

电子邮件

理论课程

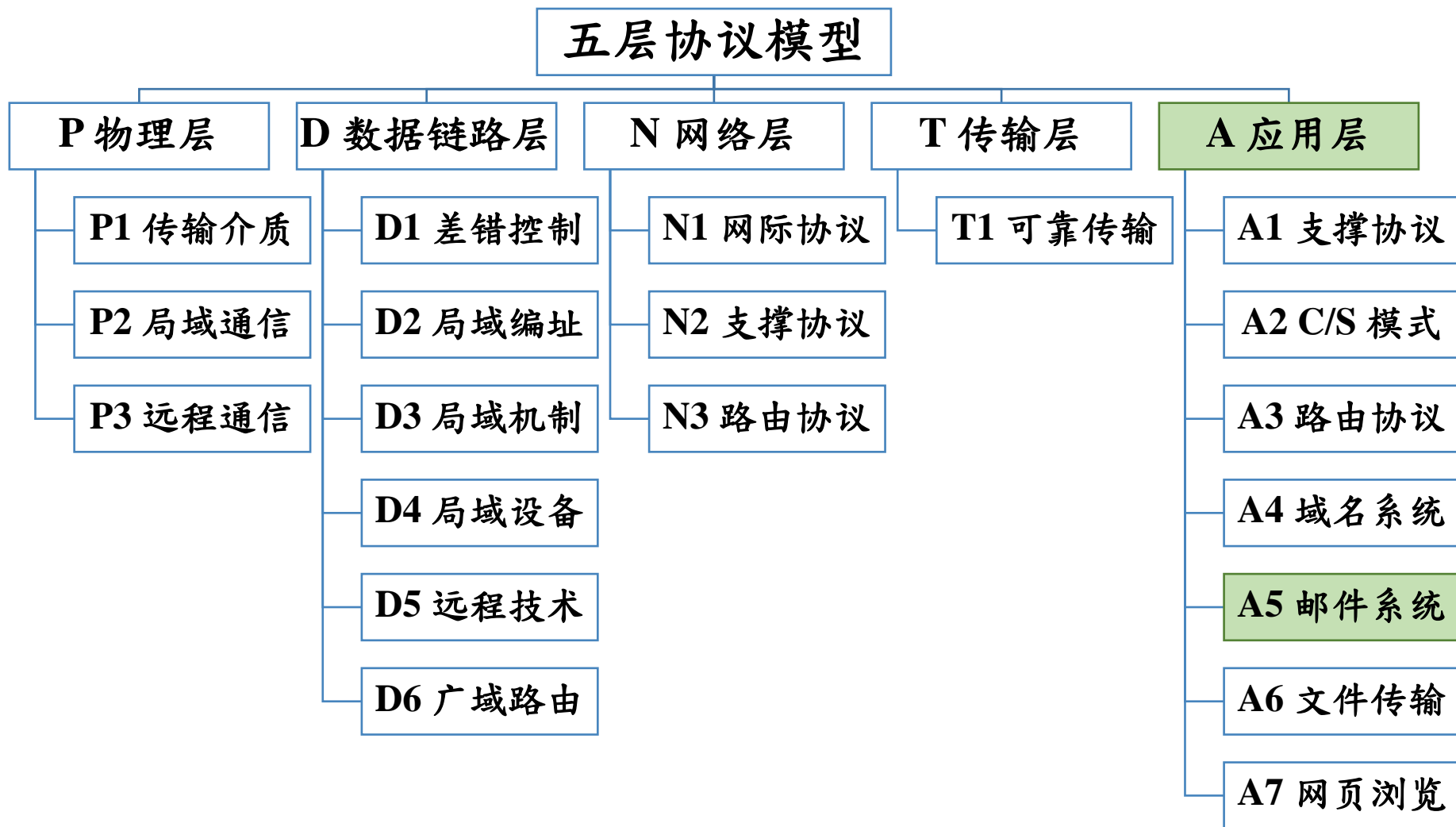


廈門大學
XIAMEN UNIVERSITY



信息学院 黃 燁
(国家示范性软件学院) 博士, 副教授
School of Informatics Dr. Wei Huang

知识框架



主要内容

- 电子邮件的格式
- 主要构成：MTA、MUA、MDA
- 主要协议（作用、原理、端口号）
 - 电子邮件的传输：SMTP
 - 电子邮件的传输扩展：MIME
 - 电子邮件的访问：POP3，IMAP

对应课本章节

- **PART I Introduction And Internet Applications**
 - **Chapter 4 Traditional Internet Applications**
 - **4.12 Electronic Mail**
 - **4.13 The Simple Mail Transfer Protocol (SMTP)**
 - **4.14 ISPs, Mail Servers, And Mail Access**
 - **4.15 Mail Access Protocols (POP, IMAP)**
 - **4.16 Email Representation Standards (RFC2822, MIME)**

内容纲要

1	电子邮件概述
2	电子邮件格式
3	邮件代理和协议
4	SMTP协议
5	MIME标准

电子邮件概述

- 电子邮件 (e-mail)
 - 1972年BBN的Ray Tomlinson发明，并采用 @ 符号。
 - 把邮件发送到收件人使用的邮件服务器，并放在其中的收件人邮箱中，收件人可随时上网到自己使用的邮件服务器进行读取。
- 特点
 - 使用方便，传递迅速和费用低廉。
 - 现在电子邮件可传送文字信息，还可附上声音和图像。

常用的邮件服务器和客户端

- 常用邮件服务器（包括但不限于）
 - 国际：GMAIL；163，126，TOM
 - 单位：XMU，IEEE
- 常用邮件客户端（包括但不限于，非广告）
 - 国外：Outlook，Thunderbird
 - 国内：Foxmail，QQMail，尚邮
 - Android手机：K9Mail

内容纲要

1	电子邮件概述
2	电子邮件格式
3	邮件代理和协议
4	SMTP协议
5	MIME标准

电子邮件地址

- 电子邮件系统要求每个用户有一个地址

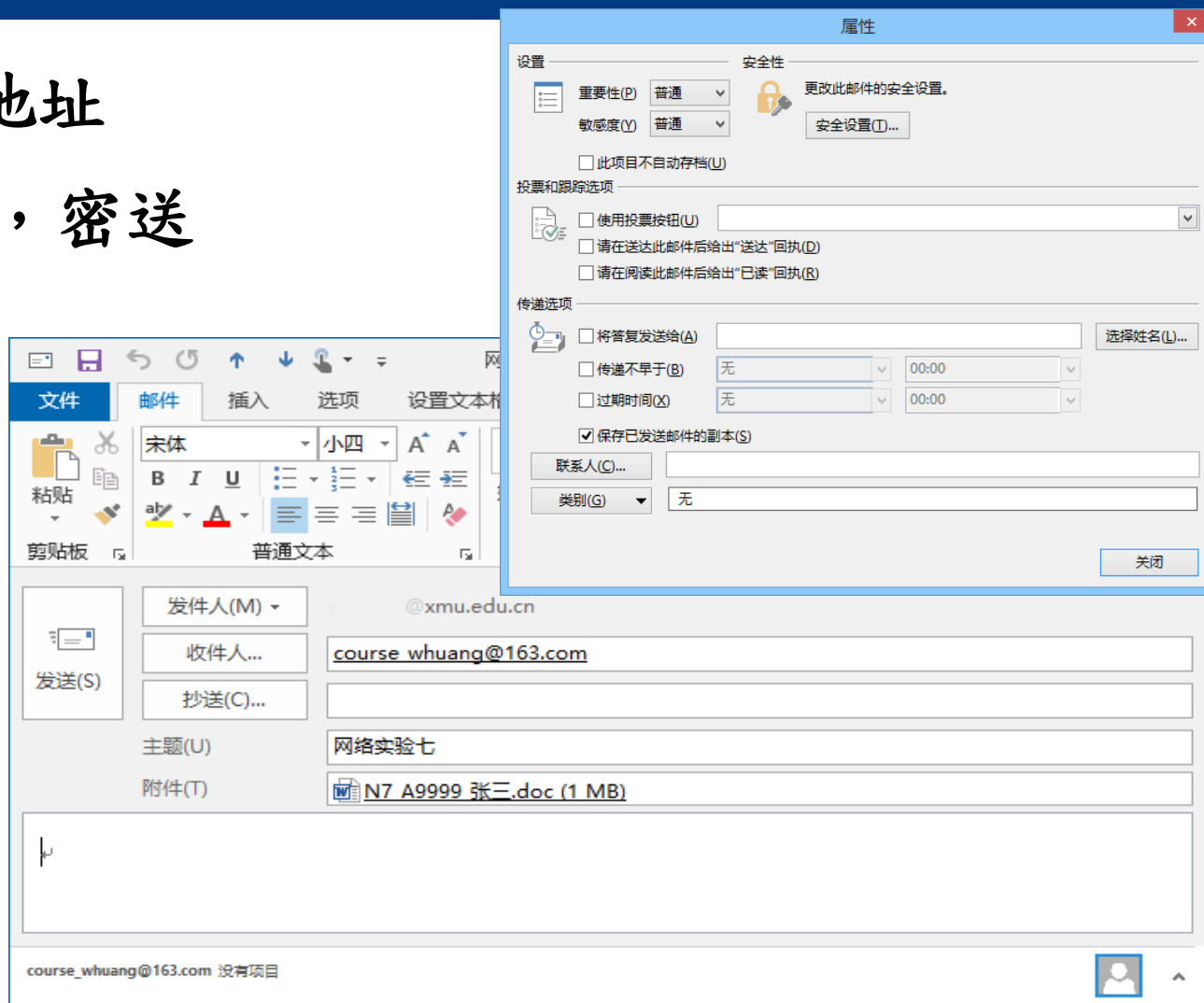
– 形式是：用户名@主机域名

- 这里@念作“at”，意思为“在”。
- 左侧为用户名，在该域名的范围内是唯一的。
- 右侧为邮件服务器所在主机的域名，在全世界必须是唯一的。
- 邮箱所在主机通常称为邮件服务器。

whuang@xmu.edu.cn

电子邮件的组成

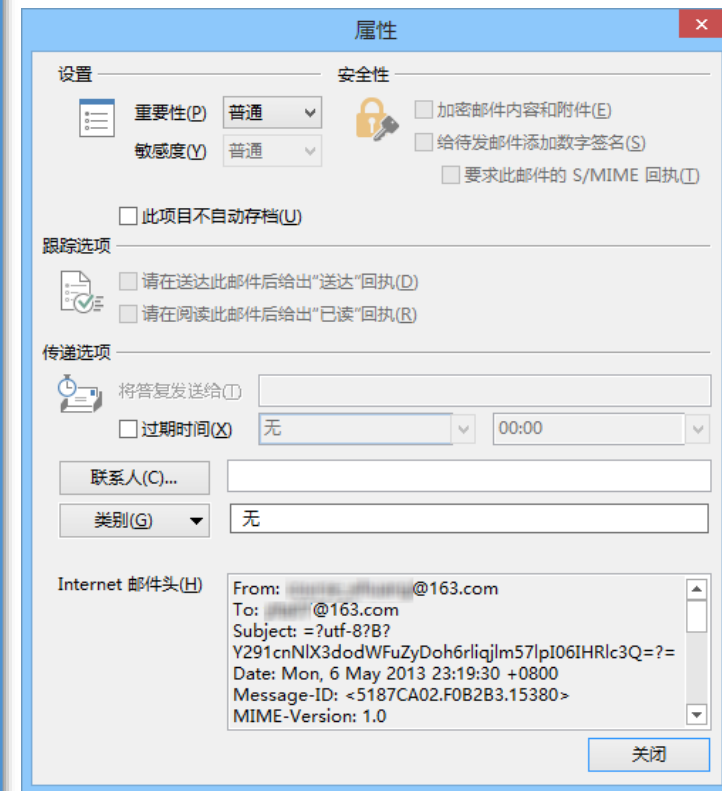
- 发件人，回复地址
- 收件人，抄送，密送
- 时间
- 主题，内容
- 附件
- 扩展属性
 - 是否要求回执
 - 重要性



查看Internet邮件头

- 如图，邮件头说明了什么？

From: *****@163.com
To: *****@163.com
Subject: =?utf-8?B?Y291cnNlX3dodWFuZyDoh6rliqjlm57lpI06IHRlc3Q=?=
Date: Mon, 6 May 2013 23:19:30 +0800
Message-ID: <5187CA02.F0B2B3.15380>
MIME-Version: 1.0
Content-Type: Text/HTML;
charset="utf-8"
X-CM-TRANSID: EMCowEB5p3z_yYdRYrLXAg--.25871S2.re
X-CM-SenderInfo: xfrx22xhbzx3xdqjqiywtou0bp/
Delivered-To: *****@163.com
Errors-to: *****@163.com
Return-path: *****@163.com
References: <000001ce4a6d\$1c1482c0\$543d****\$@163.com>
In-Reply-To:
<000001ce4a6d\$1c1482c0\$543d****\$@163.com>
Content-Transfer-Encoding: base64



内容纲要

1	电子邮件概述
2	电子邮件格式
3	邮件代理和协议
4	SMTP协议
5	MIME标准

三种代理

- 邮件用户代理 (Mail User Agent , MUA)
 - 帮助用户读取、编写和回复邮件，再将这些信息转给MTA发送；有Outlook、Foxmail等。
- 邮件传输代理 (Mail Transport Agent , MTA)
 - 把邮件由一个服务器传到另一个服务器或邮件投递代理。
- 邮件投递代理 (Mail Delivery Agent , MDA)
 - 将MTA接收的邮件，根据收件人地址投放到用户的邮箱里。在投放过程中，还可以进行邮件过滤、自动回复等功能。

邮件协议

- 常见的电子邮件协议
 - SMTP (Simple Mail Transfer Protocol , 简单邮件传输协议)
 - POP3 (邮局协议)
 - IMAP4 (Internet邮件访问协议)
- 表示邮件的协议
 - MIME : 扩展了电子邮件标准 , 使其能支持非ASCII字符、二进制附件等多种格式的邮件消息。
- 上网访问邮箱则通过HTTP或HTTPS协议

电子邮件最主要的组成构件

- 用户代理

– MUA, MTA, MDA

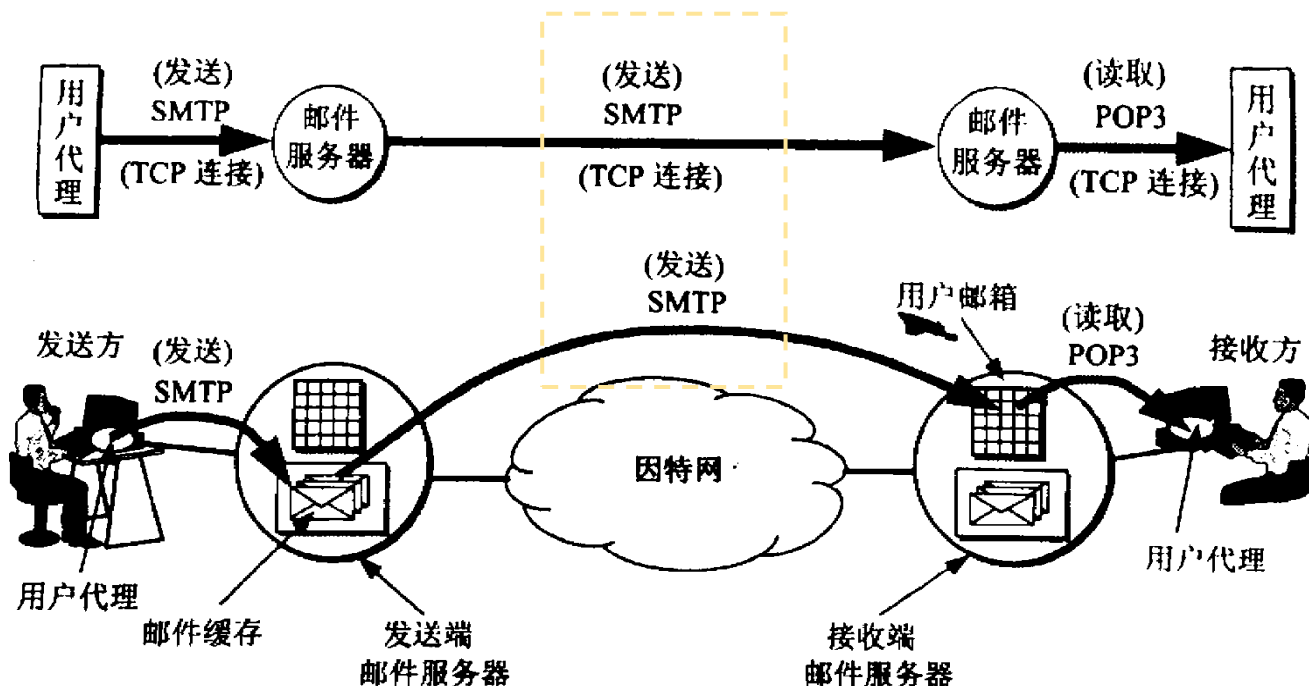


图 12-8 电子邮件的最主要的组成构件

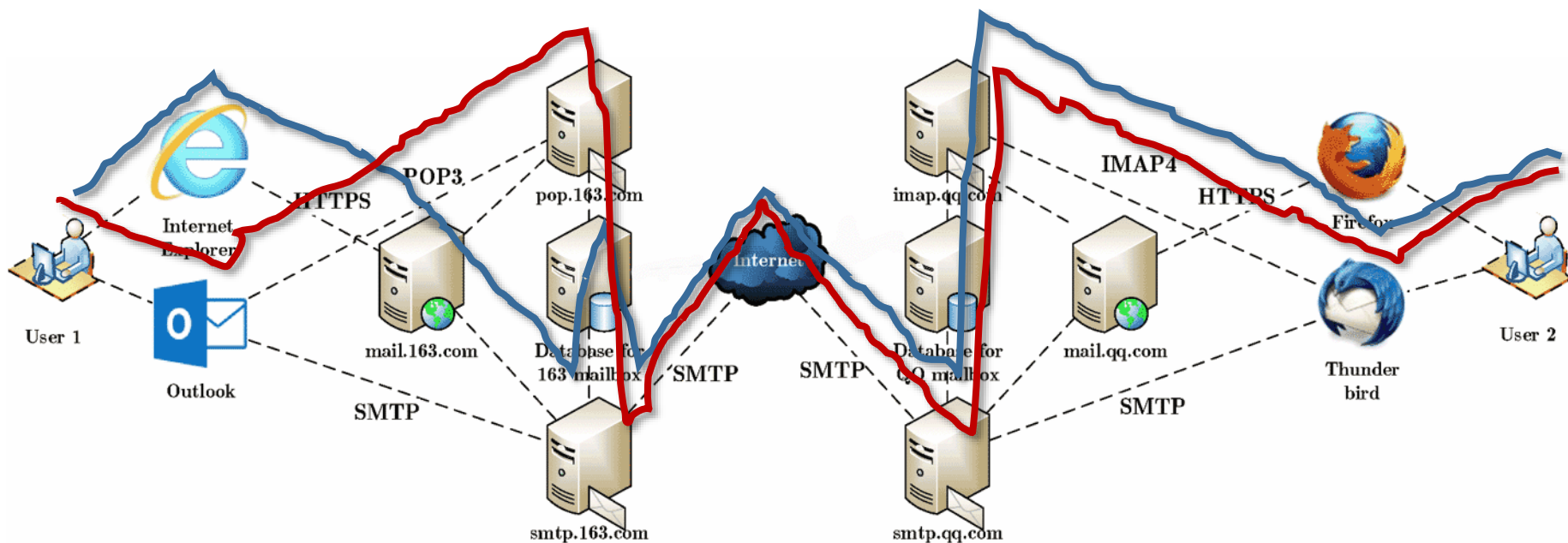
电子邮件软件

- 电子邮件软件分为两个概念模块：
 - 电子邮件接口应用
 - 一种机制，为用户撰写和编辑传出的消息，以及读取和处理传入的电子邮件
 - 邮件传输程序
 - 作为客户端向目标计算机上的邮件服务器发送邮件；
 - 邮件服务器接受传入消息并将每个邮件存放在相应的用户邮箱中。
- 电子邮件的规格可分为三大类：传输、访问和表达

邮件服务器

- 邮件服务器需要使用发送和读取两个不同的协议
 - 功能：发送和接收邮件，同时还要向发信人报告邮件传送的情况（已交付、被拒绝、丢失等）。
- 一个邮件服务器既可以作为客户，也可以作为服务器。
 - 当邮件服务器 A 向另一个邮件服务器 B 发送邮件时，邮件服务器 A 就作为 SMTP 客户，而 B 是 SMTP 服务器。
 - 当邮件服务器 A 从另一个邮件服务器 B 接收邮件时，邮件服务器 A 就作为 SMTP 服务器，而 B 是 SMTP 客户。

电子邮件如何传输



发送和接收电子邮件的重要步骤

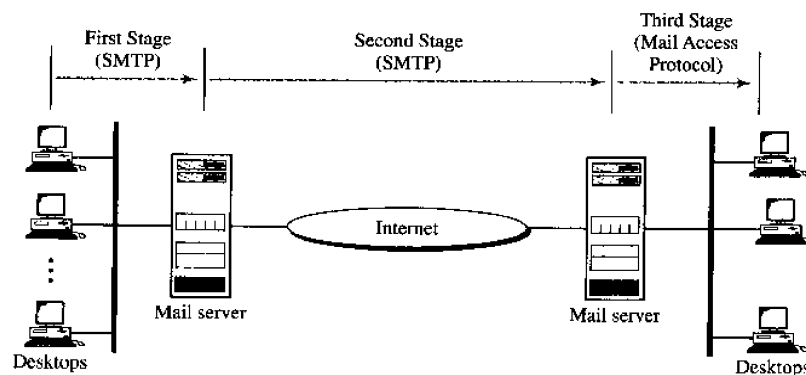
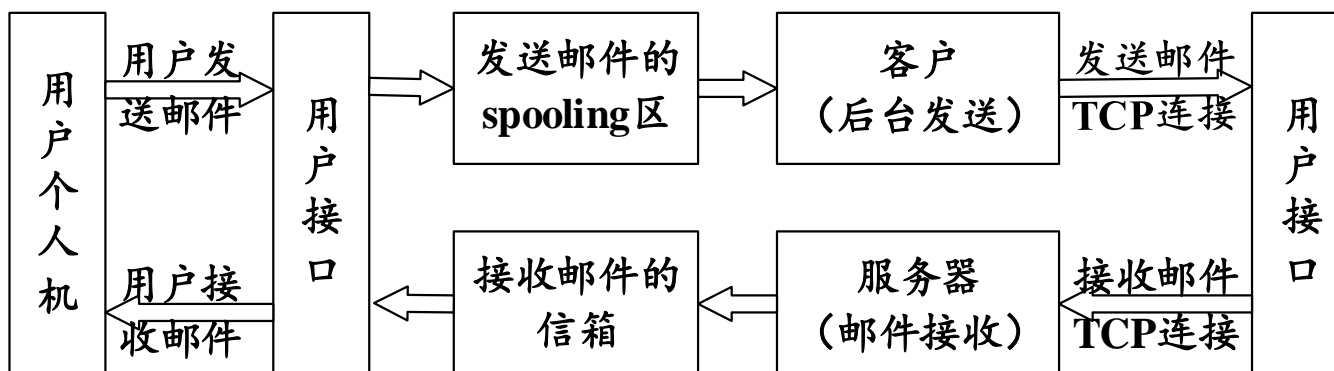
- 发件人调用用户代理撰写和编辑要发送的邮件。
- 发件人的用户代理把邮件用 SMTP 协议发给发送方邮件服务器，
- SMTP 服务器把邮件临时存放在邮件缓存队列中，等待发送。
- 发送方邮件服务器的 SMTP 客户与接收方邮件服务器的 SMTP 服务器建立 TCP 连接，然后就把邮件缓存队列中的邮件依次发送出去。

发送和接收电子邮件的重要步骤

- 运行在接收方邮件服务器中的SMTP服务器进程收到邮件后，把邮件放入收件人的用户邮箱中，等待收件人进行读取。
- 收件人在打算收信时，就运行PC机中的用户代理，使用POP3（或IMAP）协议读取发送给自己的邮件。
- 请注意，POP3服务器和POP3客户之间的通信是由POP3客户发起的。

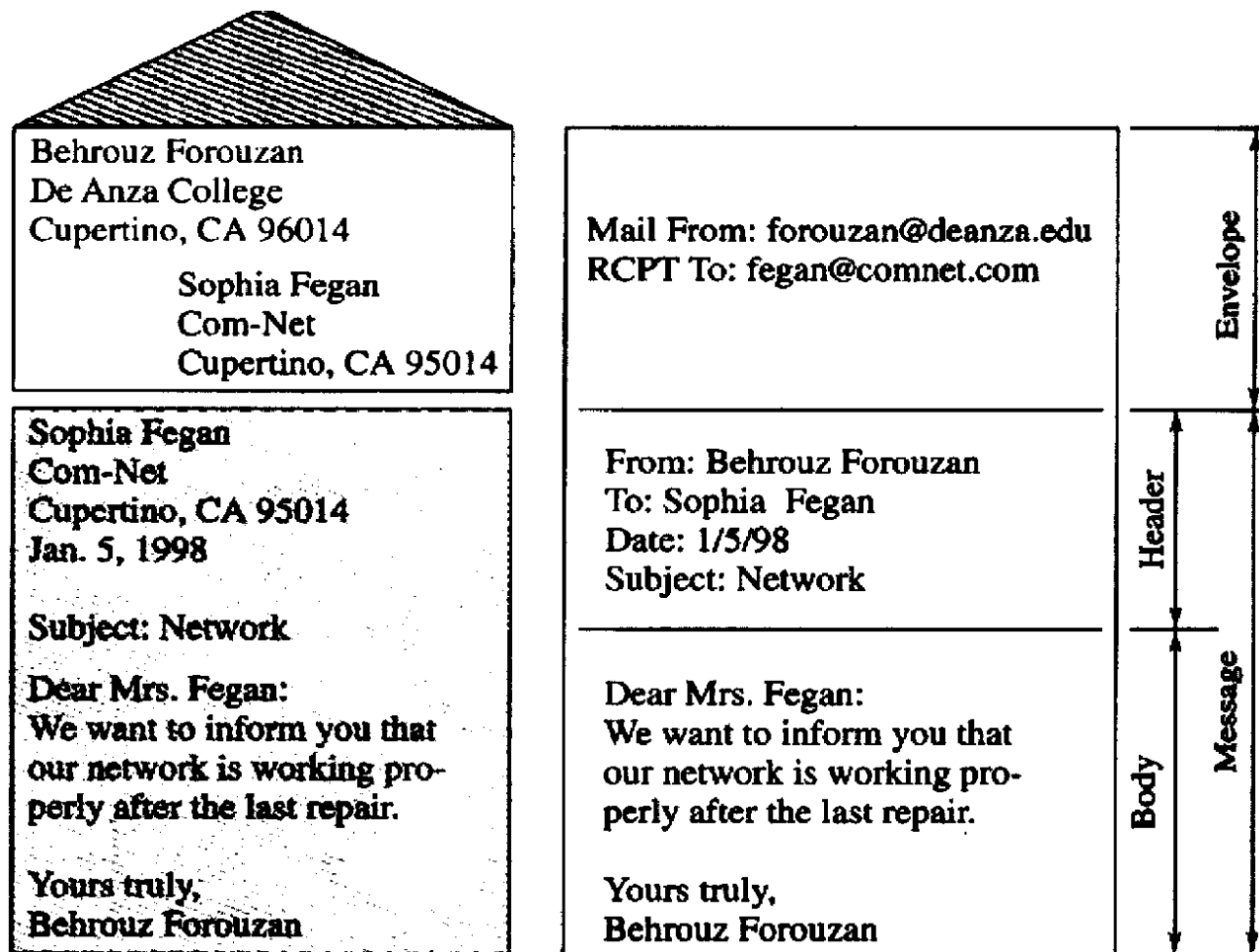
TCP/IP电子邮件系统的模型

- TCP/IP电子邮件系统采用端到端传输方式
 - 发送方的MTA负责将邮件传送到接收方的MTA



邮件格式

- 信封
- 消息
 - 邮件头
 - 邮件主体



内容纲要

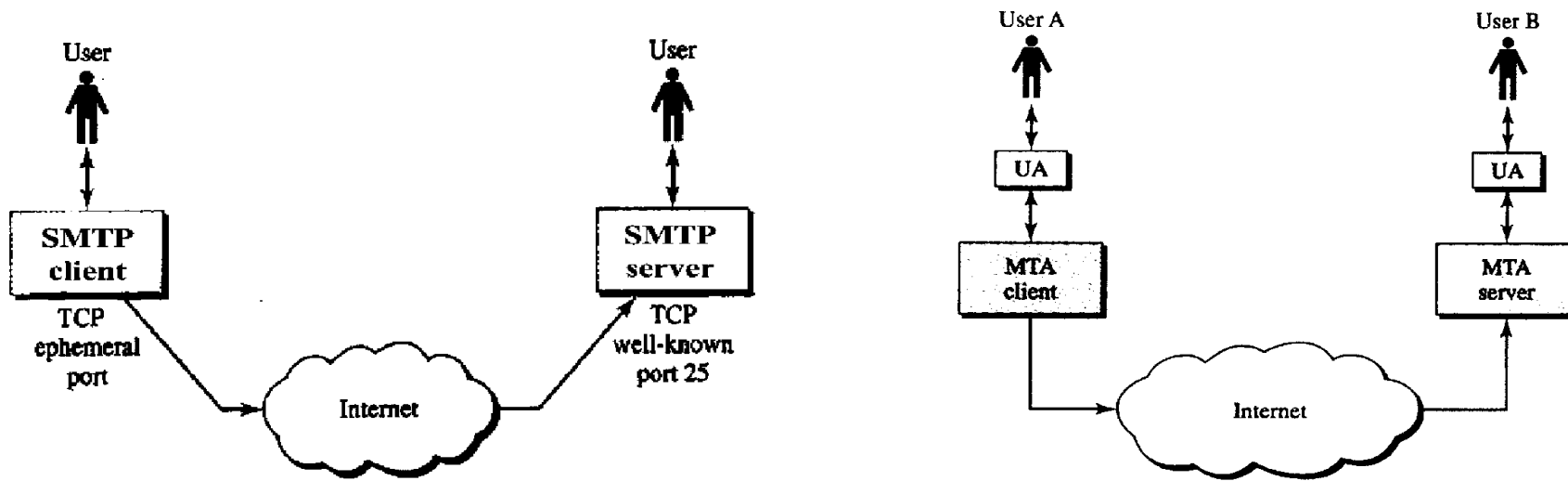
	2	电子邮件格式
	3	邮件代理和协议
	4	SMTP协议
	5	MIME标准
	6	POP3协议

简单邮件传输协议 (SMTP)

- 简单邮件传输协议 (Simple Mail Transfer Protocol)

- 邮件传输程序使用的标准协议，规定在两个相互通信的 SMTP 进程之间应如何交换信息。

- 端口号：25 (明文), 465 (SSL加密)



SMTP

- SMTP可以被描述为
 - 遵循流范式
 - 使用文本控制消息
 - 只传送文本消息
 - 发送一个给定消息的副本
 - 允许客户端列出用户，然后向列表中的所有用户发送消息的单个副本

SMTP通信的三个阶段

- 连接建立

- 连接是在发送主机的 SMTP 客户和接收主机的 SMTP 服务器之间建立的。
- SMTP不使用中间的邮件服务器。

- 邮件传送

- 邮件接收后，收件人并不需要马上阅读

- 连接释放

- 邮件发送完毕后，SMTP 应释放 TCP 连接。

SMTP

• 使用客户服务器方式

— 负责发送邮件的
SMTP 进程是客户

— 负责接收邮件的
SMTP 进程是服
务器

```
Server: 220 somewhere.com Simple Mail Transfer Service Ready
Client: HELO example.edu
Server: 250 OK
Client: MAIL FROM:<John_Q_Smith@example.edu>
Server: 250 OK
Client: RCPT TO:<Mathew_Doe@somewhere.com>
Server: 550 No such user here
Client: RCPT TO:<Paul_Jones@somewhere.com>
Server: 250 OK
Client: DATA
Server: 354 Start mail input; end with <CR><LF>.<CR><LF>
Client: ...sends body of mail message, which can contain
Client: ...arbitrarily many lines of text
Client: <CR><LF>.<CR><LF>
Server: 250 OK
Client: QUIT
Server: 221 somewhere.com closing transmission channel
```

SMTP命令列表

命令	描述
DATA	开始信息写作
EXPN<string>	验证给定的邮箱列表是否存在，扩充邮箱列表，也常被禁用
HELO<domain>	向服务器标识用户身份，返回邮件服务器身份
HELP<command>	查询服务器支持什么命令，返回命令中的信息
MAIL FROM<host>	在主机上初始化一个邮件会话
NOOP	无操作，服务器应响应OK
QUIT	终止邮件会话
RCPT TO<user>	标识单个的邮件接收人；常在MAIL命令后面可有多多个rcpt to：
RSET	重置会话，当前传输被取消
SAML FROM<host>	发送邮件到用户终端和邮箱
SEND FROM<host>	发送邮件到用户终端
SOML FROM<host>	发送邮件到用户终端或邮箱
TURN	接收端和发送端交换角色
VRFY<user>	用于验证指定的用户/邮箱是否存在；由于安全方面的原因，服务器常禁止此命令

别名扩展

- 邮件转发器允许报文副本再发到多个目的地
 - 通常，服务器会查询一个小型的邮件别名数据库，把传入的收件方地址映射为一组地址 A，然后再向 A 中的每个地址转发一份副本。
- 别名映射可以是“多对一”或“一对多”

命令和响应

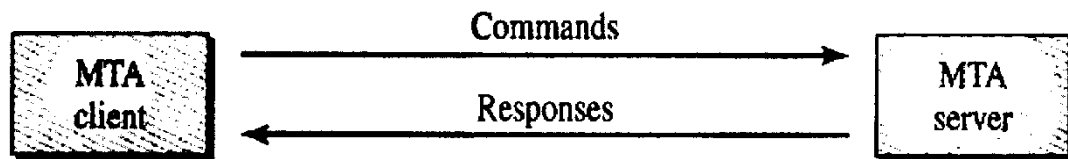
- 用命令和响应在MTA客户端和MTA服务器传输消息
- 命令格式：KEYWORD: INFO，命令或答复由回车结束

- 示例：

HELO: xyz.xmu.edu.cn

MAIL FROM: abc@xyz.xmu.edu.cn

RCPT TO: def@ghijk.com



- 响应从服务器发送到客户端。
 - 响应是三位数的代码，后面可能有额外的文本信息。
- 三个阶段：连接建立，邮件传输和连接终止。

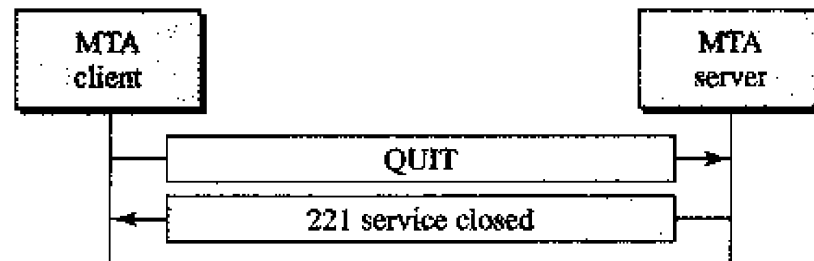
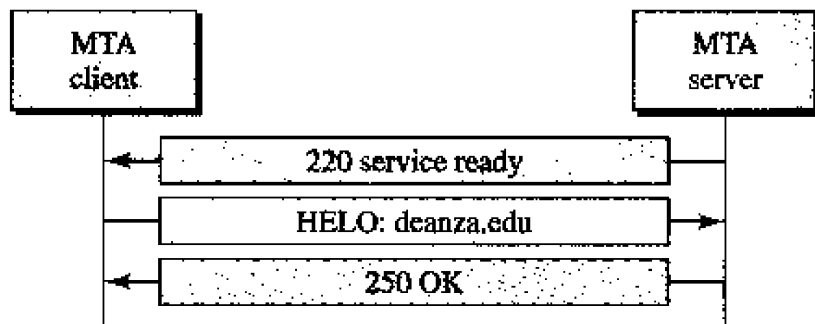
邮件传输阶段

- 建立连接

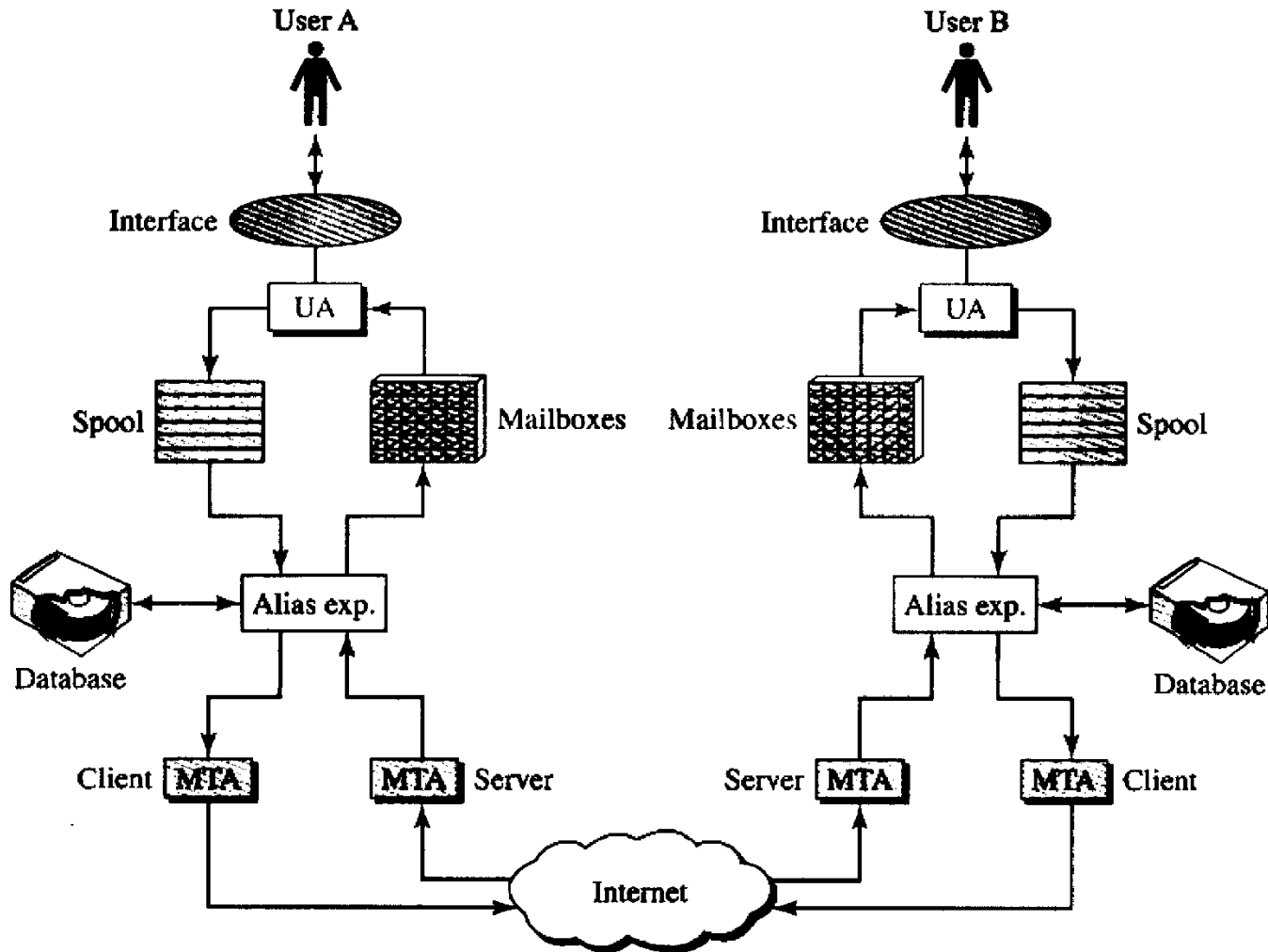
- 在客户端与已知端口25进行TCP连接后，SMTP服务器启动连接阶段。

- 连接终止

- 消息成功传输后，客户端终止连接。



完整的邮件系统



内容纲要

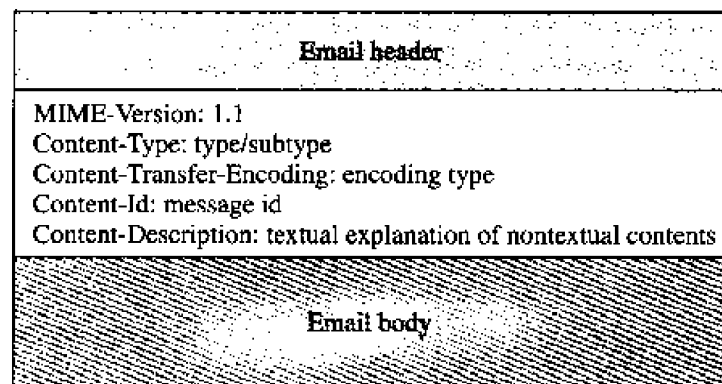
	3	邮件代理和协议
	4	SMTP协议
	5	MIME标准
	6	POP3协议
	7	IMAP4协议

多用途互联网邮件扩展 (MIME)

- **Multipurpose Internet Mail Extensions**
 - MIME是允许SMTP发送的非ASCII数据的补充协议。
 - 不是邮件协议，不能替代SMTP。
 - 仅对SMTP扩展。在接收端服务器的SMTP接收NVT（网络虚拟终端）ASCII数据并将其传递到MIME被转换回原始数据。

MIME头

- 在原始SMTP头部分添加五个标头以定义传输参数
 - MIME-Version 标识MIME的版本
 - Content-Type 说明邮件的性质
 - Content-Transfer-Encoding 主体是如何编码的
 - Content-Id 邮件的唯一标识符
 - Content-Description
内容描述



MIME头的参数值

- MIME-Version: 1.0
- Content-Type: [type]/[subtype]; parameter
 - Text: 用于标准化地表示的文本信息，可以多种字符集；
 - Multipart: 用于连接消息体的多个部分构成一个消息；
 - Application: 用于传输应用程序数据或者二进制数据；
 - Message: 用于包装一个E-mail消息；
 - Image: 用于传输静态图片数据；
 - Audio: 用于传输音频或者音声数据；
 - Video: 用于传输动态影像数据，音视频。

MIME头的参数值

- Content-Type: [type]/[subtype]; parameter
 - multipart/mixed : 存在附件
 - multipart/related : 存在内嵌资源
 - multipart/alternative : 只有纯文本与超文本正文
 - application/xhtml+xml : XHTML文档
 - image/jpeg
- Content-Transfer-Encoding: [mechanism]
 - 7bit, 8bit, binary, quoted-printable, base64

MIME的编码

- MIME编码格式
- 支持的文件类型

Type	Description
7bit	NVT ASCII characters and short lines
8bit	Non-ASCII characters and short lines
binary	Non-ASCII characters with unlimited-length lines
Base64	6-bit blocks of data are encoded into 8-bit ASCII characters
Quoted-printable	Non-ASCII characters are encoded as an equal sign followed by an ASCII code

Table 22.3 Data types and subtypes in MIME

Type	Subtype	Description
Text	Plain	Unformatted text
Multipart	Mixed	Body contains ordered parts of different data types
	Parallel	Same as above, but no order
	Digest	Similar to Mixed, but the default is message/RFC822
	Alternative	Parts are different versions of the same message
Message	RFC822	Body is an encapsulated message
	Partial	Body is a fragment of a bigger message
	External-Body	Body is a reference to another message
Image	JPEG	Image is in JPEG format
	GIF	Image is in GIF format
Video	MPEG	Video is in MPEG format
Audio	Basic	Single channel encoding of voice at 8 KHz
Application	PostScript	Adobe PostScript
	Octet-stream	General binary data (eight-bit bytes)

Base64 编码

- Base64

- 当最高比特不一定为零时发送字节数据的一种解决方案。
- 将数据转换为可打印字符。
- 过程将二进制数据分成24位块。
- 每个块分为4个部分，每一个由6位。
- 每一个6位部分解释为一个可打印字符。

Base64 编码过程

- 传递24bit : 01001001 00110001 00111001

- 先划分为4个6bit组，即：

010010 010011 000100 111001

(18:S) (19:T) (4:E) (57:5)

- 对应的base64编码为：STE5。其ASCII码为：

01010011 01010100 01000101 00110101

- 三个字符用四个字符传，适用于ASCII码不多的情况。

- 码表

- 0-25 : A-Z ; 26-51 : a-z ; 52-61 : 0-9 ;

- 62 : + , 63 : /

MIME示例

```
From: “=?gb18030?*****=?=” <*****@***.com>
To: “=?gb18030?B?*****3dodWFuZw===?=” <*****@***.com>
Subject: =?gb18030?B?u9i*****do=?=
=?gb18030?B?*****k=?=
Mime-Version: 1.0
Content-Type: multipart/mixed;
    boundary="----=_NextPart_*****" (Mime_separator)
Content-Transfer-Encoding: 8Bit
Date: Sat, 11 May 2013 13:16:33 +0800
X-Priority: 3
Message-ID: <*****@*****.com>
*****
This is a multi-part message in MIME format.
-----=_NextPart_*****
Content-Type: multipart/alternative;
    boundary="----=_NextPart_*****";

-----=_NextPart_*****
Content-Type: text/plain;
    charset="gb18030"
Content-Transfer-Encoding: base64
```

内容纲要

	3	邮件代理和协议
	4	SMTP协议
	5	MIME标准
	6	POP3协议
	7	IMAP4协议

Email 访问

- 互联网服务供应商开始提供电子邮件服务
- 电子邮件访问如下两种形式之一
 - 专用电子邮件接口应用访问电子邮件
 - 网页的网络浏览器
- 访问协议与传输协议不同
 - 访问只涉及单个用户与单个邮箱交互
 - 传输协议允许用户向其他用户发送邮件

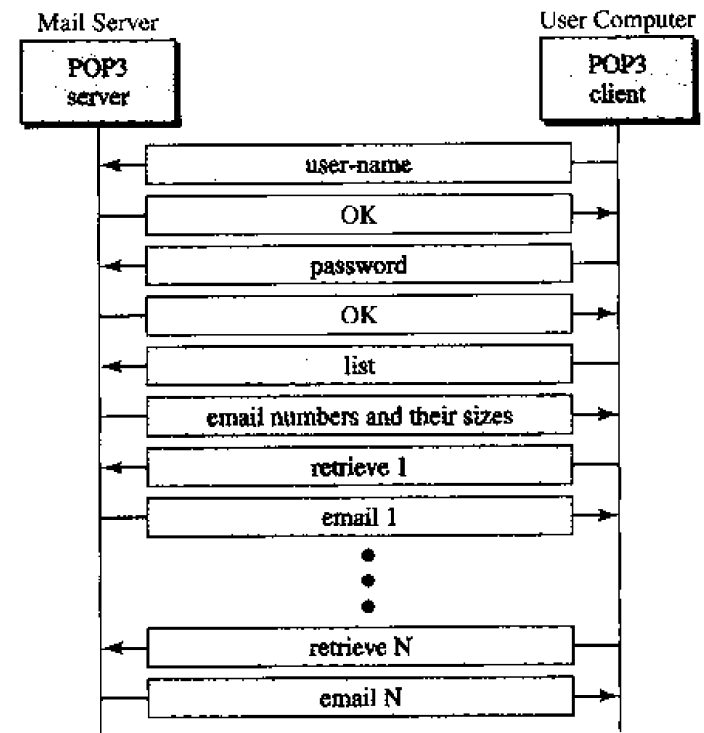
Email 访问

- 接入协议的特点
 - 提供对用户邮箱的访问
 - 允许用户查看标头、下载、删除或发送消息
 - 客户端运行在用户的个人计算机上
 - 服务器运行在存储用户邮箱的计算机上
- 查看没有下载邮件内容的邮件列表
 - 特别是在双方连接缓慢的情况下
 - 例：手机用户可查看头部和删除垃圾邮件而不必下载内容

邮局协议3 (POP3)

- Post Office Protocol (POP)

- POP3允许用户通过PC机动态地检索邮件服务器上的邮件
- 下载邮件，删除邮件
- 端口号：110（明文），
995（POP3S，with SSL）



POP3收邮件过程

- POP3服务器向客户发送一行欢迎词，进入授权状态。
- 授权状态
 - 客户发送USER命令给出用户在邮件服务器上的邮箱名，若是合法用户，服务器回答“+OK”。
 - 客户再发送PASS命令给出口令。POP3服务器确定用户是否有权访问该邮箱，若有权访问，服务器再次回答“+OK”，若是非法用户，服务器回答“-ERR”。
 - USER和PASS命令用口令方式对用户进行授权验证。

POP3收邮件过程

- 事务状态

- 若对用户的授权验证成功，则服务器申请资源与用户的邮箱关联，进入事务(transaction)状态
- 在事务状态，服务器将存储的邮件，从1开始编号；客户可发送命令检索(RETR)、删除(DELE)(作删除标记)项目等。

- 更新状态

- 客户发QUIT命令，进更新(update)状态，删去标记的邮件
- 关闭TCP连接，服务器释放资源，POP3会话结束

POP3 会话实例

```
+OK Welcome to coremail Mail Pop3 Server (163coms[...  
AUTH  
-ERR Not support ntlm auth method  
CAPA  
+OK Capability list follows TOP USER PIPELINING...  
USER y****7  
+OK core mail  
PASS *****  
+OK 7 message(s) [231904 byte(s)]  
STAT  
+OK 7 231904  
LIST  
+OK 7 231904 1 9854 2 140937 3 10224 4 17139 ...  
UIDL  
+OK 7 231904 1 1tbiMAkaCFEAER4pHgAAsC 2 1tbiNQuc...  
RETR 7  
+OK 1698 octets  
Received: from [192.168.7.128] (unknown [119.233.1...  
f0uMueDze6FWdWBSy3zL/K1nBCoAwR8FFrAhNS4gN8AhsIy++3...  
QUIT  
+OK core mail
```


POP3 会话实例

Received: from [192.168.7.128] (unknown [119.233.192.167])
by smtp13 (Coremail) with SMTP id *****fVdw9pFRM59yBA--.***S2;
Tue, 14 May 2013 16:31:45 +0800 (CST)
X-Coremail-DSSMTP: 119.233.192.167
Message-ID: <5***F670.5*****@163.com>
Date: Tue, 14 May 2013 16:31:44 +0800
From: W** H**** <*****@163.com>
User-Agent: Mozilla/5.0 (Windows NT 5.1; rv:17.0) Gecko/20130328 Thunderbird/17.0.5
MIME-Version: 1.0
To: *****@163.com
Subject: Test a rar attach
Content-Type: multipart/mixed;
boundary="-----*****20706"
X-CM-TRANSID:*****fVdw9pFRM59yBA--.***S2
X-Coremail-Antispam: *****9EdanIXcx71UUUUU7v73
***** YCTnIWIEvJa73UjIFyTuYvjxU4mFCUUUUU
X-CM-SenderInfo: ***** biJRIIdCE9o8Tv-YAAAsQ

This is a multi-part message in MIME format.

-----*****20706

Content-Type: text/plain; charset=GB2312

POP3 会话实例

Content-Transfer-Encoding: 7bit

-----*****20706

Content-Type: *****;

name="result.rar"

Content-Transfer-Encoding: base64

Content-Disposition: attachment;

filename="result.rar"

*****AAC0ZRtLPytq0IdNQoAIAAAAHJ1
c3VsdC50eHQAsOhgB*****dNHhnB4bx0x86eD5
I+CoS7d*****V4H8e3v5a/WubX52V88wwBSxn5eKz
0sy50mimnT26*****0ab8/VX1YLQJruTj/5jjHmqkKkguvW5kUICb20jP3
VdHj1Ij*****++3MGE3QrE90wHD
R6Ay*****1qVqWp5y9cTE5eQATF*****UhPkYvygn
S5gHFg*****A==

-----*****20706--

内容纲要

	3	邮件代理和协议
	4	SMTP协议
	5	MIME标准
	6	POP3协议
	7	IMAP4协议

因特网邮件访问协议4 (IMAP4)

- **Internet Mail Access Protocol (IMAP)**
 - 三种工作模式：离线、在线和断连方式。
 - 端口号：143 (明文) ， 993 (IMAPS, with SSL)
 - 先身份验证，鉴定登录名和口令，后获得访问权
- **IMAP协议适合使用多台计算机的用户**
 - 让邮件服务器维护一个中心数据库，能够被多台机器访问
 - 不允许用户将邮件下载到自己的计算机上，只能在线访问邮箱，但可以只读邮件的某一部分

IMAP与POP3 的区别

- IMAP4只下载邮件的主题，并不下载内容
- 邮件客户端软件阅读邮件时才下载邮件内容
- 支持维护自己在服务器上的邮件目录
- 支持直接抓取邮件的特定部分（如文本）
- POP3是“脱机”协议，IMAP是联机协议

垃圾邮件（SPAM）

- 不请自来、强行塞入信箱的垃圾邮件
- SPAM的主要特性包括：
 - 未经消费者的同意，与消费者需求不相关
 - 以诈骗的方式骗取邮件地址
 - 攻击性的广告：夸张不实、不健康、钓鱼网站
 - 散布的数量庞大
- 阻拦垃圾邮件的方法
 - 发信来源，关键词、特征匹配

选作作业

- 下载邮件服务器软件，以localhost为域名，新建admin@localhost邮箱，尝试搭建邮件服务器，并用邮件客户端下载
 - Ipswitch IMail Server，WinMail
- 探究邮件服务器和客户端软件各有什么功能
- 用Omnipeek监听收发邮件的数据流

谢谢观看



廈門大學
XIAMEN UNIVERSITY



信息学院 黄 焯
(国家示范性软件学院) 博士, 副教授
School of Informatics Dr. Wei Huang