

M

现代IP技术丛书

MODERN IP TECHNOLOGY

# 垃圾邮件 与反垃圾邮件

曹麒麟 张千里 编著  
李 星 审

## 技术

人民邮电出版社  
POSTS & TELECOMMUNICATIONS PRESS

现代 IP 技术丛书

# 垃圾邮件与反垃圾邮件技术

曹麒麟 张千里 编著

李 星 审

人民邮电出版社

## 图书在版编目 (CIP) 数据

垃圾邮件与反垃圾邮件技术/曹麒麟, 张千里编著. —北京: 人民邮电出版社, 2003.2

(现代 IP 技术丛书)

ISBN 7-115-10818-8

I. 垃... II. ①曹... ②张... III. 电子邮件—基本知识 IV. TP393.098

中国版本图书馆 CIP 数据核字 (2002) 第 104483 号

### 内 容 提 要

本书针对目前国内关注的垃圾邮件问题进行了深入的讨论, 介绍了垃圾邮件的起源与历史, 重点分析了国内垃圾邮件的特点和常用的垃圾邮件控制方法, 并对有关的法律问题和反垃圾邮件组织工作的开展进行了讨论。本书特别深入地讨论了常用邮件服务器的安全配置, 并且给出了详细的操作步骤。本书的末尾还提供了内容比较丰富的附录。

本书是邮件服务器管理员实用的技术参考书, 其中的一些数据是经过大量研究得到的第一手资料, 可以供有关部门制定政策及从事有关法律研究的人员参考, 同时也可供那些正在想方设法控制垃圾邮件的 ISP、从事网络安全方面工作的人员以及其他感兴趣的读者参考。

现代 IP 技术丛书

### 垃圾邮件与反垃圾邮件技术

- 
- ◆ 编 著 曹麒麟 张千里  
审 李 星  
责任编辑 陈万寿
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
读者热线 010-67129258  
北京汉魂图文设计有限公司制作  
北京隆昌伟业印刷有限公司印刷  
新华书店总店北京发行所经销
  - ◆ 开本: 787×1092 1/16  
印张: 10.75  
字数: 256 千字  
印数: 1-4 000 册
  - 2003 年 2 月第 1 版  
2003 年 2 月北京第 1 次印刷

---

ISBN 7-115-10818-8/TN · 1951

定价: 20.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

## 序 言

电子邮件为人们的工作和生活带来了极大的便利，甚至某种程度上改变了人们的沟通方式。然而，作为其发展的副产品——垃圾邮件，却给 Internet 用户、网络管理员和 ISP 带来了无尽的烦恼，据统计，全世界因为垃圾邮件每年要损失 1000 万美元。垃圾邮件起源于美国，在 20 世纪 90 年代曾经一度泛滥。经过不懈的努力，包括技术上和法律上的，目前美国的垃圾邮件正在逐年减少。然而问题并没有得到彻底解决，垃圾邮件的源头正逐渐地转到了中国及东南亚的一些国家和地区。

垃圾邮件不仅浪费了宝贵的网络资源，还带来严重的社会问题。甚至有些国内的 IP 站点被列入了国外某些组织的“黑名单”，来自某些 IP 地址的邮件在国际出口被屏蔽等，影响了国内用户对电子邮件的正常使用。垃圾邮件对国内还是一个新出现的问题，国内有关的资料还非常缺乏。国外的专著也不多。

中国教育科研网紧急响应组（CCERT）是一个隶属于中国教育科研网（CERNET）、服务于国内广大的 Internet 用户的网络安全组织。从 1998 年成立之初就把反垃圾邮件作为一个重要的研究课题。该书的两位作者都是该小组的成员。张千里参与了 CCERT 的创建工作，也是最早开始研究垃圾邮件问题的人员之一，曹麒麟的硕士论文的研究方向就是垃圾邮件及反垃圾邮件技术。他们在反垃圾邮件方面都积累了丰富的经验。我相信这本书会对广大的邮件服务器管理员，对 ISP 和有关的政府部门提供及时有益的参考。希望这本书也会再次推动国内业界反垃圾邮件活动，纯洁我们的网络，推动我国的 Internet 的健康发展。

李 星

## 前 言

垃圾邮件是 Internet 发展的一个副产品，在国外曾经泛滥很长一段时间，由于各方面的努力，目前在欧美国家已经基本得到控制。一部分公司和个人开始转向还没有采取有效控制措施国家和地区，利用他们的邮件服务器的安全漏洞转发垃圾邮件，掩藏了真实来源的同时转移了成本，中国和韩国成为世界主要的垃圾邮件的源头。随着 Internet 在中国的普及，国内一些商业机构也看到了这种广告方式的商机，从 2001 年下半年开始从国内直接发出的垃圾邮件迅速增加。部分国外公司甚至利用中国没有相关的法律限制，直接在中国设立公司从事此类商业活动。这些行为带来的后果是中国成了 Internet 世界的众矢之的，国外的部分反垃圾邮件组织和公司开始大量屏蔽中国的 IP 站点和来自中国部分域名的邮件。

2002 年 3 月，随着《南方周末》以“互联网的分裂”报道开始，国内媒体对垃圾邮件问题及由此产生的后果进行了长时间的讨论，中央电视台、齐鲁电视台、文汇报和北京晨报等都进行了深入的讨论。

什么样的邮件是垃圾邮件？如何有效地控制垃圾邮件？个人用户、ISP 和立法部门都能做些什么？目前有关的中文资料还不多见，尤其缺乏必要的数据库。国外的网站上有一些资料，但这方面的系统的专著也寥寥无几。一本专门讨论垃圾邮件及反垃圾邮件技术的书无论对网络管理员、ISP 还是有关的官员和法律工作者都是有一定参考价值的。

本书针对目前国内关注的垃圾邮件问题进行了深入的讨论。重点分析了国内垃圾邮件的特点和常用的垃圾邮件控制方法，并对有关的法律问题和反垃圾邮件组织工作的开展进行了讨论。本书特别深入地讨论了常用邮件服务器的安全配置，并且给出了详细的操作步骤，对于不熟悉邮件服务器的管理员很有帮助。全书共分为 10 章，第 5 章到第 7 章由张千里执笔，其余章节由曹麒麟执笔，最后本书经过李星教授审阅。

本书的编写过程中，得到了 CCERT 全体工作人员的大力支持和帮助，李星教授在百忙中审阅了全书，并撰写了序言，在此一并表示感谢。希望本书的出版能够对我国的反垃圾邮件事业有所帮助。

曹麒麟  
张千里

# 目 录

<b>第 1 章 电子邮件的工作原理</b> .....	1
1.1 电子邮件的发展简介 .....	1
1.2 TCP/IP .....	2
1.3 电子邮件的工作原理 .....	3
1.3.1 邮件的格式 .....	3
1.3.2 邮件的传送[4] .....	5
1.3.3 POP 与 IMAP .....	6
1.4 SMTP 协议的基本结构 .....	7
1.4.1 SMTP 的基本模型 .....	8
1.4.2 SMTP 的基本命令[5] .....	9
1.5 SMTP 的安全缺陷 .....	10
<b>第 2 章 垃圾邮件及其危害</b> .....	14
2.1 什么是垃圾邮件 .....	14
2.2 垃圾邮件的历史 .....	15
2.3 垃圾邮件的危害 .....	17
<b>第 3 章 国内垃圾邮件状况分析</b> .....	19
3.1 国内邮件服务器的安全状况 .....	19
3.2 国内垃圾邮件的特点分析 .....	22
3.2.1 垃圾邮件的内容分布 .....	22
3.2.2 垃圾邮件的来源分析 .....	23
3.2.3 垃圾邮件的变化趋势 .....	25
3.3 国内 IP 被国外屏蔽状况 .....	26
3.4 病毒邮件 .....	29
3.5 结论 .....	30
<b>第 4 章 反垃圾邮件的一般技术</b> .....	32
4.1 如何保护邮件地址 .....	32
4.2 追踪垃圾邮件 .....	33

4.2.1	邮件信头生成过程 .....	33
4.2.2	邮件信头分析 .....	36
4.2.3	信头分析中需要注意的几个问题 .....	36
4.3	几种常见的邮件服务器客户端读取邮件信头的方法 .....	40
4.4	如何查找管理员 .....	46
4.4.1	使用 whois 查询 .....	47
4.4.2	使用 Nslookup 查询 .....	50
4.4.3	使用 dig 查询 .....	52
4.5	几种可行的做法 .....	54
<b>第 5 章</b>	<b>邮件服务器的安全配置 .....</b>	<b>56</b>
5.1	Sendmail 服务器安全配置 .....	57
5.1.1	Sendmail 服务器的安装和配置步骤 .....	57
5.1.2	提高 Sendmail 系统的安全性能 .....	59
5.2	Exchange 服务器安全配置 .....	64
5.2.1	Exchange 的安全特性 .....	64
5.2.2	Exchange 2000 相对于 Exchange 5.5 的改进[3] .....	66
5.2.3	安全配置 Exchange 服务器的一些建议[2] .....	67
5.3	其他常用邮件服务器的安全配置方法 .....	70
5.3.1	Qmail 的安装[5] .....	70
5.3.2	Qmail 的配置 .....	73
5.3.3	Postfix 的安装和配置[1] .....	75
<b>第 6 章</b>	<b>邮件发信权限控制技术 .....</b>	<b>80</b>
6.1	Open Relay 与垃圾邮件 .....	80
6.2	Open Relay 的关闭方法 .....	82
6.2.1	Sendmail[2] .....	82
6.2.2	Netscape Messaging Server[9] .....	84
6.2.3	Exchange .....	87
6.2.4	Qmail .....	90
6.2.5	Postfix .....	90
6.2.6	IMail [4] .....	91
6.2.7	Lotus Notes[5] .....	91
6.3	常见邮件系统发信认证配置 .....	92
6.3.1	Sendmail 发信认证的安装方法 .....	93
6.3.2	在 postfix 中实现基于 cyrus-sasl 的 SMTP 认证 .....	95
6.3.3	基于 qmail 的 smtp 用户认证设置方法 .....	98
6.3.4	客户端的配置 .....	100



第 7 章 垃圾邮件的过滤 .....	104
7.1 基于 IP 地址的过滤技术 .....	106
7.1.1 基于网络的 IP 地址过滤技术[4] .....	106
7.1.2 基于主机的 IP 地址过滤技术 .....	107
7.1.3 DNSBlockList——RBL .....	108
7.2 MTA 过滤 .....	112
7.2.1 Sendmail[8] .....	113
7.2.2 Netscape Messaging Server[2] .....	115
7.2.3 Postfix[7] .....	119
7.3 使用 Procmail 设置邮件过滤 .....	123
7.3.1 Procmail 介绍[1] .....	124
7.3.2 配置 Procmail .....	124
7.3.3 利用 Procmail 进行邮件过滤应当考虑的事项 .....	129
7.4 MUA 过滤 .....	130
第 8 章 垃圾邮件的法律问题 .....	133
8.1 技术上的局限性 .....	133
8.2 垃圾邮件的法律问题 .....	133
8.3 垃圾邮件的立法状况 .....	135
8.3.1 美国的立法状况 .....	135
8.3.2 欧洲的立法状况 .....	135
第 9 章 反垃圾邮件组织 .....	137
9.1 国外著名反垃圾邮件组织 .....	137
9.1.1 MAPS——Mail Abuse Prevention System .....	137
9.1.2 Spamhaus .....	137
9.1.3 CAUCE .....	139
9.2 CCERT 反垃圾邮件小组 .....	139
9.2.1 反垃圾邮件的政策 .....	140
9.2.2 技术支持与应急响应 .....	140
9.2.3 反垃圾邮件工作的基本要素 .....	142
第 10 章 反垃圾邮件工作的未来 .....	144
附录一 SMTP 协议中用到的代码和含义 .....	146
附录二 CERNET 关于制止垃圾邮件的管理规定 .....	147
附录三 与 E-mail 有关的 RFC .....	149
附录四 反垃圾邮件 Internet 资源 .....	160



# 第 1 章 电子邮件的工作原理

## 1.1 电子邮件的发展简介

Internet 问世后的最初应用就是电子邮件。虽然今天 Internet 的应用范围得到了极大的拓展，电子邮件仍然是它最为广泛的应用之一。在过去的若干年里，电子邮件的功能已经丰富了很多。据 2000 年 3 月份 Messaging Online 的一份数据显示，Internet 上有 5.69 亿个邮箱，平均每个 Internet 用户有 1.8 个。有趣的是中国的 Internet 也同样是从电子邮件开始的，《中国互联网发展大事记》中记载“1987 年 9 月 20 日，钱天白教授发出我国第一封电子邮件‘越过长城，通向世界’，揭开了中国人使用 Internet 的序幕”。今天，电子邮件已经成为商业、政府、教育等行业最基本的通信工具。

在 Internet 电子邮件流行之前已经有几种形式的电子邮件存在了。从 20 世纪 60 年代后期人们就开始使用计算机共享文件的方式来传递信息。到了 70 年代开始出现基于 APARnet 传输的电子邮件，大多数人认为 1971 年使用 APARnet 发出的电子文本信息是第一封 Internet 电子邮件。1972 年 Ray Tomlinson 写了第一个电子邮件程序，叫做 SNDMSG，在 APARnet 上使用。为了创建一个世界范围的电子邮件系统，在 1984 年 ISO（国际标准化组织）和 ITU（国际电信联盟）发布了一组新的信件传递标准，这就是 X.400。在 1988 年和 1993 年又做了两次更新。在欧洲这个标准在较为广泛的范围内被接受，但由于它的规模和复杂性，并没有在全球范围内流行。相反，基于 TCP/IP 的电子邮件从一开始就显示出了强大的生命力。1982 年 Internet 协会发布了基于 TCP/IP 的 SMTP（简单邮件转发协议），即 RFC 821（Request for Comments，Internet 的标准文档）。稍后发布的 RFC 822 定义了 ASCII 代码的纯文本的信件结构。由于 Internet 的快速发展和巨大成功，这种 Internet 邮件也迅速被广为接受，特别是在 Internet 发源地美国。在开始的时候这些纯文本的格式就可以满足当时的需要了。但很快人们希望使用更复杂的功能，如多媒体文件格式，于是 Internet 协会 1996 年又发表了一系列的关于 MIME（Multipurpose Internet Mail Extensions）格式的定义，支持用户的这种需求。今天人们使用的大部分邮件是 Internet 邮件，其他一些形式的邮件要么采用了 Internet 的标准，要么被 Internet 邮件系统替代。虽然 X.400 在欧洲还在使用，但已经可以与 Internet 邮件系统比较好地对接了。

在本书中我们只讨论 Internet 电子邮件。

## 1.2 TCP/IP

Internet 的原形是 1969 年建立的 APARnet。在互联网发展史上具有决定意义的一件事是在 1983 年 1 月 1 日, APARnet 正式转换成 TCP/IP 网络。正是 TCP/IP 的出现, 才使得互联网得以在全世界的范围内迅速发展并具有今天的规模。根据 TCP/IP, 互联网协议分为 4 层, 加上最底层的硬件层一共是 5 层 [1], 如图 1.1 所示。

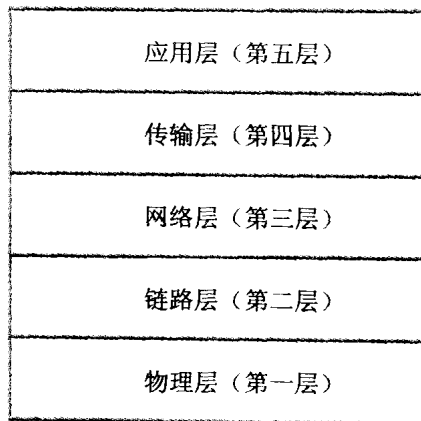


图 1.1 TCP/IP 的分层结构

- 物理层: 对应于网络的基本硬件, 这也是 Internet 物理构成, 即我们可以看得见的硬件设备, 如 PC 机、互联网服务器、网络设备等, 必须对这些硬件设备的电气特性作一个规范, 使这些设备都能够互相连接并兼容使用。

- 链路层: 它定义了将数据组成正确帧的规程和在网络中传输帧的规程, 帧是指有一定结构的一串数据, 它是数据在网络中传输的单位。

- 网络层: 本层定义了互联网中传输的“信息包”格式, 以及从一个用户通过一个或多个路由器到达最终目标的“信息包”转发机制。

- 传输层: 为两个用户进程之间建立、管理可靠而又有效的端到端连接。

- 应用层: 它定义了应用程序使用互联网的规程。电子邮件的 SMTP 就建立在这一层。

Internet 的核心层是网络层和传输层, 相应的核心协议是 IP 和 TCP。IP 的主要功能包括无连接数据报传送、数据报寻径以及差错处理三部分。IP 协议的特点是点到点的, IP 对等实体间的通信不经过中间机器, 对等实体所在的机器位于同一物理网络, 对等机器之间有直接的物理连接。IP 层的主要功能是屏蔽下面物理层的差别, 向上一层提供一致的数据格式。所有要传输的数据, 被按照一定的格式分组封装成 IP 数据报, 数据报单元通过寻径等机制进行传输, 在接收方进行重组, 得到最初要传送的数据。IP 是不可靠的数据传输协议, 由于网络的拥塞而发生的的数据丢失等情况是不可避免的, 因此 Internet 还必须有一定的控制重传机制, 这就是差错与控制报文协议(ICMP)。

尽管计算机通过安装 IP 软件, 从而保证了计算机之间可以发送和接收数据, 但 IP 还不

能解决数据分组在传输过程中可能出现的问题。因此,若要解决可能出现的问题,还需要 TCP 来提供可靠的并且无差错的通信服务。TCP 被称作一种端对端协议,这是因为它为两台计算机之间的连接起了重要作用:当一台计算机需要与另一台远程计算机连接时,TCP 会让它们建立一个连接、发送和接收数据以及终止连接。传输控制协议 TCP 利用重发技术和拥塞控制机制,向应用程序提供可靠的通信连接,使它能够自动适应网上的各种变化。即使在 Internet 暂时出现堵塞的情况下,TCP 也能够保证通信的可靠。Internet 是一个庞大的国际性网络,网路上的拥挤和空闲时间总是交替不定的,加上传送的距离也远近不同,所以传输数据所用时间也会变化不定。TCP 具有自动调整“超时值”的功能,能很好地适应 Internet 上各种各样的变化,确保传输数值的正确。

IP 只保证计算机能发送和接收分组数据,而 TCP 则可提供一个可靠的、可控的、全双工的信息流传输服务。虽然 IP 和 TCP 这两个协议的功能不尽相同,也可以分开单独使用,但它们是在同一时期作为一个协议来设计的,并且在功能上也是互补的。只有两者的结合,才能保证 Internet 在复杂的环境下正常运行。凡是要连接到 Internet 的计算机,都必须同时安装和使用这两个协议,因此在实际中常把这两个协议统称作 TCP/IP。

TCP/IP 除了 TCP 和 IP,还包含物理接口和 IP 层之间的 ARP/RARP、应用层的 FTP、SMTP 和 BOOTP 等,所有的这些协议构成 Internet 的 TCP/IP 协议族。

## 1.3 电子邮件的工作原理

### 1.3.1 邮件的格式

为了让邮件能够顺利地上传送,能够被不同的邮件服务器正确识别,并且在各种各样的终端上显示出来,就需要遵循一定的格式。邮件的格式在 RFC 822 中加以定义[2],为了支持多媒体文档的传输,1996 年又发表了一系列的关于 MIME(Multipurpose Internet Mail Extensions) 格式的定义。

RFC 2045: MIME 第一部分 Internet 信体格式

RFC 2046: MIME 第二部分 媒体类型

RFC 2047: MIME 第三部分 非 ASCII 文件信体的扩展

RFC 2048: MIME 第四部分 注册过程

RFC 2046: MIME 第五部分 符合标准和举例

RFC 822 在 2001 年 4 月更新为 RFC 2822 (目前为建议稿),以反映近年来的实践 [3]。

了解信件的基本格式会帮助我们理解邮件的传输和邮件的信头,这里主要讨论基于 RFC 822 的基本信件格式。

在最高层信件是非常简单的,它含有一系列的文本,每一行以回车(CR)和换行(LF)组成。信件由信头、信件体和之间的空行组成。信头有定义的格式,以使得 MTA、MDA 和 MUA 能对它进行程序分析。信头是必须的,信件体是可选的。下面是一个简单的例子:

From: admin@ccert.edu.cn .....信头部分  
To: staff@ccert.edu.cn  
Subject: Meeting Notice  
Date: 2002-5-6

.....信头和信件体之间的空行

Hi Everyone, .....信件体  
There is a meeting this afternoon.  
Thanks.  
Admin

RFC 822 为信头定义了 20 多个标准的字段，包括 Date、From、To、CC 等一些必须的字段和一些非必须的字段。另外，在信件的传输过程中，MUA 和 MTA 还会在信头上加入一些路径信息，它们合在一起构成了收到的邮件的信头部分。下面是一个完整的信头，我们以此为例子介绍一些关键字段的含义。

```
Received: (eyou send program); Tue, 09 Apr 2002 11:02:03 +0800
Received: from unknown (HELO dns.ccert.edu.cn) (unknown@202.112.57.6) by
166.111.8.16 with SMTP; Tue, 09 Apr 2002 11:02:03 +0800
Received: from kylinlp ([202.112.50.23]) by dns.ccert.edu.cn (8.10.2+
Sun/8.10.2) with ESMTP id g393EPo04562 for <caoql00@mails.tsinghua.edu.cn>;
Tue, 9 Apr 2002 11:14:26 +0800 (CST)
Reply-To: caoql@dns.ccert.edu.cn
From: caoql@dns.ccert.edu.cn
To: caoql00@mails.tsinghua.edu.cn
Subject: =?gb2312?B?08q8/rXEveG5uQ==?=
Date: Tue, 9 Apr 2002 11:02:35 +0800
Organization: CCERT
Message-ID: <000001c1df73$0378dc50$173270ca@kylinlp>
MIME-Version: 1.0
Content-Type: text/plain; charset="gb2312"
Content-Transfer-Encoding: base64
X-Priority: 3 (Normal)
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook, Build 10.0.2616
Importance: Normal
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2600.0000
```

这里是信件的信体。

信件到此结束。

(1) From:

From: caoql@dns.ccert.edu.cn 表示生成该信件的人。

(2) To:

To: caoql00@mails.tsinghua.edu.cn, 指出收件人。

(3) Subject:

邮件的主题。

(4) Reply-To:

标识发件人希望的回复地址。

(5) Message-ID:

Message-ID: <000001c1df73\$0378dc50\$173270ca@kylintp>唯一地标识一个信件, 该字段由 MUA 或者第一个 MTA 产生。

(6) Received:

Received 字段含有信件的一个特定的 MTA 处理记录。处理信件的每个 MTA 必须在每个信件头的上面加入这个字段, 这个信息对于跟踪信件非常有用。

(7) 其中以 X 开头的字段不是 RFC 822 中要求的字段, 是 SMTP 服务器扩展的字段, 由软件厂商自行定义的。

### 1.3.2 邮件的传送[4]

电子邮件与普通邮件有类似的地方, 发信者注明收件人的姓名与地址 (即邮件地址), 发送方服务器把邮件传到收件方服务器, 收件方服务器再把邮件发到收件人的邮箱中, 如图 1.2 所示。

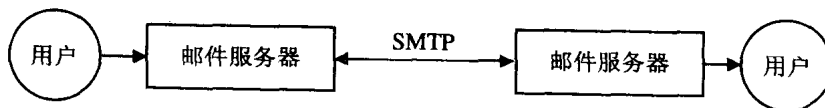


图 1.2 高层 SMTP 接口模型

下面解释邮件传送中涉及到的几个概念:

MUA (Mail User Agent), 邮件用户代理, 帮助用户读写邮件。

MTA (Mail Transport Agent), 邮件传输代理, 负责把邮件由一个服务器传到另一个服务器或邮件投递代理。

MDA (Mail Delivery Agent), 邮件投递代理, 把邮件放到用户的邮箱里。

整个邮件传输过程如图 1.3 所示。

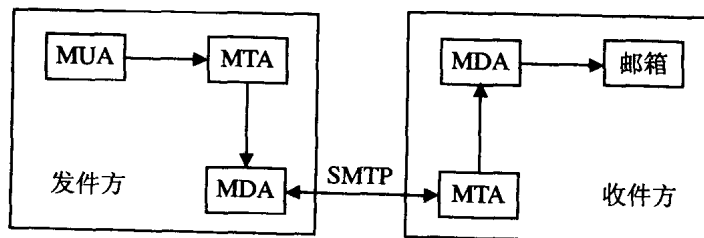


图 1.3 邮件传输中的代理

目前使用的 SMTP 协议是存储转发协议,意味着它允许邮件通过一系列的服务器发送到最终目的地。服务器在一个队列中存储到达的邮件,等待发送到下一个目的地。下一个目的地可以是本地用户,或者是另一个邮件服务器,如图 1.4 所示。

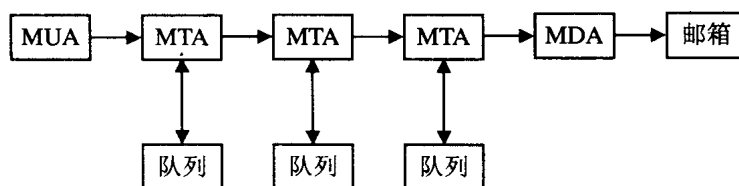


图 1.4 存储转发

如果下游的服务器暂时不可用,MTA 就暂时在队列中保存信件,并在以后尝试发送。

### 1.3.3 POP 与 IMAP

当我们具备了两个 SMTP 服务器,就可以实现异地的邮件通信了。实际使用中,为了方便,我们通常还需要用到 POP 和 IMAP 服务器,以及供用户编辑和阅读邮件的应用程序,如 Outlook、Foxmail 等。

POP 服务器主要是为了解决用户的终端不能总连接在网上等问题。为了随时接收可能发来的邮件和传出邮件,SMTP 服务器必须时刻处于工作状态。个人终端,如便携机或者拨号上网,可能无法满足此要求。而且也没有必要为每个用户的终端都安装一个复杂的 SMTP 服务器。另一方面,随着网络的发展,许多用户不愿意学习使用复杂的服务器来从服务器端直接读取邮件。他们希望能够把信件下载到本地,使用简单方便的应用程序编辑和阅读。POP 就是为了解决这些问题而设计的。MTA 把邮件投递给 POP 服务器,暂时存放所有收到的邮件,等待用户来取。用户取信时使用 POP 客户端,把信件下载到本地机器上。简单的讲,SMTP 服务器好比一个完整的邮政系统中的各个邮局,负责传送信件;POP 服务器仿佛一个收发室,负责暂时保管信件和发送信件。

最新的 POP 协议是在 RFC 1939 (Post Office Protocol - Version 3) 中加以详细定义的,比较完整的 POP 接口模型如图 1.5 所示。

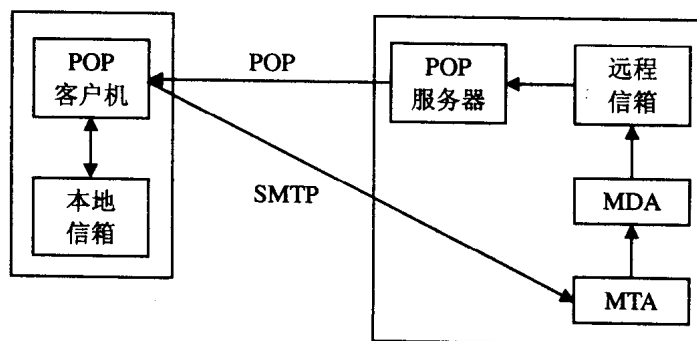


图 1.5 POP 接口模型

考虑了 POP 之后,一个实际的邮件传输的过程如图 1.6 所示。

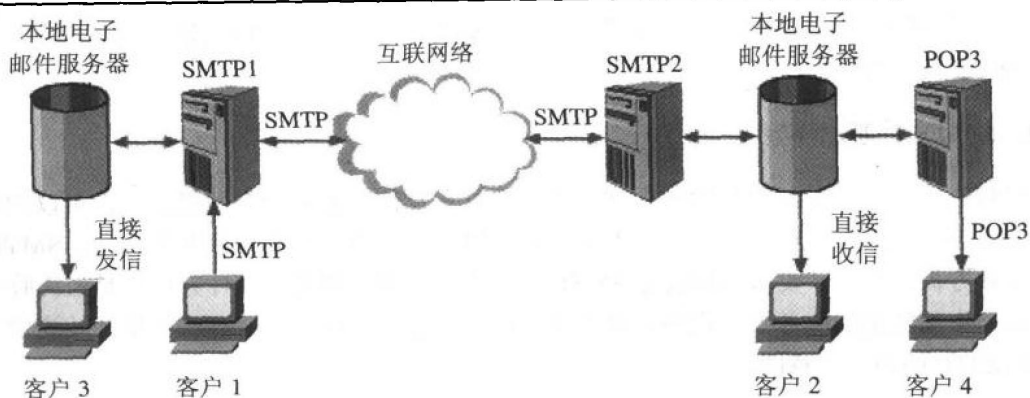


图 1.6 邮件传输示意图

IMAP 比 POP 提供了更多的功能，也更复杂一些。POP3 缺乏严格的邮件处理能力，通常邮件被从服务器下载到客户端后，邮件就会从服务器上删除。这有利于节省服务器硬盘空间，但对于可能使用多个终端的用户却不够方便。例如用户在家、办公室和旅行时可能需要使用不同的终端，他需要从不同的客户端上看到相同的内容，但使用 POP 邮箱就会被分割成 3 个不同的部分。IMAP 通过维护服务器上的邮箱并允许客户端的计算机对服务器上的邮件进行操作，允许使用者从多个地点访问邮箱而不会出现邮件被分割的情况。IMAP 的接口模型如图 1.7 所示。

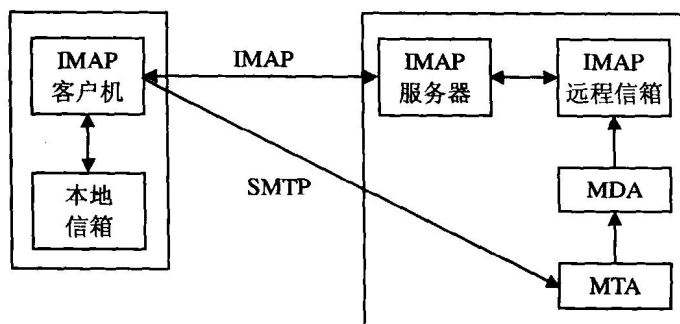


图 1.7 IMAP 接口模型

与 POP 一样，MTA 收到电子邮件，把它提交给 MDA，由它把电子邮件投递给一个邮箱。IMAP 客户机管理它本地的邮箱，并与 IMAP 服务器通信管理它的远程邮箱。IMAP 服务器提供一个远程客户机与存放在服务器上邮箱的接口。IMAP 也是通过与 MTA 的对话发送信件。

## 1.4 SMTP 协议的基本结构

SMTP (Simple Mail Transfer Protocol) 定义了保证电子邮件的可靠和高效传送的机制，最初的内容包含在 RFC821 中 [5]。在 2001 年 4 月 RFC 2821 对该协议进行了更新，取代了旧的 RFC821 [6]。TCP/IP 协议的应用层中包含有 SMTP 协议，但事实上它与传输系统和机制无关，仅要求一个可靠的数据流通道。它可以工作在 TCP 上，也可以工作在 NCP、NITS 等



协议上。在 TCP 上，它使用端口 25 进行传输。SMTP 的一个重要特点是可以在可交互的通信系统中转发邮件。

### 1.4.1 SMTP 的基本模型

SMTP 提供了一种邮件传输的机制，当收件方和发件方都在一个网络上时，可以把邮件直传给对方；当双方不在同一个网络上时，需要通过一个或几个中间服务器转发。SMTP 首先由发件方提出申请，要求与接收方 SMTP 建立双向的通信渠道，收件方可以是最终收件人也可以是中间转发的服务器。收件方服务器确认可以建立连接后，双方就可以开始通信。图 1.8 是 SMTP 的模型示意图。

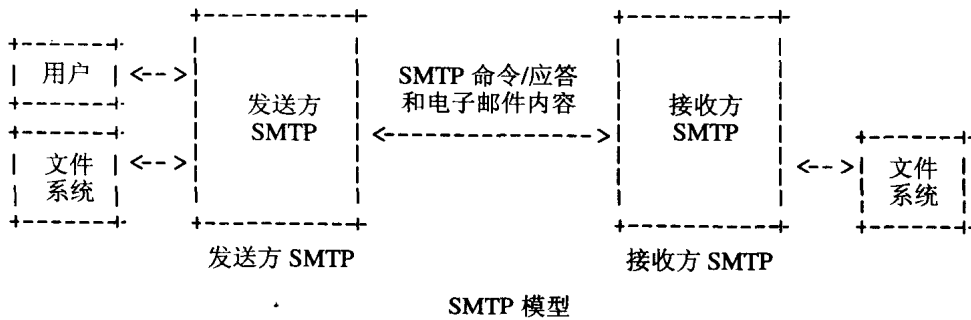


图 1.8 SMTP 模型

发件方 SMTP 向收件方发出 MAIL 命令，告知发件方的身份；如果收件方接受，就会回答 OK。发件方再发出 RCPT 命令，告知收件人的身份，收件方 SMTP 确认是否接收或转发，如果同意就回答 OK；接下来就可以进行数据传输了。通信过程中，发件方 SMTP 与收件方 SMTP 采用对话式的交互方式，发件方提出要求，收件方进行确认，确认后才进行下一步的动作。整个过程由发件方控制，有时需要确认几回才可以，如图 1.9 所示。

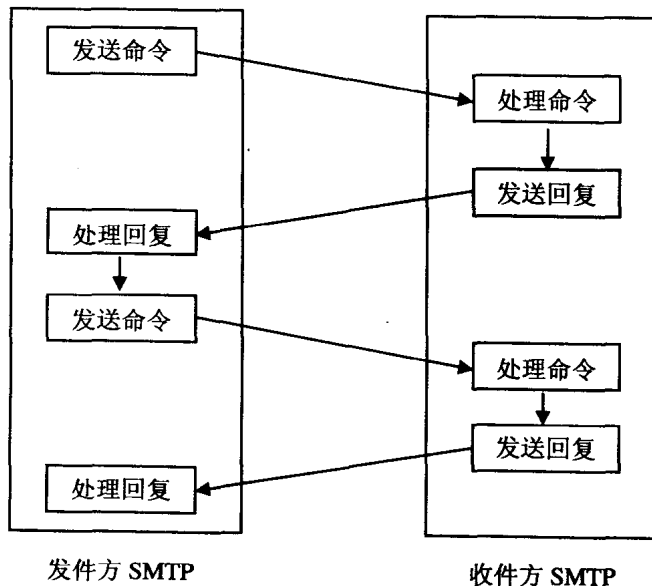


图 1.9 SMTP 会话过程

为了保证回复命令的有效，SMTP 要求发件方必须提供接收方的服务器及邮箱。邮件的命令和答复有严格的语法定义，并且回复具有相应的数字代码。所有的命令由 ASCII 码组成。命令代码是大小写无关的，如 MAIL 和 mail、mAIL 是等效的。

### 1.4.2 SMTP 的基本命令[5]

SMTP 定义了 14 个命令，它们是：

```
HELO <SP> <domain> <CRLF>
MAIL <SP> FROM:<reverse-path> <CRLF>
RCPT <SP> TO:<forward-path> <CRLF>
DATA <CRLF>
RSET <CRLF>
SEND <SP> FROM:<reverse-path> <CRLF>
SOML <SP> FROM:<reverse-path> <CRLF>
SAML <SP> FROM:<reverse-path> <CRLF>
VRFY <SP> <string> <CRLF>
EXPN <SP> <string> <CRLF>
HELP [<SP> <string>] <CRLF>
NOOP <CRLF>
QUIT <CRLF>
TURN <CRLF>
```

其中使 SMTP 工作的基本的命令有 7 个，分别为：HELO、MAIL、RCPT、DATA、RSET、NOOP 和 QUIT。分别介绍如下。

- **HELO**——发件方问候收件方，后面是发件人的服务器地址或标识。收件方回答 OK 时标识自己的身份。问候和确认过程表明两台机器可以进行通信，同时状态参量被复位，缓冲区被清空。

- **MAIL**——这个命令用来开始传送邮件，它的后面跟随发件方邮件地址（返回邮件地址）。当邮件无法送达时，它也用来发送失败通知。为保证邮件的成功发送，发件方的地址应是被对方或中间转发方同意接受的。这个命令会清空有关的缓冲区，为新的邮件做准备。

- **RCPT**——这个命令告诉收件方收件人的邮箱。当有多个收件人时，需要多次使用该命令，每次只能指明一个人。如果接收方服务器不同意转发这个地址的邮件，它必须报 550 错误代码通知发件方。如果服务器同意转发，它要更改邮件发送路径，把最开始的目的地（该服务器）换成下一个服务器。

- **DATA**——收件方将该命令之后的数据作为发送的数据。数据被加入数据缓冲区中，以单独“<CRLF>.<CRLF>”行结束数据。结束行对于接收方同时意味着立即开始缓冲区内数据传送，传送结束后清空缓冲区。如果传送接受，接收方回复 OK。

- **RSET**——这个命令用来通知收件方复位，所有已存入缓冲区的收件人数据、发件人