

摘要

光纤通信技术的迅速发展，特别是密集波分复用 (DWDM) 技术的发展，使单波长上的数据传输率达到 80Gbps，单根光纤接近 Tbps，从而为利用计算机网络实现有线电视网络、电话交换网络和计算机网络服务的集成提供了通信基础。然而，现有的 Internet 在体系结构上存在着用户数据传输平台层次结构复杂、传输效率低下和服务质量较难得到保障的缺点，使之难以满足高速综合数字业务传输服务的需要。

针对现有的 Internet 体系结构存在的不足，西南交通大学四川省网络通信技术重点实验室提出了“单物理层用户数据传输与交换平台体系结构” (SUPA—Single Physical layer User-data transfer & switching Platform Architecture)，以 SUPA 体系结构构建的网络称作 SUPANET。SUPANET 采用带外信令控制思想将 Internet 现有的协议栈保留在其“信控管理平台” (S&M-platform) 上，以保持与 Internet 的互联互通能力，而利用面向以太网的物理帧时槽交换 (EPFTS – Ethernet-oriented Physical Frame Timeslot Switching) 技术将用户数据传输与交换平台 (U-platform) 简化为单物理层结构，为用户提供有服务质量保证的、面向连接的业务传输服务。

SUPANET 是采用新型体系结构的高速交换网络，其用户平台的数据帧称为面向以太网的物理帧 (EPF—Ethernet Physical Frame)，以最大 MAC 帧长度作为 EPF 帧的数据承载字段长度。随着以太网在局域网市场地位的不断巩固，现有的普通以太网用户能否在已经投资的硬件和软件环境里不做或尽可能少做改动，方便、高效地接入 SUPANET，享受 SUPANET 提供的高速的、有服务质量保证的业务传输服务，是 SUPANET 最终能否被用户接受的最重要因素之一。

本论文提出的以太网接入 SUPANET 的基本思想是：利用 SUPANET 面向以太网和采用带外信令的特性，通过为同一台以太网主机上的每个不同业务流分配一个虚拟的 MAC 地址，从而虚拟出多个“用户”，每个业务流在链路层形成以太网帧时，其源 MAC 地址域的值即为分配的虚拟 MAC 地址，当此帧到达 SUPANET 边缘交换机时，

边缘交换机根据此虚拟的 MAC 地址来查找此业务对应的虚连接，从而可直接将 MAC 帧封装为 EPF 帧，在 SUPANET 网内传输。

论文中的接入技术需要在用户和接入交换机之间的 UNI 信令的支持。UNI 信令主要用来定义以太网用户和接入交换机之间的信息交互的规则和过程，完成以太网用户到 SUPANET 的高效接入。

论文的主要工作可以概括为：

- (1) 详细阐述了接入思想以及接入过程。
- (2) 设计了接入技术中的 UNI 信令。
- (3) 对接入技术的接入流程进行了仿真。

关键词：SUPA；SUPANET；虚拟 MAC 地址；UNI 信令

Abstract

When DWDM technique raised the raw bit rate in a single fiber towards 80Gbps (enough to support all the telephone or cable TV traffic in a big city), it had motivated the idea of merging three separate networks (i.e. telephone, cable TV and computer networks) into an integrated, packetized digital network. However, Internet was not catered for such converged service because of its' complicated three-layer user-data translate platform ,low transmission efficiency and absence of QoS insurance.

For the very reason, SC-Netcom Lab(Sichuan Network Communication Technology Laboratory) has introduced the Single User-data transfer & switching Platform Architecture Network (SUPANET). The SUPANET simplifies User-data transfer & switching platform (U-platform) into a single physical layer platform by adopting the out-band signaling concept, while keeps existing protocol stacks of Internet in the Signaling & Management platform (S&M-platform) to enable interoperation with Internet. The key technique enabling the single U-platform is called EPFTS(Ethernet-oriented Physical Frame Timeslot Switching). SUPANET provides multimedia traffic with a QoS-insured, connection-oriented transmission service.

SUPANET is a high bitrate switch network with a new architecture. The frame in User-data platform is called Ethernet-oriented Physical Frame, its' payload length is the max length of Ethernet MAC frame. With the enhancement of the Ethernet network on the LAN market, whether the Ethernet user can access to SUPANET conveniently and efficiently without any more software or hardware investment is one of the key factors that SUPANET is admitted by Ethernet users.

The basic idea of Ethernet access to SUPANET is to simulate a "user host" for each application traffic by distributed a virtual MAC for that application . When Ethernet frame is created for a application in data-link layer, it's SOURCE MAC ADDRESS field was filled by the

virtual MAC. As that Ethernet frame was transferred to SUPANET access switch, It can be encapsulated directly to a EPF and switched to a correct virtual path by the map table of virtual MAC and VLI (Virtual Line Index) without the need of more decapsulation and encapsulation operation.

A User-Network Interface (UNI) between the Ethernet host and SUPANET access switch should be designed. The UNI signaling defines the roles and process of the interaction between the Ethernet user and SUPANET access switch.

The author's main contributions can be summarized as:

- A .The main idea and access process of Ethernet access to SUPANET are expatiated.
- B .A UNI signaling based on access technology was designed.
- C .The access process was simulated by OPNET.

Key Words: SUPA;SUPANET; virtual MAC address;UNI signaling

第1章 绪论

1.1 本课题的研究的背景

本课题的研究的大背景源于四川省网络通信技术重点实验室关于下一代 Internet 体系统结构的研究工作。通信技术高速化和网络应用数据的多媒体化传输使传统的 Internet 技术难以满足实时音、视频数据的传输服务质量需求。因此，实验室提出了单物理层用户数据传输与交换平台体系统结构(SUPA – Single physical layer User-data transfer & switching Platform Architecture)。根据 SUPA 所设计的高速交换网络称作 SUPANET。SUPANET 最主要的特点有如下两点：第一，采用带外信令技术，所有的信控管理(S&M)平台的信息均是通过传统的 Internet 来进行传输的，和数据传输平台完全分离；第二，用户数据平台在 DWDM 的基础上采用单物理层实现物理帧的高速交换。SUPANET 是面向综合数据业务传输的高速核心网络，能否与现有他类型网络的互连互通是 SUPANET 未来能否走向成功的关键因素。局域网是现有用户应用网络的主要模式，而现有的局域网市场中，以太网占据了绝对的优势地位，现有的普通以太网用户能否在已经投资的硬件和软件环境里不做或尽可能少做改动，方便地接入 SUPANET，享受 SUPANET 提供的高速的、有服务质量保证的业务传输能力，是 SUPANET 最终能否被用户接受的重要因素之一。SUPANET 的设计初衷就是面向以太网的，所以以太网接入 SUPANET 也是整个 SUPANET 系统研究工作中的重要的组成部分。本论文研究的对象正是出于上述问题而提出的。

1.2 SUPA 网络体系统结构与 SUPANET

四川省网络通信技术重点实验室在研究下一代城域网技术时，提出了三维以太城域网(3D-EMAN)^[2]体系统结构，后来又将 3D-EMAN 体系统结构推广到包括广域网在内更广的网络范围，产生了通用的“单物理层用户数据传输平台体系统结构”。采用 SUPA 体系统结构的网络被称

为 SUPANET。

1.2.1 SUPA 借用带外信令技术将用户平台简化为单层

通常被传输的信息可分为用户数据和控制管理信息。最初两类信息在同一信道内传输，即使用同一频带，故称之为带内（in-band）信令传输；考虑到两类信息的性质不同，在某些情况下两类信息的重要性和实时性的要求也不同，于是在某些通信系统中将两类数据分别在不同的信道（频带）内传输，这就是带外（out-band）信令传输。

在 OSI/RM 中，由中间（中继）系统组成的网络部分通常被称为“通信子网”（Sub-network），是为端系统（End-system）提供传输与交换服务的网络部分。出于类似的考虑，网络界将通信技术中的带内、带外传输的概念延伸到通信子网中。由于网络使用层次结构来描述的，习惯上将通信子网中的数据传输与交换的层次结构称为平台。如果两类数据在通信子网中的传输与交换都是在同一平台进行的，则称之为采用“带内信令”传输与交换技术；否则称为“带外信令”技术（注：此时的“带”已不再意味着“通信频带”，而意味着“平台”）。用户数据传输交换平台简称用户平台/平面（U-Platform - User-data transfer & switching Platform）；控制与管理信息使用的称为“信控、管理平台”（S&M-Platform - Signaling & Management Platform）^[4]。

最早采用带外信令思想的网络是电话交换网（相当于通信子网），开始是物理上独立通信线路组成的两个网络，用户数据网用于传输话音；信令和管理网（TMN - Telephone Management Network），用于传输拨号音和网络管理信息。当在后期的发展中，当局间干线被数字通信技术（PDH）取代后，在 2Mbps 的 E₁ 信道中，30 个 64Kbps 的 E₀ 信道被用于传输话音，代表 U-platform；两个 E₀ 信道用于传输拨号与管理信号，代表 S&M-platform^[1]。后来的宽带综合业务数字网（B-ISDN）^[2]也采用了带外信令控制的思想，其结果是将 U-platform 简化为由 ATM 信元交换层和物理层组成的两层平台。MPLS^[6]也采用了带外信令控制的思想，将 U-platform 简化为增强型的两层平台（也有人称之为 2.5 层平台）。

出于对上述两层平台结构的分析，四川省网络通信技术实验室提出可以利用带外信令技术进一步将 U-platform 简化为单物理层平台，

以便解决 Internet 和 MPLS 在服务质量保障问题上由于必须依靠现有数据链路层技术和物理层的服务质量保障机制所面临的窘境^[26]。

1.2.2 SUPANET 的协议层次模型和接口

SUPA 的体系结构将 OSI/RM 和 Internet 中的通信子网的 3 层传输与交换平台的简化为单物理层结构，但在信控、管理平台上保留了 Internet 的协议栈，以便保持与 Internet 的互联互通和保护全球在 Internet 上的投资。SUPANET 可以工作于两种模式：缺省 Internet 模式和 SUPA 模式。缺省模式与现有的 Internet 工作过程完全相同，无需做进一步说明。图 1.1 为工作于 SUPA 模式下 SUPANET 域内的接口与协议层次结构示意图。

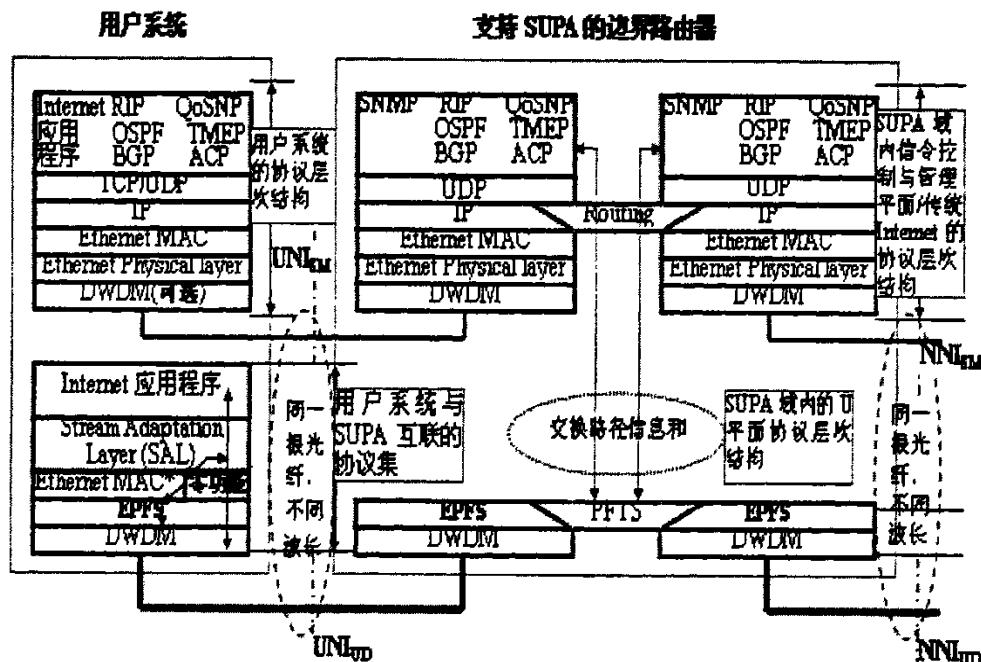


图 1-1 单物理层用户数据传输平台的体系结构工作示意图

如图 1-1 所示，工作于 SUPA 模式下的 SUPA 端系统既可能是支持 SUPA 的计算机系统，也可能是 SUPA 域与传统 Internet 域内的设备互联的网关中支持 SUPA 接口协议的软硬部件。这类网关能够为传统 Internet 设备之间借用 SUPANET 实现互联提供一种高效的“渡船服务”（又称“装载服务”或“隧道服务”）。在这种情况下，由于用户数据的传输是在高速、高效的物理层 EPFS (EPFS- Ethernet-oriented

Physical Frame Sub-layer)平台(图1-1中的EPFS子层)上进行的,对传统的Internet设备来讲,SUPANET相当于不可见高效传输“管道”,因此,其互联效果就像Internet设备直接互联一样。SUPA中间系统是SUPA域中的节点设备,SUPA中间系统间的互联将通过“网络—网络接口”(NNI)来完成;而与SUPA端系统互联时,则通过“用户—网络接口”(UNI)来完成。

图1-1的上半部分表示用户系统和SUPA节点的信令控制与管理平台(S&M platform)中的协议层次结构和信控、管理接口(UNISM和NNISM);下半部为用户数据传输平台的协议层次结构和接口(UNIUD和NNIUD)。在SUPA域内,用户数据传输平台被简化为单物理层,由“类以太网物理帧子层”EPFS和DWDM子层组成。SUPA节点利用“类以太网物理帧时槽交换”技术(EPFTS—Physical Frame Timeslot Switching)实现在多个端口间数据的高速转接。

SUPANET设备也可以工作于缺省Internet模式,此时,接口将按带内信令方式工作。换言之,UNI或NNI的信控管理接口与用户数据传输用接口将合并为一个接口。此时,图1-1中下半部分不再参与数据的交换,因此,上半部的SUPA专用协议(QoSNP、TMEP等)将处于不活动状态,所提供的服务也就退化为传统的Internet服务。为了使SUPA节点能自动识别用户希望使用的服务,可使用IP报头中的特殊字段来标识(IPv4的TOS和IPv6中的Traffic Class)。

1.2.3 SUPANET的SUPA模式下的工作过程

当SUPANET与SUPA端系统互联时,所有节点将工作于SUPA模式。在此模式下,SUPANET将为用户数据流提供虚线路(VL—Virtual Line)服务,信令控制和管理平台内的服务质量协商协议(QoSNP-Quality of Service Negotiation Protocol)将扮演连接建立和服务质量协商的角色。QoSNP将在连接建立过程中充分利用Internet的IP路径信息作为选择输出端口参考,输出端口内输出波长的选择,则需要根据用户请求的传输速率、优先级别以及相关波长剩余的传输服务保障能力等因素进行。只有当一对用户之间所选择的虚通路(VP—Virtual Path)上所有节点都能提供要求的服务质量参数(数据传输率、优先级别、抖动、数据丢弃率等)时,一条单向的虚线路连接VL才能够建立。因此,在

双向通信情况下，需分别建立两条虚连接来提供相应的服务。一条虚通路（VP）由沿着该通路上各节点间的虚线路标识符(VLI – Virtual Line Identifier)共同标识。而 VLI 则是一对 EPFTS 节点间沿传输方向的唯一标识符。

图 1-1 中的流监控信息交换协议(TMEP – Traffic Monitoring information Exchange Protocol)用于在节点间交换网络流量分部与拥塞状况，为节点控制拥塞、统计分析和入网控制提供基本数据。入网控制机制在 UNI 接口上根据局部和全网拥塞状况，设置和修改入网控制子状态控制字(ASCW – Admission State Control Word)，SUPANET 的边界节点的用户数据传输平台（EPF 子层）将根据 ASCW 状态将入网速率限制在承诺速率 CIR、承诺突发数据率 CBR 和额外突发数据率 EBR 之间某值，必要时丢弃超限帧。当拥塞严重时，SUPA 边界节点将利用 ACP 通知 SUPA 端系统，SUPA 端系统将相应地修改 EPF 子层的输出状态控制字(OSCW – Output State Control Word)状态，作为 EPF 子层限制输出速率的依据，同时应当在网关处利用另一侧的流控手段反馈到数据源，以抑制数据发送率。

图 1-1 中的用户数据流适配层 SAL 是 SUPA 端系统中应用层数据流与 EPF 子层之间的长度和速率适配功能层。当 SUPA 端系统为 SUPA 计算机时，该层将按允许发送的速率从用户数据源中获取 1530 字节数据，装入 EPF；在相反方向上将 EPF 中的 1530 字节提交相应的用户进程。当 SUPA 端系统为网关时，SAL 将来自非 SUPA 域的 IP 报文截为 1530 字节装入 EPF，并保证 IP 报文在离开 SUPANET 时能够恢复报文边界。在接收方，SAL 将按界恢复出 IP 报文。

1.2.4 面向以太网的物理帧时槽交换 (EPFTS) 技术

由于“以太网的介质访问控制（MAC）帧”在 95% 以上的用户网络和端系统（用户机或服务器）中广为使用，因此，SUPANET 的物理帧设计必须充分考虑其需要。因此，“面向以太网的物理帧时槽交换”（EPFTS）技术的物理帧（EPF）“净荷部分”的固定长度等于的 MAC 帧（1530 字节），以避免用户（用户网）的数据链路层在 EPFTS 传输与交换过程中的分段和重组。EPF 帧头为 16 字节，帧总长度为 1546 字节。在任一波长数据吞吐率中，单个 EPF 帧的传输时间被定义为用

户数据传输和交换的基本时槽 (Timeslot)^[8]。根据单个波长的数据率很容易计算出该速率下单位时间内可传输的最大帧数，因此，只要单位时间内为用户预留的“时槽数”足够，就很容易满足用户的数据吞吐率需求。至于传输时延和传输时延的变化（即抖动）这类服务质量需求，还进一步需要 EPFTS 交换节点内的服务质量保障机制来满足。EPF 在单波长的光波内和在交换节点内的波长之间，作为独立的数据单元进行传输或交换（中继）。为了避免像 SDH 那样的全网时钟同步需求，物理帧时槽之间以异步方式进行同步。

EPFTS 采用对物理层的传输能力资源进行预留，为不同性质的数据流动态分配最多可用时槽数、即配额 (Quota)，采用有利于减少查表时延的半步超前交换 (HFS - Half-step Forward Switching) 技术^[22,25]，采用新型的交换结构和多优先级别的面向输出的排队与仲裁调度机制等，以满足用户服务质量需求。

EPF 是“面向以太网的物理帧”的简称，由图 1-2 所示的具有组播功能的物理帧头和 1530 字节的用户数据字段（净荷）部分组成。因此，EPF 的帧头开销为 1%，其理论信道利用率为 99%。

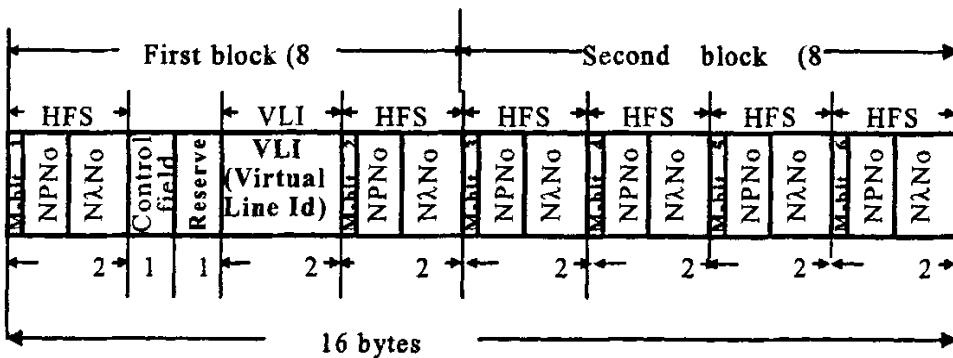


图 1-2 EPF 的交换字段的格式

图 1-2 中的 6 个两字节的 HFS（半步超前交换）字段分别标识最多 6 个组播输出端口和波长，其中 NPNo（Next Port Number）为输出端口号，NλNo（Next Lambda Number）为该端口中波长号。当第一个 HFS 字段的 M-bit 为“1”时，该帧为组播帧，且后面还有 HFS 字段；当第一个 HFS 的 M-bit 为“0”时，为单播帧，后面不再有 HFS 字段；后续的 HFS 字段的 M-bit 为“1”，则表示后面还有 HFS 字段，

为“0”表示该字段为最后一个 HFS 字段。显然，M-bit 6 永远为“0”。图 1.2 中的控制字段（Control field）由突发数据位（B-bit）、结束位（E-bit）、丢弃位（D-bit）和 3 位优先级别位（Priority-bits）组成；该字节中两比特保留未用；Reserved 字节也保留未用。图中的虚线路标识（VLI – Virtual Line Identifier）用于在一对 EPFTS 节点间的端口对内标识同一波长中 65,536 个虚连接。学界和业界发现各种网络信息流的数据到达具有自相似性（Self-Similarity），即在不同的观察尺度下获得的“数据幅度/时间”谱，在较为平滑的曲线上都存在许多随机的高峰谱线，这些谱线代表网络中的突发性数据。突发数据流的存在，可能源于单个用户的连续数据流，也可能是由于交换节点的处理特性造成的。图 1-2 中的突发标志位（B-bit -Burst bit）是为了保证同一实时性数据流的虚连接中的突发数据的连续性和整体性而设置的。

图 1-2 中的丢弃标志（D-bit - Discard bit）是为了出现紧急情况时，将部分对服务质量影响较小的实时数据或可以通过重传而恢复的非实时性数据设置的。可丢弃帧可以根据优先级别和结合应用数据性质和协商的服务质量确定。为了保障商定的服务质量，单位时间内允许丢弃的帧数是服务质量协商中的参数之一，EPFTS 节点将根据协商结果控制单位时间内丢弃的帧数，以保证不会由于丢弃一定数量的 EPF 而降低服务质量。

图 1-2 中设置的 3 比特优先级别标志（Priority Bits）与以太网定义的 8 个优先级别相适应。但是以太网的相关标准中并未定义如何使用优先级别，而留待用户处理，这可能造成用户将各自的数据都设置为最高优先级别的局面，而失去设置优先级别的意义。在 EPFTS 中，优先级别的设置将根据服务协商阶段承诺的服务质量，由 EPFTS 域内的边界路由器设置。

图 1-2 交换字段中最后一个比特称为“结束标志位”（E-bit – Ending bit）用于结束虚连接使用。当通信双方之一要停止发送数据过程时，将利用信控管理平台内的连接结束协议（CEP – Connection Ending Protocol）与通信对方协商结束事宜。当对方同意结束该方向的数据传输时，发回确认信息，此后，连接结束发起方将发送连接结束执行 PDU，EPFTS 域内的入口边界路由器将该 PDU 映射为 E-bit 置“1”的 EPF，PFTS 域内所有节点收到 E-bit 为“1”的 EPF 后，将继续前传该帧，同时将相关虚连接占用的系统资源归还系统。结束通信的过程通常仅

涉及单个方向，但 CEP 协商过程中也可以双向同时进行。

1.3 国内外研究现状与本课题的研究意义

本论文所研究的接入技术是四川省网络通信技术重点实验室在研究“单物理层用户数据传输与交换平台体系结构网络”(SUPANET)的过程中，为简化接入交换机在对进入 SUPANET 网络的以太网帧的解封装过程(从链路层到网络层，甚至到传输层)和封装过程(从网络层或传输层到链路层)、提高对以太网用户的业务传输服务质量保证能力而提出的一个新思想。目前在国内外未见类似的研究。

本论文阐述的接入思想的意义在于通过为业务虚拟一个 MAC 地址来标识一个业务流，通过此虚拟的 MAC 地址来唯一对应 SUPANET 为此业务流所建立的虚连接，从而实现以太网帧直接封装到 EPF 帧在网内传输。由于简化了以太网在接入交换机上的处理流程，从而改善了业务传输过程中的延时和抖动，削弱了接入交换机处的传输瓶颈，进一步保障了业务的服务质量。

1.4 本论文的组织

本论文主要针对 SUPANET 面向以太网这一特点，初步设计了以太网用户接入 SUPANET 的相关机制和 UNI 信令，设计目的是在 QoS/NP 协议和 CEP 协议的配合下实现“用户到用户”的“全交换通路”，以求解决局域网和广域网之间因协议类型和协议层次的转换而造成的接入交换机瓶颈问题。

全文共分为五章，第一章主要介绍了本论文的研究背景，并较详细地介绍了整个 SUPA 体系结构，以及在此体系结构下信控管理平台和用户数据平台的工作机制和工作流程。

第二章主要讲述了以太网接入现有的高速传输网络如 ATM 网络和帧中继网络的相关原理和技术，并对其接入复杂程度和数据传输性能作出了初步的分析。

第三章对本文提出的以太网接入 SUPANET 技术作出了较详细的阐述。从接入的数据传输流程到所涉及到的相关问题作出了较为详细

的讨论。

第四章主要讨论了普通以太网用户和 SUPANET 接入交换机之间的用户—网络接口信令。论文对信令的消息类型、信息元素类型进行了定义，并对消息格式、信息元素格式作出了详细的设计，同时对信令的相关流程也进行了详细的阐述。最后对与所设计的 UNI 信令和 SUPANET 域内所使用的 NNI 信令的转换也作出了探讨。

第五章针对接入技术的工作流程进行的模拟，主要验证了工作过程中消息传递的正确性，以及 UNI 信令和 NNI 信令的配合的正确性。初步验证接入技术的可行性。

最后是本文的结论与展望。

第 2 章 以太网接入高速网络分析

2.1 以太网接入高速网络的必要性

2.1.1 以太网的发展历程及应用

以太网标准是一个古老而又充满活力的标准。自从 1982 年以太网协议被 IEEE 采纳成为标准以后，已经历了 20 年的风风雨雨。在这 20 年中，以太网技术作为局域网链路层标准战胜了令牌总线、令牌环、wangnet、25M ATM^[12]等技术，成为局域网事实标准。以太网在当前在局域网范围市场占有率超过 90%。以太网也成为用户应用网络一种主要应用模式。

在这 20 年中，以太网由最初 10M 粗缆总线发展为 10Base5 10M 细缆，其后是一个短暂的后退：1Base5 的 1M 以太网，随后以太网技术发展成为星形的双绞线 10BaseT。随着对带宽要求的提高以及器件能力的增强出现了快速以太网：五类线传输的 100BaseTX、三类线传输的 100BaseT4 和光纤传输的 100BaseFX。随着带宽的进一步提高，千兆以太网接口粉墨登场：包括短波长光传输 1000Base-SX、长波长光传输 1000Base-LX 以及五类线传输 1000BaseT。2002 年 7 月 18 日 IEEE 通过了 802.3ae：10Gbit/s 以太网标准。10Gbit/s 以太网是一种城域以太网技术，其设计目标就是为了将以太网技术应用到城域骨干网之中，以构建一个从用户网—接入网—骨干网的全以太网域网络。

千兆以太网接口基本应用在点到点线路，不再共享带宽。碰撞检测，载波监听和多重访问已不再重要。千兆以太网与传统低速以太网最大的相似之处在于采用相同的以太网帧结构。万兆以太网技术与千兆以太网类似，仍然保留了以太网帧结构。通过不同的编码方式或波分复用提供 10Gbit/s 传输速度。所以就其本质而言，10G 以太网仍是以太网的一种类型。

在以太网技术中，100BaseT 是一个里程碑，确立了以太网技术在桌面的统治地位。而千兆以太网以及随后出现的万兆以太网标准则将

局域网技术延伸到校园网以及城域网的汇聚和骨干。

以太网以其传输速率高、技术简捷的技术特点和建网速度快、维护和升级容易的应用特性在网络领域获得了广泛的应用。可以预见，在将来相当长的一段时间内，以太网仍将是网络应用的主要模式。

2.1.2 广域网的历程

计算机的远距离通信引发了对广域网技术的研究，早期的广域网以线路切换(Line Switch)为其主要的技术支撑。在这种方式下，通信双方需要在通信期间独占一条由多个交换机转接多段线路形成的一条物理通路，因此线路的利用率较低，最典型的代表是早期的电话网络网络。

报文存储转发思想的诞生是广域网技术发展史上的一个里程碑，也是现代分组交换网的起点。X.25 网络正是基于分组交换的公用数据网。由于当时通信线路传输速率较低，且误码率高，而且传输业务对数据的准确性要求较高，故 X.25 网络采用三层结构的通信子网，同事在第二层和第三层逐段进行确认和流控，传输的效率低下。

随着数字通信线路误码率大大降低和智能终端的出现，诞生了两层结构的通信子网的公用数据网—帧中继网络^[14,15,16]。帧中继网络采用带外信令，实现信令信号和用户帧传输信道的分离。在用户平面采用只有物理层和数据链路层的两层结构，并取消了数据帧在传输过程中的逐段确认操作，大幅提高数据转发的速率。

随着多媒体业务在网络上传输的不断增多，帧中继网络越来越难以满足这类业务传输的服务质量需求，综合业务数字网技术（B-ISDN）应运而生，作为 B-ISDN 网络支撑的 ATM 技术得到了较好的发展。ATM 保留了帧中继网络面向连接的传输特性，采用了 53 字节定长信元作为基本的交换单元，进一步提高了数据转发的速率，从而提高了传输业务的服务质量。

可以预见，随着通信技术和电子科学技术的发展以及光纤链路的使用，将会出现具有更高交换速率的交换技术。

2.1.3 广域网和局域网的融合统一趋势

局域网技术和广域网技术一直在沿着两条不同的道路向前发展，在发展的过程中注重自身技术参数和网络性能的提高，但对网络本身与其他网络的融合则缺乏足够的思考，而从而产生过各种各样不同技术类型的网络。局域网中出现了以太局域网、令牌环局域网以及令牌总线局域网；而在广域网中随着通信技术和终端的发展，产生了 X.25 网络、帧中继网络以及 ATM 网络。多种类型网络的存在因应用环境的不同和技术发展层次差异有其历史的必然性和合理性。但是，网络类型的不同意味着在网络互连互通的过程需要作出更多的努力：

- (1)不同网络协议的转换对传输效率的影响。
- (2)不同的网络控制机制和策略使得传输业务受到影响。
- (3)不同的网络的服务质量保障机制对业务传输服务质量的影响。

异种网络间的差异造成的影响远远不止上述三点，其对用户业务的服务质量的影响是显而易见的。不同类型的网络是否会很快被一种统一成一种网络技术呢，答案是否定的。首先，在技术层面上现在还没有一种网络能够适合局域网、广域网应用要求，以统一局域网和广域网；其次，不同类型的网络在现实生活中均有应用，大量的固定设备投资使得各种网络不可能很快退出。

虽然网络类型统一这个梦想很遥远，但某些网络技术还是在这方面作出了努力，如 ATM 网络本身是作为广域网技术而发展起来的，但在一段时间内 ATM 局域网发展也非常迅速，当时有人甚至认为 ATM 可能统一广域网和局域网，但最后没有成功。而现在以太网在局域网领域已经确定其垄断地位的情况下，随着 10G 以太网的出现，渐渐渗透到城域网领域。虽然我们现在就下结论以太网技术将会是统一所有网络的技术还为时尚早，但却体现了网络将逐渐融合的趋势。

随着多媒体业务流在局域网和广域网业务流上的比重不断增加，对业务传输的服务质量要求也不断提高。局域网和广域网之间的高效、简捷的实现互连互通，是需要重点考虑的问题。

本章就以太网接入帧中继网络和 ATM 网络的相关技术作出了初步的探讨，最后对其不足作出了简要的分析。

2.2 以太网接入帧中继网络

2.2.1 帧中继网络

帧中继技术是在 OSI 第二层上用于简化的方法传送和交换数据单元的一种技术，它是在分组技术充分发展、数据与光纤传输线路逐渐替代已有的模拟线路、用户端日益智能化的条件下诞生并发展起来的。帧中继仅仅完成 OSI 物理层和链路层核心层的功能，将流量控制、纠错等留给只能终端去完成，大大简化了节点机之间的协议；同时，帧中继采用虚电路技术，能充分利用网络资源，具有吞吐量高、时延低、适合突发性业务传输等特点。应用在广域网（WAN）中，支持多种数据型业务，如局域网（LAN）互连、文件传输、图像业务查询、图像监视等。

2.2.2 以太网接入中继网络

利用帧中继网络进行局域网互连是帧中继最典型的一种应用。在已建成的帧中继网络中，其用户数量占 90% 以上，这主要是因为帧中继很适合为局域网用户传送大量突发数据。

帧中继网络在业务量少时，通过带宽的动态分配技术，允许某些用户利用其他用户的空闲带宽传送自己的突发数据，实现带宽资源共享，降低了用户的通信费用；在网络业务量大甚至发生拥塞的情况下，由于已为每个用户分配了 CIR，网络按照用户信息的优先级及公平性原则，将某些超过 CIR 的帧丢弃，并尽量保证未超过 CIR 的帧的可靠传送，因此用户不会因拥塞造成数据的不合理丢失。上述特点使得帧中继网络很适合为 LAN 用户提供服务。

以太网一般通过带有帧中继接口的路由器接入到帧中继网络中，此时，路由器相当于以太网和帧中继网络中的网关^[17]。

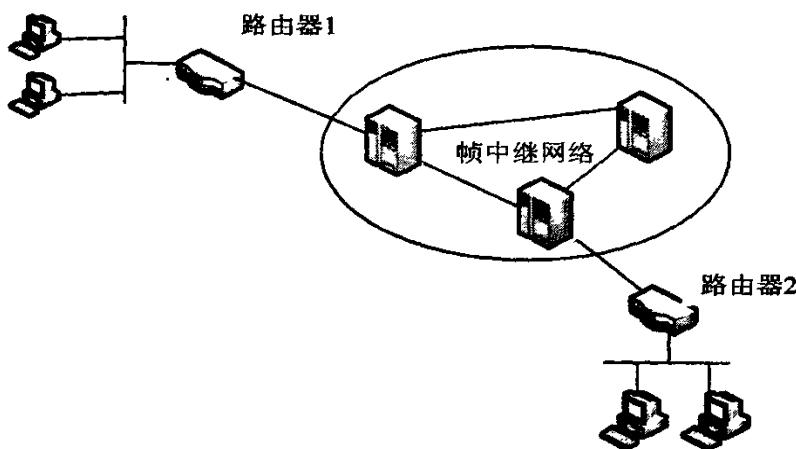


图 2-1 以太网接入帧中继网络

如图 2-1 中, 路由器 1 通过帧中继交换机之间的单条物理链路, 可以和路由器 2 建立一条逻辑连接, 一般都是永久虚连接, 由帧中继网络管理员事先配置, 也可以由路由器 1 使用控制平面的 LAPF 和 Q.933 协议, 通过呼叫建立交换虚电路。逻辑连接建立之后, 在各个帧中继交换机建立关于这条虚连接的交换表, 以太网中的数据就可以送到路由器中在帧中继网络中传输了。

当路由器收到一个来自以太网的数据帧时, 提取其中的 IP 报文, 根据其目的 IP 地址来确定对应的虚连接在帧中继接口上的 DLCI。这种映射关系在建立虚连接时就已经确定, 然后将 IP 报文封装到 LAPD 帧中, 根据交换表在帧中继网内转发。

当此 LAPD 帧到达目的端路由器 2 后, 把网络层报文从帧中分离出来, 重新封装成 MAC 帧, 发往对应的以太网主机, 过程如图 2-2 所示。

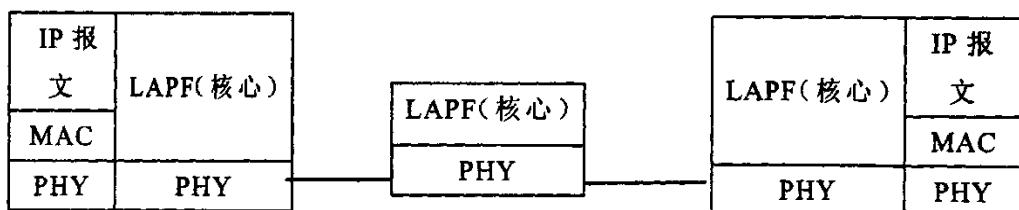


图 2-2 用户数据传输过程协议栈

2.3 以太网接入 ATM 网络

2.3.1 ATM 网络

ATM 是“异步转移模式”的英文缩写。所谓“异步转移模式”^[26]，是一种采用统计时分复用技术“面向分组”的传送模式；在 ATM 中，信息流被组织成固定尺寸的块（称为“信元”）进行传送，信元长度为 53 字节；信元的传送是“面向连接”的，只有在已经建立好的虚连接（“虚电路”）上才能接收和发送信元。ATM 网络特点是在通过 QoS 级的支持对多种业务进行无缝集成，并通过统计复用来保证资源的合理分配。在交换技术上使用利于硬件交换的定长信元交换，有利于实现高速交换。同时在管理上采用带外信令，实行动态带宽管理。

ATM 网络因其对综合业务传输具有较好的服务质量保证的特性，使得其在广域网领域占有了垄断地位。但其和局域网的融合过程中，进展就不如其在广域网中的发展速度。现有的以太网接入 ATM 网络技术和以太网接入帧中继网络相似，也是通过一个多协议路由器接入^[27]，ATM 的虚连接是从多协议路由器开始发起的，中止于 ATM 网络的另一端的多协议路由器，在此不作过多的介绍。

为了使得传统的以太网和令牌环网络同 ATM 网络有更好的互操作性，使得以太网用户能够享受到 ATM 网络的优异的传输性能，而隐藏 ATM 网络复杂的管理和控制机制，产生了基于 ATM 的局域网仿真 LANE 技术^[21,29]。

2.3.2 ATM 局域网仿真

ATM 局域网仿（LANE）提供了在 ATM 网络上运行现有的 LAN 协议（IEEE802.3）的手段，它定义了怎样使用 ATM 网络作为骨干网来互连现有的 LAN，以及现有的以太网上的主机怎样在对等的环境下与具有 ATM 接口的主机以及符合 ATM UNI 标准的服务器、路由器、桥接器等各种 ATM 终端设备组成仿真 LAN 进行高速通信。其主要的目标有：

- (1) 将 ATM 的众多优点 (如低延时抖动、高带宽等) 引入 LAN 中。
- (2) 允许 LAN 终端用户接入 ATM 骨干网。
- (3) 在 ATM 上进行 LAN 桥接。

ATM 技术与传统局域网在某些重要的特征时相冲突的, 例如传统的以太网在链路层采用变长的 MAC 帧, 帧长为 (46—1500 字节)^[35], 而 ATM 采用 53 字节的定长信元。ATM 采用面向连接的传输模式, 数据传输之前需要建立虚连接, 并且采用单独的信令和编址方式。而以太网采用无连接的方式, 网内采用 MAC 地址来标识各个用户, 所以在 LANE 中必须要经过复杂的地址转换才能够实现 MAC 帧在 ATM 网络上透明传输。

2.3.2.1 局域网仿真的配置结构

LANE 服务的体系结构基于客户机/服务器的查询和响应模型。其中使用下列客户机和服务器:

——局域网仿真客户机 (LEC): 是所仿真的局域网的端系统, 它向现有的局域网 (例如以太网) 提供 MAC 层的服务接口。它代理有 LAN 的所有终端与 ATM 网络传输数据, 完成地址解析 (将 MAC 映射成 ATM 地址) 实现整个 ELAN 中所有终端的通信。

——局域网仿真配置服务器 (LECS): 负责保存仿真局域网中的 LEC 的配置信息。

——局域网仿真服务器 (LES): 提供 MAC 地址与 ATM 地址的映射功能, 每个 ELAN 中只有一个逻辑 LES。

——局域网仿真广播和未知服务器 (BUS): 负责处理广播 MAC 地址, 从而在仿真的局域网内提供广播和多传送通信能力。

LAN 仿真并没有规定上述服务器的位置, 任何具有和 ATM 连通性的一个和多个设备都行, 然而, 为了保证可靠性和高性能, 厂商一般都在网络设备上实现这些服务器。

2.3.2.2 局域网仿真的协议结构

LANE 仿真局域网的媒体访问控制子层 (MAC:Media Access

Control), 用户可以利用该功能而不必对逻辑链路层以上的应用作任何改变, 因此 LANE 可以支持任何高层软件。

LANE 的协议栈如图所示, 以 ATM 代替 MAC 子层, 成为 ISO 网络七层模型中物理层与数据链路层中逻辑链路控制子层 (LLC) 之间的一层, 由 ATM 提供传统的 LAN 数据链路层的 MAC 子层的服务。

来自现有的局域网的 MAC 帧在 ATM 桥接上终接, 利用 LANE 功能把 MAC 帧按图所示进行封装, 去掉帧校验序列并加上一个标志头, 然后 AAL5 协议进行处理, 在 ATM 网络上传输。对现有的局域网上的用户来说, 整个系统 (ATM 呼叫建立、信元拆装等) 是透明的, 其运行环境和原有的局域网相同, 但是网络的速率和带宽却大大提高了。

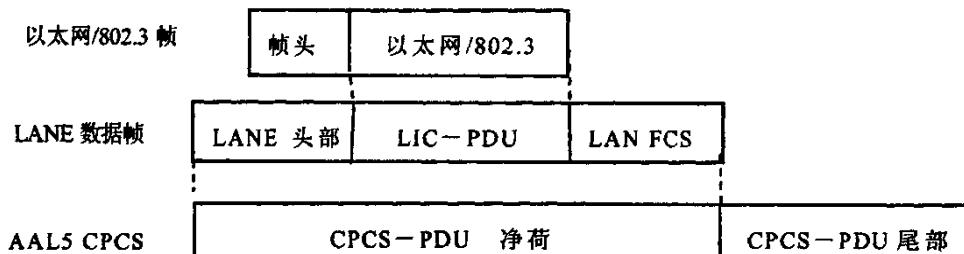


图 2-3 LANE 的帧封装过程

2.3.2.3 局域网仿真工作过程概述

当仿真的局域网中安装新的 LEC 时, 该 LEC 将访问 LECS 以获得 LES 的 ATM 地址, 然后 LEC 建立与 LES 的 ATM 连接, LES 将其 MAC 地址与 ATM 地址映射表中对新安装的 LEC 的 MAC 地址和 ATM 地址进行登记。接下来, LEC 向 LES 发送局域网仿真地址解析协议 (LEARP: LAN Emulation Address Resolution Protocol) 请求, 其中包括广播 MAC 地址, 从而获得 BUS 的 ATM 地址。

当主机要发送数据报时, 必须首先进行地址解析, 以获得被叫主机的 ATM 地址。下面以将要发送 IP 数据报为例, 说明 LANE 地址解析的步骤。如图所示, 当所仿真的局域网中的主机要发送 IP 数据包时, 它首先使用在 LEC 安装阶段所获得的 BUS 的 ATM 地址建立与 BUS 的 ATM 连接, 然后利用该连接向 BUS 发送 IP ARP, 其中包括被叫主机的 IP 地址, 接着 BUS 向仿真的局域网内所有的端系统广播该 IP

ARP。被叫主机接收到 IP ARP 后, 将自己的 MAC 地址进行回应。但主机真正需要的是 ATM 地址, 而不是 MAC 地址, 当主机侧接收到被叫主机的 MAC 地址后, 它将向 LES 发送 LE ARP 请求, 其中包含被叫主机的 MAC 地址。LES 将进行被叫主机的 MAC 地址和 ATM 地址的映射并通知主叫主机。主叫主机将使用该 ATM 地址建立与被叫主机 ATM 的 SVC 连接, 并在其上发送 IP 数据报。

这种两步地址解析方法会影响 IP 数据报的延时性能并会增加额外的广播业务量。同时每台主机必须侦听每个广播数据报, 这样会浪费每个主机的 CPU 周期, 并且会消耗网络带宽。

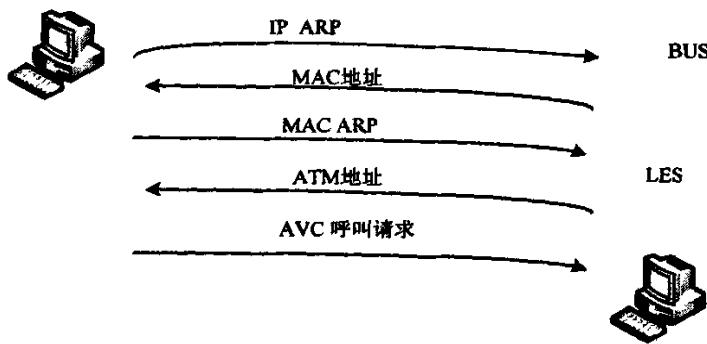


图 2-4 LANE 的地址解析方法

2.3.2.4 LANE 的优缺点分析

LANE 实际上是对 ATM 层屏蔽了网络层及以上层的协议, 从而能够在 ATM 上支持多协议。这是它的优点, 同时也是它的缺点, 除了带来额外的协议开销外, 由于没有相应的机制把网络层的 QoS 请求映射到 ATM 层, LANE 用户不能充分利用 ATM 的 QoS 保障机制。另外, LANE 要求所仿真的局域网内所有设备都必须使用同样的最大帧长度, 例如, 以太网的 MTU 为 1500 字节, 而 ATM 主机能使用比这大得多的 MTU (MTU 比较大意味着较大的吞吐率), 但是如果以太网与 ATM 主机位于同一个仿真 LAN 中, 则 ATM 主机只能使用较小的以太网 MTU。

使用 LANE 时, 网络规模也会受到限制, 当前使用的 LANE 1.0 版本, 不允许多个虚 LAN 产生的业务共享一条 ATM 虚电路, 这样, 当虚拟 LAN 的数量增加后, 虚电路的数量也相应增加了。由于每条虚

电路都要求一定数量的建立开销，并且 ATM 交换机支持的虚电路数量也有一定限制，所以网络的规模也会受到限制。

2.4 本章小结

从上面的介绍中知道，以太网接入到帧中继网络的过程中，以太网用户本身和帧中继网络是没有联系的，完全是通过多协议路由器来实现数据的转发，通过判断 IP 报文的目的地址来确定对应的虚连接。这样，一对异地以太网用户之间的不同业务在帧中继网络上是完全同等对待（将 IP 报文映射到同一条虚连接上），这对于服务质量需求较高的业务是及其不利的。同时，由于以太网用户没有机制来提出业务本身的服务质量要求，网桥也就不能根据用户业务的服务质量需求来建立一条“适当”的虚连接。

LANE 将异地的局域网通过 ATM 网络来组织成一个更大的“局域网”，实现现有的以太网和 ATM 网络的融合，但由于其烦杂的地址解析机制，并没有得到过多的青睐。

本论文要讨论的主题就是在以太网在接入 SUPANET 时，如何根据用户业务本身的服务质量^[31]需求来建立虚连接，实现基于业务的而并非基于用户的细粒度的、有服务质量保证的传输服务。

第 3 章 以太网接入 SUPANET 技术

3.1 以太网接入 SUPANET 思想的产生

3.1.1 优化接入过程

从第二章的分析可知，广域网技术和局域网技术在两个不同的方向向前发展。局域网的应用环境也从当初的文本传输、共享打印机等简单的应用为主向音频、视频等流媒体传输为主的应用。另一方面，在局域网技术发展过程中高速率的局域网技术标准不断出现，如 100Mbps 的快速以太网和 1Gbps 的超高速以太网，以及新近出现的 10Gbps 的甚高速以太网等，这些标准的确立为局域网中流媒体传输打下了坚实的技术基础。同时，广域网传输技术得到了很大的发展，在光纤通信技术的支撑下，带宽已经不再成为制约广域网传输速率的“瓶颈”；另外，随着交换机和骨干路由器不断采用硬件技术，转发速率已经达到线速要求。可以预见，在未来更高速率的交换技术和更高带宽的链路会不断涌现，广域网段不会对业务传输的过程产生不可接受的延时和抖动。那么，在局域网和广域网中各种问题的解决是否意味着最终问题—用户到用户之间的业务传输的服务质量得到保证了呢？很明显，答案是否定的，在局域网内传输的业务，服务质量始终是令人满意的；随着 ATM 技术在广域网内普遍应用，在 ATM 网络提供的网络边缘到边缘各种业务的服务质量得到了业界的认可，然而一种技术的成功不意味着全部问题的解决，ATM 在广域网内实现有较好服务质量保证的综合业务传输，但用户的业务必须从局域网经广域网传输到达目的端，在局域网和广域网的结合处，成为整个业务传输服务质量的瓶颈。或许有人会说，现行的 ATM 主机可以直接接入 ATM 网络，在 ATM 用户端到端之间完全利用 ATM 网络的优异性能保证业务传输的服务质量。这些人可能忽视了一个基本的事实：以太网现在以及在可以预见的未来将会一直占据着巨大的市场份额，ATM 也不可能大规模延伸到用户桌面。所以，ATM 不可能成为最终的解决方案。

问题的关键在于局域网和广域网在存在太多的差异，各种广域网

技术在提出的时候将更多的技术指标放在了更高的交换和传输速率上，而缺少对其与现有的局域网的衔接和融合的思考，以至于在各种技术标准成熟之后，发现问题仍然没有得到解决。

面对现有网络对业务服务质量保障存在的问题和未来“三网合一”的需求，四川省网络通信技术重点实验室提出了“单物理层用户数据传输平面网络体系结构”(SUPANET, Single User-data transfer Platform Architecture Network)。首先，SUPANET 设计了单物理层的用户数据传输平台，在协议层次上保证了用户数据传输的高效性；其次，在交换结构上采用了全新的交换技术 HFS(半步超前交换技术)，保证了交换的高速率。更为重要的是 SUPANET 在设计之初就判定是以太网将会成为主要的用户应用环境，SUPANET 的数据帧设计成以太网 MAC 帧的“天然载体”，当 SUPANET 技术标准成熟之后，使得现有的以太网用户能够在不做大的软件改动就能高效接入到 SUPANET 中，真正满足用户的端到端的业务的服务质量保障需求。本文所研究的以太网接入 SUPANET 技术是整个 SUPANET 体系结构的重要组成部分。

3.1.2 以太网接入 SUPANET 的含义

本论文中，以太网接入 SUPANET 的主要含义是指在 SUPANET 平台下，充分利用 SUPANET 面向以太网帧的特性，使得现有的广泛应用于以太网内用户能够在做尽量少的改动的情况下，通过 SUPANET 实现用户到用户的业务服务质量保证。其不同之处在于是从以太网用户源端来的 MAC 帧在 SUPANET 边缘交换节点上根据 MAC 帧本身携带的相关信息来确定此数据包所对应的虚连接，从而直接进行 EPF 帧的封装和 SUPANET 网内的转发。不需要在 SUPANET 边缘交换节点上做更多的解封装和封装操作。接入过程涉及到用一套用户到接入交换机之间的信令，实现以太网用户接入到 SUPANET 中，保障用户间端到端的业务传输的服务质量。

3.2 以太网接入 SUPANET 技术

在 SUPANET 设计之初，其应用环境就是建立在以太网环境之上

的，它同时支持两种不同的数据传输模式：传统 Internet 模式和 SUPA 模式，传统 Internet 模式即为 IP 报文在无连接传输模式下通过路由器逐跳路由，最终到达目的地。而 SUPA 模式就是通过 SUPANET 专用协议 (QoSNP)^[43] 在数据传输源端和目的端之间建立一条虚连接，用 EPF 帧来承载用户数据，在 SUPANET 交换机内部通过 VLI (Virtual Line Index) 值的替换实现高速交换。SUPANET 的以太网接入是在 SUPA 模式下实现传统的以太网用户接入到 SUPANET 的一种技术。

3.2.1 问题的提出

在前面的 3.1.1 节的分析中知道，对于用户到用户之间业务传输的服务质量的保证问题，其结症在于局域网和广域网的结合处，不仅需要复杂的协议转换，还需要对用户数据进行不同层次的封装和解封装操作，这些操作占用了接入交换机大量的处理能力，在对已经接入的业务造成严重时延和抖动的同时，还影响了更多业务的接入请求。所有造成这种情形的最根本原因在于不能在用户到用户之间建立一条“纯粹的”交换链路。如图 3-1 所示：

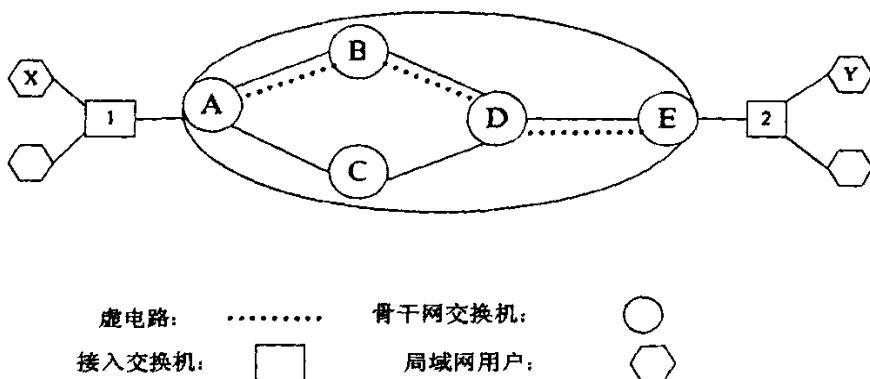


图 3-1 以太网经面向连接的广域网互连

在图 3-1 中，数据流在骨干网中采用虚电路方式交换，在用户数据开始传输前建立的虚链路上通过标签替换和转接。标签根据采用的骨干网具体组网网络技术不同而有所差异，如 ATM 的 VPI 和 VCI、帧中继的 DLCI 或 MPLS 中的标签等。下面以 ATM 作为骨干网为例，简要叙述一下局域网用户 X 和异地局域网用户 Y 之间通过

ATM 网络的数据传输过程,如图 3-2 所示。

- (1) 用户 X 将第一个数据帧 (MAC 帧) 传送到接入交换机 1, 交换机 1 解封装 MAC 帧, 取得 IP 报文。
- (2) 接入交换机 1 根据 IP 报文的 IP 目的地址 (如果要更细粒度的区分业务, 还需要将 IP 报文继续解封装到 TCP 包或 UDP 包以获得原、目的端口号), 启动虚连接建立信令。
- (3) 虚连接建立完成之后, 数据开始传输。用户 IP 数据报文经过 ATM 的 AAL 层适配 (假设采用 AAL5), 将 IP 报文封装到 AAL5 帧中。
- (4) ATM 层将上层交付的 AAL5 帧切割成 53 字节的信元, 并根据连接建立过程相关信息填充信元头部。对最后一个信元可能还需要进行相关填充。
- (5) 物理层将信元转化成电信号或光信号, 传输到下一交换节点。
- (6) 下一节点将接收到的信元根据其中输入 VCI 和 VPI 值, 查找交换表得到其输出 VCI 和 VPI 值, 替换相关字段, 从输出端口输出。
- (7) 在中间交换节点不断重复 (6) 的替换操作, 当信元到达最后一个交换节点 (接入交换机 2) 之后, 缓存信元, 直到收到此 AAL5 帧的最后一个信元 (由结束标志位标识)。
- (8) 重新将缓存的信元组装成 AAL5 帧, 并恢复出 IP 报文。
- (9) 根据 IP 报文的目的地址, 查找路由表, 并根据链路层协议 (假设为以太网链路), 重新将 IP 包封装成 MAC 帧。
- (10) 将 MAC 帧最终发送到目的用户 Y 中。

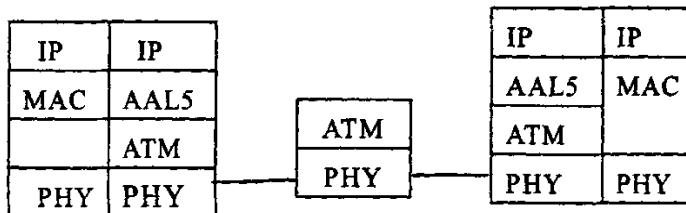


图 3-2 以太网帧经过 ATM 网络的协议栈

除虚连接建立过程之外, 对于每一个 IP 包, 都需要经历上述所有的过程。即使是在连接建立之后, 对于接入交换机, 其对每一个在以太网用户侧所接收的 MAC 都需要向上解封装至 IP 报文甚至到 TCP 或

UDP 包，然后根据相关信息查找交换表取得交换信息，再进行适配、封装以及切割操作，最后发送信元。当信元到达网络出口端接入交换机后需要对信元进行重新组装、适配等相关的逆向操作，以获取 IP 报文，最后再重新进行链路层封装，发送到目的端用户。这些操作使得只能在骨干网接入交换机之间进行建立一条“端到端”的交换链路，而不能在用户 X 和 Y 之间建立一条“用户到用户”的交换链路。

之所以会造成上述的费时操作，原因有如下点：第一，ATM 网络和 IP 网络不管是在网络体协结构还是数据传输模式上均存在巨大的差异；例如 IP 网络采用变长 IP 报文逐跳路由，而 ATM 网络采用定长信元通过虚连接交换，这使得用户数据承载协议的转换和数据的适配操作必不可少。第二，传输数据的协议层次多次“变迁”使得过多费时的封装和解封装操作存在。在上例中，用户的 MAC 帧处于数据链路层，而 MAC 帧头信息不足以区分不同的传输业务流（即不能根据 MAC 帧的信息来确定数据所对应的虚连接），从而需要将其解封装到 IP 层或传输层，而其后又需要将其封装成 IP 报文，再继续封装、适配、切割成信元。

3.2.2 虚拟 MAC 地址

建立一条“用户到用户”的全交换链路，尽量减少用户数据在协议层次上的“变迁”，省去接入交换机上的一些费时操作，应该是解决接入交换机“瓶颈”的一种合理的思维方式。在全通路交换链路建立之后，接入交换机接收到用户端传来的 MAC 帧，根据其中的相关地址信息，直接将其封装到 SUPANET 的 EPF 帧中，而不需继续解封装。然而，在骨干网中，是采用标签来作为链路的局部标识符的，不同的标识符标识不同的应用业务流，即一系列的标签的串接标识了一个应用业务流的一条传输路径；而在用户到接入交换机之间，是采用目的 MAC 地址来作为此段链路的标识符的，又因为网络接口卡的 MAC 地址是固定的，所以在用户和接入交换机之间标识符是固定的，故不可能通过 MAC 地址来区分出不同的应用业务流。

为了解决上述矛盾，在论文的接入技术中，我们提出了“虚拟 MAC 地址”的概念，其主要思想是在以太网用户端虚拟出多个用户，这种虚拟只是建立在逻辑链路层上，其实现方式是给每个用户业务请求分

配不同的虚拟 MAC 地址，每个虚拟 MAC 标识一个业务流。这样从同一个物理以太网主机上来的多个不同业务流对于 SUPANET 接入端交换机来说，类似于来自多个不同的以太网用户。如下图所示：

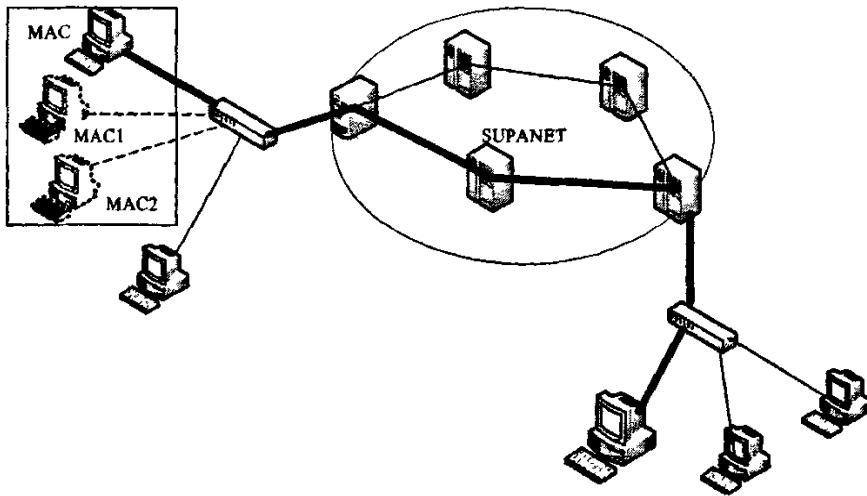


图 3-3 由“虚拟 MAC”地址虚拟出的用户

在图 3-3 中，一个网络接口卡地址为 MAC 的以太网用户同时有两个应用层业务通过 SUPANET 与远端以太网用户通信，SUPANET 的接入交换机分别给两个业务流分配的“虚拟 MAC”地址为 MAC1 和 MAC2；两个业务流在链路层生成的 MAC 帧的源 MAC 地址分别为 MAC1 和 MAC2，SUPANET 的边缘交换机则认为是来自不同的以太网用户，通过这种方式为每一个应用业务流虚拟出一个用户，使得“虚拟 MAC”地址成为业务的标识符。

由于在 SUPANET 网络内部采用 VLI (Virtual Line Index) 值进行替换转发，所以在入口接入交换机需要建立一个“虚拟 MAC—VLI”映射表；而在出口接入交换机需要建立一个“VLI—目的 MAC”映射表，其中的目的 MAC 地址为用户的 MAC 地址。这样当用户某个业务数据在链路层形成 MAC 帧时，将“源 MAC 地址”字段填充为入口接入交换机给此业务分配的虚拟 MAC 地址。当入口交换机收到 MAC 帧后，只需要提取源 MAC 地址，查找“虚拟 MAC—VLI”映射表，得到相关 VLI 交换信息，生成 EPF 帧头，然后直接封装 MAC 帧，生成 EPF 帧，在 SUPANET 网内传输；而当 EPF 帧到达出口交换机时，根据 EPF 帧头部的 VLI 值来检索“VLI—目的 MAC”映射表，恢复出 MAC 帧，将其目的 MAC 地址字段替换为检索到的目的用户 MAC 地

址, 然后直接将 MAC 帧发送到目的用户。这样就极大地简化了入口交换机和出口交换机的协议处理层次和操作流程, 类似建立一条用户到用户的全交换通路, 有利于保证用户到用户之间业务传输的服务质量。

“虚拟 MAC 地址”存在两个较大的优势: 首先, 它能在较低的协议层次上(链路层)行使“标签”功能, 接入交换机将其作为“虚拟 MAC-VLI”映射表的索引, 实现低层次交换; 其次, 它能唯一标识一个高层的用户业务流(因为它本身就是在业务请求时由接入交换机分配的), 故能实现细粒度的交换, 使得一对异地以太网主机上不同服务质量需求的业务能够得到服务质量保证。

需要指出的是, 上述思想主要是基于 SUPANET 传输平台的两大特性而提出的, 第一, SUPANET 采用带外信令技术, 所有的控制信息和管理信息都是通过传统的 Internet 平台通过 UDP 协议来传输的, 而并不是像 ATM 网络那样通过专用的复杂信令来传输的, 这就为接入技术信令在普通的以太网主机上实现提供了可能性; 第二, SUPANET 面向以太网的传输特性, 所承载的数据类型为 MAC 帧, EPF 帧格式和 MAC 帧报文格式接近, 容易转换(只需要一次封装操作或者一次解封装操作), 这是实现全通路交换的有利条件, 这也是以太网接入 SUPANET 技术的最初出发点。所以, 本论文中的接入技术是针对 SUPANET 这个特定的平台的, 而并不是面向所有的骨干网传输技术的。

上述接入技术需要一整套用户和 SUPANET 接入交换机之间的 UNI 信令来支撑。信令的主要功能是完成以太网用户从发起连接请求到数据传输完成, 释放连接整个过程中相关消息的传递和应答, 建立和管理用户到用户间的连接。具体的 UNI 信令将在论文第四章进行详细的设计。

3.3 接入技术的数据传输过程

本节以一个具体的实例来描述采用接入技术的数据传输过程。在第一章关于 SUPANET 的介绍中知道, SUPANET 是采用面向连接的虚电路交换模式来传输数据的, 一次数据传输包括三个过程: 连接建立过程、数据传输过程以及连接拆除过程。

3.3.1 连接建立过程

在以 SUPA 模式传输数据之前，需要根据用户业务的传输服务质量需求建立一条虚连接。由于 SUPA 体系结构的交换模块采用半步超前交换技术，交换表中包括本节点的输出端口号、波长号以及 VLI 值和下一交换节点的输出端口号、输出波长号和 VLI 值，如表 3-1 所示。源端用户利用 UNI 信令将用户请求发送到接入交换机，接入交换机启动 QoSNP 协议，根据传统的路由协议将请求信息发往下一交换节点并协商，如果此节点能够满足所要求的服务质量需求，则建立相应的交换表项，预留相关资源，并将业务请求信息继续传向下一节点直到目的端接入交换机，如果中途任一节点不能满足服务质量需求，则向上一节点发送否认信息，直至源端接入交换机，则此次协商失败。当目的端接入交换机收到请求后，沿途各交换机上关于此业务的交换表项均建立完成，然后将请求信息发送到目的端用户。如果目的端用户接受此次请求，则发送一个肯定确认信息，当这个确认信息回送到目的端接入交换机后，目的端接入交换机依据协商时生成的交换表项和目的端用户的 MAC 地址，建立“VLI-MAC”映射表，其中的 VLI 值是目的端接入交换机关于此业务的输入 VLI 值，而 MAC 值是目的端的物理 MAC 地址。

目的端接入交换机继续将确认信息传送到源端接入交换机。此时源端交换机分配一个 48 位的虚拟 MAC 地址，并建立一个“虚拟 MAC-VLI”的映射表；其中的“虚拟 MAC”即是源端接入交换机为此次业务所分配的 MAC 地址，而 VLI 即为在连接建立时所分配的 VLI 值。源端接入交换机继续将确认信息和分配的虚拟 MAC 值前传到源端用户。源端用户收到确认信息后，发送一个反向的确认信息，直至目的端用户，至此，用户到用户的连接建立完成。

表 3-1 交换表表项格式

本跳输出			下一条输出		
端口号	波长号	输出 VLI	端口号	波长号	输出 VLI
OutPort	OutλN	VLIout	NHOP	NOλN	NHVLI

当连接建立完成之后，各节点生成关于此次业务传输的交换表项如下图所示：

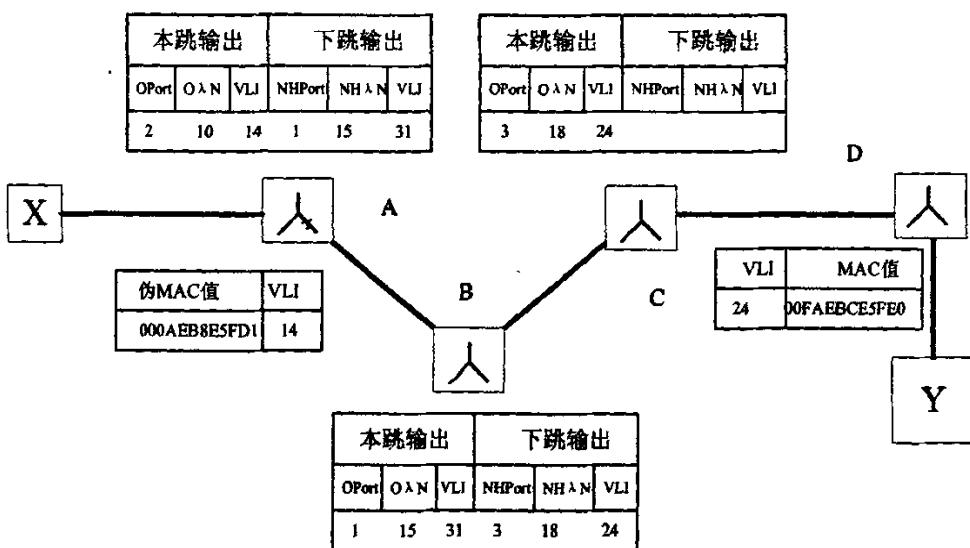


图 3-4 虚连接建立完成

3.3.2 数据传输过程

在连接建立之后，源端和目的端可以开始数据传输过程。先假设 A 端应用层数据流经传送层和网络层封装后形成一个 IP 数据报文，IP 报文继续下传至链路层形成以太网帧。填充 MAC 帧的源地址并不是 X 的网络接口卡的物理地址，而是 SUPANET 的接入交换机在连接建立过程中为此业务分配的“虚拟 MAC 地址”。当 MAC 帧形成之后，将发送到接入交换机。接入交换机收到此 MAC 帧后，提取其源 MAC 地址，并依此地址来查找在连接建立时所创建的“虚拟 MAC—VLI”映射表；当在此映射表中获得该交换机节点中输出端口号 Oport(2)、输出波长号 O λ N(10) 以及输出 VLI(14)。当接入交换机获得这些信息后，及可启动交换模块进行交换；由于采用半步超前技术^[25]，所以在交换的同时，查表部件依靠本节点的输出端口、波长以及 VLI 值，查找交换表，获取下一节点的输出端口号、波长以及 VLI 值。当查表操作和交换操作均完成之后，将下一条输出信息填充到 EPF 帧的相关字段，并从输出端口将帧发送到下一跳节点。当 EPF 帧到达交换节点 B 时，交换机能直接提取 EPF 中的本节点输出信息，启动交换部件和查表部件，实施半步超前交换过程；查表部件和交换部件的并行操作，这是半步超前交换的主要优势体现。

[25]。

当EPF帧传递到核心交换机C时，交换机C提取EPF帧头的本跳输出信息进行交换；但由于其下一跳节点D就是SUPANET的边界节点，其输出信息对交换机C来说不再重要，故在C节点进行交换完成后不再需要填写EPF帧中的下一跳输出信息字段，EPF帧直接从C节点的输出端口输出到下一交换节点D。

边界节点D接收到EPF帧后，由于其为边界节点，其操作与核心节点的处理过程有所不同；D先提取EPF帧中的本跳输出端口号、波长号以及VLI值，来查找在连接建立时所创建的“VLI-MAC”映射表，查得本EPF帧中所承载的MAC帧的目的MAC地址（00FAEBCE5FE0）。然后将EPF帧解封装，提取MAC帧，将刚查询所得到的MAC值填入MAC帧的目的MAC字段。然后由对应的以太网端口发送到目的端主机。

上面描述了一个源端MAC帧经SUPANET传送到目的端的完整过程。目的端所接收到的MAC的源端MAC地址是业务业务的虚拟MAC地址，目的MAC地址为真实的目的端MAC地址，虽然传输过程经过了广域网（SUPANET），但对于目的端来说，其类似于一个局域网内的传输。

3.3.3 连接释放过程

当数据传输完成之后，源端主机会发送一个传送完成的信令消息到接入交换机，接入交换机启动 CEP 协议，逐渐释放各交换机上为此连接预留的相关资源，删除边缘交换节点的交换表项和映射表项以及中间核心节点的交换表项。一次数据传输过程完整完成。

3.4 接入技术所需要考虑的问题

论文论述的接入思想致力于通过 SUPANET 平台创建一条用户到用户的全交换数据通路，但用户所在以太网和高速传输网络 SUPANET 还是在相关特性上有较多的不同，使得一些问题不可避免。

3.4.1 帧长与填充问题

本文的接入思想是基于 SUPANET 所提出的，在第一章关于 SUPANET 的叙述中知道，SUPANET 一个重要的特性就是采用定长 EPF 帧传输，帧长度为 1546 字节，数据负载长度为 1530 字节，依据以太网 MAC 帧的最大长度而确定的。定长帧利于采用硬件来进行高速交换，提升核心交换机的转发速度。SUPANET 在面向以太网的同时，有一点是与现行的以太网特性相背的，即以太网的 MAC 帧长是可变的，其长度在 46 字节—1500 字节之间。普通用户接入到 SUPANET 中，由于 SUPANET 的 EPF 帧的负载字段的长度定为 MAC 帧的最大长度，所以避免了在接入交换机上进行耗时的切割操作，但有人或许疑问，还是需要进行 EPF 帧的填充操作，因为不可能所有的 MAC 帧都是最大长度的。

用户数据通过网络来传输时，必须在协议栈上进行多层封装，如图 3-5 所示：

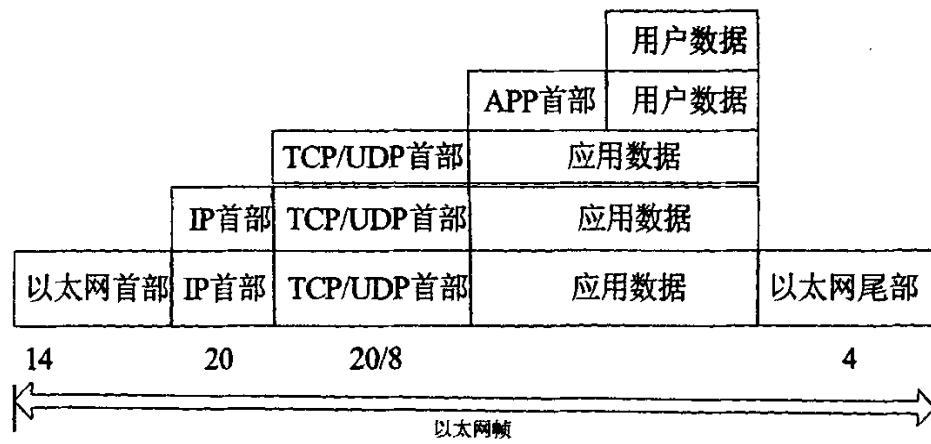


图 3-5 数据在协议栈时的封装过程

以太网的数据帧的负载最大长度 1500 字节是由以太网的链路特性所决定的，其称为链路层 MTU^[36] (Max Transmisstion Unit)，但这并不意味着所有的以太网帧的负载长度都是 1500 字节的。在 IP 层收到一份要发送的 IP 数据报时，它要判断向本地的哪个接口发送数据，并查询该接口获得 MTU，IP 层把接口 MTU 与数据包的长度进行比较，如果需要则进行分片。由上图得知，帧的最终长度还是取决于其上层

递交的 IP 包的长度。

的确如此, SUPANET 的应用背景使得上述问题的影响变得不重要: SUPANET 主要是针对需要一定服务质量保证的大数据量业务传输而设计的, 例如多媒体业务传输。根据上面的分析, 对于需要 SUPANET 传输的业务, 在 MAC 层形成的帧人负载长度绝大部分都是 1500 字节的, 在接入交换机上并不需要进行填充操作, 对于一些需要填充到 1500 字节的 MAC, 是在以太网主机的网卡实现好还是在接入交换机上实现比较好呢?

在用户主机上实现:

需要对网卡作出改动, 即需要在网卡计算 MAC 长度时, 如其长度不是 1500 字节, 则填充至 1500 字节。优势在于接入交换机不涉及到任何填充和切割操作。但其需要对以太网卡作出改动, 因为传统的以太网接口卡对大于 46 字节而小于 1500 字节的 MAC 帧并没有做任何填充操作 (对不够最小帧长的 MAC 帧有填充)。

在接入交换机上实现, 对用户主机不需要作出任何改动, 用户更容易接受, 但会消耗接入交换机的处理能力。由上面的分析知道, 需要填充的帧很少, 这种消耗是微不足道的。故在接入交换机上实现填充比较合适。

3.4.2 虚拟 MAC 地址的分配原则

“虚拟 MAC 地址”由接入交换机分配给连接发起端用户, 用来标识一个用户的业务流。分配的原则主要需要考虑如下问题: 只需要其不与已经存在的真实的 MAC 地址重合就能满足要求。当在此次数据传输过程结束之后, 接入交换机要能够“回收”此虚拟 MAC 地址, 以便为后面的业务请求分配。

MAC 地址也叫物理地址、硬件地址或链路地址, 由网卡设备制造商生产时写在硬件内部。IP 地址与 MAC 地址在计算机里都是以二进制表示的, IP 地址是 32 位的, 而 MAC 地址则是 48 位的。MAC 地址的长度为 48 位 (6 个字节), 地址的最高位为 0 时是普通地址, 为 1 时是组地址, 组地址允许多个站点使用同一个地址。当把一帧送到组地址时, 组内所有站点都会收到该帧。MAC 地址的次高位用来区分全局地址和局部地址。局部地址是网络管理员指定的, 在这个局域网之

外就没有任何意义。相反，全局地址是由 IEEE 指定的，以保证世界上没有两个地址相同的站点。由于有 $48-2 = 46$ 位可用地址，因此 MAC 地址大约有 7×10^{43} 个。

在 SUPANET 的边缘交换机给以太网主机分配“虚拟 MAC”地址用来虚拟出一个用户，在前面的介绍中知道，边缘交换机是采用“虚拟 MAC”地址来检索到业务流所对应的虚连接标识符，一旦产生冲突，则可能在边缘交换机错误地转发了并非属于此业务流的 MAC 帧。如何生成此“虚拟 MAC”地址以保证不和局域网内现有的真实 MAC 地址冲突呢？有两种相应的解决方法：

第一种方法是通过 MAC 地址的分发机构 IEEE。当 SUPANET 整个研究工作得到充分认可之后，可以向 IEEE 组织申请一个专用的 MAC 地址块（类似于网络接口卡生产商所申请地址块）作为 SUPANET 专用，这样接入交换机所分配的“虚拟 MAC”地址就能同传统的真实 MAC 地址所区分。由于 SUPANET 仍在研究阶段，这种方式的实施需在后期继续讨论。

第二种方式就是让边缘交换机随机生成一个普通的本地 48 位的 MAC 地址（最高位为 0，次高位为 1）分配给以太网的用户，这是一种最简单实现方式。对于这种情况，产生和局域网内其他用户相冲突的概率为 2^{46} 分之一，故随机分配一个 MAC 地址虽然在理论上存在和现有的 MAC 地址冲突的可能，但概率极小。

在上面的分析中，我们可以将采用第二种即随机分配的方式来生成“虚拟 MAC”地址的一种实用方式。如果要实现理论意义上的无冲突解决方案，还需在后期继续研究。

3.5 本章小结

本章主要介绍了以太网接入 SUPANET 接入思想的产生以及主要原理，并用一个具体的传输过程来描述了接入的数据传输流程；另外对接入技术需要解决的问题作出了相关的分析。

第 4 章 以太网接入 SUPANET 的信令设计

SUPANET 是面向以太网的高速传输网络，以太网接入 SUPANET 技术的目的在于使得以太网用户能够在不作大的改动的情况下高效地接入到 SUPANET，它是 SUPANET 核心网络和以太网用户之间的纽带。接入涉及到连接建立、数据传输和连接释放等过程，这些过程牵涉到源端用户和接入交换机、接入交换机和目的端用户以及源端用户和目的端之间的信息的交互。信息的交互是通过信令来完成的，SUPANET 核心网络中虚连接的建立是通过 QoSNP 协议来实现的，而连接的释放是通过 CEP 协议来实现的。本章将对和接入紧密相关的用户和接入交换机间的 UNI 信令进行具体的设计。

4.1 信令

信令的概念最初起源于电话网，在电话网中，为了在任意两个用户之间建立一条话音通路，相关的电话交换机必须进行相应的话路接续工作，并把接续的结果或进一步要求以信令的方式送至另以相关交换机或用户，在接续过程中，必须要遵循一定的协议或规约，这些协议或规约称之为信令。信令的主要功能是完成连接的建立、保持和释放。

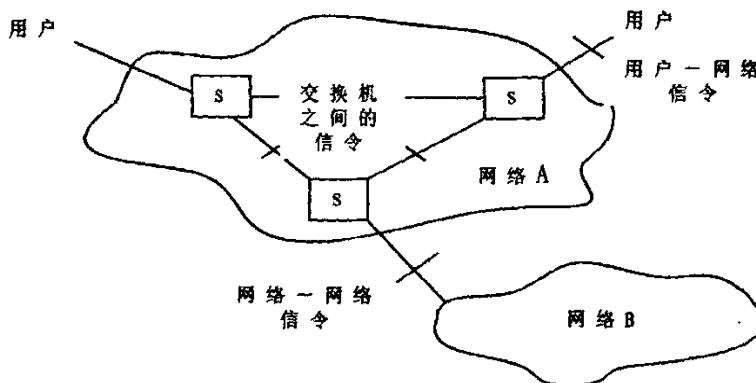


图 4-1 不同类型的信令

信令^[40]按其工作的区域，可分为用户信令和局间信令。用户信令

是用户与网络之间进行交互的信令，它们在用户线上传递，完成用户到网络的接入工作。在现代通信中也称为用户网络接口 UNI (User-to-Network Interface)；局间信令是网络内部交换机之间传输网络控制、管理信息的信令，它在中继线上传输，故称为网络节点接口 NNI (Network-to-Network Interface) 信令。NNI 信令包括线路信令、选择信令和管理信令。线路信令完成占用信号、应答信号以及正反向拆线信号的传送；而选择信令则传送局间的地址码信号；管理信令传送网络的拥塞信号、计费信号和网络维护信号；整个完整 NNI 信令系统由上述三部分组成。

4.2 SUPANET 中的信令

在第一章对 SUPANET 的描述中知道，SUPANET 需要一套完善的信令机制来对整个网络实现控制和管理。SUPANET 为了信令传输的效率和对现有的 Internet 的兼容，采用了以传统的 IP 网络来作为 SUPANET 网络内用于传输控制、管理等信息的承载网络。换句话说，即 SUPANET 网络的信令信息是通过传统的 IP 网络来传输的，而用户数据是使用高速的 EPFTS 交换平台来传输的，从而实现了信令信息传输平台和用户数据传输平台的分离。

SUPANET 内部的用于控制网络的协议有：服务质量协商协议 (QoSNP)、连接释放协议 (CEP)、流量监控协议 (TMEP)；CEP 用于业务流传输结束后各交换节点连接释和资源的回收；TMEP 用于 SUPANET 网络内部交换节点的流量的监测和控制；QoSNP 是基于服务质量约束的连接建立协商协议。TMEP 属于网络内部交换节点之间交互信息的专用协议，本文将不再赘述。而对和接入技术紧密相关的 QoSNP 协议将做较详细的介绍。

4.2.1 QoSNP 综述

QoSNP 工作在 SUPANET 的信控、管理平台上，是一个应用层协议。它利用在传统路由器中已有的传送层协议 UDP 来传输 QoSNP 协议数据单元 (PDU)。在请求服务时，SUPANET 的端系统将启动 QoSNP

协议，并将用户的 QoS 参数写入到 QoSNP-PDU 的相关字段，利用信控、管理平台上的 IP 路径选择功能选择输出端口，在根据服务质量需求在该端口上选择可支持该连接的波长，如果该节点具备支持该服务质量请求的能力，则为该连接预留相关资源，并将协商过程推进到下一节点，直到 SUPA 域的另一个单系统。如果在 IP 路径选择经过的任一节点不具备支持请求的服务质量的能力，则协商失败，相关信息将依次被反馈到前面已协商成功的节点，释放已预留的资源，宣布虚连接建立过程失败^[44]。

4.2.2 QoSNP 中的 QoS 参数

服务质量参数既是用户向网络请求所需的服务性能高低的量度值，也是网络服务提供者向用户宣传的可提供的或承诺的服务性能的量度值。各种不同类型的网络都会根据需要为自己定义相应的 QoS 参数，例如 ATM 定义的 QoS 参数包括：峰值信元速率(PCR)、持续信元速率(SCR)、最小信元速率(MCR)、信元延迟变化极值(CDVT)、信元丢失比率(CLR)、信元传送延迟(CTR)、信元错误比率(CER)等，ATM 的 QoS 机制能够提供较完善的 QoS 保证，但业务划分种类繁多，成本较高，不利于 ATM 网络的发展。

对网络设备设计者来说，其关心的是单节点的如数据吞吐率、传输时延、抖动、数据丢失率等参数的确定性和变化范围；而对于用户来说，其关心的是如端到端时延、可靠性、安全性等服务参数。故对 QoS 参数的确定需要综合的规划。SUPANET 是面向将来的三网合一的需求而设计的高速交换网络，在对传输业务的特性需求和 QoS 参数繁杂度进行权衡后，初步定义了如下参数^[43]：数据吞吐率、端到端时延、端到端抖动以及 EPF 丢失率等 QoS 参数。

4.2.3 QoSNP 的 PDU 格式

在 SUPANET 网络中，所有的专用的应用信令协议（TMEP、ACP 和 QoSNP 等）都是采用 UDP 协议来作为其传送层的承载协议，端口号为 400（今后需经 IETF 批准）。其通用 PDU 格式如下：

表 4-1 SUPANET 专用应用层协议通用 PDU 格式

地址字段前头部(IPv4 为 12 字节, IPv6 为 8 字节)			IP 分组/ 报文头	
源 IP 地址 (IPv4: 4 字节; IPv6: 16 字节)				
目的 IP 地址(IPv4: 4 字节; IPv6: 16 字节)				
可选部分(变长 IP 头部分, 为 4 字节倍数)				
源端口(2 bytes)		目的端口(2 bytes)	UDP 头	
报文长度 (2 bytes)		校验和 (2 bytes)		
协议编号 (1 byte)	PDU 类型 (1 byte)	PDU 参数字节总数 (PDU Para-bytes)	SUPANET 协议 PDU 部分	
虚通路源端 Id		虚通路目的端 Id		
PDU 参数组 1				
...				
...				
PDU 参数组 N.....				

上述 SUPANET 的协议头中的“协议编号”字段用来区分不同的 SUPANET 专用协议, 如“01”代表 QoSNP; “02”代表 TMEP; “03”代表 ACP; “04”代表 CEP 等。“PDU 类型”用来标识不同的消息类型, QoSNP 定义的消息类型有:

表 4-2 QoSNP 在 TSDN 协商模式下定义的消息类型

PDU 二-十 六进制码	记忆码	名称
01	E-I-E-TSDN Request	TSDN 端到端逐级协商请求消息
02	L-TSDN Acknowledge	TSDN 局部肯定性确认信息
03	L-TSDN Neg Acknowledge	TSDN 局部否定性确认信息
04	E-I-E TSDN Dis-Connection	TSDN 端到端逐级断开连接消息
05	E-E QN-Success	端到端协商成功消息
06	E-E QN-Success Acknowledge	端到端协商成功确认信息

4.2.4 QoSNP 的协商流程

QoSNP 是 SUPANET 端系统之间的应用层协议。端系统包括支持

SUPA 体系结构的计算机和 Internet 与 SUPANET 之间的应用层网关, 前者是 SUPA 域中的服务器或客户机; 后者则是 Internet 与 SUPANET 之间的应用层网关, 是 Internet 数据源与最终接受者借用 SUPANET 过渡的“渡口”。协商过程包括路由选择、消息确认、资源预留、VLI 的分配等操作。协商从 SUPANET 的一个端系统开始, 在中间交换节点逐级 (hop-by-hop) 进行, 直到 SUPANET 的两个端系统所有的服务质量能够为提供传输服务的节点满足为止, 其协商流程示意图^[43]如下:

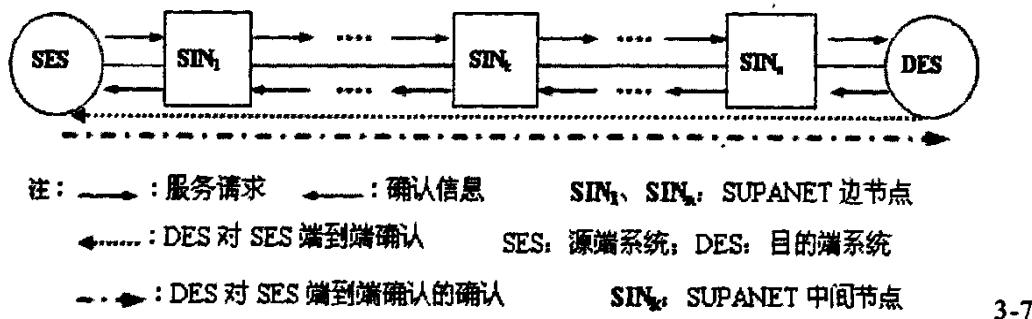


图 4-2 QoSNP 协商流程

图 4-2 中实线右向箭头代表服务请求, 实线左向箭头代表确认信息; 虚线左向箭头代表 DES 对 SES 的端到端确认, 虚线右向箭头代表 SES 对 DES 的端到端确认的确认。SUPANET 的端系统 (SUPANET 主机或 SUPANET 与 Internet 间的网关) 的服务请求信息从源端边节点开始经过各个中间节点, 逐级进行服务质量的协商资源。若中间节点能够满足用户请求, 则协商成功, 相关的资源被预留, 临时交换表 (见 3.3.2 节) 内增加新的数据项。当前节点将此协商请求传向下游节点, 当收到下游节点的确认后, 将确认中相关的信息存入相应的临时交换表内。这一过程一直推向目的方的 SUPANET 的端系统, 如果路径上所有节点内的协商都成功, 则建立起一条服务质量可保证的端到端的虚通路; 否则, 只要路经的节点中有一个不能满足, 该协商失败, 已预留的资源被退还给该节点, 临时交换表内相关的数据项被删除。

而中间节点的协商过程如下:

- ① 根据目的端系统地址和 IP 选径功能, 找到下一节点后, 利用 U-MIB 在相应的端口上选择能够满足服务质量的波长。确定能否满足相应的服务质量请求的判断原则是: 能否在选定端口中找到一个波长, 该波长中指定的优先级别的剩余输出配额足以满足所请求的服务质

量。如果找到，则在该波长的指定优先级别的剩余配额中，减去需要的配额，然后将协商请求前传至下游节点；否则向上游节点发出否定性确认。

② 在选定的端口和波长内，为该虚连接指派虚连接索引字段标识符(VLI)，作为今后在用户数据传输平台上与下游节点之间标识该连接的标识符 VLI 的一部分，也作为下游节点查找交换表的索引。

③ 由于 EPFTS 采用了半步超前交换策略，协商成功后返回上游节点的信息中，包含该通路在本交换节点的输出端口号、输出波长号和（输出）虚连接标识符 VLI。而与服务质量相关的信息如传输配额依旧保存在本节点的相应表如 U-MIB 中，供数据传输的入网控制等操作使用。因此等到下游节点返回肯定性确认后，将返回的确认信息中的下游端口号和波长号与 VLI 一道，存入相应的 EPFTS 交换表中。

④ 如果下游节点发回否定性确认，则将预留的资源退还给系统（分配给该 VLI 的配额），删除交换表和 U-MIB 中的相应数据项，并向上游节点发回否定性确认。此否定性确认过程将沿着请求协商的反方向一直进行反向确认的接力，直至请求协商的 SUPANET 端系统。至此，协商过程失败。

⑤ 为了防止因否定性确认丢失而造成预留资源的浪费，各节点内将设置连接建立等待时钟(T-NW – Negotiation Waiting Timer)。当端到端的协商确认在两个方向上都完成后，在用户数据传输平台即可传输 EPF，一旦收到 EPF，T-NW 比特将被复位为“0”；因此，如果信控平台的 T-NW 时钟超时，而 PFTS 交换表中的 T-NW 比特为“1”，则表明协商不成功，相应的 SUPANET 中间节点将调用 CEP 协议进行协商不成功处理。

4.3 以太网接入技术的 UNI 信令设计

以太网接入技术考虑的接入对象是普通的以太网用户，连接请求是从源端以太网用户发起的，建立过程由以太网用户到 SUPANET 接入交换机、源端接入交换机到目的端的接入交换机、目的端接入交换机到目的以太网用户三个阶段组成。源端以太网用户通过用户网络接口(UNI)信令和接入交换机进行建立连接所需的消息通信；SUPANET 网络内部再通过 QoSNP 完成网络内部的连接建立协商过程；最后，目

的端接入交换机通过 UNI 接口和目的端以太网用户完成信息交互。在这个过程中，牵涉到 UNI 信令和 NNI 信令的转换^[39]和交互。一条“用户到用户”的连接建立需要 UNI 信令和 QoSNP 密切配合。

4.3.1 UNI 信令消息与信令元素

SUPANET 的连接建立过程类似于 PSTN 中的通话电路的建立。在 PSTN 中，用户摘机后听到拨号音，然后拨号直到听到对方振铃，这个过程称为连接建立过程；对方摘机后通话开始，即为数据传输过程；通话完毕后，用户挂机，呼叫连接释放，称为释放过程。呼叫的连接和释放过程是由用户和交换机通过一系列的命令或响应，也称作消息的交互作用来实现的。在此主要讨论点到点的连接控制的消息类型，点到多点的连接控制在此没有涉及。

连接建立消息：

连接请求(SETUP)

连接肯定确认(CONNECT ACKNOWLEDGE)

连接否定确认(CONNECT Neg ACKNOWLEDGE)

反向确认(CONNECT REACKNOWLEDGE)

连接清除消息：

释放(RELEASE)

释放完成(RELEASE COMOLETE)

表 4-3 UNI 信令中的消息类型

报文类型	编号	当主机发送时的意义	当网络发送时的意义
SETUP	01	请求建立连接	有一个连接请求到来
CONNECT ACK	02	接受了入呼叫	网络接受呼叫请求
CONNECT REACK	03	对 CONNECTACK 报文的确认	对 CONNECT ACK 报文的确认
RELEASE	04	请求释放连接	一端发出了连接释放请求
RELEASE COMPLETE	05	对 RELEASE 的确认	对 RELEASE 的确认

不同类型的信令消息包含有一些可选(O)和必选(M)的信息元素(IEs)。其中连接建立消息 SETUP 中包含的信息元素最为复杂，如下：

- 对方用户的 IP 地址

- 端到端延时
- 端到端抖动
- 数据丢失率
- 承诺平均速率
- 突发数据率
- 一次双向协商(可选)
- 安全属性(可选)

RELEASE 消息中包含信息元素为：

- 释放原因

其他消息类型属于一些确认消息，其中可以不包含或仅仅包含简单的信息元素，在此不在赘述。

4.3.2 信令过程

信令过程规范了交互的消息序列、检验参数的一致性的规则以及连接建立和释放的序列行为，差错情况处理是规范的重要成分，包括无效消息、不一致参数以及其他一些意外情况。这些功能非常重要，因为信令协议必须十分可靠地支持用户应用。下面以一个点到点的连接和释放过程示意图来说明信令的交换过程：

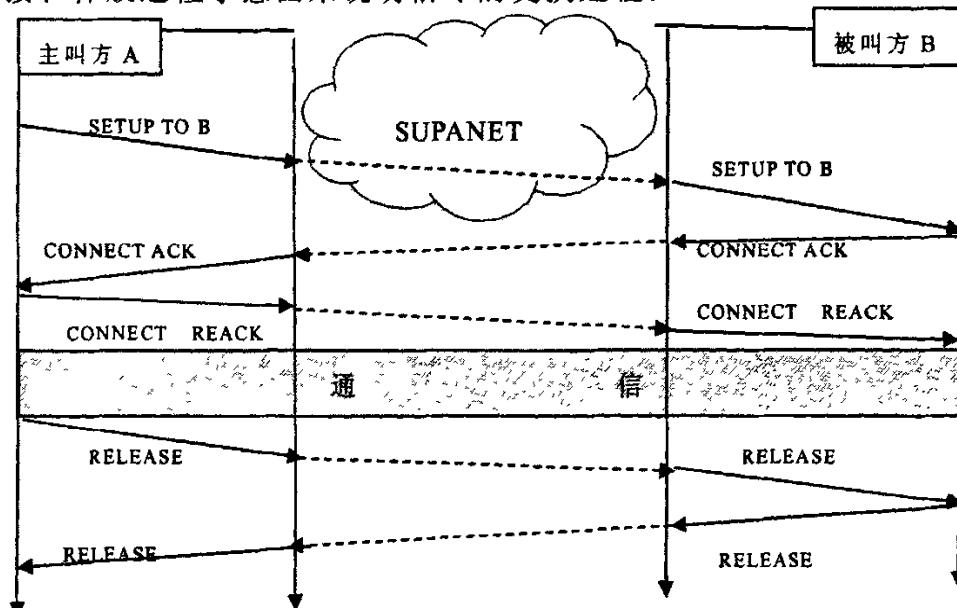


图 4-3UNI 信令连接建立、保持与释放过程

点到点连接建立、保持与释放过程如图 4-3 所示，时间是从上到下运行的。图左为普通一以太网主机用户 A，它通过 UNI 信令接口的 SETUP 消息以希望与另一以太网用户 B 通过 SUPANET 建立一条虚连接，在此处 SUPANET 抽象为网络云，具体的网内操作过程由 QoSNP 协议和 CEP 协议完成。SETUP 消息包含了与其传输业务相关的服务质量需求参数（传输速率、延时、抖动、安全等）、目的地 B 的 IP 地址以及其他需要协商的参数。当 SETUP 消息传送到 SUPANET 的接入交换机后，接入交换机启动 QoSNP 协议，并从 SETUP 消息中提取相关的信息单元，连接接纳控制(CAC)机制能根据用户的合法性检查和 SETUP 消息中的 QoS 参数需求以及当前的网络承载能力判定能否接纳此用户。当接入交换机接受此连接请求后，将相关信息单元填入到 QoSNP 协议的 PDU 中，实现 UNI 信令到 NNI 信令的转换。启动 QoSNP 协议开始网络内部协商后，各个节点将依据能否满足服务质量需求决定是向上一节点发送否认信息还是将协商请求前传。若 SUPANET 网络内部所有节点均能满足需求，协商信息到达 SUPANET 的目的端边缘交换机，边缘交换机将 QoSNP 中的相关信息提取出来，重新启动 UNI 信令，填写相关信息单元，将 NNI 信令转换为 UNI 信令，重新生成 SETUP 消息，并将消息传送到目的用户 B。用户 B 根据 SETUP 消息中的业务描述、QoS 参数等信息单元决定是否接受此次连接建立请求。如果用户 B 不接受此次连接请求，则直接发送一个否定性确认消息 CONNECT Neg ACKNOWLEDGE，消息中包括拒绝的原因等信息元素。网络侧收到此消息后，在目的端接入交换机上启动 CEP 协议，在网络节点上逐级释为此虚连接所预留的资源，并在源端接入交换机重新生成 CONNECT Neg ACKNOWLEDGE 消息，发送到用户 A，告知其连接失败，连接建立过程结束。

如果用户 B 接受该连接请求，将生成 CONNECT ACKNOWLEDGE 消息，发送到目的端接入交换机。接入交换机将其转化为 NNI 信令消息，传送到源端接入交换机，源端接入交换机再次将其转化为 UNI 信令，回送到主叫用户 A。用户 A 接收到此确认消息后，回复一个对用户 B 的 CONNECT ACKNOWLEDGE 消息的确认消息，过程类似于 SETUP 消息的发送过程，中间依然牵涉到 UNI 信令和 NNI 信令的转换。当 B 接收到此确认消息后，整个连接建立过程就全部完成。最后阶段用户 A 的反向确认信息是用来保证连接确实有效的三方握手通信，即采用从网

络到被叫和主叫到网络的消息实现三方握手。

连接建立完成之后，数据传输过程开始，虚连接一直保持到数据传输过程完成。主叫用户和被叫用户都可以启动连接释放过程，正如电话中主叫和被叫都可以首先挂断电话一样。上图中是由主叫来发送 RELEASE 消息来启动连接拆除过程。RELEASE 消息发送到源端接入交换机后，接入交换机启动 CEP 协议，CEP 协议在链路的交换节点上逐跳释放为此虚连接所保留的资源，删除相关的交换表项，当 CEP 协议释放目的端交换机的相关资源和表项后，重新生成 RELEASE 消息，并发送到目的用户 B，用户 B 回送一个 RELEASE COMPLETE 消息，此消息在目的端接入交换机上转换成 CEP 的确认消息，传送到源端接入交换机，源端接入交换机重新转换为 RELEASE COMPLETE 消息，发送到用户 A。到此，连接释放过程完成，一次数据传输的完整过程结束。

4.3.3 UNI 信令消息格式

4.3.2 节中定义的相关消息就是呼叫连接建立和连接拆除时的命令或响应，其一般格式如下：

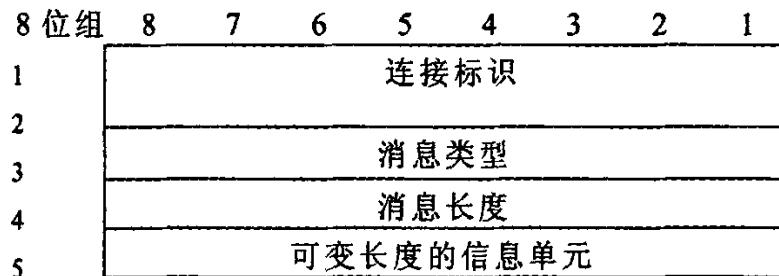


图 4-4 UNI 信令消息的一般格式

连接标识(CONNECTION ID): 标识本地用户网络接口上的呼叫。因为同一个用户可以为不同的传输业务发起多个连接建立请求，呼叫标识用于将不同的消息与不同的连接相对应。

消息类型(MESSAGE TYPE): 消息类型用于区分不同消息，其中的 7-8 位用来区分不同的消息组，00 位呼叫建立请求消息组，01 为呼叫清除消息组。其余编码备用。1-6 位用于区分同一消息组中的不同消息。例如，连接建立消息 SETUP 的编码为 00000001，连接确认消息(CONNECT ACK)的编码为 00000011；连接释放请求 RELEASE 消息的编码为 01000000，

消息长度(MESSAGE LENGTH): 用来指明消息长度后面的可变长度信息单元的八位组组数。

可变长度信息单元(INFORMATION ELEMENT): 前面呼叫标识、消息类型、消息长度是每个消息都含有的固定长度的信息单元；此外，按照各类消息不同的功能，还含有数量不等、长度可变的信息单元，每个 IE 包含 IE 标识、IE 内容长度和 IE 内容等信息，其一般格式如图 4-5 所示。

IE 一般格式中的 IE 标识符类似消息标识，是不同 IE 类型的编码，IE 内容长度指示为此信息单元内容的长度，单位为八位位组；其后为 IE 内容。

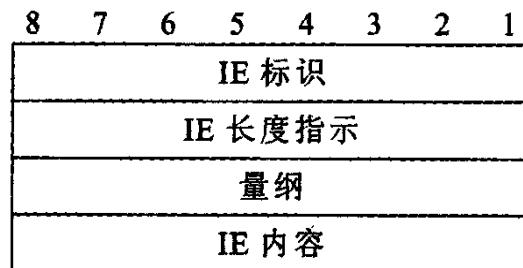


图 4-5 信息单元的一般格式

信息单元中的“IE 长度指示”是 IE 内容的长度值；而 IE 标识则是根据不同 IE 的编码来区分不同的 IE。“量纲”是值 IE 内容中数值的单位。“IE 内容”则是 IE 的数值。

各种不同的消息应该或可包含不同的 IE，如连接请求消息 SETUP 包括目的端地址、QoS 参数等 IE；而连接释放消息 RELEASE 则包括原因等 IE。信息单元的编码和说明见下表：

表 4-4 信息单元及其标识

IE 编码	说明
0000 0000	主叫 IP 地址
0000 0001	目的 IP 地址
0000 0010	承诺平均速率
0000 0011	突发数据率
0000 0100	端到端延时
0000 0101	端到端抖动
0000 0110	数据丢失率
0000 0111	一次双向协商（可选）

0000 1000	安全属性 (可选)
0000 1001	连接释放原因

4.3.4 连接状态

以下说明用上述 UNI 信令进行连接建立请求和释放时信令的用户侧和网络侧状态以及状态转移。

4.3.4.1 UNI 信令用户侧的连接状态

UNI 用户侧定义的呼叫/连接状态如表 4-5 所示

表 4-5 UNI 信令用户侧的呼叫/连接状态

状态名称	说明
U0: 空	无呼叫存在
U1: 呼叫启动	用户发出连接请求
U2: 呼叫存在	用户收到连接建立请求但未响应
U3: 连接请求	用户已应答
U4: 连接建立	用户发送连接建立的证实
U5: 激活	进入数据传输阶段
U6: 释放请求	用户请求网络释放连接
U7: 释放指示	用户已经收到切断指示

其中, U1、U4 用于连接请求建立端, U2、U3 用于连接目的端
用户状态转移简图

如图 4-6 所示, 在空状态收到用户发送的 SETUP 消息后, 进入呼叫启动状态, 并启动定时器。如果定时器超时而仍未收到确认消息, 可允许重试, 再发送 SETUP 消息, 如果仍然超时, 则取消此次连接请求。当定时器未超时而收到 CONNECT ACK 消息, 则进入连接建立状态, 此时用户发送一个对 CONNECT ACK 消息的证实消息。当对端用户收到此消息后, 连接建立过程完成, 进入激活状态, 开始数据传输。

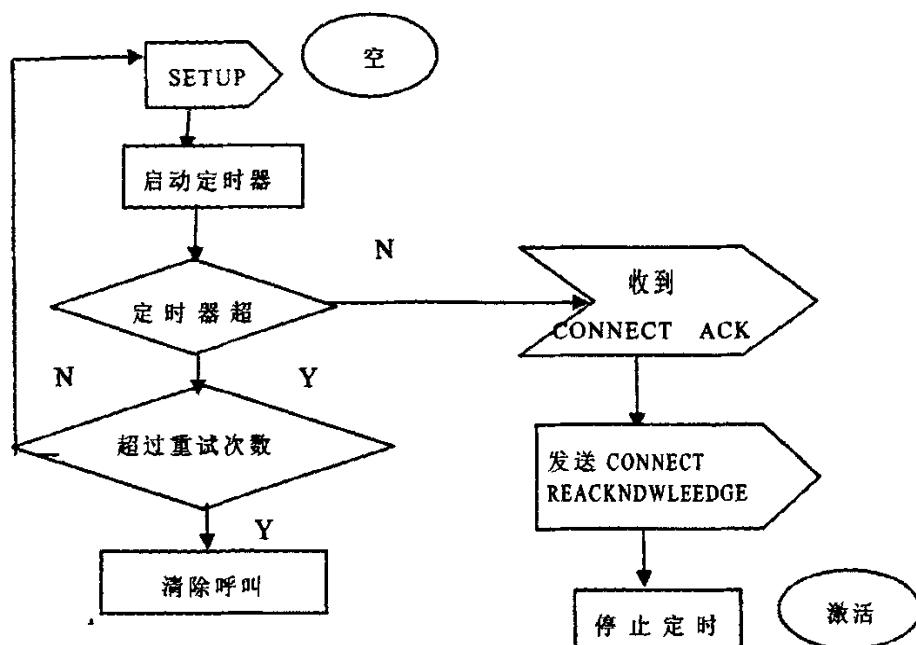


图 4-6 用户侧呼出呼叫流程图

4.3.4.2 UNI 信令网络侧的呼叫连接状态

网络侧同样定义了相关的呼叫转移状态，具体见表 4-7

表 4-6 UNI 信令网络侧的呼叫/连接状态

状态名称	说明
N0: 连接启动	网络侧收到连接建立请求
N1: 连接进行	网络侧已发送连接建立请求
N2: 连接接受	网络侧收到应答
N3: 连接证实	网络侧收到连接证实消息
N4: 激活	数据传输
N5: 释放请求	网络收到用户来的释放请求
N6: 释放指示	网络已切断连接

网络侧呼出状态转移简图，从空状态收到 SETUP 消息后进入呼叫启动状态，如果有足够的资源满足呼叫请求且被叫接受此次呼叫，则网络发送 CONNECT 消息给用户，并进入激活状态。

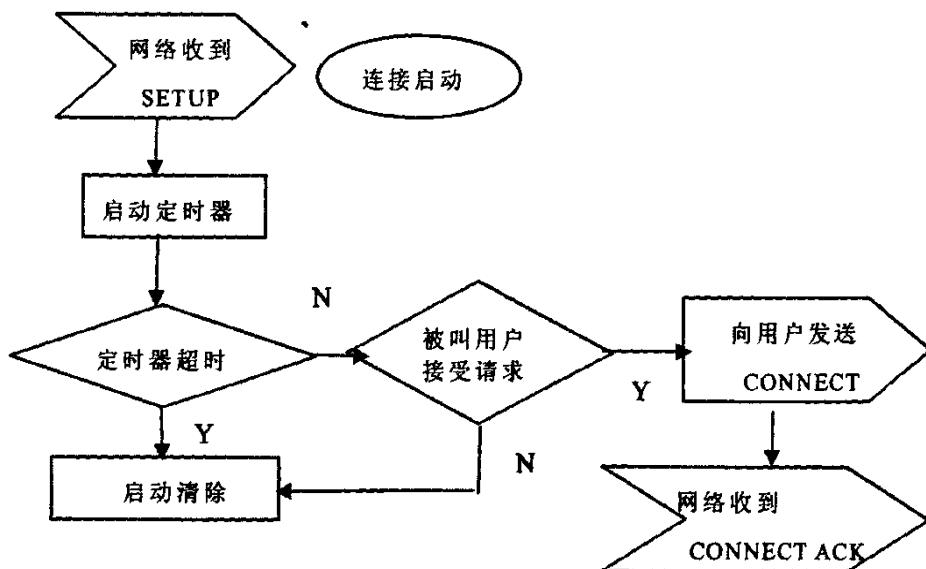


图 4-7 UNI 信令网络侧呼出流程图

整个 UNI 信令在传送层采用 UDP 协议传递, PDU 格式采用表 4-1 SUPANET 专用应用层协议通用 PDU 格式, 端口号依旧采用 400。采用通用的 PDU 格式更有利于 UNI 信令和 NNI 信令之间的交换和转换。其中协议编号采用“11”, 用来区分 SUPANET 专用的其他协议。而“PDU 类型”根据表 4-3 中的报文类型来编号。

4.4 UNI 信令和 NNI 信令的转换

SUPANET 网络定义的相关控制协议如 QoSNP、CEP 实现网络内部信息的传递, 而 UNI 信令初步实现了普通以太网用户接入 SUPANET 的梦想。但由于 UNI 信令和 SUPANET 网络内节点间采用的 QoSNP 以及 CEP 无论是在 PDU 格式还是信息元素格式上都有较大的区别, 所以用户和用户之间的信令传输在与网络的结合处需要进行转换, 故转换主要发生在如下两个地方:

- 源端接入交换机将接收到 UNI 网络侧信令转化为 NNI 信令。
- 目的端接入交换机将接收到的 NNI 信令转化为 UNI 网络侧信令。

由于消息格式和信息单元的格式不尽相同, 故 UNI/NNI 信令的之间的转换包括消息之间的转换、信息单元之间的转换以及呼叫对应这

几部分

4.4.1 消息之间的转换

由图 4-4 知道, UNI 信令的消息由消息头和可变长度的信息单元组成, 其呼叫标识、消息类型和消息长度在消息确定之后, 值都是固定的。而由图 4-2 知道, QoSNP 定义的消息中的协议编号、PDU 类型等字段也是固定的, 信令转换时只需要进行相关的消息对应。由于 QoSNP 的局部类型消息如 L-TSDN Acknowledge、L-TSDN Neg Acknowledge 属于 SUPANET 域内节点之间的交换消息, 故不牵涉到消息的转换, 转换只是牵涉端到端的消息的转换。具体的消息转换序列图如图 4-8 所示:

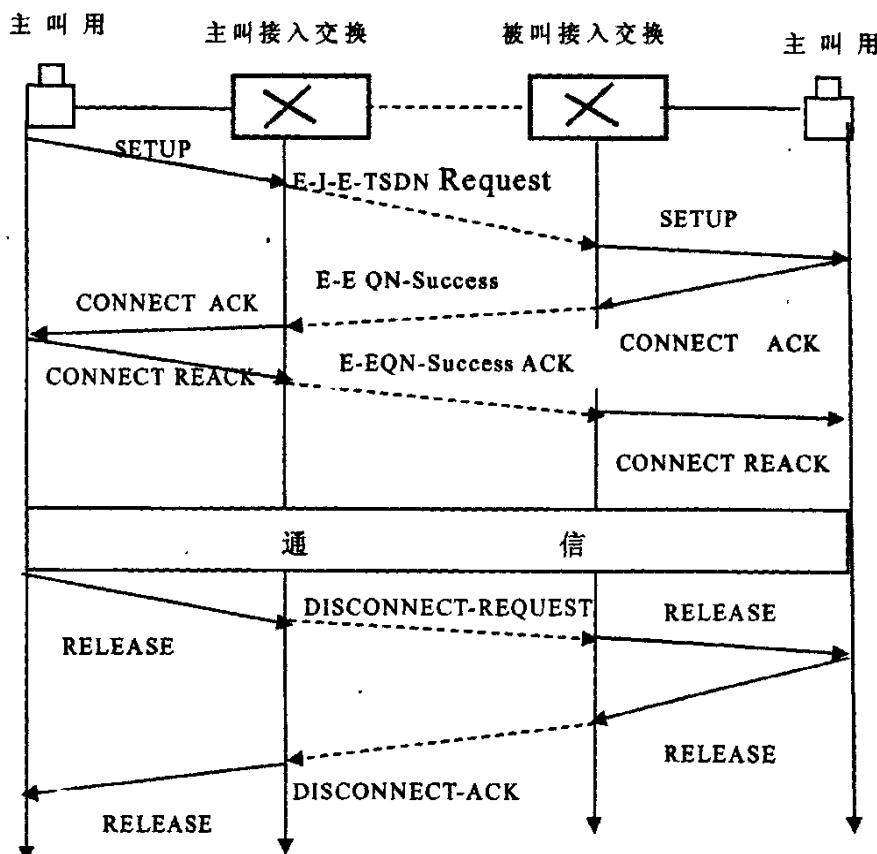


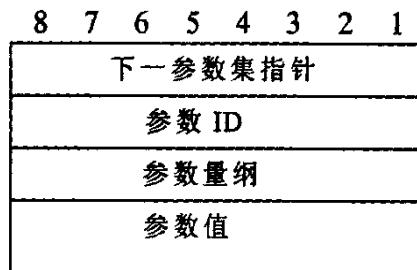
图 4-8 UNI/NNI 消息转换序列图

图 4-9 中, 当用户端生成一个 SETUP 请求消息后, 若此次呼叫通

过 CAC 机制的检查，确定为合法的呼叫，主叫接入交换机将启用信令转化模块，SETUP 消息被转换为 E-I-E-TSDN Request 消息，QoSNP 协议被激发，开始 SUPA 域内各节点的逐级协商。当 E-I-E-TSDN Request 消息到达被叫接入交换机后，同样需要启动信令转换模块，将 E-I-E-TSDN Request 消息重新转化为 UNI 信令的 SETUP 消息，发送到被叫用户。其他消息的转化过程类似、在此不再赘述。

4.4.2 信息单元之间的转换

QoSNP 协议中的消息的信息单元的一般格式如表 4-9 所示，其参数单元格式类似于表 4-5 中 UNI 信令中定义的信息元素的一般格式，格式的吻合有利于 UNI 信令信息元素到 QoSNP 参数单元的转换。



4-9 QoSNP 参数单元格式

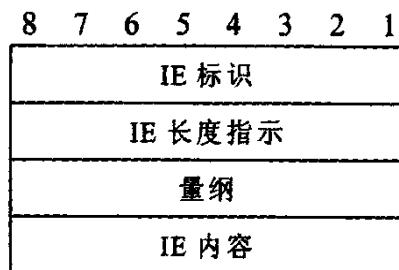


图 4-5 信息单元的一般格式

在转换的过程中，信令转换单元根据信息单元的 IE 标识符和 QoSNP 的参数单元进行对应，量纲也进行相应的对应转换。UNI 信令中的消息根据“消息长度”域以及“IE 长度指示”来确定此消息中在本信息元素后面是否还有信息元素；而 QoSNP 参数单元则是通过“下一参数集指针”域来确定是否还有后继参数，上述的转换过程比较简单。

需要注意的问题是参数的兼容。在 UNI 信令中定义的相关信息单元，例如“安全属性”，在 QoSNP 中可能没有相应的定义，即 SUPANET 可能现行还不能支持业务传输安全属性需求，这种情况下 QoSNP 将会忽视此信息单元。而对于一些 UNI 信令信息单元类型和 QoSNP 中定义的参数类型可能会出现不兼容的现象，故可能需要做一些参数的转化操作，将 UNI 信令中定义的相关信息元素“最大限度”地转换为 QoSNP 中已经定义的参数类型，以求在 QoSNP 协商参数中尽量准确

地反映出用户的服务质量需求。

4.4.3 呼叫的对应

UNI 信令的消息中通过两字节的“连接标识”来唯一标识一次连接请求，“连接标识”用于本地接口识别所属的消息对应的连接，由连接的请求端分配，且在请求端侧是唯一的。而 QoSNP 中用于唯一标识一条虚连接的是采用“源/目的端虚通路 ID”来标识的，在进行端到端服务协商请求时，开始只有“虚通道源端 ID”，由主叫接入交换机中的 NNI 信令生成。在源端用于标识唯一的一条虚连接。当请求消息到达被叫端接入交换机后，目的端也会生成一个两字节的虚通道标识符。当协商成功后发送连接确认消息时，被叫接入交换机将其生成的虚通道标识符填入“源端虚通道 ID”，“虚通道目的端 ID”为请求端虚通道 ID。故在源端和目的端，均可使用“虚通道标识符”来唯一的虚连接。在 UNI 和 NNI 信令的转换过程时，UNI 信令中的“连接标识”替换为 QoSNP 中的“虚通路源端 ID”即可完成 UNI 信令的呼叫和 NNI 信令的呼叫的转换。

至此，UNI 信令和 NNI 信令的转换全部完成。

4.5 本章小结

本章主要对以太网接入 SUPANET 技术最主要的实现基础：用户—网络接口信令（UNI 信令）进行初步的设计。其中包括信令消息的设计、信息单元的设计以及整个信令流程的介绍；最后对于 UNI 信令和 NNI 信令的转换也作出了初步介绍，基本实现了用户接入 SUPANET 的信令需求。

第 5 章 接入流程仿真

以太网接入 SUPANET 技术紧密结合 SUPANET 的特性而设计的，整个接入的过程和 SUPANET 信控平台的 QoSNP 协议、CEP 协议需要密切地配合，组成一个完整的传输系统。为了验证论文中所设计的 UNI 信令和 QoSNP 协议之间的转换和配合工作的可行性和正确性，本章利用 OPNET Modeler^[46,47,48]软件对接入流程、各节点之间消息传递过程的进行了仿真。

由于 OPNET Modeler 采用了三层建模机制（网络模型、节点模型、进程模型），最低层为进程模型，以状态机来描述协议；其次是节点模型，由相应的节点模型构成，反映设备特性；最上层为网络模型。三层网络模型和实际网络的网络、设备、协议层次完全对应。故仿真过程需要分别对仿真环境的网络域、节点域和进程域进行建模。

5.1 仿真模型设计

5.1.1 网络拓扑结构建模

整个仿真网络拓扑结构由六个节点组成，其中包括源用户节点 Source、目的用户节点 Dest、两个 SUPANET 边界网关 SUPAGetway 以及两个核心交换节点 PFTS1 和 PFTS2，节点之间采用全双工链路连接。Source 和 Dest 均执行第四章中设计的 UNI 信令；SUPAGetway 用来实现 UNI 信令和 NNI 信令之间的转换；PFTS1 和 PFTS2 分别用来模拟 SUPANET 中的核心交换节点。在整个网络拓扑中并没有考虑到 QoS 路由的问题（因为这并不是本次仿真过程的着重点），而是直接在各个节点之间进行业务协商和相关消息的传递。整个网络拓扑结构如下图：

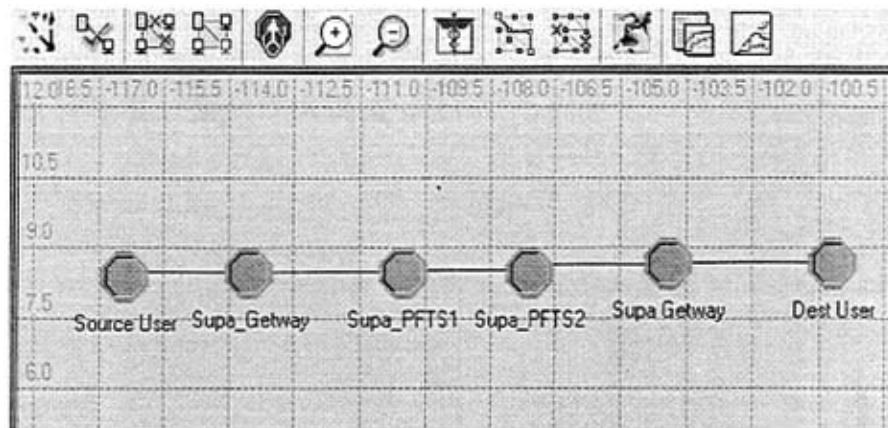


图 5-1 网络域拓扑结构图

5.1.2 节点建模

由上述的网络拓扑图知道，整个仿真环境中一共有三种类型的节点：普通以太网主机、SUPANET 边缘交换节点以及 SUPANET 核心交换节点。不同类型的节点的处理行为和完成的任务各不相同，故需要分别建模。

5.1.2.1 源节点模型

源节点是一个普通的以太网主机，其底层通信模式和传统的以太网内的主机没有任何不同。由于仿真的目的在于验证第四章所设计的 UNI 信令同 QoSNP 协议的协同配合工作流程，故对源节点的仿真重点在于对 UNI 信令的仿真。UNI 信令是在应用层实现的，其相关消息格式和信息单元格式在论文第四章中有详细的设计说明，在具体的仿真过程中对相关的消息格式做了简化。

源端节点包括四个模块，一个连接建立请求数据发起模块，其作用是模拟以太网用户在应用层发起连接建立请求，产生数据包的格式按照第四章设计的 SETUP 消息的格式生成；一个 proc 处理模块，它是源节点的核心模块，其作用是在收到 SRC 模块产生的 SETUP 数据包后，填充相关的数据域，并将其发送出去，另外，proc 模块还需要

对其收到的相关消息进行处理和回应，包括目的端的连接确认以及反向确认消息的发送。另外源端节点还需要一对发信机和收信机来实现数据包的传递，整个节点模型如下：

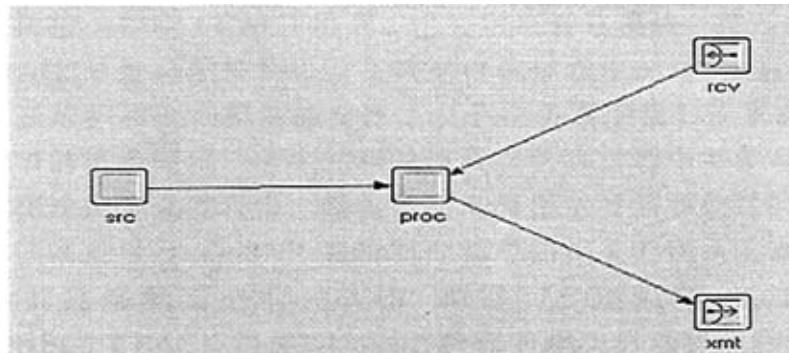


图 5-2 Source 节点模型示意图

src 进程模型的有 OPNET 自带的 simple source 模型改造而来，让其在固定时段内产生 SETUP 请求消息。而 proc 模块则需要根据在第四章中设计的 UNI 信令，通过有限状态机的转换，由进程完成信令相关行为，处理各种消息报文，proc 模块的状态变迁图如下图所示：

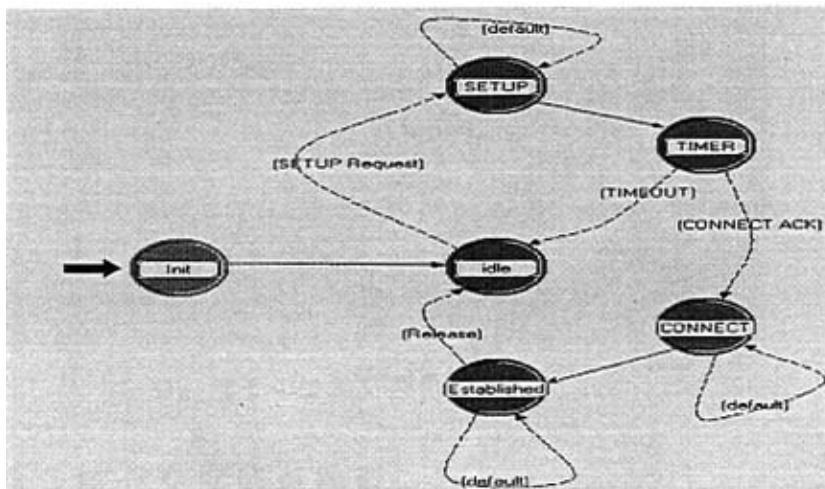


图 5-3 Proc 模块进程状态变迁图

5.1.2.2 接入交换机模型

在接入流程中，SUPANET 接入交换机是 SUPANET 与以太网之间

的网关，其作用是从源端用户发送来的业务请求中提取相关的 QoS 参数，实现 SUPANET UNI 信令到 NNI 信令的转换，启动 QoSNP 协议，开始 SUPANET 网络内部节点的连接建立协商。在接入交换机上还需要连接接纳控制（CAC）机制、相关用户管理模块、接入技术中的“虚拟 MAC”地址分配模块以及在数据传输过程中的对“虚拟 MAC—VLI”映射表维护和查询模块的支持，由于在仿真中我们注重的是连接建立过程中的消息流程，故上述相关模块在仿真过程中并没有体现。

在连接建立过程中，接入交换机模型主要需要处理的 UNI 信令的网络侧消息以及 QoSNP 在协商过程中的相关的确认消息。其中 UNI 信令的网络侧消息有 SETUP 消息、CONNECT ACK 消息以及 CONNECT REACK 消息。QoSNP 协商的相关消息有协商请求消息、局部确认消息、端到端协商成功消息以及成功反向确认消息。

接入交换机的节点模型比较简单，有一个 Supa_GateWay 模块和两对发射机和接收机组成，Supa_GateWay 模块主要用来上述相关的消息，而发射机和接收机用来分别与用户和网络核心节点通信。其节点模型如下：

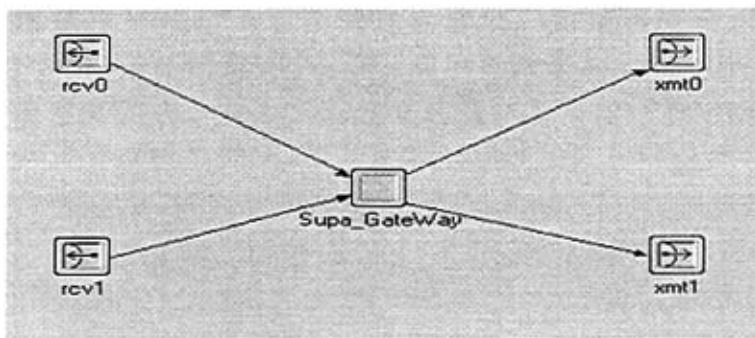
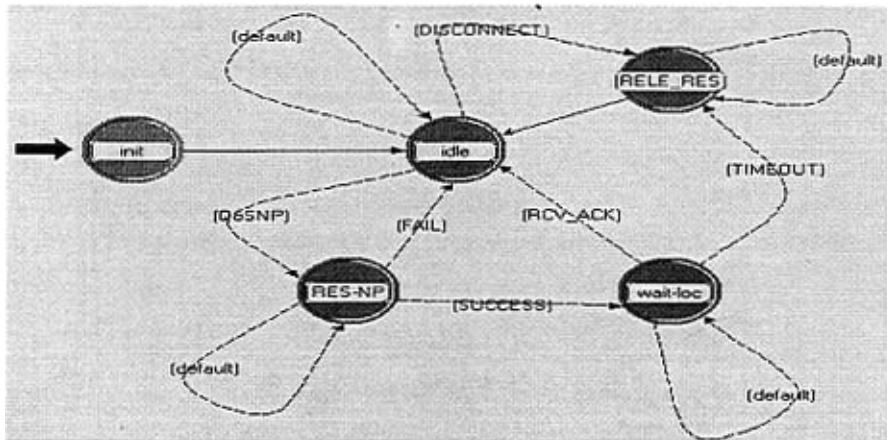


图 5-4 接入交换机节点模型

5.1.2.3 核心交换节点模型

SUPANET 的核心节点主要是在 U 平台完成 EPF 帧的高速交换，同时在 S&M 平台完成 QoSNP 协商以及其他网络内部的控制、管理信息的传递。在此次仿真过程中没有牵涉到 U 平台数据的传输，故核心交换节点只需要处理 QoSNP 协商过程相关的消息的传送，节点模型由一个处理模块和两对收、发信机组成。其中的处理模块的状态变迁图

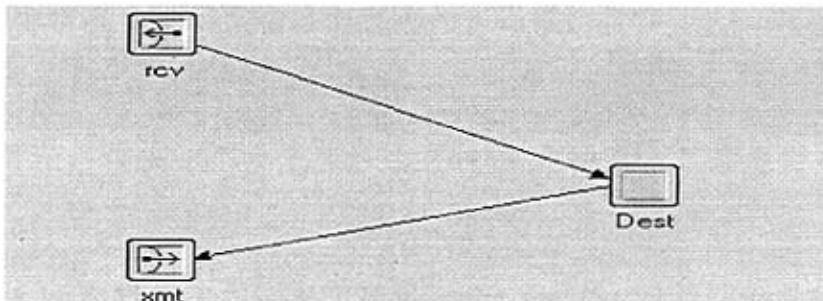
如下所示：



5-5 核心交换节点的状态变迁图

5.1.2.4 目的节点模型

目的节点收到源端的 SETUP 消息后, 根据业务请求的 QoS 参数和业务流描述来决定是否接受连接请求。如果接受请求, 则发送一个 CONNECT ACK 消息, 当收到源端用户回送的 CONNECT REACK 消息后连接建立完成。若不同意建立连接, 则发送 CONNECT Reg-ACK 消息。其节点模型如下:



5-6 目的节点模型

目的节点中的 Dest 模块的状态变迁图如下:

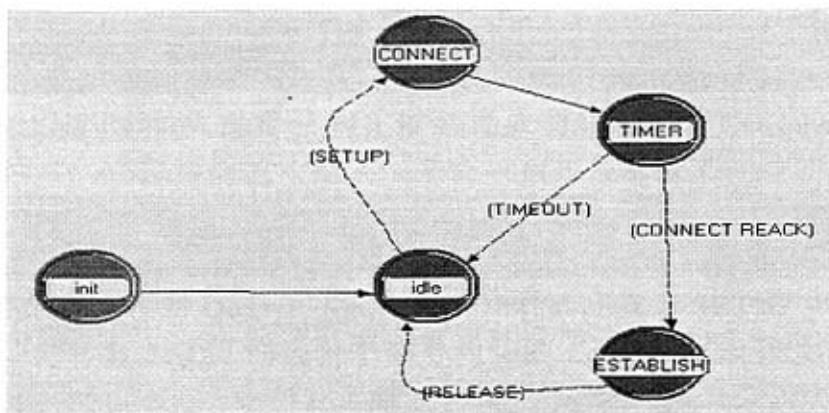
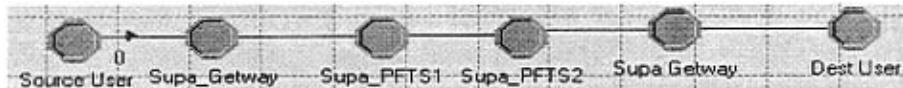


图 5-7 Dest 模块的状态变迁图

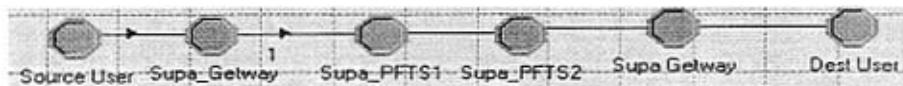
5.2 接入技术流程仿真

5.2.1 接入流程仿真

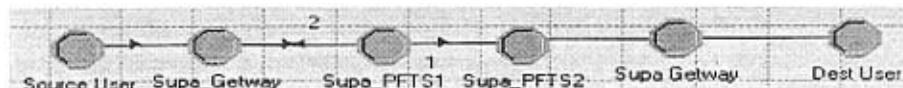
在此次仿真过程中，主要是检验第四章定义的 UNI 信令和 QoSNP 协议的配合工作过程。这个过程主要是通过 OPNET 中的动画回放功能来回放连接建立过程中消息的流动，同时在 OPNET 的控制台上输出相关的消息来体现消息传输的正确性。另外，通过统计核心节点上资源的预留情况来说明整个协商过程的正确性。



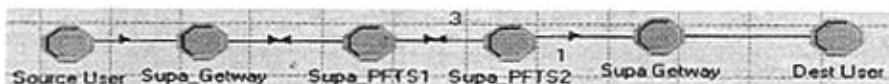
首先，Source User 生成一个 SETUP 请求包，其编号为 0；



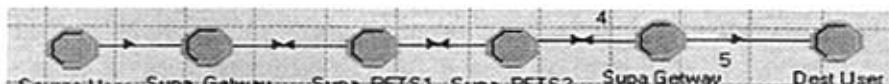
SETUP 包传送到接入交换机 supa_Gateway, 经过信令转换后，生成 QoSNP Request 报文，其编号为 1；



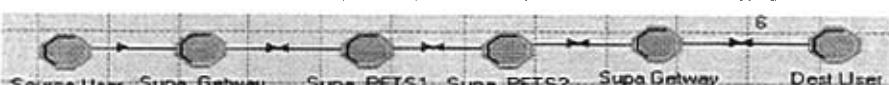
QoSNP Request 报文前传到 PFTS1 节点，经协商完成之后，向 supa_Gateway 回送一个的编号为 2 的 Local-ACK 局部确认报文。



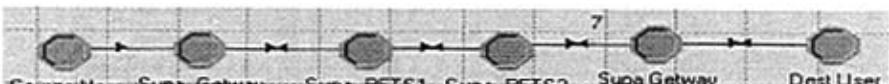
QoSNP Request 报文经 PFTS2 协商成功后，向 PFTS1 回送编号为 3 的 Local-ACK 局部确认报文，并继续前传请求报文。



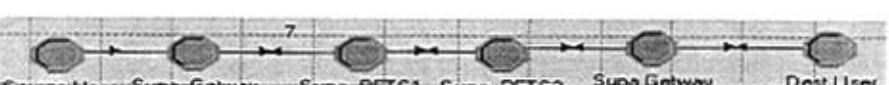
目的网关收到 QoSNP Request 后，将其还原为 SETUP 报文，编号为 5，同时向 PFTS2 回送一个编号为 4 的 Local-ACK 报文。



DEST 收到 SETUP 报文后，同意建立连接，回送一个 CONNECT ACK 报文，编号为 6。



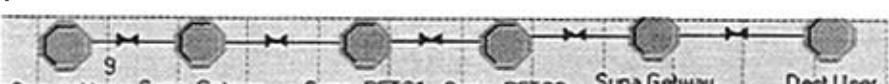
目的网关将收到的 CONNECT ACK 报文后，转换为 E-E QN Success 报文，其编号为 7。



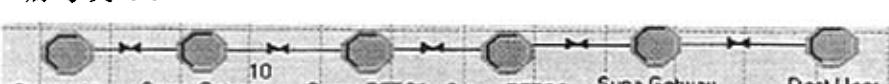
E-E QN Success 报文回送到源端网关



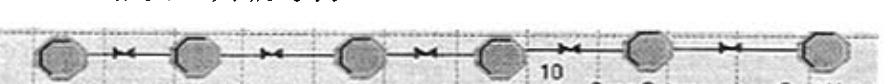
源端网关将 E-E QN Success 报文还原为 CONNECT ACK 报文，其编号为 8。



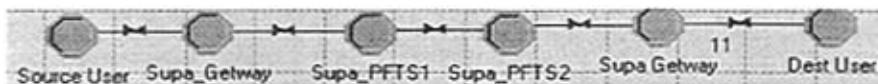
源端用户收到 CONNECT ACK 报文后回送一个 CONNECT REACK 报文，编号为 9。



源端网关收到 CONNECT REACK 报文后，生成一个 E-E QN Success ACK 报文，其编号为 10。



E-E QN Success ACK 报文前传到目的端网关, SUPA 域内协商过程完成。



目的端网关将 E-E QN Success ACK 转换为 CONNECT REACK 报文, 编号为 11。DEST 收到此报文后, 连接建立完成。

上述过程体现了从源端用户到目的用户连接建立的完整过程。从源端用户连接请求和应答, 到接入交换机 UNI/NNI 信令转换以及 QoSNP 协商、相关确认信息的传递, 均完全符合第四章设计的 UNI 信令过程和 UNI\NNI 信令转换流程。

5.2.2 对节点资源预留情况的仿真

仿真实验一验证了连接建立流程的正确性, 实验二主要是验证连接建立失败情况下消息的流程。因为用户在发起连接建立请求时, SETUP 消息携带了相关的 QoS 参数, 实际协商过程中 SUPANET 网关会提取相关的 QoS 参数并依据这些参数来在中间节点进行协商。在进行仿真的过程中, 为了简化协商的过程, 仿真时并不是对每个参数进行协商, 而是将这些参数综合转化为一个时槽数目, 用时槽数来等效资源的需求量, 通过时槽数目的预留量来模拟节点资源的预留情况。

表 5-1 各节点时槽数分配表

Src Gateway	PFTS1	PFTS2	Des Gateway
550	420	300	230

实验参数设置:

用户请求时槽数: 60

连接请求发送间隔时间 : 5s

运行时间: 30s

随机种子数: 65535

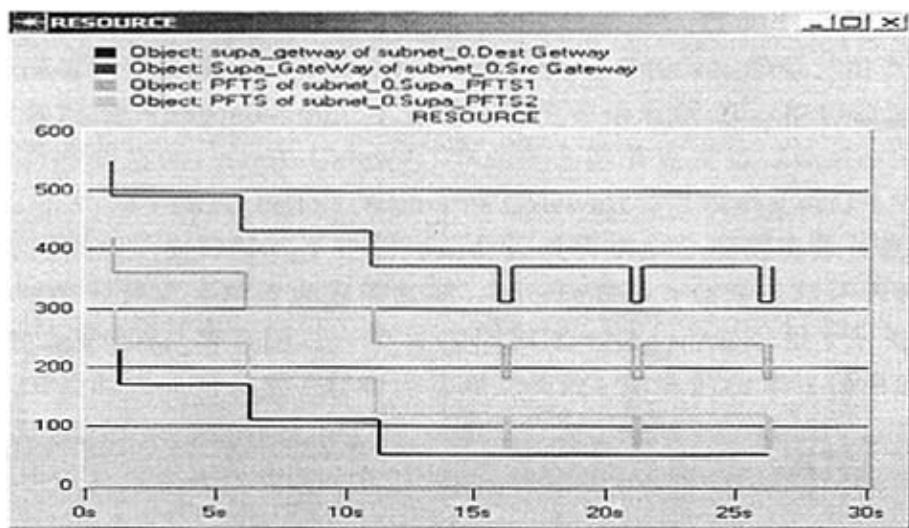


图 5-8 资源预留统计

在仿真过程中，源端用户一共产生了 6 次 SETUP 请求。在前三次中，由于时槽数最少的 Dest Gateway 节点能满足业务需求 ($230 > 3 \times 60$)，故各个节点均能预留资源，连接建立成功。下图是通过在仿真节点收到报文后，通过 OPNET 提供的 API 函数 `op_pk_format()` 函数判断报文的类型，然后在 OPNET 的 console 控制台中输出相关的报文消息，整个完整体现了一次成功协商的流程。

```
[d:\PROGRA~1\OPNET\10.0A\sys\pc_intel_win32\bin\op_runsim.exe]
: Beginning Simulation.

Source User create a SETUP MESSAGE
GateWay received a SETUP MESSAGE from Source User, and start the QoSNP process
PFTS1 received a QoSNP_Request MESSAGE from Source Getway
Source GateWay received a Local ACK MESSAGE from the PFTS1
PFTS2 received a QoSNP_Request MESSAGE from PFTS1
PFTS1 received a Local_Ack MESSAGE from PFTS2
DEST GetWay received a yu_QoSNP_Request MESSAGE from PFTS2
PFTS2 received a Local_Ack MESSAGE from Dest Getway
Dest received a SETUP MESSAGE from Source User
Dest Getway received a CONNECT_ACK MESSAGE from DEST
PFTS2 Received a E-E QN-Success MESSAGE from Dest Getway
PFTS1 Received a E-E QN-Success MESSAGE from PFTS2
Source GateWay received a E-E QN-Success MESSAGE from PFTS1
SOURCE received a CONNECT ACK MESSAGE from DEST
Source GateWay received a CONNECT RE-ACK MESSAGE from User?
PFTS1 received a E-E QN-Success_ACK MESSAGE from Source Gateway
PFTS2 received a E-E QN-Success_ACK MESSAGE from PFTS1
Dest Getway received a E-E QN-Success_ACK MESSAGE from PFTS2
Dest received the CONNECT REACK MESSAGE from the Source, Connection established!
```

图 5-9 在 console 中输出的报文流程

当开始第四次业务请求协商后, Source 节点、PFTS1 和 PFTS2 节点均成功预留资源(线条下降), 而对于 Dest Gateway 节点, 由于其资源剩余量 50 ($230 - 3 \times 60$) 小于业务的资源请求量 60, 故协商失败, 并没有预留资源。Dest Gateway 节点向上游节点发送 Local-Neg-ACK 报文, PFTS2、PFTS1 和 Source Gateway 节点收到 Local-Neg-ACK 报文后。立即释放了先前为此次连接请求所分配的资源(线条从下降处回升到第 4 次请求前的高度)。Source Gateway 节点归还资源后, 向 Source User 节点回送一个 CONNECT-Neg-ACK 报文, Source User 节点收到此消息后, 宣布连接建立失败。消息的传递过程图如下:

```
[d:\PROGRA~1\OPNET\10.0A\sys\pc_intel_win32\bin\op_runtimed.exe]
Source User create a SETUP MESSAGE
Gateway received a SETUP MESSAGE from Source User, and start the QoSNP process
PFTS1 received a QoSNP_Request MESSAGE from Source Getway
Source Getway received a Loacal ACK MESSAGE from the PFTS1
PFTS2 received a QoSNP_Request MESSAGE from PFTS1
PFTS1 received a Local_Ack MESSAGE from PFTS2
DEST GetWay received a yu_QoSNP_Request MESSAGE from PFTS2
PFTS2 received a Local-Neg-ACK MESSAGE
PFTS1 received a Local-Neg-ACK MESSAGE
Source Getway received a Local-Neg_ACK MESSAGE
Source received a CONNECT-Neg_ACK MESSAGE, Connection request failed

-----
: Simulation Completed - Collating Results.
: Events: Total <439>, Average Speed <1,326 events/sec. >
: Time: Elapsed <0.33 sec. >, Simulated <30 sec. >
-----
Press <ENTER> to continue.
```

图 5-10 第四次业务请求的消息传输过程

5.3 小结分析

在上述的流程仿真过程中可以看出:

- (1) 仿真过程体现了从用户到源端接入交换机、源端接入交换机到目的端接入交换机, 再到目的用户的消息传送过程。协商成功和不成功两种情况的消息流程在实验中均得到了体现, 并符合论文第四章中定义的消息流程。

(2) 仿真过程中论文设计的 UNI 信令和 QoSNP 协议的配合以及 QoSNP 的资源预留能力均得到了体现, 过程符合试验前预期的结果

论文中所阐述的接入思想是结合 SUPANET 和以太网的特性而提出的, 其思路新颖, 要具体实现还需要大量具体的工作需要去做。而本章的工作无疑对后期的具体实现有较好的指导作用。

结论与展望

1. 研究总结

作为四川省网络通信技术重点实验室对于 SUPANET 系统研究工作的一部分，本文就以太网接入 SUPANET 作出了相关的研究，归纳起来有如下几点：

- (1) 分析了以太网用户接入高速网络的相关技术。包括以太网接入 ATM 网络和帧中继网络的相关技术。
- (2) 根据 SUPANET 和以太网的特性，提出了以太网接入 SUPANET 的一种新思想。SUPANET 最重要的特性在于其单物理层用户数据传输平台、采用带外信令技术、面向以太网帧以及采用虚连接进行定长帧交换，论文根据上述特性提出了以太网接入 SUPANET 的接入思想，力求简化接入交换机的工作过程，保障用户到用户间业务传输的服务质量。
- (3) 根据接入思想和实验室关于 SUPANET 信控平面相关协议的研究，设计了普通以太网用户接入 SUPANET 的 UNI 信令。具体定义了信令的消息格式、信息单元格式、信令流程以及 UNI 信令和 NNI 信令的转化过程。
- (4) 采用 OPNET 仿真工具对所定义的接入流程进行了仿真。

2. 前景与展望

SUPANET 是为满足未来“三网合一”需求而设计的新型高速骨干传输网络，在设计之初就确定了其面向以太网帧的用户数据特性。随着以太网技术在现有局域网中市场地位的不断巩固，SUPANET 与传统以太网的融合问题可能是决定 SUPANET 能否被以太网用户所接受，最终能否走向成功的关键因素之一。论文中所提出的接入技术虽然对相关问题作出了初步探讨，但这是远远不够的，今后还需要进一步开展工作，这包括：

- (1) 接入技术的性能分析。虽然我们可以定性分析得出文中的本文所提出的接入技术能够简化接入交换机的处理流程，有利于保障用户

到用户间业务传输的服务质量，但缺乏定量分析，最终的结果有待详尽的实验方案。

- (2) 具体细化论文中所定义的 UNI 信令。由于时间紧迫，本论文仅对 UNI 信令进行了框架设计，其中的相关消息的信息单元设计并不完整。另外对于所设计的 UNI 信令正确性、健壮性、安全性以及可扩展性都有待进一步研究
- (3) 对以太网用户接入 SUPANET 的接入管理问题在论文中没有任何涉及，还需下一步研究。

致谢

当这篇论文完成时，心中的有太多的感慨！它不仅是我三年研究生学习阶段的一次总结，凝结着我在求学路上的努力和探索，更蕴涵着三年来老师辛勤的指导、亲人无私的关怀和同学的帮助。

首先要深深感谢我的指导老师—窦军副教授。窦老师渊博的学识、严谨的工作态度以及对学术孜孜不倦的探索精神深深的影响了我，他时时刻刻激励我去努力探索，追求进步。从论文的立项到撰写完成，窦老师给了我无数悉心的指导，他进行经常性、阶段性地把关，提出了许多宝贵的指导性意见。老师的关怀和指导是本论文完成的最重要的保证。同时，在我读研期间~老师给我提供了许多学习和实践的机会，拓展了我的思路和知识面，使我的科研能力和实践能力得到了提高。他对我的影响和教育将是我一生享受不尽的宝贵财富。在此我向 老师表示由衷的感谢和诚挚的敬意。

另外，也要感谢曾华燊教授，曾老师对 SUPA 体协结构的宏观上的把握和指导为我的论文研究工作指明了方向，使我获益良多！

还要感谢西南交通大学网络通信技术重点实验室所有同学和朋友给予的帮助，特别是李季博士、汪海鹰博士以及徐超、陈美燎、罗正武、段旭哲、毛磊雅、陈华等同学的热心帮助。你们的鼓励和支持为我论文撰写过程中克服各种困难提供了巨大的勇气！

最后，我要感谢我的家人，他们所给予的支持和关怀使我专注于课题研究，并顺利完成硕士期间学业的重要保证。

参考文献

- [1] Huaxin Zeng Jun Dou & Dengyuan Xu, "Single physical layer U-platform Architecture (SUPA) for Next Generation Internet", IEC (2004) Annual Review of IP Applications and Services.
- [2] HuaXin Zeng, Jun Dou, DengYuan Xu, "On Three-Dimension Ethernet based MAN (3D-EMAN) Architecture", in PDCAT'03, IEEE press, August 27 – 29, 2003, Chengdu, China.
- [3] 曾华燊, 朱怀芳, “现代网络通信技术”P377-386,西南交通大学出版社.2003
- [4] Jun DOU, Huaxin ZENG, Haiying WANG, "Single User-Platform Architecture and its QoS Provisioning Mechanisms in Signaling and Management (S&M) Platforms", in: Proceedings of 5th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT2004), Singapore, December 2004, Lecture Notes in Computer Science (LNCS 3320), published by Springer, pp. 429-440.
- [5] 徐民鹰, 刘信圣, “三网合一技术基础”, 中国国际广播出版社. 2003
- [6] HuaXin Zeng,Jun Dou, Dengyuan Xu,“Replace MPLS with EPFTS to Build a SUPANET. 2005 Workshop on High Performance Switching and Routing”, HongKong, China, 2005.5
- [7] HuaXin Zeng, Jun Dou, DengYuan Xu, "Single physical layer U-platform Architecture (SUPA) for Next Generation Internet", IEC Comprehensive Report(2003) on Internet Protocol(IP) Applications and Services, IEC Press, December 2003
- [8] Huaxin Zeng, Dengyuan Xu, Jun Dou, "On Physical Frame Time-slot Switching over DWDM, " in: Proceedings of PACAT03,Chengdu, China,2003. IEEE press: 286-291
- [9] 曾华燊, 许登元, 李季. “SUPANET中的物理帧时槽交换技术”, 计算机应用, 2004, 24(6): 6-9
- [10] 许登元, “类以太网物理帧时槽交换技术研究”, 西南交通大学博士论文.2005
- [11] 曾华燊, “论下一代Internet体系结构”, 计算机应用, 2004, 24 (6)
- [12] 赵慧铃, 张国红, 胡琳, 石友康, “ATM、帧中继、IP 技术与应用”, P135-139,

电子工业出版社.1998

- [13] M.D.Prycher, "Asynchronous Transfer Mode: Solution for Broadband ISDN,2ed, "New York : Eills Horwood,1993.
- [14] Liza Henderson,Tom Jenkins. "Frame Relay Internetworking ", 陈圣琳, 王欣 译, 电子工业出版社.2000
- [15] 王昕,“帧中继—LAN/WAN 互连的新途径”.电信科学.第 11 卷第 8 期 1995
- [16] Chen K, Ho K,Saksena V. Analysis and Design of a Highly Reiable Transport Architecture for ISDN Frame Relay Networks. IEEE Journal on Selected Areas in Communication, October 1989.
- [17] 田波,“帧中继技术在局域网互连中的应用”, 通信技术.第总 92 期, 1996
- [18] 李成忠, 张新有. “计算机网络原理与设计” .电子工业出版社.2003
- [19] David E.McDysan, Darren L.Spohn, "Hands-On ATM",McGraw-Hill Companies,Inc
- [20] N.finn and T.Mason, "ATM LAN Emulation", IEEE Communications,June 1996
- [21] ATM Forum, "LAN Emulation over ATM—LUNI Specification", Version 2,aflane-0084.000,July 1997.
- [22] 全首义, “ATM 宽带技术及应用”, P87-91,北京邮电大学出版社,1998.
- [23] 刘文杰. “半步超前交换技术及其改进研究 ”.西南交通大学硕士研究生学位论文.2006.
- [24] 邱智亮. “ATM 接入 IP 交换技术研究 ”.西安电子科技大学博士论文.1999
- [25] 刘文杰, 李季, 许登元, “半步超前交换技术及其性能分析”. 2005年信息、电子与控制技术学术会议论文集, 成都, 2005年: 25-29
- [26] William Stallings, " ISDN、B-ISDN与帧中继和ATM", 第 4 版, 机械工业出版社.2001
- [27] 王俊芳, 温日红. “接入型ATM交换机的实现 ”.无线电通信技术.第26卷, 第4期, 2000
- [28] 黄锡伟, 朱秀昌, “宽带通信网络”, 人民邮电出版社.1998、17;
- [29] D.Mcdysan and D.Spohn,ATM : Theory and Application ,McGraw-Hill,1995
- [30] ATM Forum ,User-Network Interface Specification,Version 3.1,September 1994
- [31] 林闯, 单志广, 任丰原, “计算机网络的服务质量 (QoS)”, 清华大学出版社.2004

- [32] 陈锡, “ATM 交换技术”, 人民邮电出版社.2000
- [33] RFC 3035 - “MPLS using LDP and ATM VC Switching”, Jan 2001
- [34] Dengyuan Xu, Xinyou Zhang, Jun Zhao, “ Extended DWRR Scheduling Algorithm in PFTS”, in 2004 IEEE Workshop on IP Operation and Management Proceedings(IPOM 2004), Beijing, October 2004:191-197
- [35] Andrew S. Tanenbaum. “Computer Networks”. 熊桂喜, 王小虎 译. 清华大学出版社.2002.7
- [36] W.Richard Stevens. “TCP/IP Illustrated Volume 1:The Protocols”.Addison Wesley,1994
- [37] 张继军, 高鹏, “基于分组网络的服务质量保证”, 北京邮电大学出版社, 2004
- [38] 曾华燊, “广域网技术纵横”, 计算机应用, 1994, 第4期
- [39] 李炳彰, “用户信令与网络信令的转换”, 无线电工程.Vol. 29 ,No.3.
- [40] 傅明, 王喆. “现代通信网的信令类型及ITU建议的各种信令系统”.甘肃科技,第18卷,第11期.2002.11
- [41] 石晶林, 丁炜等. “MPLS宽带网络互联技术”, 人民邮电出版社.2001
- [42] 卢华斌,韩庆棉,方强,段喜风,李炳彰. “实用化的ATM用户信令(UNI)系统的实现”, 无线电通信技术.第27卷, 第5期.2001
- [43] ITU-T Recommendation Q.2931 Broadband Integrated Service Digital Network(B-ISND)-Digital Subsenser Signaling Ststem No.2 (DSS2)- User- Network (UNI) Layer 3 specification for Basic Call/Connection Control
- [44] 赵君. “SUPANET中的服务质量协商协议研究”, 西南交通大学硕士研究生学位论文.2006
- [45] 赵君, 高雨, “SUPANET中QoS协商流程的研究”. 计算机科学.2004.8, Vol31, 增刊.
- [46] 刘韵洁, 张智江等, “下一代网络”, 人民邮电出版社.2005
- [47] 马丁·德·普瑞克, “异步传递方式--宽带ISDN技术(修订本)”, 人民邮电出版社.1999
- [48] Thomas M.chen, “ATM Switching System”, Artech House ,London,1996
- [49] 陈敏, “OPNET网络仿真”, 清华大学出版社.2004
- [50] 网络论坛, <http://netforum.com.cn/>
- [51] 王文博, 张金文, “OPNET Modeler与网络仿真”, 人民邮电出版社.2003
- [52] ITU X.641 - Information technology - Quality of service: Framework,

December 1997.

- [53] A. Ralf Steinmetz and Klara Nahrstedt, "Multimedia Computing, Communications, and Applications", Prentice Hall International, Inc., 1995.
- [54] 曾华燊, 高雨, "论高速接入网技术", 计算机应用.第26卷第8期,2006
- [55] HenShall J, Shaw S. "OSI Explained: End-to-End Computer Communication Standards", ISBN 07458-0253-2, Ellis Horwood, 1988

攻读硕士学位期间发表的论文

- [1] 喻雷锋, 窦军, 张洪. SUPANET中的流量控制方法探讨. 计算机科学, 2005.08:58-61
- [2] 陈美燎, 窦军, 喻雷锋. SUPANET中的连接允许控制和流量整形研究. 计算机科学, 2005.08: 23—26