

职业网络管理员培训教材
第四册 (共八册)

TCP/IP与 网络体系结构

Introduction to TCP/IP & Networking Architectures

5.04-43

马 凯 译

O'REILLY
FUNDAMENTALS

中国电力出版社

TN915:04-43
011

TCP/IP与网络体系结构

Introduction to TCP/IP & Networking Architectures

马 凯 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

中国电力出版社

图书在版编目 (CIP) 数据

TCP/IP 与网络体系结构 / (美) 韦斯特耐特技术培训公司 (WestNet Learning Technologies) 编著; 马凯译. —北京: 中国电力出版社, 2000.7

(职业网络管理员培训教程丛书)

书名原文: Introduction to TCP/IP & Networking Architectures

ISBN 7-5083-0387-3

I .T … II .①韦 … ②马 … III .①计算机网络 - 通信协议 ②网络结构 IV .TP393

中国版本图书馆 CIP 数据核字 (2000) 第 38439 号

北京市版权局著作权合同登记

图字: 01-2000-1698 号

Simplified Chinese Edition, jointly published by O'Reilly & Associates, Inc. and China Electric Power Press, 2000. Authorized translation of the English edition, 1999, WestNet Learning Technologies, the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

简体中文版由 O'Reilly & Associates, Inc. 授权中国电力出版社出版 2000。英文原版的翻译得到 WestNet Learning Technologies 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——WestNet Learning Technologies 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式复制。

书 名 / TCP/IP 与网络体系结构

书 号 / ISBN 7-5083-0387-3

责任编辑 / 胡顺增

封面设计 / Ellie Volckhausen, Hanna Dyer, 张健

出版发行 / 中国电力出版社

地 址 / 北京三里河路 6 号 (邮政编码 100044)

经 销 / 全国新华书店

印 刷 / 北京市地矿印刷厂

开 本 / 787 毫米 × 1092 毫米 16 开本 19.5 印张 318 千字

版 次 / 2000 年 8 月第一版 2000 年 8 月第一次印刷

印 数 / 0001-5000 册

定 价 / 45.00 元 (册)

前言

介绍

本书对计算机网络业中所涉及到的四种网络结构进行了全面的介绍和讲解。我们希望本书的内容对你有所帮助并易于使用。

本书的安排

本书被分为四个主要部分。这四个部分是：

第一部分：网络体系结构介绍

第一章 网络体系结构概论

第二部分：网络体系结构

第二章 TCP/IP 介绍和概论

第三章 TCP/IP 协议和路由

第四章 TCP/IP 网络应用和管理

第五章 SNA 体系结构

第六章 AppleTalk

第七章 DECnet 体系结构

第三部分：网络体系结构集成

第八章 在 TCP/IP 网络中集成 IBM 大型机

第四部分：总结

第九章 网络体系结构实例研究

更多的信息

如您希望得到更多的信息，可以用电子邮件同我们联系。我们很高兴为您提供更多的关于产品和服务方面的信息。

奥莱理软件（北京）有限公司

<http://www.oreilly.com.cn>

E-mail: info@mail.oreilly.com.cn

第一章

网络体系结构概论

体系结构是指一个特定实体的设计和建造方法。如果将某种建筑结构称作“哥特式”，那它必定具备哥特式建筑所共有的特征。对于计算机网络来说，同样是为了不同的目的开发和建造具备共同体系结构特征的网络系统。无论早期的开发者是如私营的IBM公司或政府的代理机构，都是为了特定的目的开发了相应的网络体系结构。本课程将分析讨论通常能在今天的公司里找到的网络体系结构。其中主要有四种被广泛接受的网络结构，它们是：

1. TCP/IP
2. SNA
3. AppleTalk
4. DECnet

我们将对以上四种网络体系结构进行讨论，并且对其中的一些内容作详细的讲解。在所有的体系结构中，TCP/IP体系结构是非常有特点的，我们将作主要的讲解。TCP/IP协议被认为是“开放的”，因为从其最初的版本直到目前的最新版本都是公开的，并且是不收费的。它已成为现在网络环境下的最重要的网络体系结构。因为TCP/IP日渐重要和流行，所以我们将课程中花大部分时间对其进行讲解。

TCP/IP被广泛地接受和使用的原因并非仅仅是因为其开放的体系结构，更重要的原因在于它是Internet的基础。同时几十年来它也是作为网络协议随UNIX操作系统发行的配套产品。因为Internet的流行和基于UNIX的系统的广泛应用，促使TCP/IP成为所有体系结构类型中的主导者。

尽管TCP/IP获得了广泛的接受,并同一些操作系统进行捆绑销售,但其他的一些网络体系结构也同样重要。因为多数公司并非只是由单一的网络体系结构组成,一般是有几种不同的网络体系结构,我们对这些体系结构都应该深入了解。因此我们将对所有主要的体系结构进行讲解,但将把重点放在TCP/IP协议及相关的应用层协议的讲解上。

问题讨论

1. 体系结构和协议之间有何区别?
2. TCP/IP 是一种体系结构吗?
3. 你认为 IBM 的 SNA 产品的使用将会增加还是减少? 请给出你的理由。
4. AppleTalk 能被看作“开放”的体系结构吗?
5. 广泛使用 TCP/IP 的操作系统是哪种操作系统?

第二章

TCP/IP 介绍和概论

TCP/IP 介绍

TCP/IP 是 Transmission Control Protocol/Internet Protocol (传输控制协议 / Internet 协议) 的缩写, 虽然它代表了为早期的 ARPANET 所开发的两个协议, 但是它的含义已经超出了协议本身。其实在 TCP/IP 的“世界”中包含了三个主要的组件:

- 网络, 将网络集中到一起称为 Internet (因特网)。
- 一整套的协议、软件和构成许多操作系统的标准组件的应用程序, 如 FTP (File Transfer Protocol, 文件传输协议) 和 TELNET (虚拟终端应用程序)。
- 实际的连网协议, 例如传输控制协议和 Internet 协议。

本章将讲解协议及建立于它们之上的应用, 这些应用在逻辑上构成了应用层。本章和随后的章节将对 TCP/IP 和自治系统进行讲解, 所谓自治系统, 即指在同一个管理实体控制之下, 并基于 TCP/IP 的网络。本章将介绍“子网”互联以及建立 TCP/IP 网络的概念。完成本章的学习后你将能够:

- 认识 TCP/IP 的协议层以及它的组件和功能, 并能指出和每一个 TCP/IP 协议层相对应的 OSI 协议层的名称。
- 理解子网的概念, 并且知道为何需要子网及怎样使用子网。
- 识别以带点十进制数形式表示的 B 类 Internet 地址的各组成部分。

- 给出一个以层次结构表示的 Internet 域名的一般形式。
- 描述 IP 和 TCP 所提供的网络传输服务的特征。
- 识别一个应用需要使用 UDP 的原因。

图 2-1 表明了 TCP/IP 和 OSI 模型的关系。请记住，TCP/IP 协议在 OSI 模型出现的 10 年前就存在了，所以两种体系结构的协议层并非相互一一对应。TCP/IP 仅有两个精确定义的协议层，即传输控制协议层 TCP 和 Internet 协议层 IP。其应用层非常广泛，并未精确定义。一些由 OSI 的会话层、表示层和应用层所执行的任务，如不同处理器之间数据格式的转换、对话管理被留给了单独的 TCP/IP 应用去处理。

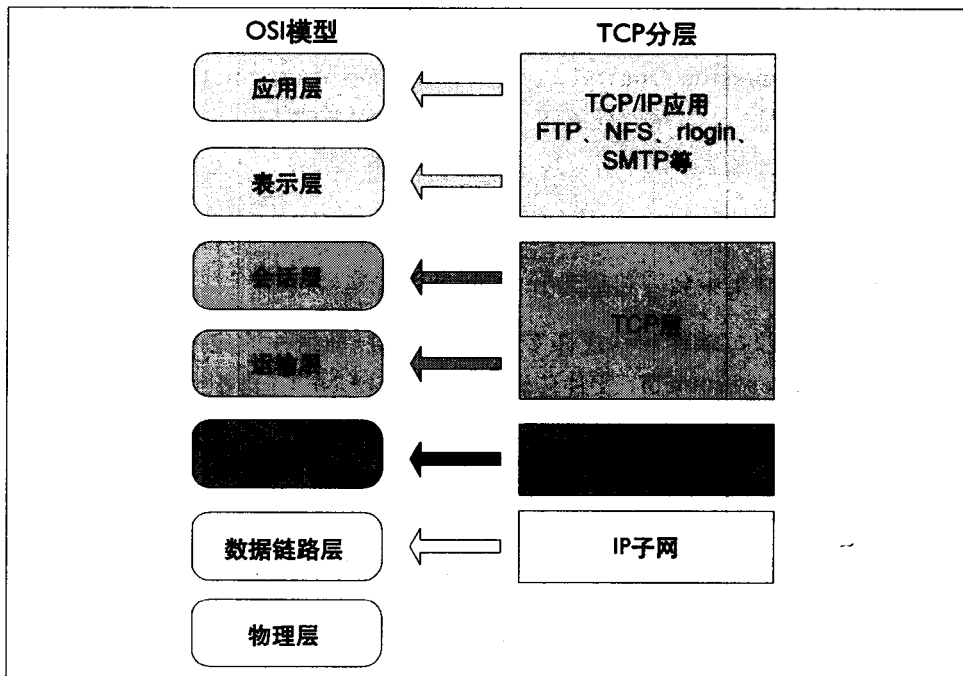


图 2-1 TCP/IP 和 OSI 模型

因为TCP/IP早在个人电脑和工作站出现之前就已经存在，当时的网络计算机都是带多个哑终端的小型机或大型机，一些TCP/IP的术语和其他网络体系结构中所使用的术语是有差别的。因为每一个网络节点，即是一台带多个用户所使用的哑终

端的“主机”，TCP/IP所使用的术语“主机”和术语“网络节点”有着相同的含义，可以互换使用。如果你在其他写有有关TCP/IP的资料中遇到术语“主机”，则你可以用术语“节点”来替换它，反之亦然。图2-2展示了网络、子网和节点（即主机）之间的关系。请注意子网是用路由器来连接的。

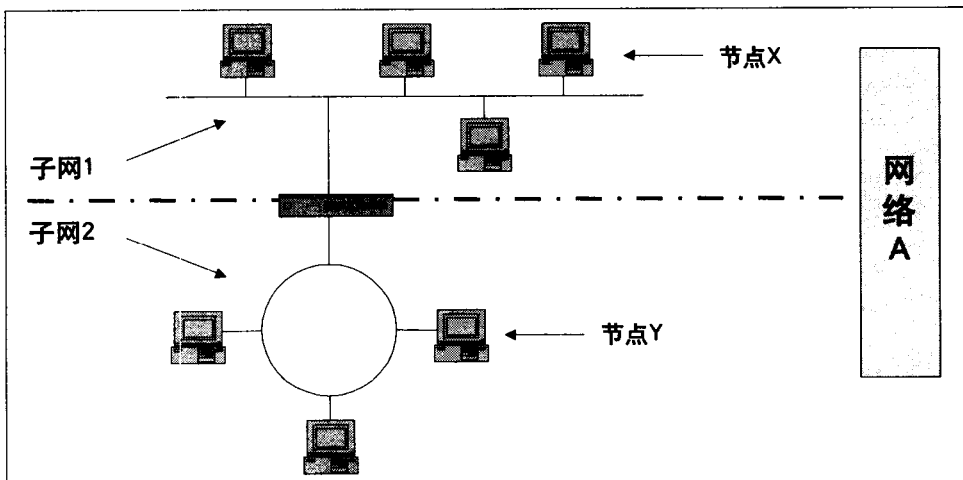


图2-2 TCP/IP 和子网

早期的 ARPANET 设计和建造者并未预料到网络的发展速度会如此之快。当时 LAN（局域网）并未出现。所设计的 Internet 编址方案中，Internet 地址是由网络地址和节点地址组成的，该编址方案无法表示今天如此多的局域网。为了解决这一问题，出现了子网的概念。本质上子网是一个独立的 LAN，即 Internet 的一个组成部分。子网的使用使得网络可由一个到多个物理网络构成。这样在建立网络时，可对网络中可管理的用户数、网络带宽和流量进行充分考虑。

一个网络可看作由多台主机组成的集合，即由 WAN（广域网）连接的独立节点。它可以是一个有独立网络 ID 的 LAN。也可以是 LAN 子网的集合，该子网集合由连接一个以上 LAN 的节点组成，或是以上几种方式的组合。

图2-3展示了TCP/IP协议层更加详细的组成情况，其中包括了一些已经介绍过的部分和未介绍的部分。我们再一次给出了对应的OSI协议层。本章将讲解在一个自治网络中，有关这些层的协议和服务。

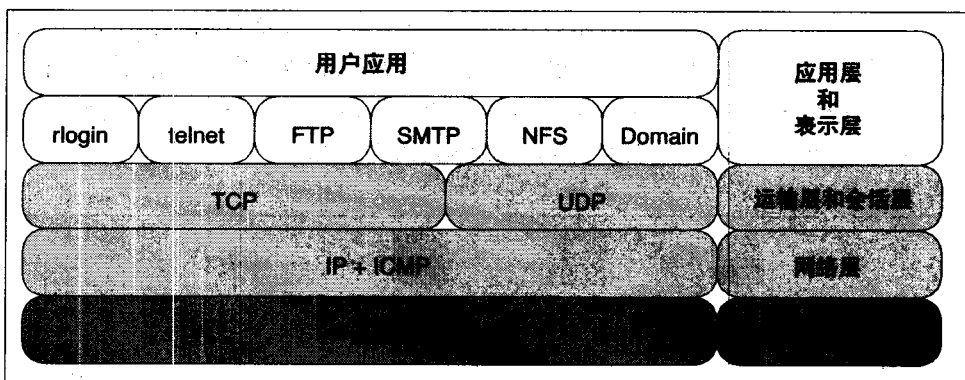


图 2-3 TCP/IP 协议层

我们将对这些层先进行简单的描述，然后进行深入的讨论：

IP 子网 + 地址解析协议 (Address Resolution Protocol, ARP): IP 的 IP 子网功能，也称为网络接口，使用多种低级别的协议，如 IEEE 802.2 逻辑链路控制 (Logical Link Control, LLC) 协议，并和以太网 (802.3) 或令牌环网 (802.5) 等协议协调工作，进行数据报的传送。低级别的 IP 子网的特征取决于该节点所连接的通信链路。如果节点连接到 LAN，IP 子网则由基于某一个协议如 IEEE 802.2 LLC 的协议来组成。如果节点连接到一个远程子系统，则 IP 子网由使用数据链路层协议如 HDLC、X.25、帧中继或 ATM (Asynchronous Transfer Mode; 异步传输模式) 的复杂子系统来构成。在以上两种情况下，IP 子网接受 IP 数据报并且经过特定的链路进行传输，必要时使用 ARP 协议确定 LAN 的地址。ARP 是用于将 IP 地址映射为 LAN 硬件地址的。

Internet 协议 (IP) + Internet 控制报文协议 (Internet Control Message Protocol, ICMP): Internet 协议层的作用是为上层提供无连接的、不可靠的、最尽力的包交换服务。该协议体现在一个 Internet 数据报中。ICMP 是 IP 的一个组成部分，产生 IP 用来控制网络数据流和报告网络错误的报文，如“目的端不可到达”和“传送超时”等报文。通常是由路由器产生 ICMP 报文的。在另一方面，主机通常产生一个 PING (ICMP 请求应答)。PING 是用来检测所给的网络目的地址是否可达的。

传输控制协议 (TCP): TCP 层的作用是提供面向连接的、端到端的、无差错的报文传送。因为它使用 IP 层所提供的不可靠传输服务, 所以它的主要任务是进行错误纠正和连接的管理。该层还提供如设置滑动窗口的尺寸 (或称为数据通过网络的字节量) 的流量控制功能。

用户数据报协议 (UDP-User Datagram Protocol): 相对于 TCP 协议的另外一种可供选择的协议是 UDP 协议, UDP 为那些并不要求提供像 TCP 所提供的全部服务的那些应用提供简单的、无连接的、不可靠的数据报传输。网络管理应用 SNMP (Simple Network Management Protocol; 简单网络管理协议) 即是这种应用的一个例子。这种情况下是由应用来纠正或恢复 UDP 消息在传输过程中的丢失或顺序错误。正如 TCP 一样, UDP 提供目的端口和源端口以便在端点之间传输信息。

TCP/IP 用户应用和应用服务包括:

- rlogin: 提供从一个 TCP/IP 系统到另外一个 TCP/IP 系统的远程登录。
- 简单邮件传输协议 (Simple Mail Transfer Protocol, SMTP): 它为邮件服务和网络新闻提供电子信息。
- TELNET: TELNET 是连接到远程终端的标准协议。
- 文件传输协议 (File Transfer Protocol, FTP): 它提供在两个系统之间进行文件传输的服务。FTP 使用 TELNET 来建立连接控制。
- 域名协议 (Domain Name Protocol, DNS): 其作用是将层次结构的 Internet 域名如 sales@mycompany.com 转换为 Internet 地址如 205.169.85.70 的。
- 网络文件系统 (Network File System, NFS): 它是由 Sun Microsystems 开发的。NFS 是在网络中实现文件共享的事实上的标准。

TCP/IP 的命名和地址分配

在 TCP/IP 中, 除非是将几个基本的概念搞清楚, 不然 TCP/IP 中有关命名和地址分配的问题是比较容易让人混淆的。管理中心 (在美国, IP 地址的申请是由 ARIN, 即美国 Internet 编码登记处来负责处理的) 是一个负责在国际 Internet 中进行域名

和 IP 地址的管理和控制的机构。个人或公司可向域名管理中心申请域名和 IP 地址，或向已经拥有很多 IP 地址的 ISP（Internet 服务提供商）申请。域名服务器需要知道域名和对应的 IP 地址，才能保证那些以名字为地址的应用（如电子邮件）到达目的地。对于公司的 Intranet 也同样应用这一策略。下面是一些重要的术语和概念：

- 网络管理员：负责对某处的一个或多个网络进行命名。
- 系统管理员：负责对子网中的节点进行命名。通常网络管理员和系统管理员是同一个人。名字和地址在分配上必须从几个层次上去考虑：
 1. 网络号：网络必须有一个由域名管理中心分配的地址，同时必须保证它在 Internet 网中是唯一的。
 2. 子网号：子网必须有一个在网络中唯一的地址，它由网络管理员进行分配。
 3. 节点号：网络中的每一个节点有一个在子网中的地址，它由系统管理员分配。
 4. 局域网号：局域网中每一个节点都有一个固化在机器（通常是网络接口卡 NIC）上的硬件地址。局域网号通常和局域网的类型有关。例如，在以太网中，局域网号指的即是以太网地址。

为了命名方便，TCP/IP 采用分级命名系统对机器和网络命名。例如，电子邮件地址就是这样命名的：

- 节点名：网络中的每一个节点有一个由字母或数字构成的名字，例如 sales。
例子：sales@abcxyz3.com
- 域名：指明一个网络或网络组的名字。域名管理中心为域名的建立作了常规约定。域名通常是由字母或数字组成，并由点进行分隔，如：nodename.xxx.yyy.zzz。最右边的部分（zzz）是域本身的名称，域名中从右向左直到 nodename 表明一种相对的层次关系。一个域名看上去应该形如：registrar@college.edu。

Internet 定义了几个高级别的域（最右边的部分）。所有的商业站点如大多数的公司都属于“com”域。每一个在 Internet 中的站点都有一个在域名管理中心注册的唯一名字。一个站点名可能覆盖多个网络号。

1. 在一个站点中，网络和节点还可以继续进行分组和命名。
2. 用户名，每一个使用节点的用户至少有一个由字母和数字组成的标识号。例如，Jane Doe 在“company”中有一个被称为“janed”的用户名，该用户名是用来使用 VAX 工作站的，该工作站有一个特定的 LAN 地址。它的主机有一个 Internet 地址 205.169.85.33。它包括了网络号、子网号（如果使用了子网）和节点号。稍后你将会学到更多有关 Internet 地址的知识。Jane 的地址可以被表示为 janed@205.169.85.33 或 janed@company.com。

IP 地址由 4 个八位组或字节（1 个八位组 = 8 位）组成，或者说由 32 位组成。每个八位组可表示的值的范围由十进制数 0 到十进制数 255。或由二进制表示为 00000000~11111111。下面表示了八位二进制数据是如何转换为十进制数的：

1	1	1	1	1	1	1	1
128	64	32	16	8	4	2	1

(128+64+32+16+8+4+2+1=255)

换句话说，255 是八位组的所有位都为 1 时所取得的十进制数值。

现在给出一个由八位组转换为十进制数据的例子：

0	1	0	0	0	0	0	1
0	64	0	0	0	0	0	1

(0+64+0+0+0+0+0+1=65)

Internet 地址的分类

Internet 地址有 5 类不同的网络地址，即 A 类地址、B 类地址、C 类地址、D 类地址和 E 类地址。在此我们将只讨论 A 类到 C 类地址，因为 D 类和 E 类地址为保留地址。要决定一个地址的类别，只要看以点分隔的十进制数的第一个八位组。

这是一个地址的例子（4 个八位组）：

10. 1. 23. 19 (十进制)
00001010. 00000001. 00010111. 00010011 (二进制)

为了适应不同规模的网络，这些八位组被从不同的地方分开。

- A类地址：1~126 （如 10.1.23.22）
- B类地址：128~191 （如 172.16.22.48）
- C类地址：192~223 （如 143.28.8.10）

对于A类地址，第一个八位组代表的是网络部分，所以上述例子中的网络地址为10。第2、3、4个八位组（余下的24位）可被网络管理员用来进行子网部分和主机部分的划分。

对于B类地址，前面两个八位组代表的是网络部分，所以上述例子中的网络地址为172.16，第3、4两个八位组（余下的16位）进行子网部分和主机部分的划分。B类地址是分配给拥有主机数在256到65536的网络的。

对于C类地址，前3个八位组是网络部分。例子中C类地址的网络地址为193.28.8。第四个八位组是分配给本地的子网和主机，该类地址对于拥有少于256台主机的网络尤其适合。

D类和E类地址是为特殊用途而保留的。例如，一种特定类型的多点传送用的是D类地址。

Internet地址习惯上被写为由点分隔的4个十进制数的形式，被称为点分十进制表示法。一个Internet地址实际上是一个32位（即4字节）的二进制数。为了便于人们使用而被写成了四个十进制整数的形式。每一个整数代表了地址的一个字节，所以最大的值是255，并且，理论上地址的范围应是0.0.0.0到255.255.255.255（16进制的“00000000”到“FFFFFFFF”）。（记住，十六进制情况下，一个“F”代表4个1，1个“0”代表4个0）。

Internet地址128.33.68.2等同于二进制的1000 0000 0010 0001 0100 0100 0000 0010，和十六进制的80 21 44 02。

Internet 地址中包含了网络地址、子网地址（如果设置了的话）和一个主机地址。三部分可通过几种方式包含于一个二进制字中。网络地址可能由第 1、2、3 个字节构成，剩下的 3 个字节、2 个字节或 1 个字节构成子网地址和主机地址。如果作更进一步的划分，即哪些位构成子网地址、哪些位构成主机地址稍后将作进一步的讲解。所幸的是在实际网络配置的过程中，我们仅仅只会用到上述复杂地址划分的一种。

子网掩码

网络中的每一台主机必须识别 Internet 地址中哪些位与其所在的物理网络或子网相对应。通常这是通过“子网掩码”来实现的，“子网掩码”通常是由软件进行设置的。稍后我们将就子网划分作详细讨论。如果用户或系统管理员设置了不正确的“子网掩码”，主机中的软件将无法识别局域网中报文的地址，从而将造成无法通信。子网掩码在整个网络中必须保持一致。除非使用了不同大小的子网划分，子网掩码是一个 32 位的 16 进制数，可以屏蔽掉地址的主机地址部分，标识出地址的网络和子网号。

- 对一个没有进行子网划分的网络来说，网络地址 205.169.85.21 的子网掩码是 255.255.255.0。
- 对于一个分成七个物理网络的网络来说，网络地址 205.169.85.21 的子网掩码是 255.255.255.240，此处为网络的增长留有余地。

TCP/IP 网络所使用的 Internet 编址方案，使得它可以定位 Internet 网中的任何一台机器。如果你的网络已接到 Internet 或计划接到 Internet，必须从域名管理中心获得地址的网络部分。这样才能确保取得的网络地址在 Internet 上是唯一的。域名管理中心对于那些并未接到 Internet 的 TCP/IP 网络的地址使用和分配是没有限制的，但是使用 TCP/IP 协议的软件必须用以点分隔的 32 位地址或以点分隔的十进制数来表示。在今天面向 Internet 的文化环境中，后一种情况已相当少见了。

无类域间路由（CIDR）

CIDR（Classless Inter-Domain Routing）是对 Internet 地址策略的改进，它比起用 A 类、B 类和 C 类使用的编址方式更有效。能够使得 IP 地址更有效地被分配。

随着网络不断地和 Internet 相联接, Internet 面临的不仅仅只是地址资源的枯竭, 而且面临路由表的崩溃问题。

例如, 让我们来考虑一个 B 类的 Internet 地址的情况。在使用 B 类地址的情况下, 能够使用的主机地址为 65000 个。但大多数拥有 B 类地址的公司, 并没有如此多的主机。或考虑这样一种情况, 某组织在考虑了目前和将来的需求情况下, 计算出需要 100 个地址, 如果他们得到的是一个 C 类地址, 则将有超过 150 个的地址被浪费掉。这些地址对于其他组织或公司来说也是无法利用的。

无类域间路由 (CIDR) 通过加入网络地址前缀的方式, 替代 A 类、B 类和 C 类地址中网络地址只能是固定的 8 位、16 位或 24 位的方式, CIDR 目前使用的是 27 位到 13 位的地址前缀形式。它允许划分到小到只有 32 台主机大到 500000 台主机的网络地址块。这种方法使得在地址的分配上更加符合一个组织或公司的实际需要。

一个 CIDR 地址包括标准的 4 字节 (32 位) 的 IP 地址和一个表明使用多少位来作网络前缀的信息。例如: 对于 205.169.85.0/26 这个 CIDR 地址来说, “/26” 即指明了前 26 位唯一标识了一个网络, 余下位用来标识某一特定的主机。这使得用户可以使用 64 个主机地址。

前缀数字越小, 可用于分配给主机的地址就越多。一些 CIDR 前缀值和对应可分配的主机地址数如下所示:

CIDR 前缀	主机地址数
/27	32
/26	64
/25	128
.	.
.	.
/20	4,096
/19	8,192
.	.
.	.
/14	262,144
/13	524,288