

11-7BCN-xx

SC 6 Nxxxxxx

**标题：未来网络：问题陈述和需求 – 第X部分：
网络模型 - TCP/IP/M**

源：中国国家成员体

项目：待定

项目编辑：待定

状态：第一稿

本文件是一份由中国国家成员体组织专家编写的技术报告。

这份报告的内容将在**2012年SC6/WG7**广州会议上被介绍和讨论，期待它能引起会议的广大关注并能成为技术报告**29181**的子项目。

ISO/TC /SC6 Nxxxxxx

Date: 2011-12-31

ISO-IEC TR29181-X (1rd WD)

ISO/JTC1/SC6/WG7

Secretariat: KATS

未来网络：问题陈述和需求 – 第 X 部分：

网络模型 – TCP/IP/M

Élément introductif — Élément principal — Partie n: Titre de la partie

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type:
Document subtype:
Document stage:
Document language:

Copyright notice

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to ISO's member body in the country of the requester:

*ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org*

as appropriate, of the Copyright Manager of the ISO member body responsible for the secretariat of the TC or SC within the framework of which the draft has been prepared]

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents

Page

1.	范围.....	1
2.	引用文件.....	1
2.1	规范性引用	1
3.	定义.....	1
3.1	网络	1
3.2	支路.....	1
3.3	电路 (CIRCUIT)	1
3.4	模拟电路.....	1
3.5	数字电路.....	1
3.6	虚电路	2
3.7	永久虚电路 (PVC)	2
3.8	交换型虚电路 (SVC)	2
3.9	虚实电路.....	2
4.	ABBREVIATIONS.....	2
5.	概述.....	3
5.1	基本结构.....	4
5.2	数据传送.....	6
5.2.1	TCP/IP 数据传送.....	6
5.2.2	M 数据传送.....	7
5.3	网络接口.....	7
5.4	IP 产生专一的逻辑网络	8
5.5	独立的物理网络	9
5.6	互相通信.....	9
6.	TCP/IP/M 三/四层模型	9
6.1	应用层	9
6.2	传输层	9
6.3	网际互联层	10
6.4	网络接口层	10
6.5	虚实电路层	10
6.5.1	虚电路交换.....	10
6.5.2	电路交换.....	10
6.5.3	分组交换.....	10
6.5.4	报文交换.....	10
6.5.5	虚实电路交换.....	10
6.5.6	虚实电路、虚电路、数据报对比.....	10
6.6	三层和四层模型的融合	12
7.	TCP/IP/M 协议簇	13
7.1	网络应用程序.....	13
7.1.1	TELNET	13
7.1.2	FTP	14
7.1.3	RSH	14
7.1.4	NFS.....	14
7.1.5	SNMP	14
7.1.6	X-WINDOW.....	14
7.2	传输层协议.....	14
7.2.1	UDP	14

7.2.2 TCP15

7.3 IP 协议15

Foreword

国际标准化组织（国际标准化组织）和 IEC（国际电工委员会）组成了世界范围的标准化专业系统。属于 ISO 或 IEC 的国家机构参加国际标准，通过技术委员会建立由有关机构，以处理特定领域的技术活动发展。ISO 和 IEC 技术委员会在共同感兴趣的领域合作。其他国际组织、政府和非政府与 ISO 和 IEC 联系，也可参加有关工作。在信息技术领域，ISO 和 IEC 已成立了一个联合技术委员会(ISO/IEC JTC1)。

国际标准是按照 ISO/ IEC 规范第 2 部分来制订的。

技术委员会的主要任务是制定国际标准。采用国际标准草案，并传阅会员国投票表决机构的技术委员会。作为国际标准发布,要求受到至少百分之七十五参加投票的成员团体的支持。

在特殊情况下，当一个技术委员会已经收集到通常作为国际标准发布的不同类型的数据（也就是说还没有达到“最佳状态”），它可以决定由参与成员以简单的多数表决方式来发表技术报告。这种技术报告完全是参考性的，发布后也不用定期进行审阅，除非它提供的数据已经被认为无效或没有用处。

可能需要注意的是，本文件的某些内容可能牵涉到专利权利。ISO 不负责识别任何这样的专利权问题。

ISO/IEC TR29181（问题陈述和需求）是由 ISO/IEC JTC1/SC6 技术委员会第七工作小组准备。

ISO/IEC TR 29181 包含以下部分

- TR 29181-1 - 第1部分: 总论
- TR 29181-2 - 第2部分: 命名寻址
- TR 29181-3 - 第3部分: 交换和路由
- TR 29181-4 - 第4部分: 移动
- TR 29181-5 - 第5部分: 安全
- TR 29181-6 - 第6部分: 媒体发布
- TR 29181-7 - 第7部分: 服务构成

Introduction

本文件是 ISO/IEC JTC1 SC6 WG7 编写的《未来网络技术报告—问题陈述和要求》的第 X 部分，主要涉及未来网络三层/四层混合架构的网络模型（TCP/IP/M）研究。TCP/IP/M 是从网络底层结构上来解决三网融合(通信网、广电网、互联网)引起的高质量的实时媒体通信问题，从而给未来网络提供一个流畅、绿色环保的网络环境。本报告的目的是讨论如何开发一个全新设计的新的网络模型（TCP/IP/M），以帮助实现未来网络的远大理想。

网络模型是通信网络的根本，所有硬件和软件的网络通信都依赖于它。网络模型的设计可以从根上改变网络结构，可以解决旧网络体系的不足，并能满足未来网络新的需求。考虑到对旧网络的兼容性，新的网络模型需要兼容现有的 TCP/IP 四层模型，同时又能提供更好的技术体系来实现未来网络。

与那种旨在逐步完善现有技术，同时保护旧网络整体结构完整性的渐进改良型网络发展路线不同，未来的网络必须产生出一个全新的网络架构。一个新的设计需要进行彻底分析、充分理解需求、精心策划和集体协调。为了达到最大的效益并找到最佳的解决方案，在开展具体技术方案的标准化工作之前，需要编制一份战略规划文件。

这份报告可以被看成是未来网络网络模型研究和标准化的战略方案。它解释未来网络网络模型研究的任务，并讨论应如何发展。

这个文件可以用来在技术发展过程中激发网络技术开发研究者对于未来网络模型研究的兴趣和创新思路，也可用其作为标准化过程中对各种命名和寻址方案评估的参考标准来使用，它也可在标准制订后实施过程中被用作对产品评估、测试的指标性参考。

未来网络：问题陈述和需求 – 第 X 部分：网络模型

1. 范围

本技术报告（缩写 TR）是关于未来网络模型的研究方案，是 ISO/IEC 未来网络标准化领域内的重要课题。它描述了未来网络模型的特征，包括概述、三/四层模型、同步时分差异、ARP、TCP/IP/M 协议簇。

2. 引用文件

2.1 规范性引用

ISO/IEC PDTR 29181-1, 未来网络：问题陈述和需求 – 第一部分：总论.

RFC 1955, 适用于下一代 IP 的互联网路由和寻址的新方案

3. 定义

本技术报告采用下列术语和定义。

3.1 网络

电路或电路的一部分。在电的系统中，由若干元件组成的用来使电信号按一定要求传输的电路或这种电路的部分，叫网络。

3.2 支路

网络的一部分,由一个或多个两端元件串联组成。

3.3 电路（CIRCUIT）

“电路”是几个组件通过导线互相连接，也可以称为“网络”。更特定地，电路是可以形成闭合回路的网络。在电路里，任意组件可以以“支路”来代表，任意两条或两条支路相交处的任何一点称为“节点”。

3.4 模拟电路

自然界产生的连续性物理自然量，将连续性物理自然量转换为连续性电信号，运算连续性电信号的电路即称为模拟电路。

3.5 数字电路

数字电路又名逻辑电路，是一种将连续性的电信号，转换为不连续性定量电信号，并运算不连续性定量电信号的电路。数字电路中，信号大小为不连续并量化的电压状态。多数采用布尔代数逻辑电路对定量后信号进行处理。典型数字电路有，振荡器、寄存器、加法器、减法等，来运算不连续性定量电信号。

3.6 虚电路

虚电路是在分组交换散列网络上的两个或多个端点站点间的链路。它为两个端点间提供临时或专用面向连接的会话。它的固有特点是，有一条通过多路径网络的预定路径。提前定义好一条路径，可以改进性能，并且消除了帧和分组对头的需求，从而增加了吞吐率。从技术上看，可以通过分组交换网络的物理路径进行改变，以避免拥挤和失效线路，但是两个端系统要保持一条连接，并根据需要改变路径描述。

3.7 永久虚电路 (PVC)

在用户主机之间建立虚拟的逻辑连接，并且保证在其上传送信包的正确性和顺序性，通信前后要进行虚电路的建立和拆除。永久虚电路是一种在网络初始化时建立的虚电路，并且该虚电路一直保持，是一种提前定义好的，基本上不需要任何建立时间的端点站点间的连接。在公共一长途电信服务，例如异步传输模式（ATM）或帧中继中，顾客提前和这些电信局签订关于 PVC 的端点合同，并且如果这些顾客需要重新配置这些 PVC 的端点时，他们就必须和电信局联系。X.25 网络和 B-ISDN 都提供 PVC 服务。

3.8 交换型虚电路 (SVC)

是端点站点之间的一种临时性连接。这些连接只持续所需的时间，并且当会话些电信局提供的分组交换服务允许用户根据自己的需要动态定义 SVC。

3.9 虚实电路

一种通过同步时分差异从统计复合时分电路分离出多个时分电路的虚电路。

4. ABBREVIATIONS

FN	未来网络 (Future Network)
IMP	互联网混合协议(Internet Mixed Protocol)
IP	互联网协议(Internet Protocol)
MAC	硬件地址 (Media Access Control)
TCP	传输控制协议 (Transmission Control Protocol)
UDP	用户数据包协议 (User Datagram Protocol)
SNMP	简单网络管理协议 (Simple Network Management Protocol)
SMTP	简单邮件传输协议 (Simple Mail Transfer Protocol)
FTP	文件传输协议 (File Transfer Protocol)
RPC	远程过程调用协议 (Remote Procedure Call Protocol)
ICMP	互联网控制报文协议 (Internet Control Message Protocol)
ARP	地址解析协议 (Address Resolution Protocol)
RARP	反向地址转换协议 (Reverse Address Resolution Protocol)
FDDI	光纤分布式数据接口 (Fiber Distributed Data Interface)

ATM 异步传输模式 (Asynchronous Transfer Mode)

TCP/IP 传输控制协议/因特网互联协议, 又称网络通讯协议 (Transmission Control Protocol/Internet Protocol))

OSI 开放式系统互联 (Open System Interconnect)

QoS 服务质量 (Quality Of Service)

M 混合协议(Mixed Protocol)

5. 概述

TCP/IP/M 是未来网络协议, 其保留了现有互联网的四层网络模型, 同时又推出了新的三层网络模型。在未来网络里, TCP/IP/M 可以利用四层模型实现常规数据的快速传输, 也可以利用三层模型实现语音、视频、广播等的流畅通信, 达到满足不同的需要的目的。

四层的分组交换的协议结构和现有 TCP/IP 相似, 但基于三层的虚拟电路交换和 TCP/IP 有本质不同。三层结构先建立复合时分虚电路, 通过同步时间差异, 为不同的时分电路预留好预定的带宽, 同时所传的数据为基于面向连接的、可靠的、可实现不间断信息流。数据传输不是现有的分组交换, 实质为电路传输, 故称之为虚实电路。它是一个根据实际需求预建的传输电路通道, 带宽独占且是固定的, 根本不需要现有体系中的 QoS 来保证数据流的性能达到一定的水准。如在线观看电影, 可根据预定的带宽, 无缝传输, 省略了将电影分解成多个小包、组装数据、达到一定缓冲再播放的过程。三层/四层混合网络架构的设计彻底解决了语音质量、传输内容与带宽分配、分组传输与电路传输、路由选择的互容问题。其具体工作原理见下图。



图 1. TCP/IP/M 工作原理图

5.1 基本结构

目前的通讯方式主要是电路传输和分组传输构成，而分组传输主要是 TCP/IP 协议，简称 IP 协议。未来网络将会涉及数据传输和视频广播及语音通话的融合需求这个问题，为此我们认为未来网络用电路的三层架构的虚实电路与 IP 传输的四层架构的融合体系能解决此问题，即三层架构用于视频广播及语音通话，而 IP 数据传输仍用四层架构，详见下图：

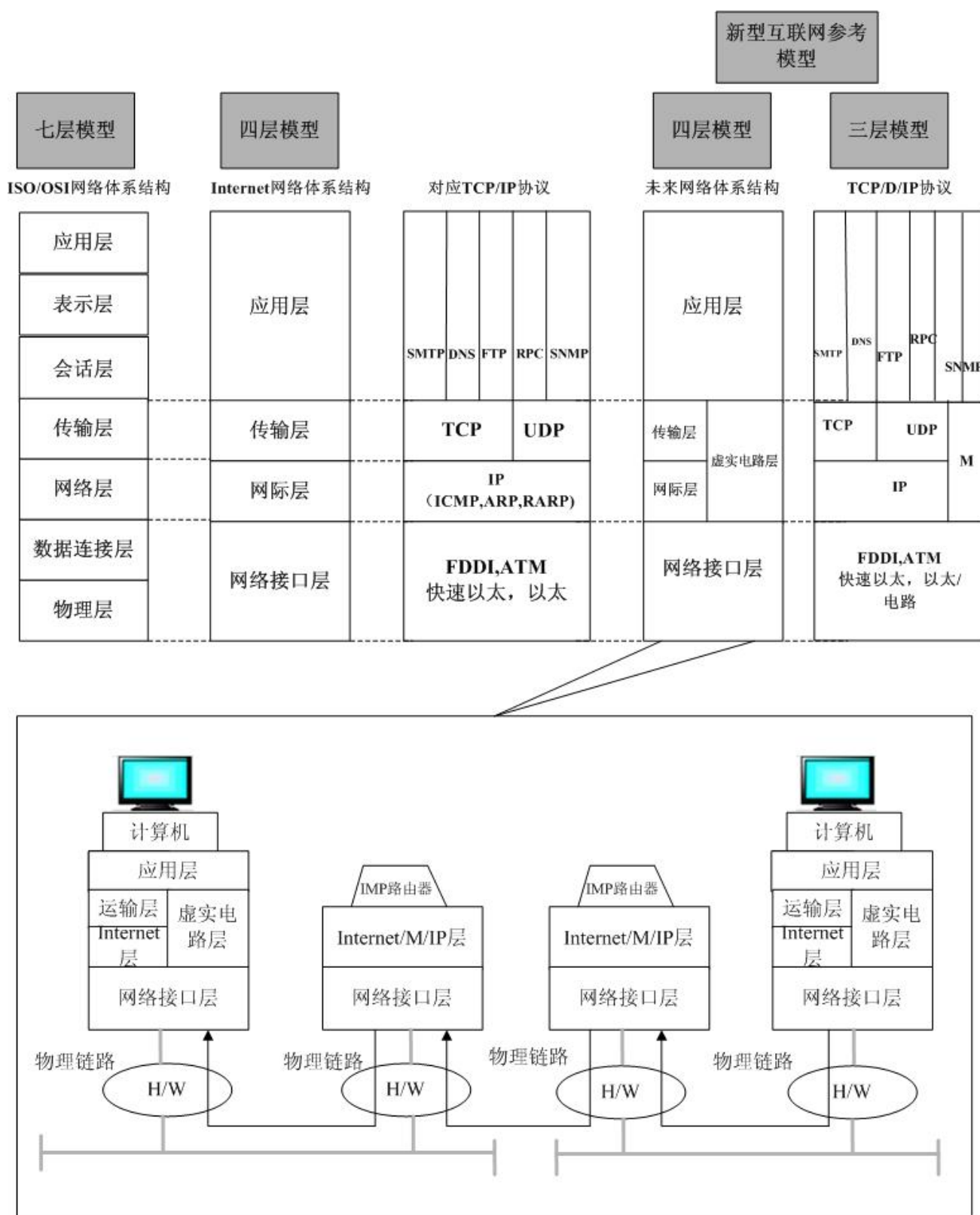


图 2. 未来网络模型

未来网络模型特征如下：

- 1) 独特的双模网络层次结构，包括现有的互联网四层模型和新的三层模型。
- 2) 将互联网传输的时分传输通过时间节点的错开形成复合时分差异，从而形成复合时分中的多通道虚电路，解决了在同一统计时分电路的多层传输难题,可以在同一链路中按需分别传输分组包和电路数据。
- 3) 虚实电路概念的提出，完全解决了现有高质量的视频、语音通信的难题。
- 4) 向下兼容，新网络和旧网络能互联互通，能从旧网络平稳过渡到新网络。
- 5) 复合时分差异电路传输结合了电路和分组交换的优点，从而适应未来网络高速、稳定节能、环保的需求。
- 6) 高效、快速、稳定、节能、环保的网络将加快人与人之间、信息与信息之间的交流，将促进经济、科技乃至军事的发展。

5.2 数据传送

5.2.1 TCP/IP 数据传送

让我们看看当它经过如图所示的 TCP 协议堆栈流动时的情况，当应用程序使用 TCP(传输控制协议)，数据在应用程序与 TCP 模块之间传递。当应用程序使用 UDP(用户数据报协议)，数据在应用程序与 UDP 模块之间传递。FTP(文件传输协议)是使用 TCP 包的典型应用，在这个例子的协议堆栈是 FTP/TCP/IP/ENET。SNMP(简单网络管理协议)是使用 UDP 的应用，在这个例子的协议堆栈是 SNMP/UDP/IP/ENET。

TCP 模块，UDP 模块，和以太网驱动程序是 n-to-1multiplexers（多路复用器）。作为 multiplexers 它们复用许多输入到一个输出。它们也是 1-to-n-de-multiplexers(分路器)，作为分路器，它们通过协议头从一个输入产生许多输出。

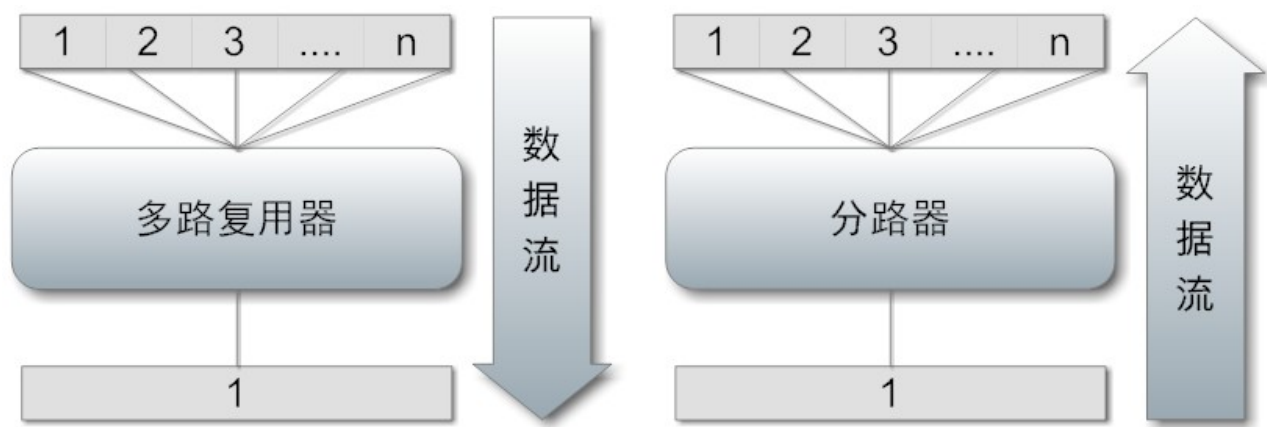


图 3. n-1 多路复用器和 1-n 分路器

如果以太网帧离开网卡进入以太网驱动程序,包能向上传递给 ARP(地址解析协议)模块或到 IP（网间协议）模块。以太网帧的字段属性决定了以太网帧是否被传递给 ARP 或 IP 模块。

如果 IP 报文进入 IP 包，它被向上传递到 TCP 或 UDP，由在 IP 头的字段属性决定。如果 UDP 报文进入 UDP，应用消息向上传递给网络应用程序，这是由 UDP 头的端口值决定的。如果 TCP 消息进入 TCP，应用消息向上传递给网络应用程序，这是由 TCP 头的端口值决定的。

向下复用很容易实现因为从每一个开始点只有一条向下的路径；每个协议模块增加它的头信息从而使包能在目的计算机上被分开。

从应用程序出来的数据通过 TCP 或者 UDP 复合到 IP 模块，然后被送到更低层。

尽管因特网技术支持许多网络媒介，在我们这里使用的例子都是以太网来讨论的因为以太网是基于 IP 的最常见物理网络。在图 1 中的计算机有唯一的以太网连接。物理地址对每一个在以太网的接口是唯一的，它们被存储在以太网驱动程序的底层接口中。

5.2.2 M 数据传送

新的混合协议（M）无需 TCP/IP 协议，它的数据传送是一条通道，没有复用和分路的概念。M 和 TCP/IP 数据通过 IP 报头信息区分，分送到不同的协议进行处理。

5.3 网络接口

如果一台计算机和 2 个独立的以太网连接，它的样子如图 4 所示。

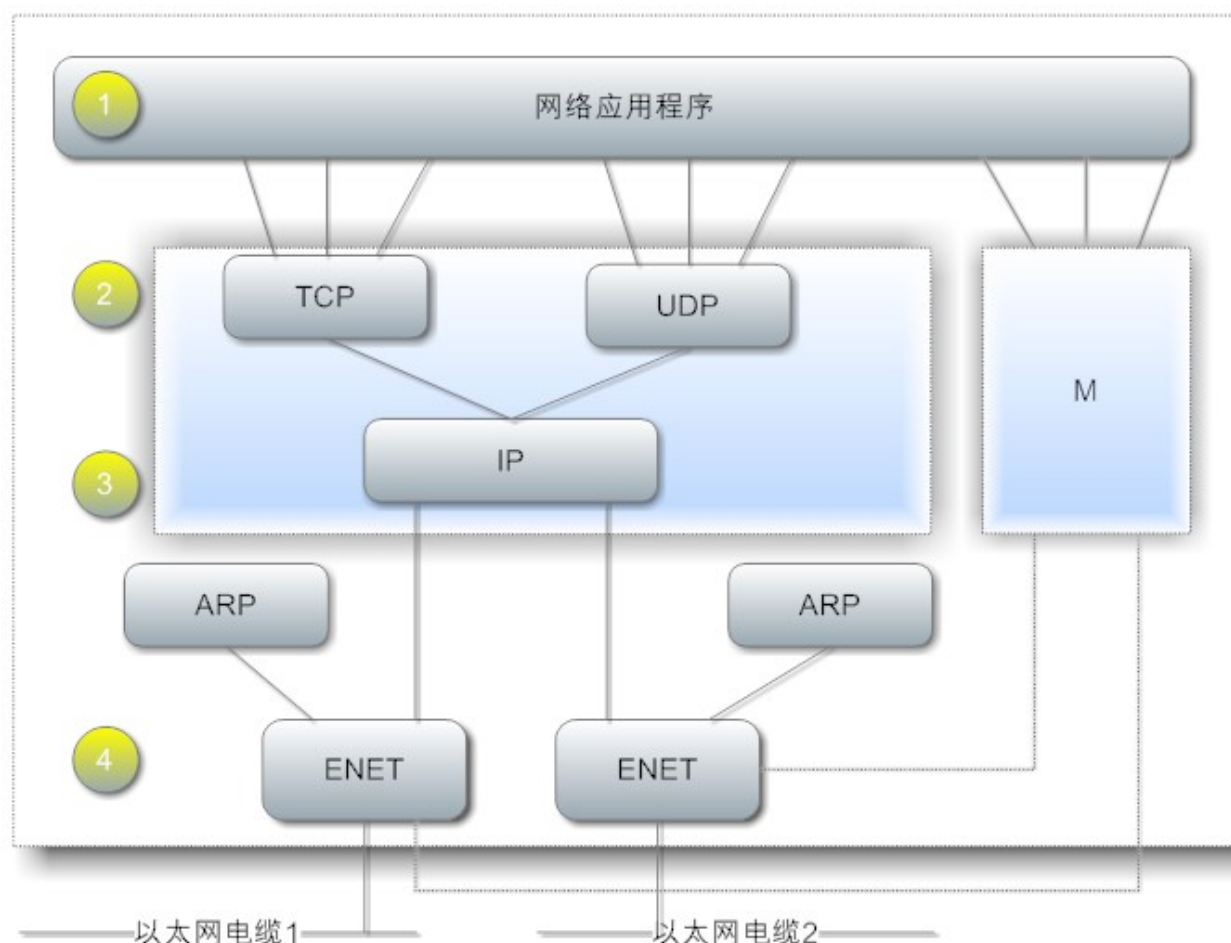


图 4. TCP/IP/M 两个网络接口

请注意这台计算机有两个物理地址和两个 IP 地址。从图 4 的结构中我们发现计算机有多于一个的物理网络接口，那么 TCP 模型中的 IP 模块就是 n_to_m 复用器和 m_to_n 分路器的结合。

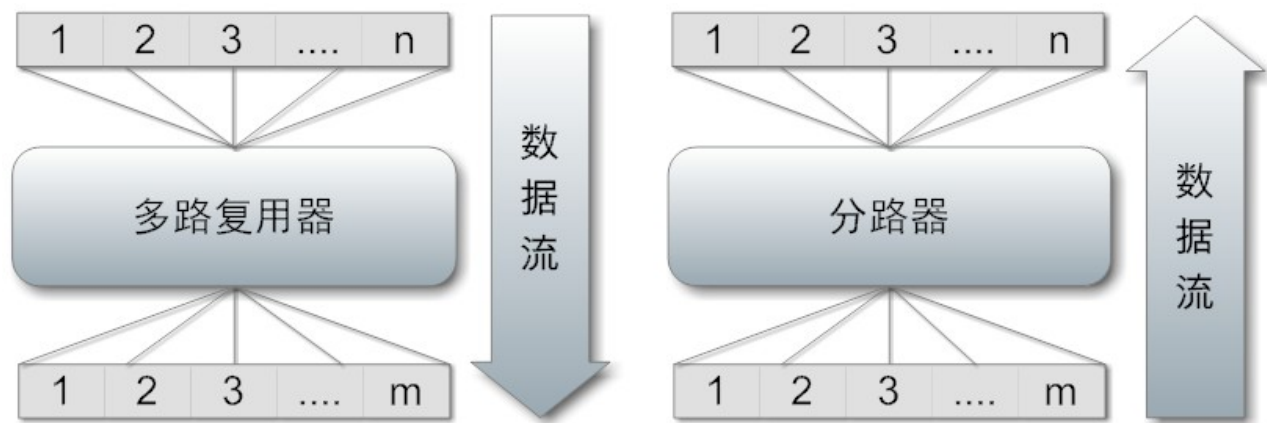


图 5. TCP n-m 多路复用器和 m-n 分路器

从上面我们可以看出这种多路技术可以从任意的方向接收和发送数据，有一个以上网络接口的 IP 模块比我们最初的把数据从一个网络送到另一个网络的例子要复杂的多，数据可以从各个网络接口传过来也可被送向网络。

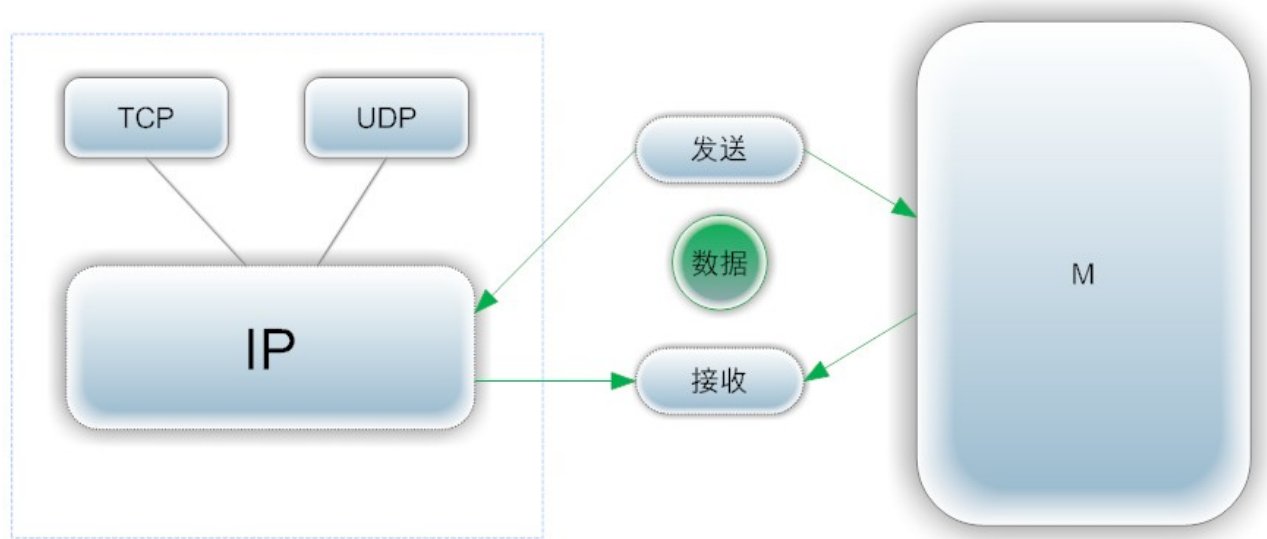


图 6. IP 包发送和接收的例子

对于四层模型，发送 IP 包到另一个网络的过程叫做传递 IP 包，一台专门用来传递 IP 包的计算机叫做“路由器”。就如你从图中看到的一样，在路由器上传递的 IP 包不涉及 TCP 和 UDP 模块，一些路由器执行时更本就不要 TCP 或 UDP 模块。

对于三层模型而言，任何计算机之间只传输 M 包。

5.4 IP 产生专一的逻辑网络

IP 模块是 interner 技术成功的中心所在，当消息向下通过协议栈时每一个模块或驱动程序加上自己的头到消息中去。每一个模块或驱动程序从消息中去掉相应的头当消息沿协议栈向上传时，IP 头包含了用以从许多物理网络中区分唯一的逻辑网络的 IP 地址，互相连接的物理网络是 internet 的组成部分，这些互相联络的物理网络就构成了 internet。

5.5 独立的物理网络

网络应用程序通过 IP 隐藏了底层的网络硬件，如果新的物理网络投入使用，只要用新的驱动程序。因此，尽管硬件技术在变化，网络应用程序仍旧是完美无缺以及不易受攻击的。

5.6 互相通信

如果两台计算机能够互相通信，就称它们是“能互操作的”。如果 internet 畅通，就称它们有“互用性”。有许多计算机的用户从互联网获益，因为在市场上计算机的互操作性。一般地，当你买了一台计算机，它就能互相通信。如果计算机没有互操作性，那这样的计算机就没有市场竞争力。

6. TCP/IP/M 三/四层模型

TCP/IP 协议栈是美国国防部高级研究计划局计算机网（Advanced Research Projects Agency Network, ARPANET）和其后继因特网使用的参考模型。ARPANET 是由美国国防部（U.S. Department of Defense, DoD）赞助的研究网络。最初，它只连接了美国境内的四所大学。随后的几年中，它通过租用的电话线连接了数百所大学和政府部门。最终 ARPANET 发展成为全球规模最大的互连网络-因特网。最初的 ARPANET 于 1990 年永久性地关闭。



图 7. TCP/IP/M 参考模型

如上图所示，TCP/IP 参考模型分为四个层次：应用层、传输层、网际互联层和网络接口层。新的 TCP/IP/M 参考模型分为三个层次：应用层、虚实电路层和网络接口层，将传输层、网际互联层合并为虚实电路层。这是 TCP/IP 协议的分层结构在互连网计算机上的表示，用互连网技术互相通信的每台计算机有这样的分层结构。这样的分层结构决定了计算机在 internet 上互相通信的方式。数据通过这样的分层结构从上层传到底层，然后通过网线把数据传送出去。

下面，分别介绍各层的主要功能。

6.1 应用层

应用层对应于 OSI 参考模型的高层，为用户提供所需要的各种服务，例如：FTP、Telnet、DNS、SMTP 等。

6.2 传输层

传输层对应于 OSI 参考模型的传输层，为应用层实体提供端到端的通信功能。该层定义了两个主要的协议：传输控制协议（TCP）和用户数据报协议（UDP）

TCP 协议提供的是一种可靠的、面向连接的数据传输服务；而 UDP 协议供的是不可靠的、无连接的数据传输服务。

6.3 网际互联层

网际互联层对应于 OSI 参考模型的网络层，主要解决主机到主机的通信问题。该层有四个主要协议：网际协议（IP）、地址解析协议（ARP）、反向地址解析协议（RARP）和互联网控制报文协议（ICMP）。

IP 协议是网际互联层最重要的协议，它提供的是一个不可靠、无连接的数据报传递服务。

6.4 网络接口层

网络接口层与 OSI 参考模型中的物理层和数据链路层相对应。事实上，TCP/IP 本身并未定义该层的协议，而由参与互连的各网络使用自己的物理层和数据链路层协议，然后与 TCP/IP 的网络接口层进行连接。

6.5 虚实电路层

6.5.1 虚电路交换

交换虚电路(SVC)是一类需要时动态设置的虚拟电路。一个 SVC 概念上能够换算到一个拨号连接。

虚电路与实电路对应，是指数据传输时对于物理链路的占用分配方式；而分组交换则是指数据传输的格式问题。

6.5.2 电路交换

- 1) 采用的是静态分配策略，经面向连接建立连接。
- 2) 通信双方建立的通路中出现故障，可重新拨号建立连接。
- 3) 线路的传输有固定的带宽，可摆脱不稳定的数据传输。
- 4) 不需要拆包、组包，传输效率更高。

目前的虚电路传输的仍是一种分组数据，而非真正的电路交换。

6.5.3 分组交换

采用存储转发技术，基于标记，在传输数据之前可不必先建立一条连接，因此，分组交换具有高效、灵活、可靠、迅速等优点，但传输时延较电路交换较大，不实用于实时数据的传输。

6.5.4 报文交换

对应无连接的服务，传递数据前不需要建立连接，每个分组选择的路线都可能不同，适合传递少量数据，可靠性不高。

6.5.5 虚实电路交换

结合虚电路和实电路的优点，三层结构先建立复合时分虚电路，为不同的时分电路预留好预定的带宽，

通过复合时分方法，实现采用虚电路建立连接，用实电路的方法传输数据。

6.5.6 虚实电路、虚电路、数据报对比

对比的方面	虚实电路	虚电路	数据报
传输方法	复合时分电路	统计时分电路	统计时分电路
传输内容	电路数据、分组包数据	分组包数据	分组包数据
同步时间	不同步	同步	同步
连接的建立	必须有	必须有	不要
备份虚电路	必须有	没有	没有
目的站地址	仅在连接建立阶段使用，一个电路数据使用一个短的虚电路号	仅在连接建立阶段使用，每个分组使用短的虚电路号	每个分组都有目的站的全地址
路由带宽选择	一旦带宽确定一般不更改，	只在第一路由器选定，其它路由器无法确定第一路由器选定的带宽	只能根据路由带宽决定
拥塞控制	没有一路畅通	有但效果不好	随机堵塞
网络结点存储空间	没有，不需要	8 字节	15 字节
传输质量	好	一般	差
传输成本	低	一般	高
传输效率	高	不高	差
路由选择	在虚实电路连接建立时进行，所有电路数据均按同一路由同一带宽传输，并有一个以上备份路由，	在虚电路连接建立时进行，所有分组均按同一路由发送，但会有堵塞现象不能保证传输质量	每个分组独立选择路由
当路由器出故障	所有通过了出故障的路由器的虚电路均可由备份路由进行工作	所有通过了出故障的路由器的虚电路均不能工作	出故障的路由器可能会丢失分组，一些路由可能会发生变化
分组的顺序	虚实电路没有分组包传输可，但分组数据可以按虚电路 QOS 和数据报定义发送	总是按发送顺序到达目的站	到达目的站时可能不按发送顺序
电路数据传输	可以传输电路的数据，但第一路由的路由器出故障，第二路	没有电路数据传输体系	没有电路数据传输体系

	由会将电路数据按发送顺序到达目的站，但分组数据可以按虚电路 QOS 和数据报定义发送		
端到端的差错处理	由通信网负责，但 QOS 数据报由通信子网负责，数据报由主机负责	由通信子网负责	由主机负责
端到端的流量控制	由通信网负责，但 QOS 数据报由通信子网负责，数据报由主机负责	由通信子网负责	由主机负责

6.6 三层和四层模型的融合

考虑到未来网络的平稳过渡和对旧网络的兼容，我们的未来网络模型同时包括了四层模型和三层模型。那么，三层模型和四层模型的共存和触发机制就是个值得探讨的问题，我们需要让应用程序通知网络框架什么时候传输 IP 包和什么时候传输 M 数据，同时网络框架可以分别使用不同的网络模型来处理 and 传输数据。完美的场景就是一台计算机和另一台计算机之间的通信可以是 IP 包和 M 数据，见下图。

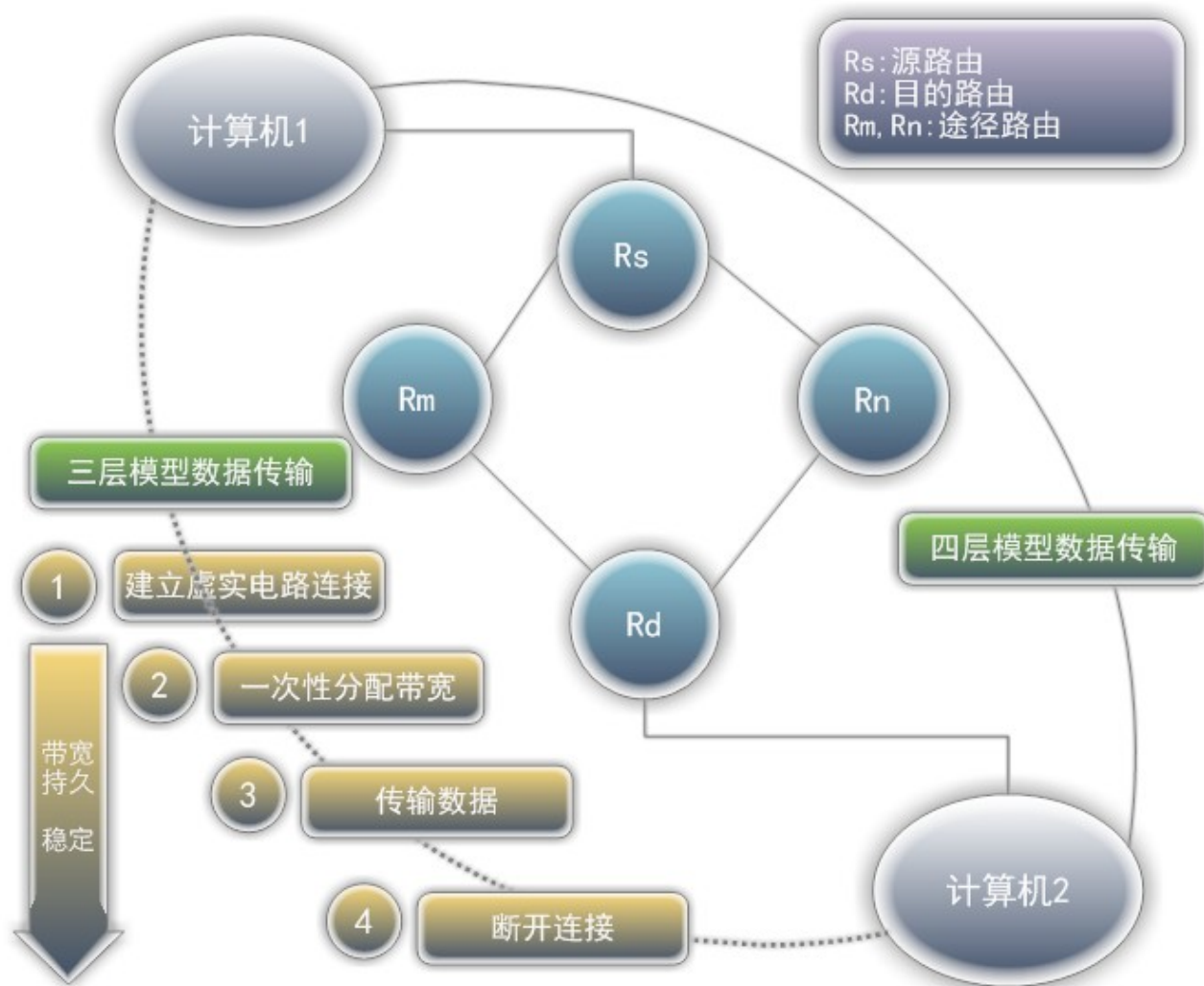


图 8. 三/四层模型的融合

7. TCP/IP/M 协议簇

7.1 网络应用程序

TCP 和 UDP 提供不同的服务，大部分的应用程序只用其中的一个。你当然选择最适合你的协议了。如果你需要可靠的数据传送，TCP 可能是最佳的选择了。如果你需要数据包服务，UDP 是最佳的。如果你需要高效的长的通路，TCP 可能是最佳的。如果你需要快的网络反应时间，UDP 可能是最佳的。如果你不想把分类，则"最佳"的选择就是模糊的。然而，应用程序能够弥补选择上的不足。举个例子，如果你选择 UDP 可是你需要可靠性，则在应用程序上加上可靠性就可以了。如果你选择 TCP 可是你需要标记的服务，那么应用程序必须在字节流里加入标记。

网络应用程序的数目是连续增长的。自从 internet 技术开始就有一些应用程序了：TELNET 和 FTP。其他的是较新的：X-WINDOWS 和 SNMP。下面是一些网络应用程序的简短说明。

7.1.1 TELNET

TELNET 提供远端登录服务。它的操作和外型同通过电话拨号是相似的。在命令行上用户敲入："TELNET DELTA"就会收到从"delta"来的登录提示。

TELNET 很好使，它使一个老的应用程序有着广泛的应用。TELNET 通常在不同的操作系统间工作。举个例子，TELNET 客户端使 VAX/VMS 系统，而服务器端是 UNIX 系统。

7.1.2 FTP

文件传送协议(FTP)，是和 TELNET 一样久远的网络应用程序，也有广泛的应用。从操作来看好象你登录上了远端的计算机。但是你必须用特殊的命令取代你习惯上的命令。FTP 命令允许你在两台计算机之间拷贝文件。

7.1.3 RSH

远端 shell (rsh 或者 remsh) 是全部远端 UNIX 类型命令族中的一个。UNIX 拷贝命令--CP，变成了 RCP。UNIX 命令"谁在登陆"WHO 变成了 RWHO。这个系列都被变成了"R"系列命令。

R*命令主要工作在 UNIX 系统和被设计成在互相信任主机间相互操作。安全性很少被考虑，但是他们提供了方便的用户环境。

为了在一台远端叫做 delta 的计算机上执行命令"cc file.c"，输入"rsh delta cc file.c"，为了拷贝文件到 delta 上，输入"rcp file.c delta"。为了登陆到 delta 上，输入"rlogin delta"，如果你在某种特定的方式上管理这台计算机，你将不会有登陆提示。

7.1.4 NFS

网络文件系统，是由美国 SUN 公司开发的，使用 UDP，它在不同的计算机之间上载 UNIX 系统文件是很出色的。一个无磁盘的工作站通过存取服务器的硬盘就好像磁盘是本地的。在主机"alpha"上的单一的数据库同样能被主机"beta"使用只要数据库文件用 NFS 上载在"beta"上。

NFS 加入大量的信息给网络，从而使连接的速度很慢，但是它的优点使主要的。NFS 客户端在内核执行，允许使用 NFS 的磁盘好象在本地一样。

7.1.5 SNMP

简单网络管理协议 (SNMP) 使用 UDP，被设计成由中心网络点来管理。它很知名，如果给它足够的数据，网络管理员就能够发觉和诊断网络问题。中心点用 SNMP 从网上的计算机收集数据。SNMP 定义了这种数据的格式。由中心点或网络管理员来解释这种数据。

7.1.6 X-WINDOW

X-WINDOW 系统使用基于 TCP 的 X WINDOW 协议来画在工作站上的位图显示窗口。X WINDOW 不仅仅是画窗口，它能够用来设计用户界面。

7.2 传输层协议

7.2.1 UDP

UDP 是在 IP 层之上的两个主要协议之一。它提供用户网络应用程序的服务，用到 UDP 的网络应用程序有：NFS (网络文件系统) 和 SNMP (简单网络管理协议)。UDP 服务只是在 IP 的基础上加了少许一点。

UDP 是无连接的数据报服务没有丢失检测。UDP 不和远端的 UDP 模块保持点到点连接，它仅仅把包发送出去而不管有没有丢失和接收来自外面的数据包。

UDP 在 IP 的基础上加了两个属性，一个是端口号，另一个是检查数据完整性的校验和。

7.2.1.1 端口

UDP 和应用程序之间的通信路径是通过 UDP 端口。这些端口是用数字表示的，从 0 开始。提供服务的应用程序用特定的端口号等待消息的进入。服务器不间断扫描客户端的请求服务。

举个例子，比如说 SNMP，叫做一个 SNMP 代理，总是在端口 161 上等待消息。每台计算机只能有一个 SNMP 代理因为只有一个 UDP 端口号 161。这个端口号是众人皆知的，它是固定的，是网络分配的唯一的号。如果 SNMP 客户请求服务，则它就发送 UDP 包到目的计算机的端口 161。

当应用程序发送 UDP 包，则远端收到的是一个单元。比如，如果应用程序发了 5 个 UDP 包，则远端就读取 5 次。当然，发的 5 个包和读取的 5 个包大小相等。

UDP 保存每一个完整的包，它不把两个应用程序消息加在一块，也不把一个包拆成两个。

7.2.1.2 校验和

在 IP 头域里显示“UDP”的 IP 包被送到 UDP 模块。当 UDP 模块收到 UDP 包它检查它的校验和。如果它的校验和为 0，则意味着在发送端校验和没有被计算，可以忽略。因此发送端的计算机的 UDP 模块产不产生校验和没关系。如果物理帧在一个网络的 2 个 UDP 模块间通信，则不需要产生校验和。但是我们推荐使用校验和因为路由表的改变可能导致数据通过不可靠的媒介。

如果校验和是正确的或为 0，目的端口就会检查它。UDP 包传向这个端口，排队等待应用程序处理它，否则 UDP 包就会被丢弃。如果 UDP 包到达的速度比应用程序能够处理的速度快或者等待的 UDP 包把队列排满，UDP 包就会被 UDP 模块丢弃。UDP 模块会一直丢弃 UDP 包直到队列有多余的空间。

7.2.2 TCP

TCP 提供和 UDP 不同的服务，TCP 提供有连接的比特流，不同于无连接的数据报服务。TCP 保证可靠传输，而 UDP 不保证。

TCP 被网络应用程序调用时保证可靠传输和不能有超时和误传。两个典型的网络应用程序是 FTP (文件传送协议) 和 TELNET。其他的流行的 TCP 网络应用程序包括：X-WINDOW 系统，rcp (远程复制)，和 rseries commands。TCP 的强大功能时要代价的：它需要更多的 CPU 和网络带宽。TCP 模块的内部比 UDP 模块要复杂的多。

和 UDP 相似，网络应用程序和 TCP 端口连接。特定的端口号对应特定的应用程序。举个例子 TELNET 服务器使用端口 23。TELNET 客户端只能通过连接特定计算机上的端口 23 才能成功。

当应用程序通过 TCP 启动，在客户端的 TCP 模块和在服务器端的 TCP 模块互相通信，这些两个端点的 TCP 模块构成了虚拟的电路。这个虚拟电路消耗两端的资源。虚拟电路是双向的，数据能够同时往两个方向传。应用程序把数据写到 TCP 端口，数据通过网络由远端的应用程序控制。

TCP 包可以分成任意大小，包与包之间没有界限。比如，如果应用程序往 TCP 端口发了 5 次，远端的应用程序也许要读 10 次，或者它就只读一次。在一端写的次数和大小与另一端读的次数和大小时没有关联的。

TCP 是一个有超时和重发的滑动窗口协议。发出去的包必须得到远端的确认。确认信息可以携带在数据包上。两个接受端能控制远端，从而防止缓冲器溢出。

对于所有的滑动窗口协议，有一个窗口的大小，窗口的大小决定了在收到确认信息以前可以发送的总的数量。对于 TCP，这个数量不是 TCP 段的数量而是字节的数量。

7.3 IP 协议

IP 网际协议 IP 是 TCP/IP 的心脏，也是网络层中最重要的协议。

IP 层接收由更低层（网络接口层例如以太网设备驱动程序）发来的数据包，并把该数据包发送到更高层---TCP 或 UDP 层；相反，IP 层也把从 TCP 或 UDP 层接收来的数据包传送到更低层。IP 数据包是不可靠的，因为 IP 并没有做任何事情来确认数据包是按顺序发送的或者没有被破坏。IP 数据包中含有发送它的主机的地址（源地址）和接收它的主机的地址（目的地址）。

高层的 TCP 和 UDP 服务在接收数据包时，通常假设包中的源地址是有效的。也可以这样说，IP 地址形成了许多服务的认证基础，这些服务相信数据包是从一个有效的主机发送来的。IP 确认包含一个选项，叫作 **IP-source-routing**，可以用来指定一条源地址和目的地址之间的直接路径。对于一些 TCP 和 UDP 的服务来说，使用了该选项的 IP 包好像是从路径上的最后一个系统传递过来的，而不是来自于它的真实地点。这个选项是为了测试而存在的，说明了它可以被用来欺骗系统来进行平常是被禁止的连接。那么，许多依靠 IP 源地址做确认的服务将产生问题并且会被非法入侵。