

1
面向 21 世纪的 IP 网络技术经典著作

路由器原理与技术

Luyouqi Yuanli
Yu Jishu

张宏科 张思东 刘文红 编著

国防工业出版社

TN915.05

15

路由器原理与技术

张宏科 张思东 刘文红 编著

国防工业出版社

·北京·

面向 21 世纪的 IP 网络技术经典著作

路由器原理与技术

Luyouqi Yuanli
Yu Jishu

张宏科 张思东 刘文红 编著

国防工业出版社

内 容 简 介

本书对路由器原理与技术作了全面的叙述,在介绍 IP 网络基本概念、特点、组成、结构及国内外发展动态与趋势等问题的基础上,主要叙述了路由器原理、实现技术、路由协议及很有实用价值的 IP 网络设计与工程应用等。

全书取材新颖、内容丰富、实用性强,反映了国内外路由器技术的现状与未来,适合于从事通信、计算机技术开发与研究的广大工程技术人员阅读,也可供大专院校通信、计算机等专业的师生和相关培训班作为教材或教学参考书。

图书在版编目(CIP)数据

路由器原理与技术 / 张宏科等编著. —北京:国防工业出版社,2003.1

ISBN 7-118-03044-9

I. 路... II. 张... III. 计算机网络—路由选择
IV. TN915.05

中国版本图书馆 CIP 数据核字(2002)第 105351 号

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

腾飞胶印厂印刷

新华书店经售

*

开本 787×960 1/16 印张 19 $\frac{3}{4}$ 374 千字

2003 年 1 月第 1 版 2003 年 1 月北京第 1 次印刷

印数:1—4000 册 定价:39.00 元

(本书如有印装错误,我社负责调换)

前 言

随着科学技术的发展,信息已成为推动社会向前发展的巨大资源。在 21 世纪,信息领域的竞争将是世界经济竞争的焦点,而信息领域的竞争不仅取决于信息技术的掌握,更取决于信息网络的建设及应用水平,特别是路由器技术作为 21 世纪网络的主要技术之一,将发挥着越来越重要作用。掌握路由器技术不仅仅是能够使用 IP 网络,更主要的是还需要掌握 IP 网络的有关技术,了解新一代 IP 网络的发展趋势。

目前,尽管 IP 网络的建设与应用比较广泛,但多数仅局限于使用或运用,至于 IP 网络究竟是怎样工作的,采用了哪些主要和关键技术,尤其是对于 IP 网络的核心部件路由器技术与协议的掌握方面,了解和精通者不多,广大读者急需这方面的一些书籍,以便系统、全面地掌握这些知识。

为了推动国内 IP 网络技术的发展,跟踪世界新一代网络先进技术,满足广大工程技术人员、科学研究人员的需要,我们在多年学习、研究与工作实践的基础上再版了这本书。重点是针对新一代 IP 网络的主要和关键技术——路由器技术,把它尽量比较全面、系统地介绍给广大读者。本书主要内容包括:第 1 章为了使读者对 IP 网络有个基本了解,首先在叙述 IP 网络基本概念、特点、组成和结构的基础上,重点对路由器技术的国内外发展动态与趋势等问题作了介绍;第 2 章主要介绍 IP 网络的核心部件路由器的原理与技术,包括路由器原理、功能、组成和最新的实现等问题,从而使读者对路由器技术有个深入的了解;第 3 章重点叙述各种 IP 路由协议,这包括 RIP、OSPF、BGP 以及 IPv4 和 IPv6 等协议,从而使读者获得较全面的 IP 路由技术知识(如设计、使用与工程实现等技术);第 4 章主要介绍 IP 网络的设计与工程应用,使读者从中获得比较全面的系统知识,并能指导 IP 网络的研究工作与网络工程的设计等;第 5 章展望了 IP 网络的发展趋势。

在本书的编写与再版过程中,得到了北方交通大学等单位的支持和帮助,以及“863”重大专项项目“高性能 IPv6 路由器协议栈研究与开发”、国家自然科学基金项目“基于 IPv6 动态服务质量理论与应用研究”等项目的资助,在此谨向他们表示谢意。此外,还得到北方交通大学简水生院士、谈振辉教授、宁滨教授、冯玉珉教授、罗四维教授、阮秋琦教授、韩臻副教授、张有根副教授、王江林老师和总参第 54 研究所龚碧秀工程师以及博士、硕士研究生姜圳、彭雪海、胡九川、周春月、沈波、卢

IV

小青、赵耀峰、刘春宁、李洪杰、刘晓璇、商超和宁科等的大力帮助和支持,特别是在这次再版过程中,刘颖、苏伟、张春青、宗瑞锐、郑祖周、郜帅、潘冬辉、宋育芳和李玉盼等作了很多具体增补和修改工作,在此谨向他们致以衷心的感谢。

由于计算机网络和通信技术发展极为迅速,路由器技术仍在发展与完善之中,加之编写时间有限,书中难免有些不妥之处,敬请广大读者指正。

目 录

第 1 章 概述	1
1.1 引言	1
1.2 IP 网络的基本概念	1
1.3 IP 网络的组成与结构	3
1.4 IP 网络的包格式	6
1.4.1 IPv4 包格式	6
1.4.2 IPv6 包格式	8
1.5 IP 路由器技术发展趋势	10
1.5.1 硬件体系结构的发展过程.....	11
1.5.2 软件体系结构的发展过程.....	15
1.6 IP 网络技术的现状与未来	20
第 2 章 路由器原理与技术	24
2.1 引言.....	24
2.2 路由器工作原理.....	25
2.3 IP 包转发的物理过程	28
2.4 路由选择策略.....	32
2.4.1 静态路由选择策略.....	33
2.4.2 距离矢量路由选择协议.....	36
2.4.3 链路状态路由选择协议.....	38
2.4.4 混合路由选择协议.....	40
2.4.5 路由协议的选择.....	40
2.5 基于 IPv6 实现的路由技术	46
2.5.1 IPv6 路由术语介绍	47
2.5.2 IPv6 的路由原理	48
2.5.3 基于 IPv6 的 LAN 技术	53
2.6 路由器的设计与实现.....	54
2.6.1 路由器基本组成与类型.....	54
2.6.2 基于 ASIC 芯片设计路由器的举例	56

2.6.3	基于网络处理器设计的路由器举例	58
第3章	IP路由技术协议	70
3.1	引言	70
3.2	路由信息协议	71
3.2.1	RIP概述	71
3.2.2	RIP基本工作原理	73
3.2.3	RIP路由表的建立过程	82
3.2.4	RIP路由表的动态维护	82
3.2.5	RIP定时器	90
3.2.6	RIPv1分组格式	90
3.2.7	基于UDP数据包格式的RIPv1	93
3.2.8	RIPv2分组格式	97
3.2.9	基于UDP数据包格式的RIPv2	105
3.2.10	RIPng分组格式	107
3.2.11	RIPng与RIPv2的比较	109
3.2.12	RIP的实现	109
3.3	开放最短路径优先协议	115
3.3.1	OSPF概述	116
3.3.2	OSPF基本工作原理	117
3.3.3	OSPF的特点	124
3.3.4	OSPFv2分组结构	125
3.3.5	OSPFv3分组结构	140
3.3.6	OSPFv3与OSPFv2的比较	152
3.3.7	接口有限状态机	153
3.3.8	邻居有限状态机	157
3.3.9	邻居交换过程	161
3.3.10	OSPF的实现	161
3.3.11	OSPF路由度量的计算	167
3.3.12	OSPF的最短路径树	169
3.3.13	OSPF小结	172
3.4	边界网关协议	173
3.4.1	BGP概述	173
3.4.2	BGP的工作原理	174
3.4.3	BGP消息的类型	175
3.4.4	BGP内部数据库和表	178

第 4 章 IP 网络设计与应用	182
4.1 引言	182
4.2 IP 网络的编址与路由	183
4.2.1 有类别 IP 编址与路由	183
4.2.2 可变长子网与路由	189
4.3 基于 LAN 和 WAN 的 IP 网络	193
4.3.1 IP 网络概述	193
4.3.2 基于 LAN(或 WAN)IP 网的基本原理	194
4.4 IP 网络工程举例	200
4.4.1 IP 网络关键设备研究开发现状	200
4.4.2 IP 网络工程编址设计	200
4.4.3 设计 VLSM 举例	202
4.5 典型应用及 IPv6 下的实现	208
4.5.1 概述	208
4.5.2 传输层协议简介	211
4.5.3 套接口编程简介	216
4.5.4 FTP 分析及在 IPv6 下的实现	230
4.5.5 Telnet 分析及在 IPv6 下的实现	235
4.6 SNMP 分析及其在 IPv6 下的实现	238
4.6.1 概述	238
4.6.2 SNMP 协议基础及 SNMPv1、v2 简介	239
4.6.3 SNMPv3 的体系结构	243
4.6.4 SNMPv3 的消息格式	245
4.6.5 基于 IPv6 的 SNMPv3 实现	246
4.7 IPSec 协议分析与实现	248
4.7.1 概述	248
4.7.2 IPSec 协议	250
4.7.3 IPSec 的实现	254
4.7.4 结论	258
4.8 监听技术的分析与实现	259
4.8.1 概念	259
4.8.2 监听与过滤技术的基本原理	260
4.8.3 监听与过滤技术的实现	261
4.8.4 网络监听与分析举例	264

第 5 章 IP 网络技术的发展趋势	266
5.1 引言	266
5.2 移动 IP 技术	266
5.2.1 移动 IP 概述	266
5.2.2 移动 IP 的工作原理	269
5.2.3 移动 IP 的发展趋势	275
5.3 安全 IP 技术	276
5.3.1 安全 IP 概述	276
5.3.2 面向 IPv6 的安全技术	278
5.3.3 高速智能的网络安全实现方案	279
5.4 多媒体 IP 网络技术	282
5.4.1 多媒体 IP 网络概述	282
5.4.2 多媒体 IP 网络技术	283
5.4.3 流媒体技术展望	288
5.4.4 多媒体 IP 网络发展趋势	290
附录 A 常用互联网 RFC 协议标准汇编	295
附录 B 常用缩略语汇编	298
参考文献	305

第 1 章 概 述

1.1 引 言

随着信息网络的迅猛发展,对网络技术的掌握和应用越来越受到人们的重视。特别是在本世纪,信息网络无疑将成为信息社会的主要传播媒介和推动社会向前发展的巨大动力,发挥出越来越重要的作用。

目前,IP(Internet Protocol,网际协议)网络作为一种最有发展前景的网络技术,已广泛地应用于整个社会,并起着十分重要的作用。但对于大多数人来说,IP究竟能做什么?它是如何工作的?它的各部分之间是怎样相互作用的?这些问题并不是每个人都能说清楚、讲明白的。虽然有很多书籍介绍了许多有关 IP 网络的知识,但它们大量讨论的还只是信息网络的使用和应用,以及如何创建时髦的网站等等。

本书为了使读者对 IP 网络有一个基本、全面的了解,在介绍 IP 网络的基本概念、特点、组成、结构及发展趋势等问题的基础上,主要讨论 IP 网络的核心部件路由器技术与原理、路由协议、IP 网络的编址与设计以及 IP 网络的应用等,从而使读者能够较为全面、深入地获得 IP 网络系统的有关知识,并能指导 IP 网络的工程建设、设计与研究工作。

1.2 IP 网络的基本概念

何谓 IP 呢?狭义地讲,IP 是计算机网络中的一种协议标准,对于在网络层采用 IP 协议的网络,我们称之为 IP 网络,其网络层的数据包格式和传输策略等都要符合 IP 协议的规范。那么,什么是 IP 包呢?IP 包实际上就是按照 IP 协议定义的格式所组成的数据包,通常由包头(Header)和随后的一些数据组成,这里的数据可以是文件、终端应用程序或一段 E-Mail 消息等。在 IP 网络上,所有的信息都将被划分为独立的 IP 包,每个 IP 包可以沿自己的路径通过网络传输。这就反映出两个问题:一个是 IP 网络是一种分组交换网络,另一个是 IP 提供面向无连接的数据报服务。下面就简要地介绍分组交换和数据报服务的概念。

分组交换(也称为包交换)是在 20 世纪 60 年代发展起来的,当时主要应用于

军事环境下的通信基础设施。不同于电路交换,在采用分组交换的网络中,用于通信的计算机把它们的数据划分为若干个分组,也就是包,然后将这些分组通过分组交换网传输。分组交换技术有很多,如 X.25、帧中继(FR, Frame Relay)和 IP 等。本书将重点讨论 IP 网络,即 IP 路由器原理与技术,因为 IP 路由器技术作为 Internet 的主要技术得到了广泛的应用,并且成为一种发展趋势。

计算机在通信时为什么要采用分组交换呢?为什么不在计算机之间建立一条临时的“电话呼叫”呢?这种差别主要来源于电话网和计算机网通信特性的不同。在电话网中,用户在通话之前首先要发一个呼叫,一旦呼叫成功,通话双方之间就建立了一条物理通路,这条通路直到通话结束才被释放,在整个通话过程中,它是一直被占用的。由于在电话网中很强调话音的质量,而且一次通话会持续较长的时间,其呼叫建立时间相对于通话的总时间来说是很短的,即呼叫建立的开销在整个通话中所占的比例不是很大,所以在电话网中采用电路交换还是比较合算的。计算机通信就不一样了,大多数计算机通信常常是短暂的,一般在秒的量级(在电话网中,一个呼叫建立的时间大约是 5s),而且经常是“突发性”的。计算机可能在很短的时间之内向多个对等体(层)发送一段较短的数据比特,强行在每秒内建立多个连接,以实现多个对等计算机之间的“会话”。很显然,如果在计算机通信中采用电路交换,那么通信的效率会大大降低。因此要采用分组交换的方式。

一般说来,分组交换有两种基本类型,即“虚电路”和“数据报”。

“虚电路”提供面向连接的服务,在每个虚呼叫的开头需要建立一条虚连接,以后的通信就在这条虚连接上进行。虚电路与传统的电路交换是有很大区别的,在电路交换中,一个呼叫在其存在期间一直占用一条端到端的物理信道;而虚电路建立在分组交换的基础之上,尽管我们感觉是占用了一条端到端的物理信道,但实际上只是断续地占用一段又一段的链路。在虚电路中,没有必要让每个包都携带目的地址,只要建立连接时的数据包携带目的地址就可以了,其他后续的包使用短的虚电路号。虚电路的优点是可靠性比较高,但效率很低。实际上,网络层采用虚电路服务不是一件很明智的事情。

“数据报”是分组交换的另一种主要形式,它提供面向无连接的服务,IP 和许多其他网络层协议都采用这种服务。在数据报的模式中,数据发送之前并不建立一条可靠的连接,而是让每个包都携带完整的目的地址,即只管将包发送到网络中,至于如何将包转发到它所指明的目的地,完全依赖于中间的每个分组交换设备。数据报服务不提供可靠性,它只是尽最大能力交付,这样可以达到最大的效率,至于可靠性则由上层协议来负责。从目前互联网的发展趋势来看,在网络层提供数据报服务是非常成功的,IP 得到广泛的应用就是一个很好的例证。

在分组交换网中,分组交换设备是一个很关键的部件,它负责数据包的转发工作。IP 网也不例外。在 IP 网络的早期,分组交换设备通常称为“网关”,这可能是

因为它们常常作为本地校园网和广域网之间“网络关口”的缘故,现在,这些设备在更多的情况下被称为“路由器”,如图 1.2.1 所示。路由器可以在其任何接口上接收数据包,并通过不同的接口把数据包转发出去,将其发送到目的地。图中的双向箭头用以指示数据包既可以从该接口进入路由器,也可以从该接口离开路由器,路由器是 IP 网络的基本和核心部件。

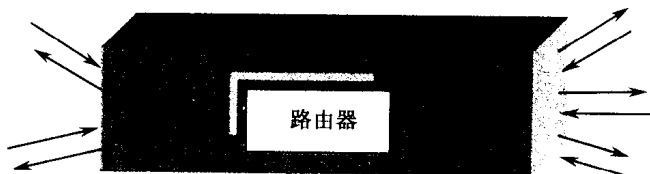


图 1.2.1 路由器的基本功能

在 IP 网络上流动的元素就是本节第一段提到的 IP 包。就 IP 包所携带的数据而言,可以是传统的计算机数据,也可以是在 IP 上使用的数字化音频和视频流。处理之后的语音和视频流其实也是数据,但是,与传统的数据传输(例如文件传输)不同,它们在传输上具有特殊的需求,因为语音和视频流对时间十分敏感,并且对延迟的容忍也十分有限。有趣的是,语音和视频流可以丢失一些数据而不会产生听觉或者视觉上的失真,但数据业务却必须保证传输的正确性,也就是说需要花费时间来重传丢失或者损坏的数据包以保证整个传输完好无损。

1.3 IP 网络的组成与结构

计算机网络中一个很基本的概念就是协议的分层。一个完整的 IP 网络可以分为 4 层,图 1.3.1 中给出了分层 IP 协议栈和 7 层 OSI(Open System Interconnect,开放系统互联)参考模型的结构比较。在计算机网络的分层模型中,4 层的 IP 参考模型和 7 层的 OSI 参考模型是两种最重要的类型。在这两种模型中,每层的功能都由一个以上的实体来执行。一个层中的实体与紧接着它的下一层中的实体直接相互作用,并为上一层提供服务。

在图 1.3.1 所示的 IP 协议栈分层结构中,子网层一般又称网络接口层,它大致相当于 OSI 参考模型中的物理层和数据链路层,负责从网络层接收 IP 报文并向物理网络发送,或从网络上接收物理帧,取出 IP 数据报并提交给网络层。子网层包含的协议主要有 ARP(Address Resolution Protocol,地址解析协议)和 RARP(Reverse Address Resolution Protocol,地址反解析协议)。网络层负责处理分组在网络中的活动,例如分组的选路等。在 IP 网络中,一般常把网络层叫做 IP 层,该层的协议主要包括 IP 协议、ICMP 协议(Internet Control Messages Protocol,因特网控

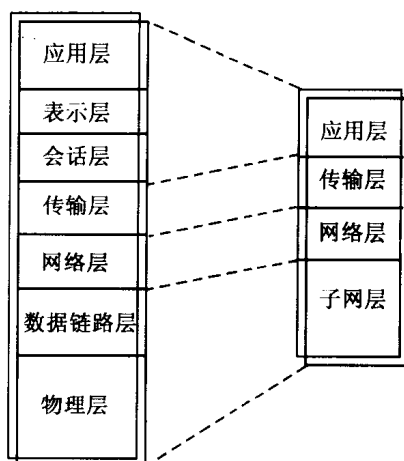


图 1.3.1 OSI 和 IP 模型比较

制报文协议)以及 IGMP 协议(Internet Group Management Protocol, 因特网组管理协议)等。传输层主要为两台主机上的应用程序提供端到端的通信。在 IP 网络中,有两个互不相同的传输协议:TCP(Transmission Control Protocol, 传输控制协议)和 UDP(User Datagram Protocol, 用户数据报协议)。应用层负责处理特定的应用程序细节,几乎各种不同的 TCP/IP 实现都会提供下面这些通用的应用程序: Telnet(远程登录), FTP(File Transfer Protocol, 文件传输协议), SMTP(Simple Mail Transfer Protocol, 简单邮件传送协议)和 SNMP(Simple Network Management Protocol, 简单网络管理协议)等。各种常见协议和 IP 网络各层次的对应关系如图 1.3.2 所示。

在 IP 协议栈的各个协议中,最重要的有两个:TCP 和 IP,因此,人们经常用“TCP/IP”来表示 IP 网络的体系结构(Internet 就使用这个体系结构)或者整个 TCP/IP 协议族。

在进行 IP 通信时,通常使用低层“子网”技术来完成相应的数据传输。在这些子网技术中,有些属于 LAN(Local Area Network, 局域网)子网,例如, Ethernet(以太网)、Token Ring(令牌环)、FDDI(Fiber Distributed Data Interface, 光纤分部式数字接口)等;还有一些属于 WAN(Wide Area Network, 广域网)子网,例如,静态和动态的点到点链路、X.25、帧中继、ATM(Asynchronous Transfer Mode, 异步传输模式)、SMDS(Switched Multi-Megabit Data Service, 交换型多兆位数据服务)等等。图 1.3.3 给出了可以运行 IP 的各种媒体介质使用路由器进行互联的示意图。这些子网中的每个子网都具有各自的内部地址格式和帧格式。有些子网技术既有头字段也有尾字段,而有些只用一个头来封装 IP。每种技术都运行在单一的速度或者速度集之上,换句话说,它们之间可以完全不同。

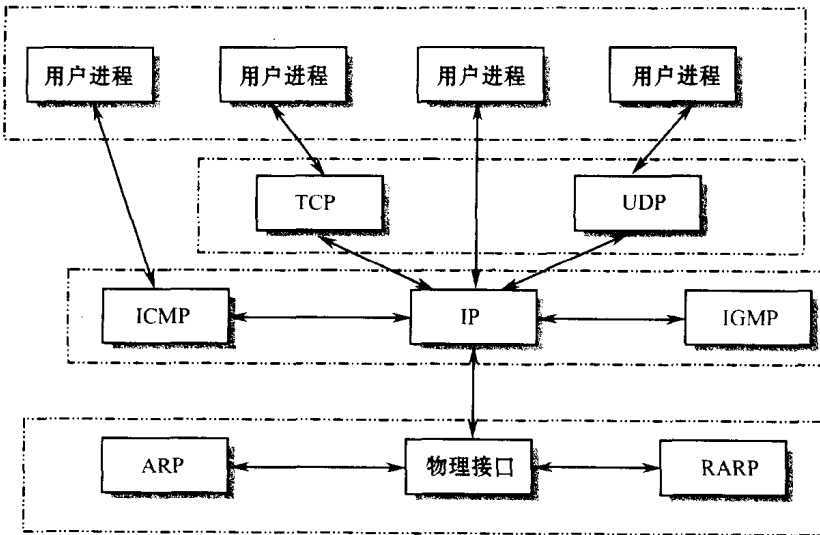


图 1.3.2 IP 网络四个层次及协议

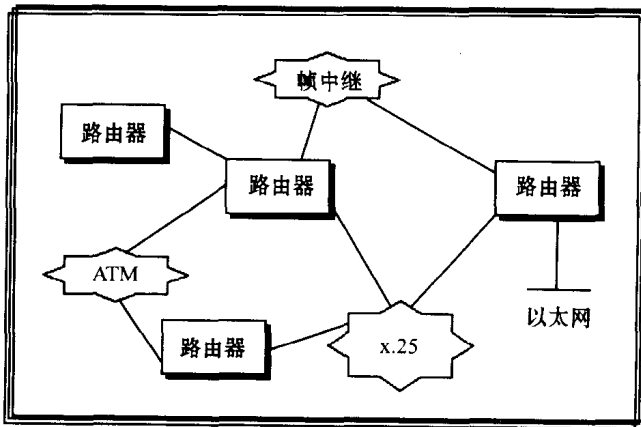


图 1.3.3 IP 运行于其他子网上的示意图

当一个路由器和另一个路由器或者端站进行通信时,需要以某种方式向其相邻节点发送数据包。一个包仅使用其 IP 地址是无法发送给其相邻接点的。IP 地址是高层地址,实际的操作是把 IP 包用特定的子网类型的帧格式封装起来,在该帧中包含有目标节点的地址,通常还包含有路由器的子网源地址。图 1.3.4 给出了 IP 协议栈中各层之间的关系以及封装的概念,这里可以用 TCP 代替 UDP,因为在分层的讨论中它们是相同的。

实际上,各种局域网和广域网都具有自己的编址方法,路由器在相互通信时必须使用相应子网的地址。因此,路由器除了需要知道其相邻节点的 IP 地址外,一

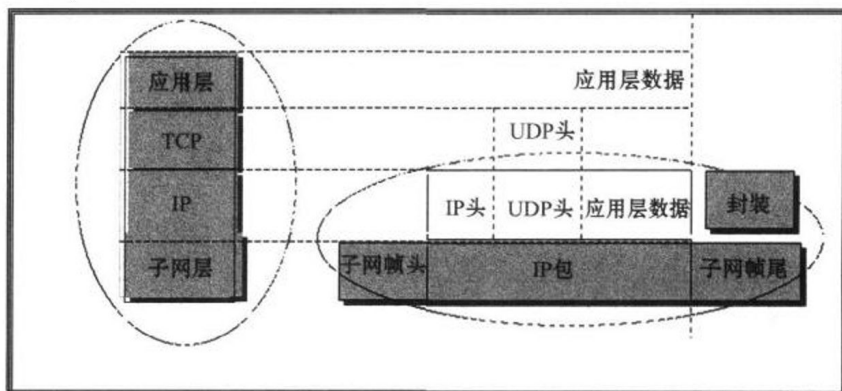


图 1.3.4 分层及封装

般还需要知道其相邻节点的子网层地址。对于每种子网介质,IP 将使用不同的技术来获取其相邻节点的子网地址。在现实生活中,具有多层地址的系统随处可见,如现实生活中的地区、街道和门牌等都是多层地址的表现。

1.4 IP 网络的包格式

1.4.1 IPv4 包格式

IP 包是由 IP 头和高层传输数据“协议”组成,如图 1.4.1 所示。它使 IP 层具有一项很重要的功能,就是“子网无关性”,正是这项功能使得 IP 可以运行在几乎所有类型的子网之上。



图 1.4.1 IP 包格式

IP 向传输层屏蔽它所支持的众多子网和所有不同底层的特征。就像多个网络层协议可以共享同一个子网层一样,也可有多个传输层协议共享 IP 层,如图 1.4.2 所示。注意,图 1.4.2 中的 ICMP 不属于传输层协议,它属于网络层,是 IP 协议的附属协议,但其报文和 UDP、TCP 类似,是封装在 IP 分组中的。

目前使用的 IP 层协议是 IPv4,图 1.4.3 中给出了 IPv4 包头的格式。图中的数字是比特位,包头的宽度为 32 比特。下面分别介绍各个字段的含义。

(1) 版本:表示本 IP 报文的版本号,占 4 比特,对 IPv4 报文来说,该字段的值为 4。

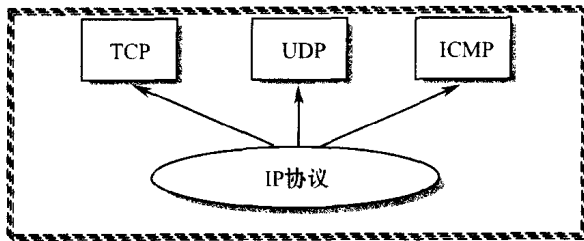


图 1.4.2 多个传输层协议共享 IP 层

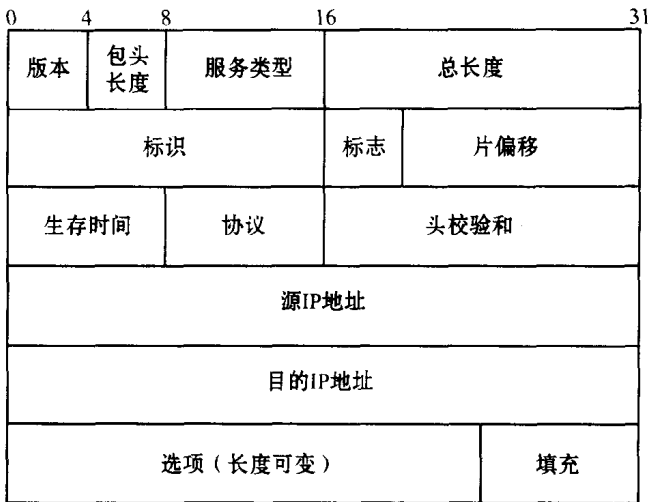


图 1.4.3 IP 包头的结构

(2) 包头长度:表示整个 IP 包头的长度,以 4 字节为 1 个单位。该字段占 4 比特,这样一个 IP 包头最多可以有 60(15×4)字节。通常 IP 包头的长度是 20 字节,即没有头部选项,这时该字段的值为 5。如果有头部选项,该字段的值就要做相应的改变。

(3) 服务类型(ToS, Type of Service):表示本报文所希望得到的服务,占 8 比特。该字段最近在业界引起了较多的关注,因为 Internet 提供商(ISP, Internet Service Provider)和顾客强烈要求能够用某种方式在他们的网络内部提供不同的“服务类”级别。当前,“区分服务(DS, Differentiated Service)”的模型定义工作组正在进展之中。在该模型中,将 ToS 字段重新定义为 DS 字段。定义区分服务的两个 RFC 文档为 RFC-2474 和 RFC-2475。而 RFC-2430 仅仅是说明如何对区分服务进行部署的一个例子,但是,随着这项新技术实践经验的累积,将有可能会出现其他的方案。

(4) 总长度:表示整个 IP 报文的长度,占 16 比特,单位为字节。这样,一份 IP