知到 CE 和 PE。
7. 网络 PE、P 路由器不需维护和管理 VPN 的所有路由信息，使 MPLS VPN 具有很强的扩展能力。

路由功能特点：

BGP-4 协议路由通告功能的扩展，实现 PE-CE、PE-PE 间路由学习。

1. 支持 BGP-4 能力协商（参见 RFC2842）。
2. 支持 BGP-4 扩展路径属性（参见 RFC2858）。
3. 支持 RD 编码格式（参见 draft-rosen-rfc2547bis-03.txt）。
4. 支持路由目标属性（参见 draft-ramachandra-bgp-ext-communities-09.txt）。
5. 支持标签信息的传送（参见 RFC2842）。

通过对 BGP-4 协议的 OPEN、NOTIFICATION、UPDATE、KEEPALIVE 消息编解码函数的修正，扩展了 BGP-4 路由通告能力，为不同的 VPN 构造各自的 VRF 表。

同一个 AS 内的 BGP 对等体之间可以采用 IBGP 全网互联或使用路由反射器（参见 RFC 1966）扩展 IBGP 路由通告能力。

PE-P 路由学习采用标准的 IGP 协议，例如 OSPF 或 RIP。

外层标签分配函数采用原有的 LDP 进程。

若采用 IP 路由器实现 MPLS-VPN，则需要采用内层标签分配函数，主要将源端 CE 发送来的数据，根据目的 CE 信息以及源端 PE 的 VRF 表，打上标签。

若采用 ATM 技术实现 MPLS-VPN，则不需要内层标签分配函数，可简单地将目的 CE VPN-IPv4 地址作为内层标签，进行寻路至目的 CE。

数据包的转发功能特点：

考虑到 ATM 交换机的硬件特性，在 ATM 网络上实现 MPLS VPN 采用单层标签模式，在 PE 端点之间建立一条隧道，VPN 的报文放在这条隧道内传送到目的地。这种模式不需要内层标签的封装，不需要标签栈的压栈、弹栈操作，骨干网络上外层标签即源端 PE 至目的端 PE 的 PVC VPI/VCI，在目的端 PE，将 ATM 信元还原为 IP 数据包后，再根据目的 IPv4 地址，在目的端 PE 的 VPN 路由信息表中确定出接口至相应目的 CE。或者，在 CE 支持 MPLS 的情况下直接由源端 CE 和 VPN 目的 CE 之间建立数据转发通路。需要注意的是在 ATM 网络实现方式下 PE 上的每一个 VPN 端口必须分配一个公网地址。

工作模式

BGP/MPLS VPN 中 2 个基本的传输流：

1. 控制流：用于 VPN 路由分发和标记交换路径（LSP）建立。
2. 数据流：传输用户数据。

各种 MPLS VPN 互联的工作模式为：

1. 单 AS 系统 BGP PEER 全网互联

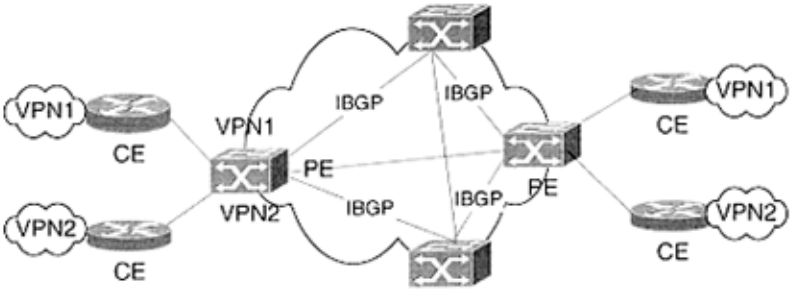


图 6-14 单 AS 系统 BGP PEER 全网互联示意图

如图 6-16 所示，骨干网上 PE/P 由多业务路由交换机组成，CE 是拥有公网地址设置的主机或路由器。AS 系统内 PE 间由扩展的 IBGP 通告路由，PE 与 CE 之间用 EBGP/STATIC ROUTE 通告路由。

2. 单 AS 系统路由反射互联

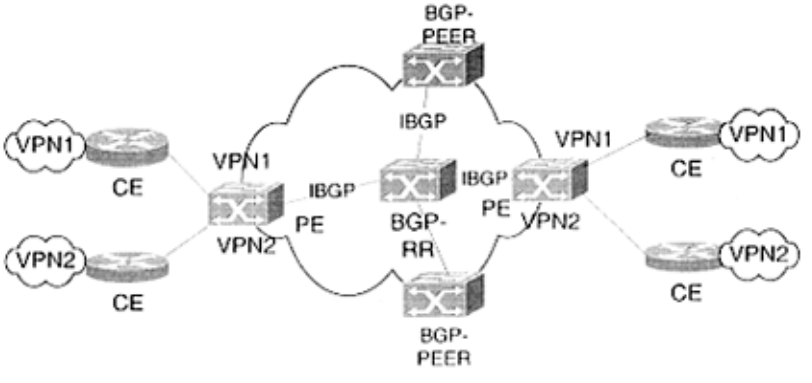


图 6-15 单 AS 系统路由反射互联示意图

在图 6-15 中，骨干网上 PE/P/RR 由多业务路由交换机组成，CE 是拥有公网地址设置的主机或路由器。AS 系统内 PE 间路由通告由路由反射器完成，PE 与 CE 之间用 EBGP/STATIC ROUTE 通告路由。

3. 支持路由反射的多自治系统

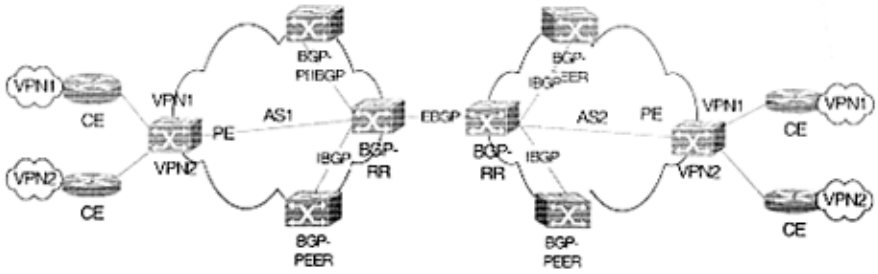


图 6-16 支持路由反射的多自治系统示意图

如图 6-16，骨干网上 PE/P/RR 由多业务路由交换机组成，CE 是拥有公网地址设置的主机或路由器。AS 系统内 PE 间路由通告由路由反射器完成，PE 与 CE 之间用 EBGP/STATIC ROUTE 通告路由。不同 AS 间由路由反射器利用扩展的 EBGP 通告路由。

多层标签体系

考虑到协议的完整性和系统的可扩展性，本方案也要能够适用于支持多层标记的交换体系。标记层数的不同主要体现在 VPN 的标签作用范围：支持多层标记，那么就可以在端点 PE 之间建立隧道，然后在 VPN 内部两个网段之间再建立一条 LSP 连接，采用多层标签实现端到端的全程 MPLS 转发。多层标记模式具有更大的灵活性，可以实现 Carrier of Carrier 等较为复杂的应用。

MPLS VPN 组网能力

MPLS VPN 的数量主要由 PE 连接 CE 的能力和 PE 之间或 PE 与 RR（路由器反射器）互联能力决定。

PE 连接的一个或多个 CE 可以属于一个或多个 VPN，每个 CE 与 PE 上的一个子接口相绑定。将 ZXB10 系列用作 PE，每个 PE 可以拥有“接口板模块数 * 单板子接口个数”个子接口，这些子接口可以连接不同的 CE（VPN）。如果采用线路接口模块，同时使用线路接口复用（ZXB10-S30）则对应每一 VC 连接都可以对应一个 CE，这时 VPN 数量可以达到 4K。因为 ZXB10 系列有着强大的模块扩充能力，一个 PE 可以组合连接多个 CE（VPN）。表 6-1 给出 PE 基于以太网接口的 CE 连接能力：

表 6-1 ZXB10 BX/AX/M1000 PE 基于以太网接口的 CE 连接能力表

类型	以太网接口模块数	最大单板接口个数	接驳 CE（VPN）个数
ZXB10—BX	6	8*100M	48
ZXB10—AX	8	8*100M	64
ZXB10—MX	4	8*100M	32

PE 之间或 PE 与 RR（路由器反射器）的连接由 PVC 构成，连接方式有 BGP PEER 全网互联和通过路由反射互联，当采用 RR 方式时可以显著减少 BGP 对等体的个数和资源占用。ZXB10 系列线路接口模块支持单板 4K 个（双向）VC 连接，并且 VC 连接个数随线路接口模块的扩充而倍数增加，适宜大规模的组网。

6.6 小结

本章主要阐述 MPLS VPN 模块是如何在 ZXB10 系统平台上实现的，从系统需求开始，给出设计依据和原则，最后对该模块子系统各功能、模块框图进行了说明，以及该模块所要求实现的功能和性能要求。

第七章 总参ATM骨干网需求及组网分析

7.1 总参综合信息网需求分析

全军综合信息网是2001年由总参提出，61所进行技术把关，经过对ATM和IP技术及总参信息网需求等多方面考虑，最终选择ATM技术来组建总参综合信息网。而我们的ZXB10设备凭借卓越的性能、良好的性价比和优良的服务系统在国内外众多设备厂商中胜出，获得了总参综合信息网8个军区中的5个军区；

全军宽带信息网络的结构、功能和运作必须符合全军的组织形式、工作职能和工作方法。它是一个高速宽带网络平台，以适应多媒体信息等不同的应用的需要。另外安全保密是全军信息化建设的核心，无论是思想上还是技术上都要树立起安全屏障，适应全军业务和科学决策的需要。选择合适的合作伙伴，借鉴国内外大型网络建设先进经验，确保网络具有良好的前瞻性和可持续发展性。

全军综合信息网网络结构如图7-1：

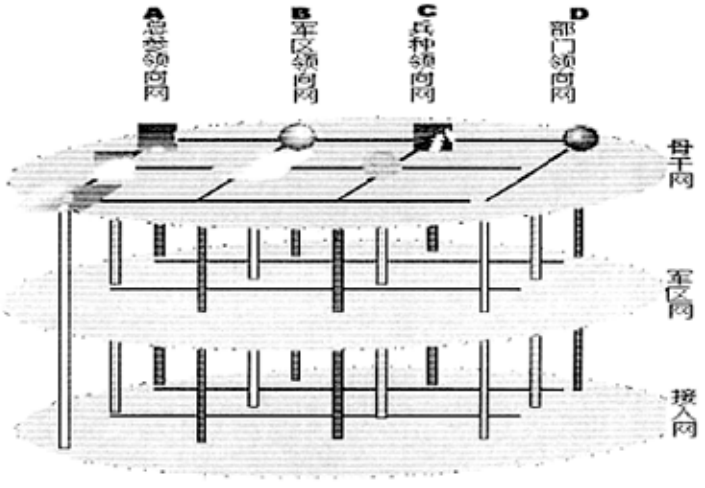


图7-1 全军综合信息网网络结构图

全军宽带信息网络需要为全军陆、海、空三军和二炮建立多级纵向网络，满足全军各军兵种、各军区、各直属单位内部联网需要。同时为总参、军区、省军区、各通讯总站多级通信部门建立横向网络，满足全军各部门间资源共享的需要。其逻辑结构是一个复杂的“格”状立体架构，对于每个独立的军队部门节点来说，它既有横向部门间的信息交互，也有纵向联网的信息交互。军队各部门和各直属单位应用系统之间既具有相对独立性，同时又存在很强的关连性。纵向看，每个军队部门内部的用户均能访问纵向网络的相应资源，横向看，各级军队单位只有授权用户才能访问横向网络资源。面对全军宽带信息网络纵横交错的立体结构，错综复杂的访问关系，种类繁多的业务应用(包括视频会议、IP语音、办公自动化、数据库查询等等)。如何建立一个专用、公共的网络平台，统一实现纵向网及横向网的信息交互，达到每个部门只需建设一个网络、通过一条通信线路，就可以实现纵向及横向的全部通信需要。

目前如果在综合信息网上实现以上要求只能把网络建立在帧中继或ATM网络基础上，通过虚电路（VC）连接各个网络节点，一般采用星形（Hub and Spoke）、树型或半网络拓扑

结构。对于全军综合信息网络的立体交叉拓扑网络来说，如果想在这种模式中实现最佳路由，any-to-any 网络结构，这意味着整个网络需要 n*(n-1)/2（n 为军队单位的数量）条 VC。而 VC 数量的剧增将进一步增加网络和路由的复杂性，这种复杂性使得对网络节点的任何变动都会对全军综合信息网的管理带来极大的痛苦。同时正确地设置 VC 需要了解端到端的业务信息，这使得流量工程也变得非常困难。也就是说，这种模式不具备适应全军综合信息系统大型拓扑结构的良好拓展性和灵活机动性；

7.2 ZXB10 MPLS VPN 打造立体化全军信息网络

ZXB10 MPLS VPN 技术在单一的基础网络设施之上，为多个军队单位，构造多个虚拟专用网络（VPN）。在全军各二级节点、三级节点分别配置一台路由器（PE），构成 MPLS 网络骨干。每个单位配置一台路由器(CE)，通过以太、FR 等汇接到本地 MPLS 网络骨干节点(PE)。PE 和 CE 路由器之间使用标准的 IP 转发。通过路由协议，PE 能够了解每个 VPN 的网络拓扑，简化 CE 间的路由，并轻松实现 VPN 内 any-to-any 的数据包转发，对全军综合信息网络随时可能调整和添加 VPN 的灵活机动性具有非常好的适应性。具体可参阅总参信息网 MPLS VPN 组网示意图 7-2。选用此网络结构具有：

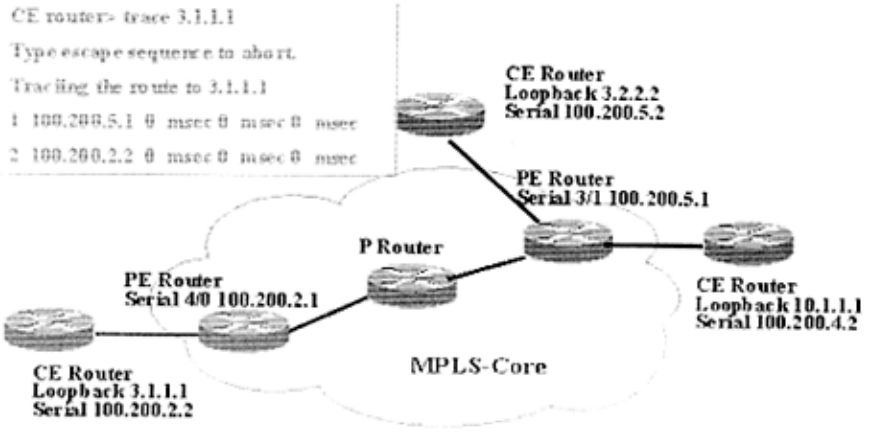


图 7-2 综合信息网 MPLS VPN 示意图

1. 3 层 MPLS VPN 的优势
   - ① 管理的优势：
     - (1) 应用简单：便于网络业务的大规模推广应用。
     - (2) 极好的可扩展性和灵活性：便于大规模部署。
     - (3) 至关重要的增值业务：提高用户的忠诚度。
   - ② 用户的优势：
     - (1) 可以使用私网 IP 地址。
     - (2) 应用简单：不需要太多的路由知识。
     - (3) 管理简单：不需要管理运营商（骨干）网络（相关的链路）。
     - (4) 接入方便：包括接入接口和接口上的路由协议的多样性。
     - (5) 构建大型网络。
2. 堪与专线媲美的安全性——和传统的 ATM 与帧中继 VPN 一样安全可靠
   - ① 地址空间隔离：MPLS 核心采用“VPN-IPv4 地址”路由，通过在 IPv4 路由上添加一个路由分辨符(RD)，确保在 VPN 中独一无二的地址在 MPLS 核心中同样是独一无二的。因此，每个军队部门纵向网具有保持自己的寻址方案的灵活性和使用公共或专用地址空间的自由。

② 路由信息隔离：PE 路由器为每个 VPN 保持一个分离的路由表（VRF），这些 VRF 不仅彼此独立，而且与全局路由表独立。即使有两个军队部门的纵向网络使用相同的地址空间，彼此之间也是完全隔离的。

③ 核心隐藏：在 MPLS 内部连接到 VPN 的接口是 BGP，没有必要透露关于核心的任何信息给用户，即使是对每个军队单位的 CE 路由器。如果在 PE 和 CE 之间使用动态路由协议，CE 唯一知道的信息是 PE 路由器的地址；如果不需要此信息，可以在 PE 和 CE 之间配置静态路由，彻底隐藏 MPLS 核心。

④ 无法伪造 MPLS 标签，伪造 IP 地址也毫无意义。

⑤ 在使用跟踪工具时，MPLS 云在输出时不显示中继段。

7.3 ZXB10 BGP MPLS VPN 解决方案

7.3.1 ZXB10 ATM MPLS VPN 的工作流程

在实际工程中，军方在综合了各种技术组网方案的基础上，在我方售前技术人员的积极引导下，选用了 MPLS VPN 方案来进行全网组网。图 7-3 是采用 ZXB10 ATM MPLS VPN 方案组网的工作流程图：

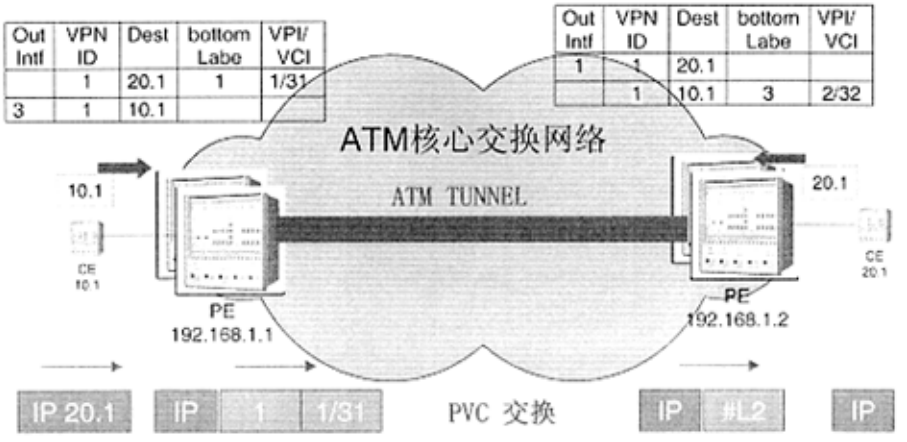


图 7-3 ZXB10 ATM MPLS VPN 工作流程图

地址 10.1 要向 20.1 发送数据，通过 CE 10.1 上的端口 3 将 IP 包发送到 PE 192.168.1.1，通过 VRF 表给该 IP 包打上内标签“1”和外标签“1/31”，该 IP 包通过 ATM 核心交换网络隧道到达 PE 192.168.1.2，这时 IP 包的外标签已经剥离，目的 PE 根据 VRF 表查找到内标签 VPN ID 为 1 的出口地址为 1，将数据从 1 端口发给 CE 20.1，在此之前将内标签剥离。

ZXB10 ATM MPLS VPN 的实现特点：

1. 接入方式灵活：PE、CE 间的路由学习可通过扩展 OSPF、RIP、BGP-4 协议或静态路由协议实现。
2. MPLS 域内同一 VPN 的不同 PE 之间路由，由 MP-EBGP 协议实现。其连接通过 MPLS LSPs 实现。
3. MPLS 域内 PE、P 之间路由学习通过传统的 OSPF/RIP 协议实现。
4. VPN 用户可以沿用原有的私有地址 IPv4，不需要作任何修改，在骨干网络采用 VPN-IPv4（RD+IPv4），可以保持全网地址的唯一性。
5. 当新的 CE 加入到 VPN 网络时，只需在 PE 上作简单配置，其余的改动信息由 IGP/BGP 自动通知到 CE 和 PE。
6. 网络 PE、P 路由器不需要维护和管理 VPN 的所有路由信息，使 MPLS VPN 具有很强

的扩展能力。

ZXB10 ATM MPLS VPN 路由功能特点：

1. BGP-4 协议路由通告功能的扩展，实现 PE-CE、PE-PE 间路由学习。为不同 VPN 构造各自的 VRF 表。
2. 支持 BGP-4 能力协商和 BGP-4 扩展路径属性。
3. 支持 RD 编码格式和路由目标属性。
4. 支持标签信息的传送。
5. 同一个 AS 内的 BGP 对等体之间可以采用 IBGP 全网互联或使用路由反射器（见 RFC1966）扩展 IBGP 路由通告能力。
6. PE-P 路由学习采用标准的 IGP 协议，例如 OSPF 或 RIP。
7. 外层标签分配采用 LDP 协议，实现与 ATM PVC 的相互映射。

7.3.2 单 AS 内全网互联 VPN

在整个综合信息网内各系统联网有多种互联方式，具体有：

1. 单 AS 内通过 MP-IBGP 互联 VPN 实例

如图 7-4，严格地讲，这种方式不会形成物理意义上的拓扑结构，也就是说它不会在站点之间控制二层连接性，而是通过控制站点之间的网络层的连接性来形成逻辑意义上的拓扑结构。

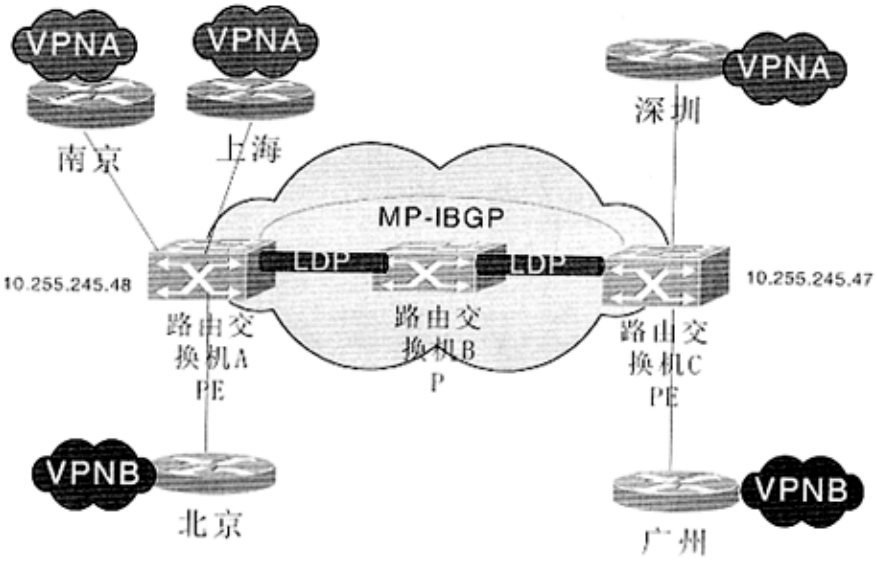


图 7-4 单 AS 内通过 MP-IBGP 互联 VPN 实例

南京、上海、北京、深圳和广州五大军区在一个自治域（AS）内，我们可以通过 MP-IBGP 进行互联，如北京、上海站点（Site）可以通过其自己的 CE 与 PE（路由交换机 A）相连，深圳站点（Site）与 PE（路由交换机 C）相连，而同一自治域内的路由交换机 A 和路由交换机 C 通过 MP-IBGP 互联，从而使上海、南京和深圳站点组成 VPN A。

这个实例说明了如何建立一个全网状网络运营商的 VPN 配置，由下列部分构成：

① 分离的 VPN 客户（VPN A 和 VPN B）。

② PE 路由器，都为 VPN A 和 VPN B 提供服务，PE 之间采用全网互连结构。

③ 信令协议资源预留协议（RSVP）。

④ 在 PE 路由器之间，通过 P 路由器建立标记交换路径（LSP）。

2. 单 AS 内通过路由反射器互联 VPN 实例

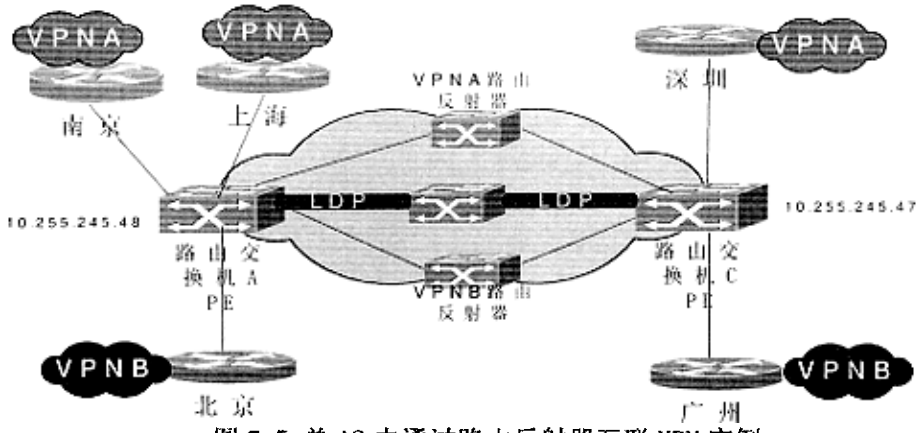


图 7-5 单 AS 内通过路由反射器互联 VPN 实例

此种如图 7-5 所示方式的互联是在同一自治域内的各站点（Site）组成不同的 VPN，各 VPN 在自治域内是通过路由反射器来进行互联，如上海、南京和深圳节点组成 VPN A，其 AS 内的互联是通过 VPNA 路由反射器实现。

7.3.3 Hub-Spoke VPN

通过 Route Target 机制控制 VPN 站点之间路由的分发可以实现站点 A 到站点 B 的数据必须经过站点 C，称之为 Hub-Spoke 结构。

1. Hub-Spoke 路由转发流程：

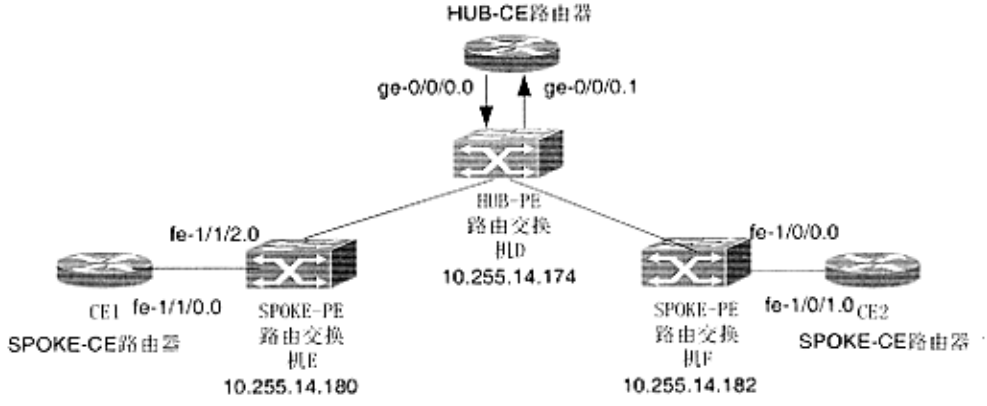


图 7-6 Hub-Spoke VPN 组网

图 7-6 这个实例说明如何建立一个 Hub-Spoke VPN，由下列组件构成：

① 一个 Hub PE 路由器（路由器 D）。

② 一个与 Hub PE 路由器连接的 Hub CE 路由器。为了是实现 Hub-Spoke VPN 拓扑的功能，在 Hub PE 路由器和 Hub CE 路由器之间必须有 2 个接口，每个接口在 PE 路由器必须有自己的 VRF 表：

(1) 一个接口（这里，接口是 ge-0/0/0.0）用于广告 Spoke 路由到 Hub CE 路由器。和这个接口关联的 VRF 表包含 Spoke PE 路由器要广告到 Hub CE 路由器的路由。

(2) 第二个接口（这里，接口是 ge-0/0/1.0）用于接收从 Hub CE 路由器广告的，为 Hub 和 Spoke 路由器定义的路由。和这个接口关联的 VRF 表包含从 Hub CE 路由器广告到 Spoke PE 路由器上的路由。

③ 2 个 Spoke PE 路由器（路由器 E 和 F）。

④ 2 个 Spoke CE 路由器（CE1 和 CE2），每个 CE 路由器连接一个 Spoke PE 路由器。

⑤ LDP 作为信令协议。

2. Spoke 路由器之间的路由分发：

如图 7-7 所示，Spoke-CE 路由器 CE1 与 Spoke-PE 路由交换机 E 相连，通过 Spoke-PE 路由交换机 E 上的 VRF 表，查找目的路由 HUB-PE 路由交换机 D 上的 Spoke to Hub VRF，HUB-PE 进行路由分发，最后通过 HUB-CE 路由器将 CE1 和 CE2 连接起来。

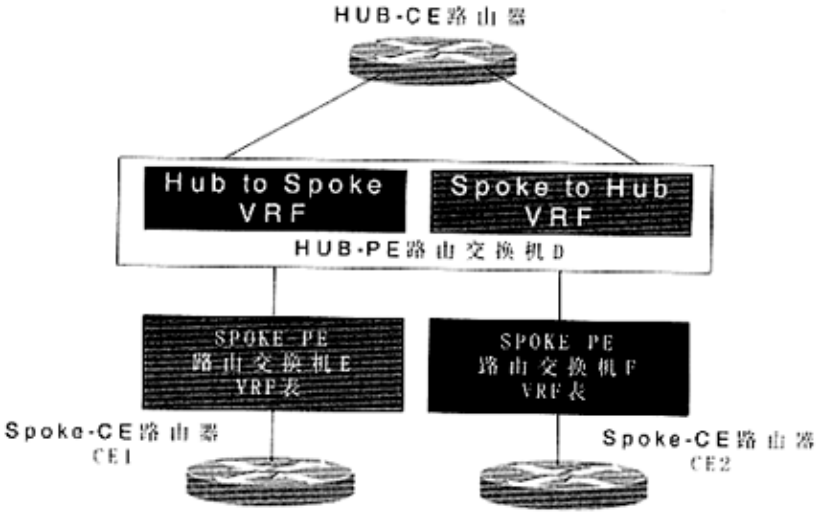


图 7-7 Spoke 路由器之间的路由分发示意图

7.3.4 跨越两个 AS 的 MPLS VPN

跨越两个 AS 的 MPLS VPN 如图 7-8 所示，主要是 ASBR 之间使用 eBGP 广播进行连接。

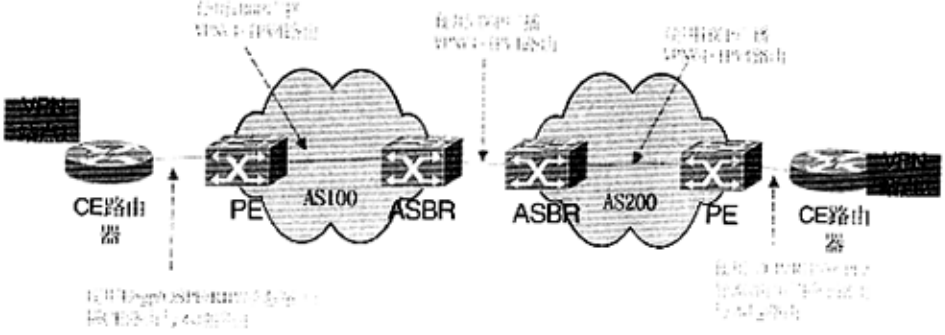


图 7-8 跨 AS MPLS VPN 示意图

注意：2 个自治系统边界路由器（ASBR，Autonomous System Border Router）之间使用 MP-eBGP 时，PE 路由交换机在通过 MP-eBGP 将路由通过另一个 PE 路由交换机之前，将给路由分配一个新标记。用于在两个 AS 之间建立 MP-eBGP 会话的接口不需与任何 VRF 相关联。因为通过 MP-eBGP 获得路由时，将会给它分配标记。每个 AS 之间的 ASBR 路由器都将分配自己的标记，该标记与相关的标记栈关联，以便达到特定的 VPN 目的地。

跨越 AS 的 VPN 路由通告如图 7-9：

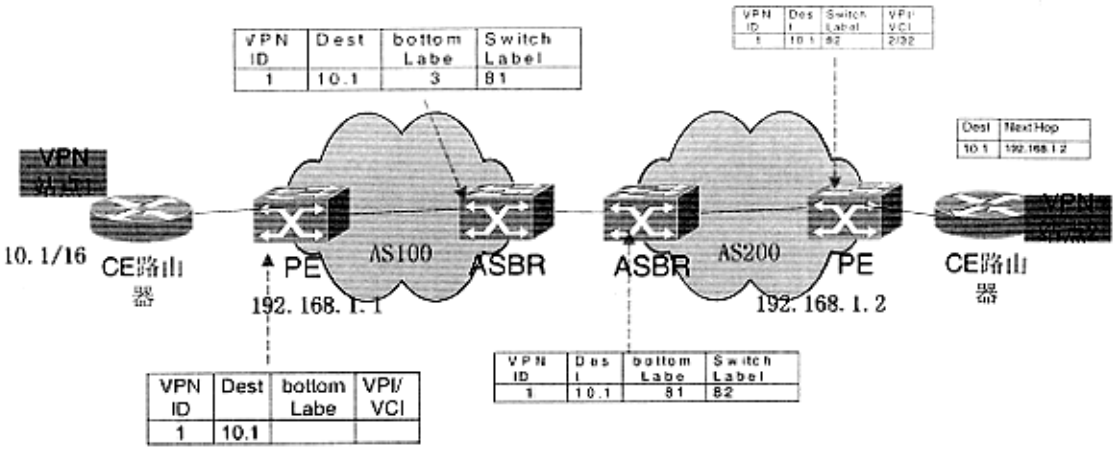


图 7-9 跨越 AS 的 VPN 路由通告

7.3.5 VPN 访问 Internet

通过该 MPLS VPN 组网方式进行组网后，系统内各节点可以各种应用业务，下面提供一些方案说明如何配置 PE 路由器，向 VPN 中的 CE 路由器提供 Internet 接入。不同的网络有不同的需求和特点，采用的方法很多。为了向第三层 VPN 提供 Internet 接入，需要在 PE 路由器进行一定配置。

1. 通过一个 VPN 访问 Internet：

在图 7-10 中，Internet 和 VPN 的接入是分离的。CE 路由器接入 Internet 时独立于 PE 路由器。

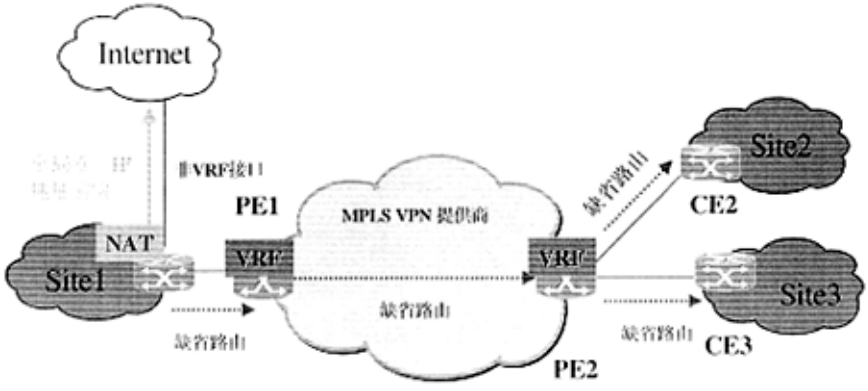


图 7-10 用户通过 VPN 访问 Internet

2. CE 提供 Internet 接入：

如图 7-11，一些 CE 路由器特别配置后，为 VPN 中的其它 CE 路由器提供 Internet 接入。在这个配置中，PE 路由器不提供 Internet 接入。CE 路由器可以将 Internet 流量发送到同一个骨干网络的不同路由器。PE 路由器只处理第三层 VPN 流量。

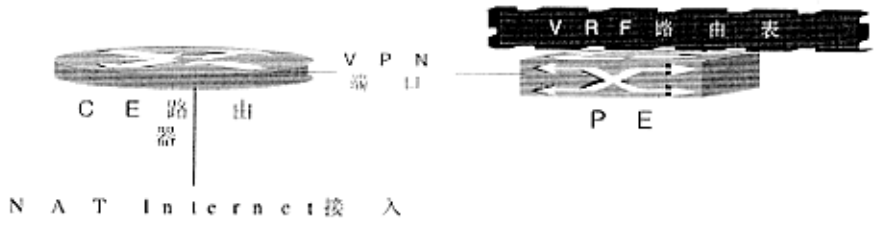


图 7-11 CE 提供 Internet 接入方式

3. PE 路由交换机提供到 Internet 的二层连接：

在这个配置中，PE 路由器扮演一个第二层设备，提供一个二层连接到其它有全部 Internet 路由的路由器。CE 路由器可以在 1 个物理接口上通过 2 个逻辑接口同 PE 路由器相连，或者使用多个物理接口同 PE 路由器相连。（见图 7-12）

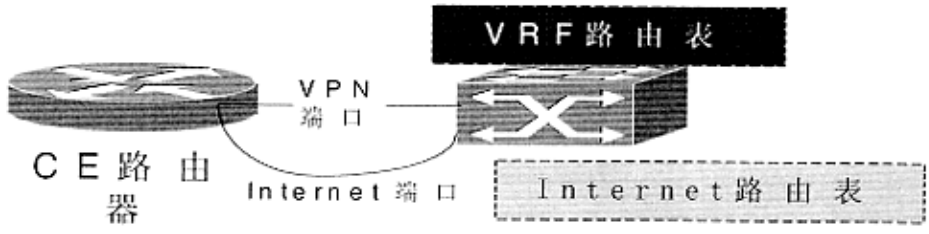


图 7-12 PE 路由器提供到 Internet 的二层连接

4. 通过相同的接口路由 VPN 和 Internet 流量：

这部分说明了如何在1 个单独的接口上做配置，以处理VPN 流量和在Internet 和CE 路由器之间的Internet 流量。这个接口即使在没有私有地址的情况下，可以处理VPN 和 Internet流量。从CE 路由器接收到的VPN 路由被加入到主路由表inet.0 中。这种方式允许 PE 路由器吸引来自Internet 的流量。（见图7-13）

在这种情况下，CE 路由器不需要执行 NAT，因为所有的 VPN 路由是公有的。CE 路由器只有 1 个接口到 PE 路由器，并通过它广播 VPN 路由。PE 路由器在 VRF 表中有 1 个缺省路由指向主路由表 inet.0。PE 路由器同时也将通过使用路由表组，从 CE 路由器收到的入口 VPN 路由。

图 7-13 通过相同接口路由 VPN 和 Internet 流量（VPN 使用公有地址）

5.　通过不同接口路由 VPN 和 Internet 流量：

在这种 Internet 接入方式中，VPN 和 Internet 流量通过不同的接口路由。CE 路由器通过 VPN 接口发送 VPN 流量，通过别在 PE 路由器主路由表对应的接口发送 Internet 流量。（参见图 7-14，CE 路由器可以使用 1 个物理接口中的 2 个逻辑接口，或者 2 个物理接口）。NAT 也发生在 CE 路由器。

PE 路由器配置为安装和广告为这个 VPN 提供的公有 IP 地址池到其它的核心路由器（为回来的流量）。VPN 流量正常路由。

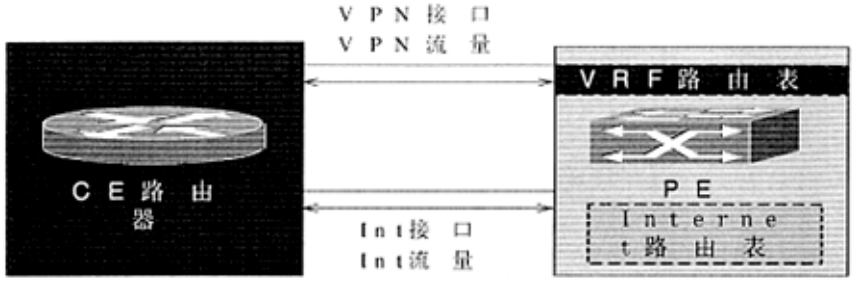


图 7-14　通过不同的接口路由 VPN 和 Internet 流量

7.4　全军宽带网信息网 MPLS VPN 建设关键技术

7.4.1　全军宽带网 BGP/MPLS VPN 跨 AS 解决方案概述

全军宽带综合业务信息骨干网组网结构如下图所示，整个网络有 8 大军区组成，分别为 L 区、C 区、G 区、B 区、S 区、J 区、N 区和 Z 区，其骨干网采用 ZXB10-BX 和 AX 进行联网，每个战区中的所有设备（CE、PE 和 P）组成一个 AS，整个网络采用 BGP MPLS VPN 跨 AS 联网的解决方案，如图 7-15 所示。

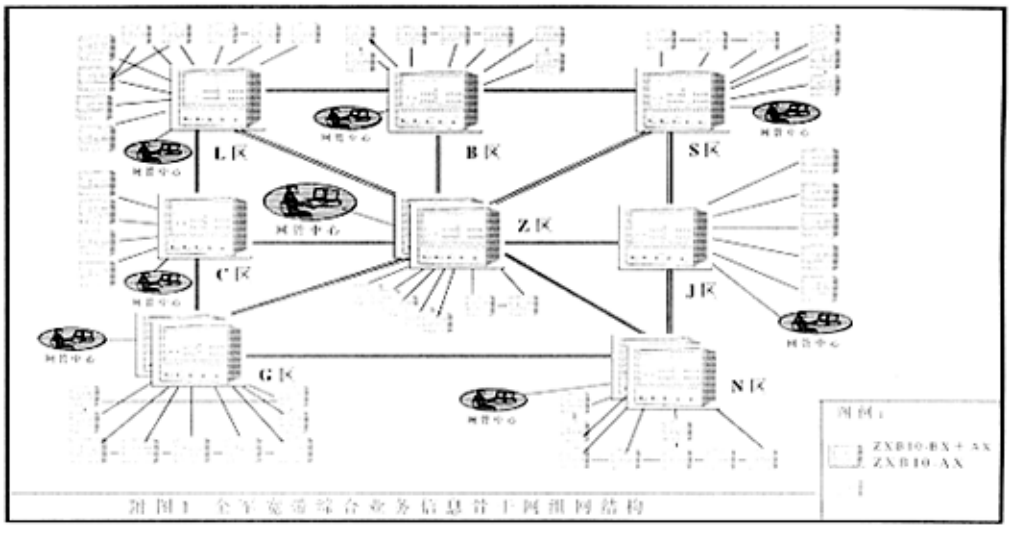


图 7-15　全军宽带综合业务信息骨干网组网结构示意

由于该网为全球最大的 ATM 网络，加上在建设该网络过程中，我们的 MPLS VPN 协议也处在一个设计、开发、调试和完善过程中，所以在协议部署方面我们建议可分两个阶段，如图 7-16 所示：

1.　第一阶段

AS 内：OSPF + LDP/CRLDP。

AS 间：eBGP + CRLDP。

2. 第二阶段

AS 内：OSPF + LDP/CRLDP + MP-iBGP。

AS 间：MP-eBGP + CRLDP。

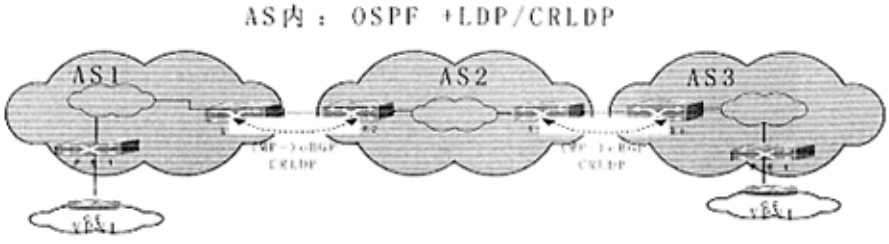


图 7-16 骨干网协议部署示意图

7.4.2 全军宽带网 VPN 用户接入骨干网方式

如图 7-17，ZXB10 系列多业务路由交换平台丰富的接口形式和业务支持，为 VPN 用户提供多种骨干网接入形式，为宽带网 VPN 多种业务开展奠定强有力的基础。

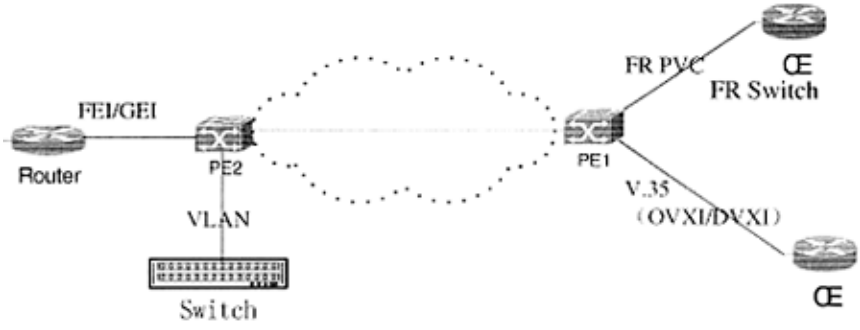


图 7-17 ZXB10 系列多业务路由交换丰富接口

7.4.3 全军宽带网 VPN 用户“Internet”访问方式

1. 集中式访问：如图 7-18，CE 上作 NAT，VPN 用户独立处理 Internet 访问。

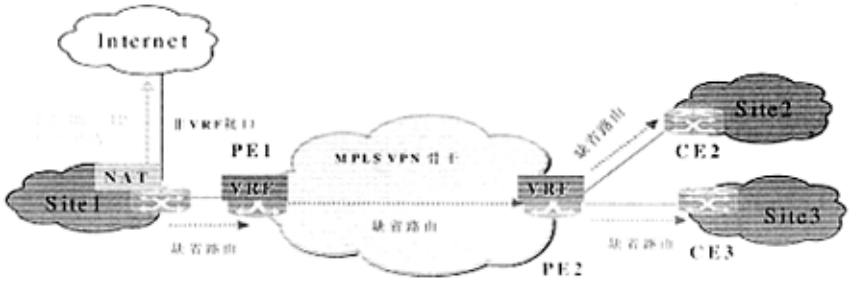


图 7-18 VPN 用户集中式访问 Internet 方式

2. 分散式访问：

如图 7-19 为分散式访问：

① 军网 Internet 实际为公共服务。

② PE 上作 NAT，VPN 提供商为所有 VPN 用户处理 Internet 访问。

③ 保证 VPN 访问与公共服务访问的同时性。

图 7-19　VPN 用户分散式访问 Internet 方式

7.4.4　全军宽带网 VPN 用户接入手段

如图 7-20 所示，全军宽带网 VPN 用户接入手段丰富多彩，具体有：

1. 专线接入：通过本地 DDN/FR/ATM 网，为用户提供本地数据专线接入。
2. LAN 接入：通过 VLAN 的方式实现 LAN 用户的接入。
3. 直接通过光 Modem 把用户的路由器（CE）接入到 PE。
4. 宽带拨号接入：通过宽带接入服务器实现 PPPoE 等宽带拨号用户的接入。

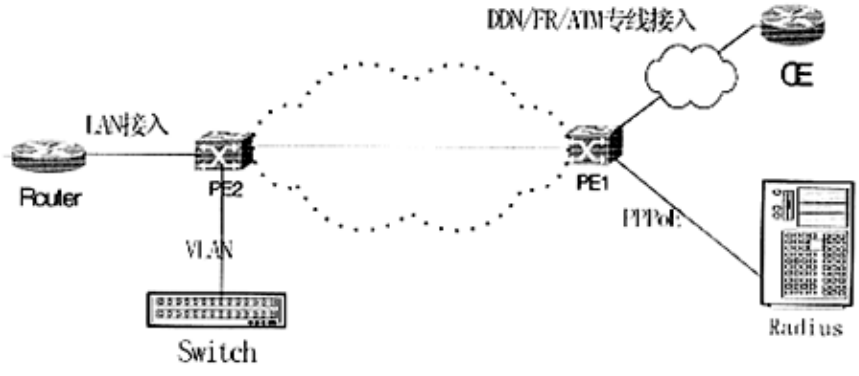


图 7-20　全军宽带网 VPN 用户接入手段多样性

7.4.5　全军宽带网 VPN RD/RT/VRF 命名使用规则

全军宽带网 VPN RD/RT/VRF 命名使用规则如图 7-21 所示。

1. RD 分配原则：
   ① 路由设别（RD，Route Distinguisher）、路由目标（RT，Route Target）和路由转发实例（VRF，VPN routing/forwarding instance）的命名是开展 MPLS VPN 业务所必不可少的号码资源。
   ② 类型区域不支持配置。
   建议：全军宽带信息网中 RD 配置形式为 <ASN>：<32 bit number>，其中 ASN 为合法的 AS 号码。
2. RT 分配原则：
   ① RT 旨在控制 VPN 站点（site）间逻辑拓扑。
   ② RT 结构上与 RD 一样，取值完全独立。
   建议：全军宽带信息网中 BGP/MPLS VPN 的 RT 格式与 RD 一致。

3. VRF 命名原则：

VRF 命名用于识别特定的 VPN。

① 所使用的字符由大小之分。

② 仅在本 VPN 的 PE 有效，属于一种助记符，本身不参与协议过程。

建议：全军宽带信息网 MPLS VPN VRF 命名由各运营实体自由选择，但能够反映服务提供者、业务性质和 VPN 用户的信息。

4. RD/RT 分级分配方案：

全军宽带信息网中 RD 配置形式 <ASN>：<32 bit number>，其中 ASN 为合法的 AS 号码。

① 战区 ASN：<32 比特数字>。

② 总参保留 ASN：<32 比特数字>。

③ 32 比特形成的数字范围为：0～4,294,967,295，共 10 位数字(digits)，取其中后 9 位数字(digits)。

④ N0 N1 N2 N3 N4 P M0 M1 M2，其中 N0～N4 共 5 位数字按照系统或者部门进行划分，M0～M2 数值区间为 0～999 按照应用进行划分，P 在第一阶段分配中恒为“1”也可进行其他形式扩展。

5. VPN IP 地址分配原则：

① 战区网骨干路由器 P 分配公有静态 IP 地址。

② MPLS VPN PE 分配公有静态 IP 地址。

③ PE 与 CE 的互联端口分配公有静态 IP 地址。

④ MPLS VPN 内的地址用户自行分配。

图 7-21 RD 各字节分配

7.5 全军宽带网信息网 MPLS VPN 典型案例

7.5.1 用户能同时访问专网资源和公网资源

1. 方案一：采用在中心节点互联一路由器，作 NAT 完成私网地址和公网地址的转换，如图 7-22 所示：

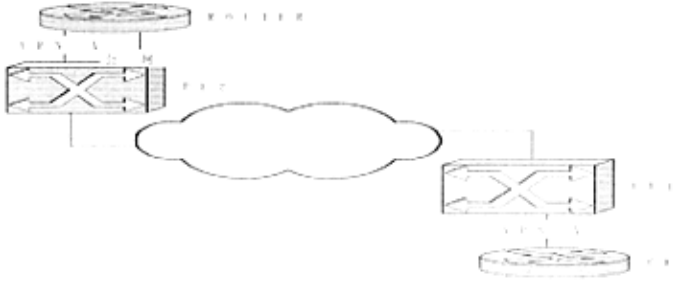


图 7-22 路由器作 NAT 方式

2. 方案二：采用在中心节点 PE 上作 NAT 转换完成私网地址和公网地址的转换，如图

7-23 所示。

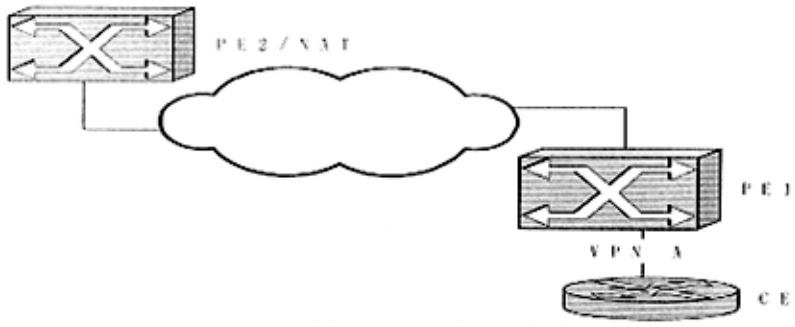


图 7-23　PE 作 NAT 方式

7.5.2　CE-PE 的以太网接入方式

图 7-24 是通过把不同的 VPN 与 VLAN 绑定提供以太网接入。

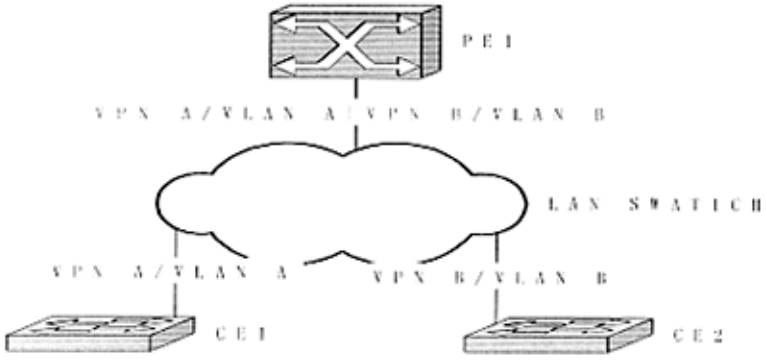


图 7-24 CE-PE 的以太网接入方式

7.5.3　用户同时访问不同的 VPN 既 EXTRANET

如图 7-25，可通过用 RT 导入功能提供不同 VPN 路由的导入完成不同 VPN 的互访。

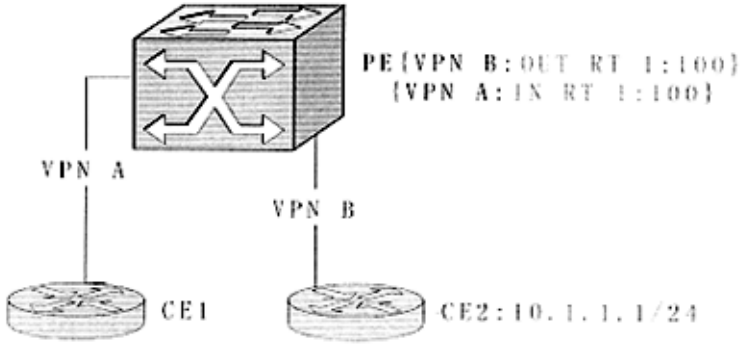


图 7-25 用户同时访问不同的 VPN 既 EXTRANET

7.5.4　提供层次化的 VPN

通过在 PE 上配置 NAT 和 ACL（Access Control List）既可完成地址转换和访问控制。

以实现总参网络能访问其他部门 VPN 而其他部门 VPN 互相隔离，具体如图 7-26 所示。

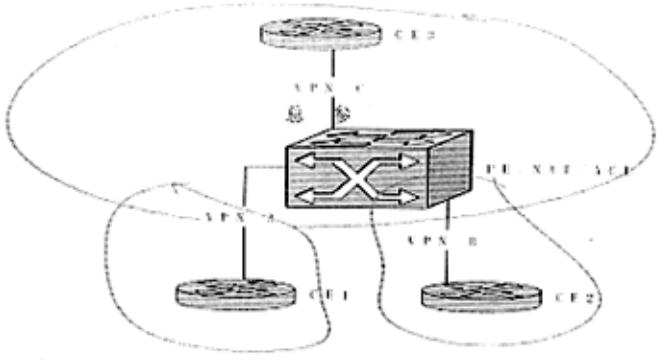


图 7-26　提供层次化的 VPN

7.5.5　站点的 Internet 访问

在图 7-27 上，通过在 CE 上配置 NAT 和 ACL 或防火墙功能完成一个 VPN 内的用户访问 Internet。

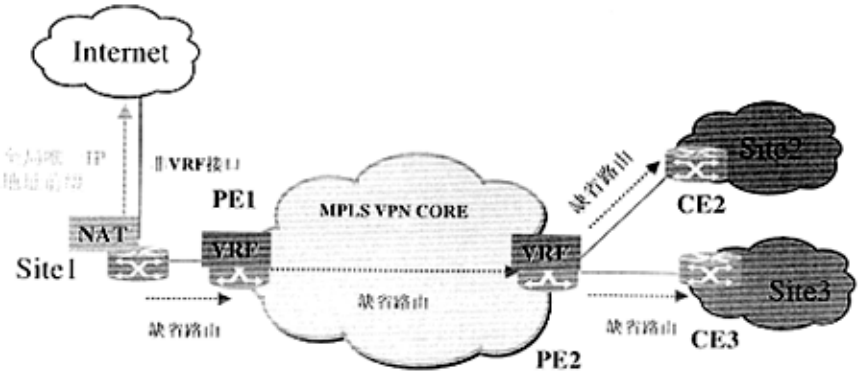


图 7-27　站点的 Internet 访问

7.5.6　基于宽带上网的 MPLS VPN

拨号用户或宽带用户通过拨号或 ADSL/PPPOE 方式连接 PE，PE 通过 RADIUS 得到相应的 VPN 信息，建立 VPN 连接完成 VPN 的移动接入，具体如图 7-28 所示。

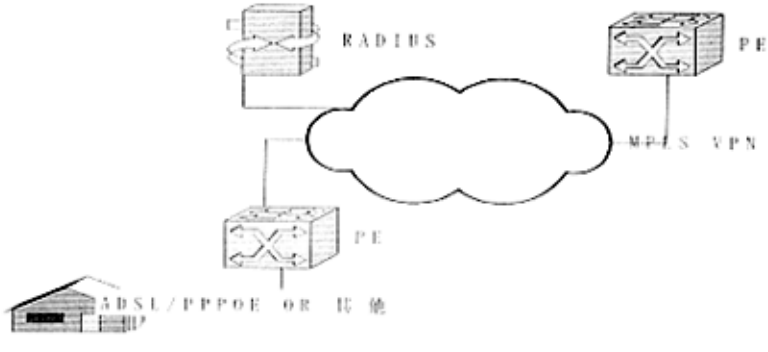


图 7-28　基于宽带上网的 MPLS VPN

7.6　小结

本章首先简要介绍和分析了总参ATM骨干网的系统需求，引出ZXB10 MPLS VPN技术可以满足该网要求，打造立体化全军信息网络。

特别是详细地分析描述如何使用ZXB10 MPLS VPN技术构建总参ATM骨干网，以及使用该技术如何实现各种业务的接入，最后阐述了全军宽带信息网MPLS VPN的典型案例。

第八章 ZXB10 MPLS VPN的应用

全军总参骨干网是我司ZXB10产品MPLS VPN首次的大规模应用，工程在开通过程中遇到不少问题，特别是五大军区设备之间MPLS、MPLS VPN的互联互通。由于各军区之间通讯方式的多样性和通讯质量的差异性，导致MPLS、MPLS VPN某些功能及性能不尽人意，我们的工程人员在现场一蹲就是几个月，我和其他系统设计人员、项目经理一起对故障情况进行分析，寻找问题和故障所在。我们的开发、测试人员根据现场反馈的故障现象和从现场设备中反馈回来的数据进行详尽分析，模拟现场故障现象。然后解决了一个又一个故障，最后系统经历了总参63所及相关专家的多次测试，MPLS VPN在总参信息网中的可使用性及各项功能、性能指标完全满足军方要求，顺利拿到了“工程初验报告”。

而国内另一知名厂家中标的另三个军区设备，不仅工程发货比我们晚了半年，而且在后续MPLS、MPLS VPN功能、性能指标测试中，三个军区之间MPLS、MPLS VPN的互联、互通，前后断断续续持续了一年多时间，互联互通测试非常不尽人意；在与我司设备之间的互联互通问题很多，而且出现问题以后，不能快速反应，尽快解决问题，提供新版本。

03年6月，军队领导在综合考虑各方面的因素之后，最终忍痛割爱，决定由我们中兴的ZXB10设备全面替换另外三个军区的设备，并且对整个总参信息网相关板件进行升级，总合同额达7500万。充分显示了我司ZXB10 MPLS、MPLS VPN的优越性能和良好的性价比，该网最终建成后将是全球最大的以ATM技术组建的骨干网。

2003年年底，系统开通替换工作全面完成，各军区（AS）间开通了MPLS VPN业务，系统经受住了军队严格的测试和业务应用，顺利完成系统初验工作。现在，全军各种业务均可运行于该网络，ZXB10 MPLS VPN必将对我军的通讯现代化建设做出应有的贡献！

现在MPLS VPN模块不仅已加入ATM设备，而且已全面加入我司开发的路由器、以太网交换机等数据产品中，并应用于各电信运营商、军网、公安、广电、学校、银行的实际系统中。

结 论

MPLS 技术的出现，结束了 ATM 和 IP 技术之争，MPLS VPN 成为各种 ATM 设备和 IP 设备之间的桥梁，随着 Internet 在企业领域应用的不断深化，MPLS VPN 作为一种廉价安全的组网方案越来越受到人们的青睐，MPLS VPN 技术已成为各种数据产品之必备功能。

该项目从 2001 年开始立项，目的是要设计开发一套 MPLS VPN 通用模块，该模块可以灵活地嵌入多业务路由交换机、以太网交换机和路由器产品中，并且要求该模块实现功能必须具有很强的市场竞争力，在国内厂家中排名第一，该模块功能首先在 ZXB10 产品上实现。作为中兴通讯一名资深的系统设计人员，本人在这一课题内的主要工作是：

1. 参与项目可行性分析：首先对 ZXB10、路由器和以太网交换机现有系统构架进行分析，并召集各项目相关人员进行讨论，明确开发一套 MPLS VPN 通用模块是可行的。其次对国内外各运营商招标标书以及各设备供应商的宣传资料分析，MPLS VPN 技术将渐成主流，以后没有 MPLS VPN 功能，许多设备将无法销售；再次，从人力、资金和时间投入上以及产品线明确 2002 年底基本系统要在 ZXB10 系统上实现的可能性进行了分析论证，最后项目通过产品线立项评审。
2. 参与项目需求分析：在确定了项目的可行性以后，对项目模块的系统需求进行收集、整理和分析：
   ① 收集国内外运营商招标标书，仔细阅读、分析标书内容，罗列出运营商所需要和关注的 MPLS VPN 各项功能。
   ② 在系统部推广人员广泛收集主要竞争对手华为、Lucent、Cisco 等公司的产品资料、应标标书后，仔细分析解剖他们产品的 MPLS VPN 已经实现了哪些功能？哪些功能尚未开发？具有哪些特有的特殊功能？
   ③ 阅读 RFC、信息产业部等有关 MPLS VPN 的协议草案、标准和书籍。
3. 参与模块设计：这是本人工作重点，在上述需求分析的基础上，确定模块究竟要实现哪些功能？哪些功能首先实现，哪些功能可以后续开发？同时要为后续功能开发的平滑升级留出接口。针对现有系统平台 ZXB10 的软硬件系统特点，确定系统模块各功能点的性能指标，以及模块与 ZXB10 原有软件系统之间的接口要求，最终形成可供开发人员进行详细设计的 MPLS VPN 模块系统设计方案和系统测试方案。
4. 开发指导：在 MPLS VPN 模块系统设计方案确定后，本人并不直接参加该模块的开发调试工作，而是跟踪该模块的开发过程，与具体开发人员一起协商讨论及时解决在开发过程中遇到的设计问题。同时与测试人员一起，根据前述系统测试方案拟制系统测试规程，针对每一 MPLS VPN 功能点，确定测试案例。
5. 参与总参项目组网设计及应标：在项目的设计开发过程中积极参与运营商和专网用户系统的应标工作，特别是总参项目的组网设计和应标，该项目为全球最大的 ATM 骨干网项目，使自己积累了丰富的大型项目的组网经验。
6. 技术积累：在模块合入 ZXB10 系统的系统调试、测试和转产过程中，以及系统开局过程中，与项目组全体同仁一起解决了一个又一个技术难题，使 MPLS VPN 模块的功能和性能指标日趋完善，积累了较多与 MPLS VPN 技术相关的关键技术。

现在 ZXB10 系统已在总参骨干网项目中全面应用 MPLS VPN 功能，并于 2003 年 12 月底全面通过军方的最终验收，投入实际使用。现在正在进行 MPLS VPN 模块向其他路由器和以太网交换机项目的移植和功能优化工作。

致　谢

在我攻读工程硕士的四年时间里，除得到学校许多老师和同学的关心外，还得到了公司许多领导和同事的支持，他们的帮助使得我的论文工作得以顺利开展。

首先，衷心感谢我学校的导师沈连丰教授。论文的许多工作，包括开题、论文撰写等，均是在沈老师的精心指导下完成的。他的言传身教使我收益匪浅。

其次要感谢我的公司——中兴通信股份有限公司，使我有幸能参加 ZXB10 MPLS VPN 这样一个优秀的课题和项目，为我创造了良好的工作环境和科研氛围。另外特别要感谢 MPLS VPN 项目经理孔勇在本项目研发过程中做了很多工作，给予了我很多的帮助和指导。

最后，还要感谢我的亲人们，他们给了我无穷的精神支持。

再次向所有帮助过我的人们表示最诚挚的谢意！

2004 年 4 月于

中兴通信股份有限公司南京研究所

参考文献

1、RFC2858 .Multiprotocol Extensions for BGP-4 [EB/OL]. http://www.faqs.org Jun. 2000.

2、RFC2842. Capabilities Advertisement with BGP-4 [EB/OL]. http://www.faqs.org May. 2000.

3、RFC3031. Multiprotocol label switching Architecture [EB/OL]. http://www.faqs.org Jan. 2001.

4、RFC1771. A Border Gateway Protocol 4 (BGP-4) [EB/OL]. http://www.faqs.org Mar. 1995.

5、darft-rosen-rfc2547bis-03.txt. BGP/MPLS VPNs [EB/OL]. http://www.atm.tut.fi Feb. 2001.

6、draft-kompella-mpls-l2vpn-02.txt. MPLS-based Layer2 VPNs [EB/OL]. http://www.atm.tut.fi Nov. 2000.

7、draft-ietf-mpls-bgp4-mpls-05.txt. Carrying Label Information in BGP-4 [EB/OL]. http://www.atm.tut.fi Jan. 2001.

8、draft-ietf-mpls-bgp4-mpls-05.txt. Carrying Label Information in BGP-4 [EB/OL]. http://www.atm.tut.fi Jan. 2001.

9、draft-nadeau-mpls-vpn-mib-00.txt. MPLS/BGP Virtual Private Network Manage- ment Information Base Using SMIv2 [EB/OL]. http://www.atm.tut.fi Nov. 2000.

10、 draft-muthukrishnan-rfc2917bis-00.txt. A Core MPLS IP VPN Architecture [EB/OL]. http://www.atm.tut.fi Nov. 2000.

11、 BGP MPLS VPN 组网技术应用规范[S]. 中华人民共和国信息产业部 2003

12、 Ivan Pepelnjak Jim Guichard. MPLS 和 VPN 体系结构 [M] 北京：人民邮电出版社，2001.8.

13、 黄锡伟、朱秀吕. 宽带通信网络[M]. 人民邮电出版社，1999.8.

附 录

术语和缩略语：

1、	ACL	Access Control List 访问控制列表；
2、	AS	Autonomous System，自治系统；
3、	ASBR	Autonomous System Border Router，自治系统边界路由器；
4、	ATM	Asynchronous transfer mode，异步传输模式；
5、	BGP	Border Gateway Protocol，边界网关协议；
6、	B-ISDN	Broadband Integrated Services Digital Network，宽带综合业务数字网
7、	CE	Customer Edge，用户网络边缘设备；
8、	CIDR	无类域间路由选择
9、	CIPOA	Classical IP over ATM，重叠模式的 IP over ATM
10、	CLNP	ISO Connectionless Network Protocol，ISO 五连接网络协议
11、	COS	Community of Service，服务群体
12、	DNS	域名服务器；
13、	EF	Expedited forwarding，加速转发；
14、	EGP	外部路由协议
15、	FEC	Forwarding Equivalency Class，转发等价类；
16、	FIFO	First in, first out，先进先出
17、	FR	Frame relay ，帧中继
18、	FTP	File transfer protocol ，文件传输协议
19、	GRE	Generic Routing Encapsulation，通用路由封装
20、	GW	Gateway router，网关路由器
21、	IBGP	Internal Border Gateway Protocol，内部边界网关协议
22、	IETF	Internet Engineering Task Force，因特网工程任务组
23、	IGP	内部路由协议
24、	IP	Internet protocol，互联网协议
25、	IPSec	Internet Protocol Security，IP 安全
26、	Ipv4	互联网协议-第 4 版
27、	IS-IS	中间系统-中间系统
28、	ISP	Internet service provider，互联网业务提供商
29、	ISDN	Integrated service data network，综合业务数字网
30、	ITU-T	International Telecommunication Union - Telecommunication Standardization Sector，国际电信联盟-电信标准化部
31、	LSP	Label Switched Path，标记交换通路；
32、	LANE	Local Area Network Emulation，局域网仿真
33、	LSR	Label Switch Router，标记交换路由器；
34、	L2F	Layer 2 Forwarding，2 层转发
35、	L2TP	Layer 2 Tunneling Protocol，第 2 层隧道协议
36、	LAN	局域网
37、	MPOA	Multiple Protocol Over ATM，基于 ATM 技术的多协议信息传送
38、	MP-BGP	Multiprotocol Border Gateway Protocol，多协议边界网关协议

39、 MSAP Multi-Service Access Point，多业务接入点
40、 MSCP Multi-Service Converge Point，多业务汇集点
41、 NAT 网络地址转换
42、 NHRP Next Hop Resolution Protocol，下一跳路由协议
43、 NSP Network service provider，网络服务提供商
44、 MPLS Multi-Protocol Label Switching，多协议标记交换；
45、 NLRI Network Layer Reachability Information，网络层（路由）可达信息；
46、 NHLFE Next Hop Label Forwarding Entry，下一跳标记转发实体；
47、 ORF Outbound Route Filter，输出路由过滤
48、 OSPF Open shortest path first，开放最短路径优先
49、 PHB Per hop hebavior，每跳行为
50、 PE Provider Edge Router，服务提供商边沿路由器；
51、 PPTP Point-to-Point Tunneling Protocol，点到点隧道协议
52、 P Provider，服务提供商核心路由器；
53、 QoS Quality of Service，服务质量
54、 R Router，路由器
55、 RD Route Distinguisher，路由区别；
56、 RTA Route Target Attribute，路由目标属性；
57、 RR Route Reflector，路由反射器；
58、 RFC Request for Comment，请求评价
59、 RSVP Resource reservation protocol，资源预留协议
60、 RT Route Target，路由目标
61、 RTP Real-time transport protocol，实时传输协议
62、 SLA Service level agreement，服务等级协议
63、 SDH Synchronous digital hierarchy，同步数字网
64、 SMDS Switched Multimegabit Data Service，交换式多兆位数据服务
65、 SNMP Simple Network Management Protocol，简单网络管理协议
66、 TCP Transmission control protocol，传输控制协议
67、 TOS Type of service，业务类型
68、 TTL Time to live，生存期
69、 UDP User datagram protocol，用户数据报协议
70、 VPN Virtual Private Network，虚拟专用网；
71、 VRF VPN Routing and Forwarding，VPN路由转发；
72、 VLL Virtual Leased Line，虚拟专线VPN
73、 VPLS Virtual Private LAN Service，基于局域网仿真VPN
74、 VTP VPN Tunneling Protocol，VPN隧道协议
75、 VLAN Virtual Local Area Network，虚拟局域网
76、 VPDN Virtual Private Dial Networks，虚拟专用拨号网
77、 VPRN Virtual Private Router Networks，基于路由的VPN