

电子邮件

理论课程

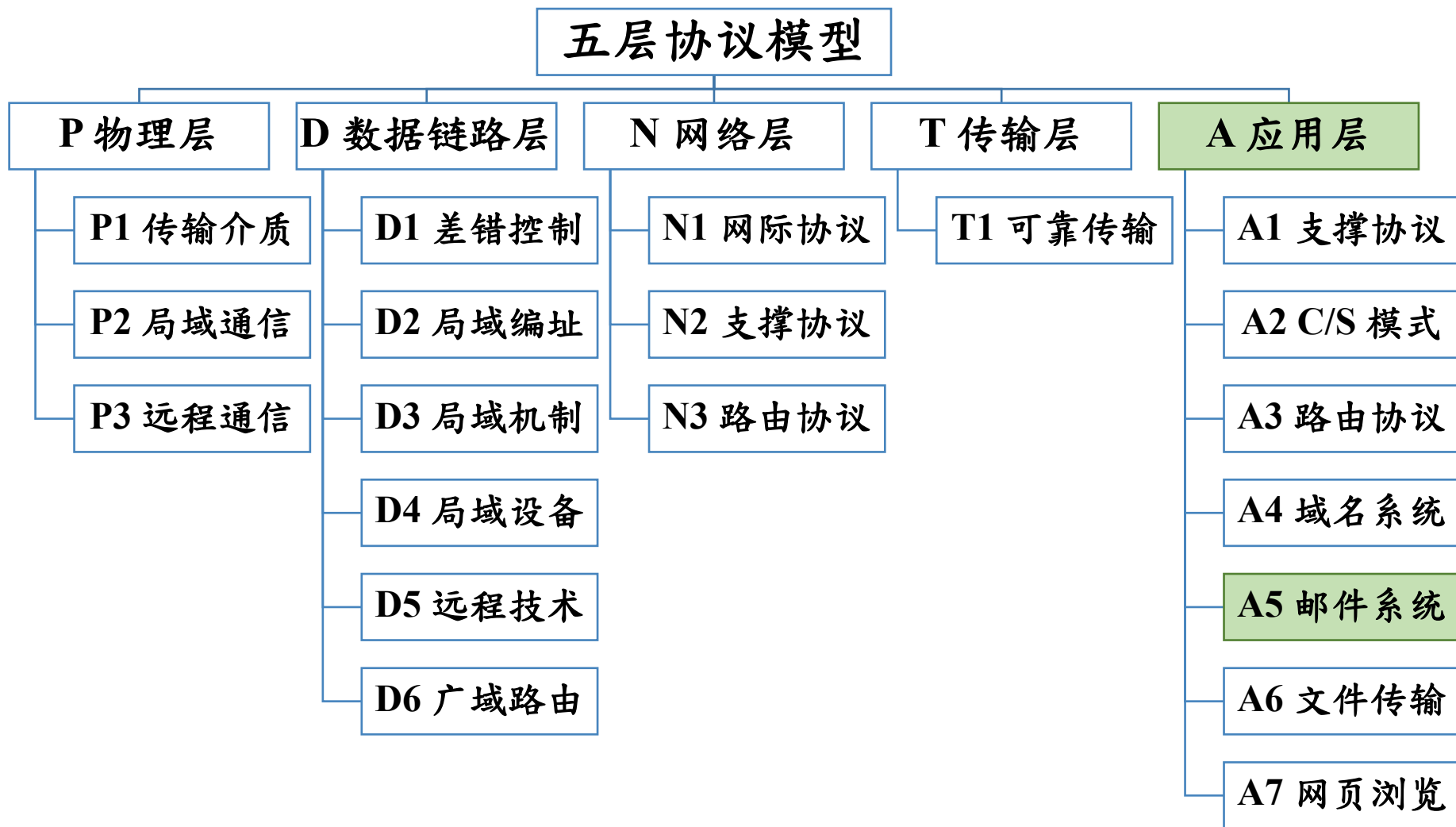


廈門大學
XIAMEN UNIVERSITY



信息学院 黄 焯
(特色化示范性软件学院) 博士, 副教授
School of Informatics 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

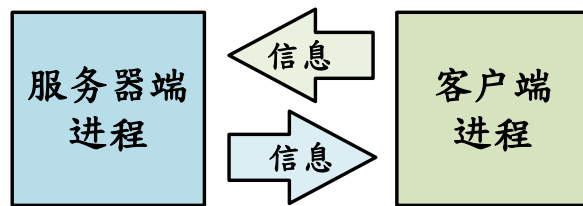
电子邮件概述

- 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的邮件服务器，放在其中的收件人邮箱中，收件人可随时上网到自己的邮件服务器读取。
- 常用的电子邮件服务器和客户端
 - 服务器：163，QQ，XMU，IEEE
 - 客户端：Outlook，Foxmail，QQ邮箱

电子邮件技术

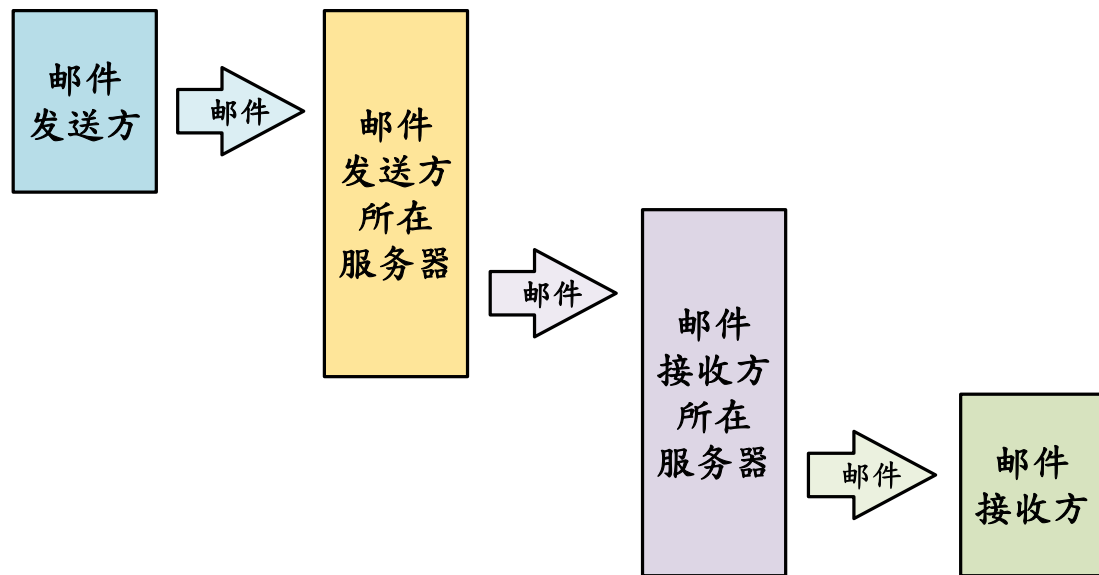
- 传统两个进程通信

- 收发双方必须同时在线



- 电子邮件通信

- 收发双方不必同时在线

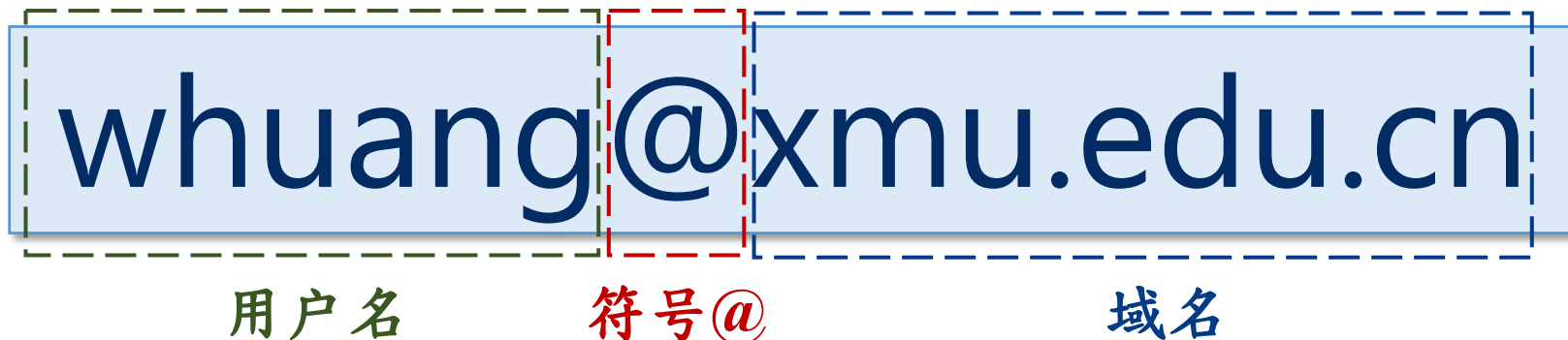


内容纲要

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

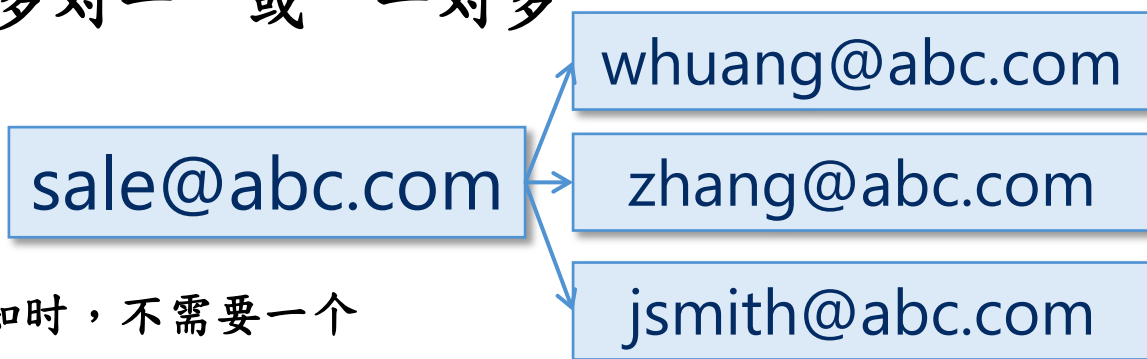
- 电子邮件系统要求每个用户有一个地址



- 邮件域名：全网唯一。
 - 邮箱所在主机，即：邮件服务器。
- 用户名：在该邮件域名的范围内唯一。
- 符号@：念“at”，意思为“在”。

别名扩展

- 别名：为同一邮件服务器上的地址设置不同名称
 - 服务器查询一个邮件别名数据库，把投递到收件方地址的邮件副本投递到别名对应的各个地址。
 - 别名映射可以是“多对一”或“一对多”

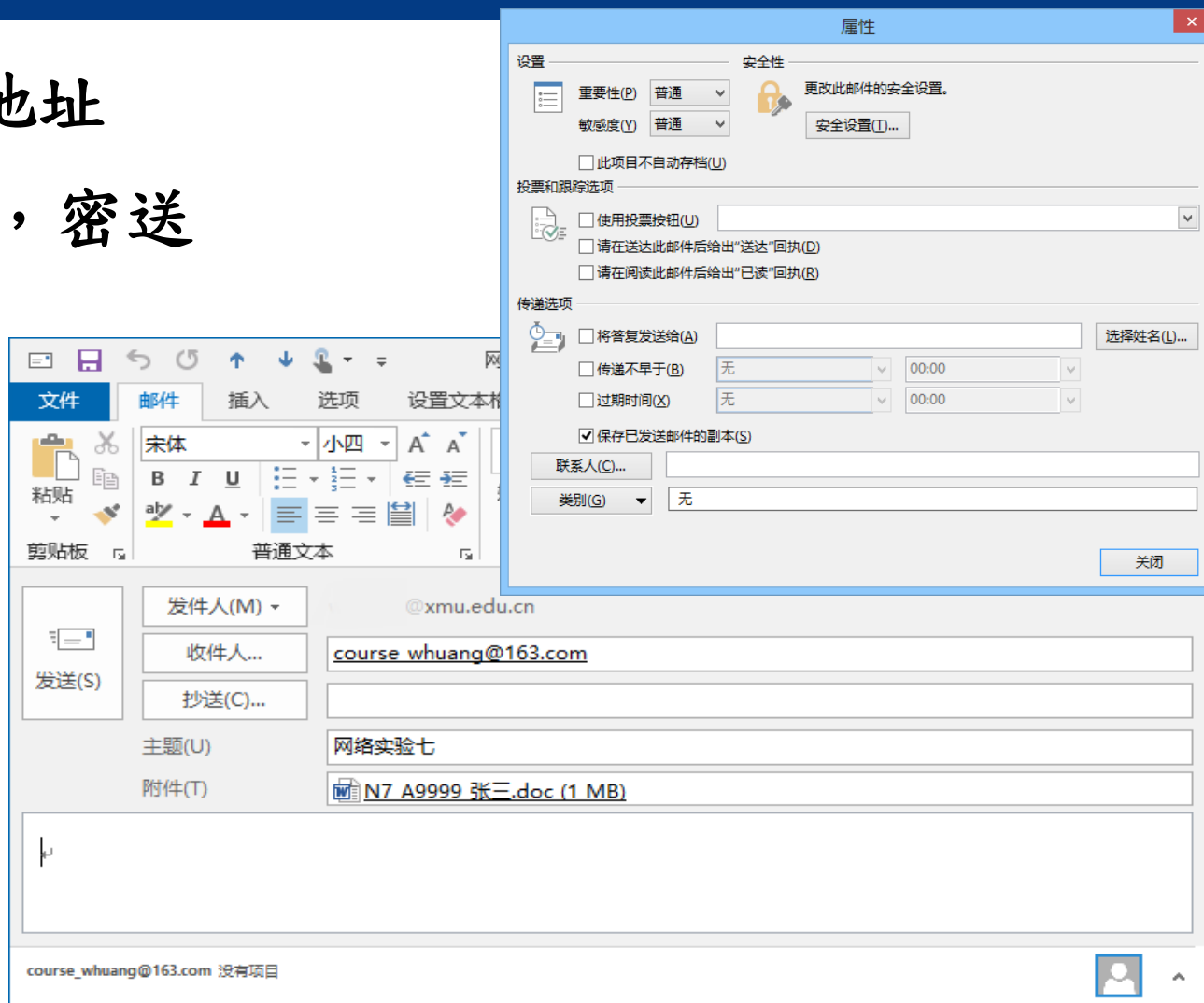


使用场景：

1. 主管对销售部门发送群通知时，不需要一个个粘贴地址，避免遗漏。
2. 客户向销售人员发送邮件时，不需要详细知道企业销售人员的离职情况。
3. 在企业规模较小存在兼职的情况下，一个人有可能暂时兼任多个职位。

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



邮件格式

• 邮件格式

消息的字符集包括ASCII码，非可打印字符使用MIME扩展。

消息头由一组头字段组成，每个头字段包括字段名和值。

消息主体可以是纯文本或HTML格式。

```
Received: from WHUANGXM (unknown [180.233.66.123])
    by mail2 (Coremail) with SMTP id
    BAAi0gAnG0ddcshkDm83BA--.12433S2;
    Tue, 01 Aug 2023 10:47:57 +0800 (CST)
From: "'Wei Huang'" <whuang@xmu.edu.cn>
To: "'John Smith'" <jsmith@xmu.edu.cn>
Cc: "'Secretary Bill Wong'" <bwong@xmu.edu.cn>
Subject: Apply for your manuscript
Date: Tue, 1 Aug 2023 10:47:58 +0800
Message-ID: <005501d9c422$94fa63f0$beef2bd0$@xmu.edu.cn>
MIME-Version: 1.0
Content-Type: text/plain;
    charset="US-ASCII"
Content-Transfer-Encoding: 7bit
Content-Language: zh-cn

Dear John,

Please send me a copy of your manuscript. Thank you.
Sincerely,

Wei Huang
Xiamen University
```

头字段

消息头

消息主体

内容纲要

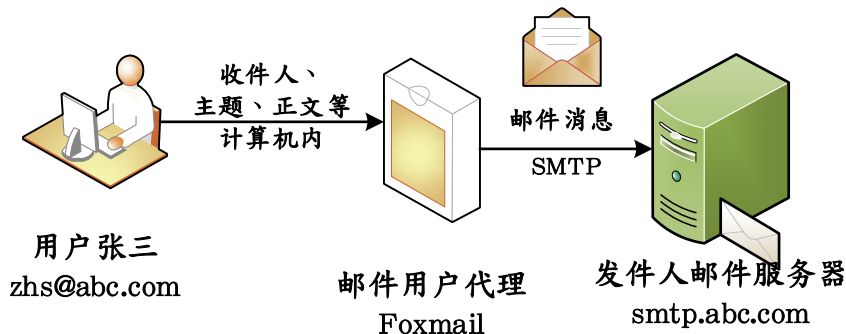
1	电子邮件概述
2	电子邮件格式
3	邮件代理和协议
4	SMTP协议
5	MIME标准

电子邮件系统的主要构件

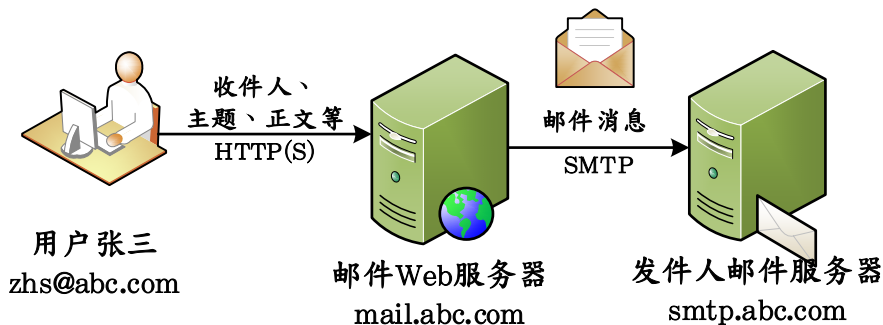
- 发件人将邮件提交至发件人服务器
 - 通过客户端提交
 - 通过网页（WebMail）提交
- 发件人服务器将邮件发送至接收人服务器
- 收件人从收件人邮件收取邮件
 - 通过客户端收取
 - 通过网页（WebMail）收取
- 收发电子邮件的协议基于TCP协议实现

发件人发送邮件

• 通过客户端提交

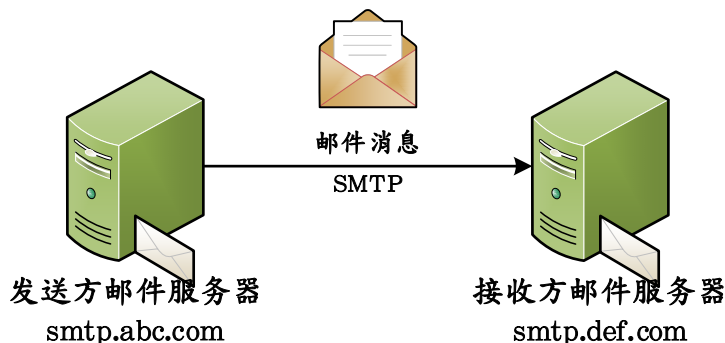


• 通过WebMail提交



发件人服务器发送至接收人服务器

- 不同服务器之间发送



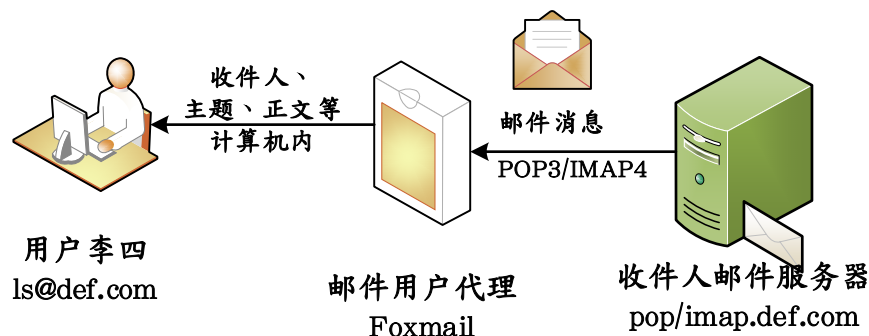
- 同一个邮件服务器内直接发送

- 内部实现

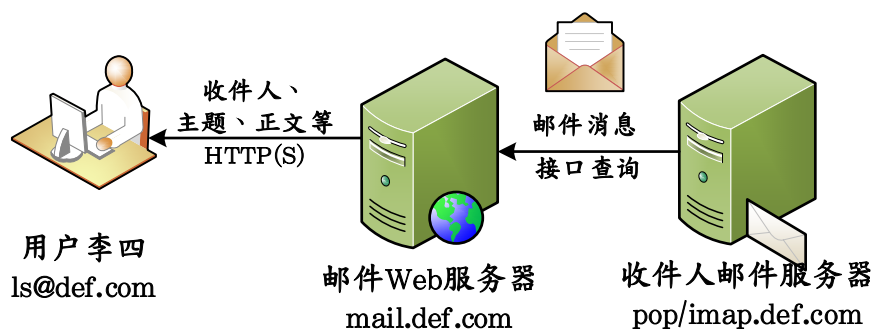
- 服务器把邮件临时存放在邮件缓存队列中，等待发送。
- 接收方邮件服务器中的SMTP服务器进程收到邮件后，把邮件放入收件人的用户邮箱中，等待收件人进行读取。

收件人从收件人邮件收取邮件

• 通过客户端接收



• 通过WebMail查看



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，再将信息转给MTA发送
 - 例如：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。

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。

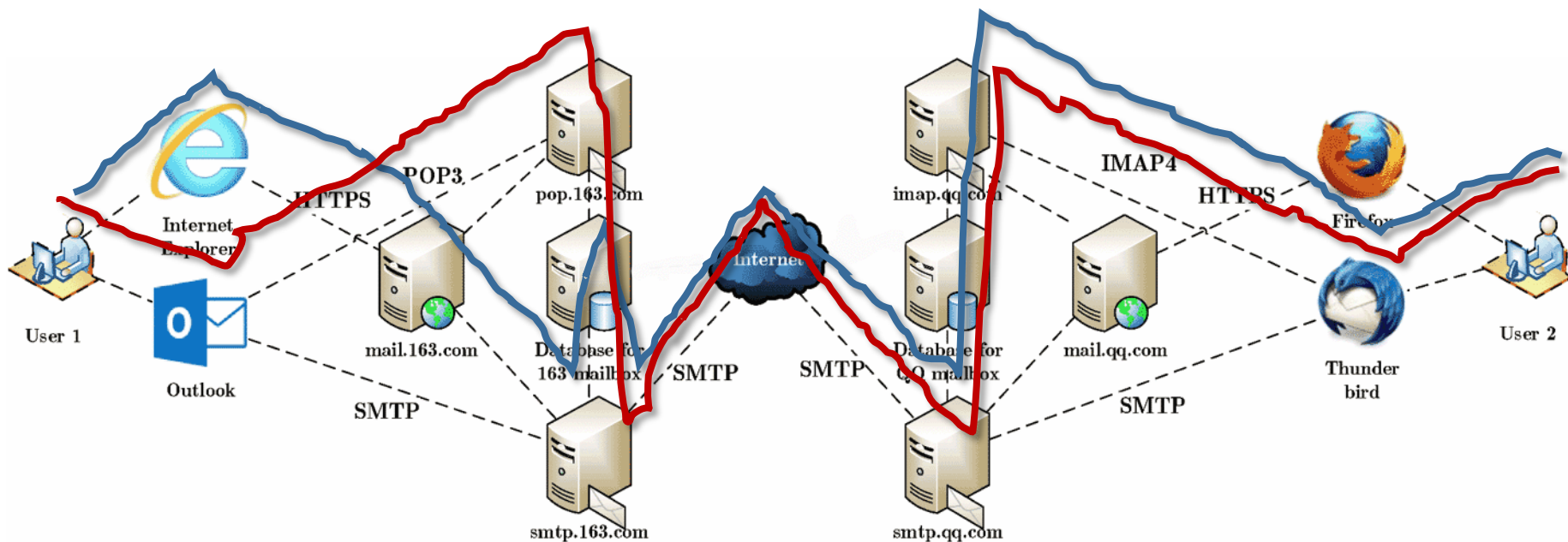
邮件协议

- 传输协议
 - SMTP (Simple Mail Transfer Protocol , 简单邮件传输协议)
- 访问协议
 - POP3 (邮局协议)
 - IMAP4 (Internet邮件访问协议)
- 表示协议
 - 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。
- 网页邮箱协议 : HTTP或HTTPS

邮件服务器

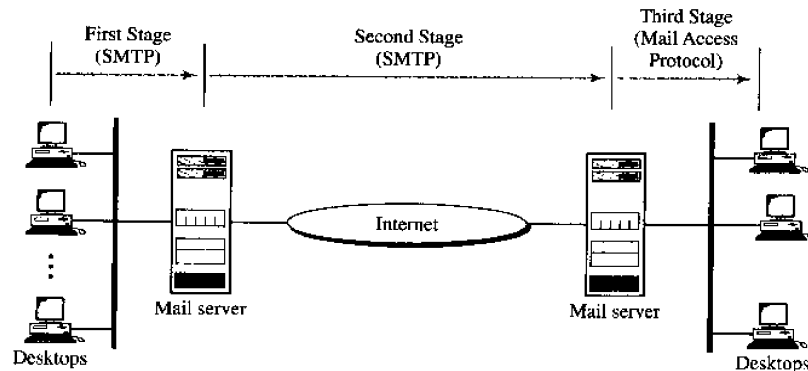
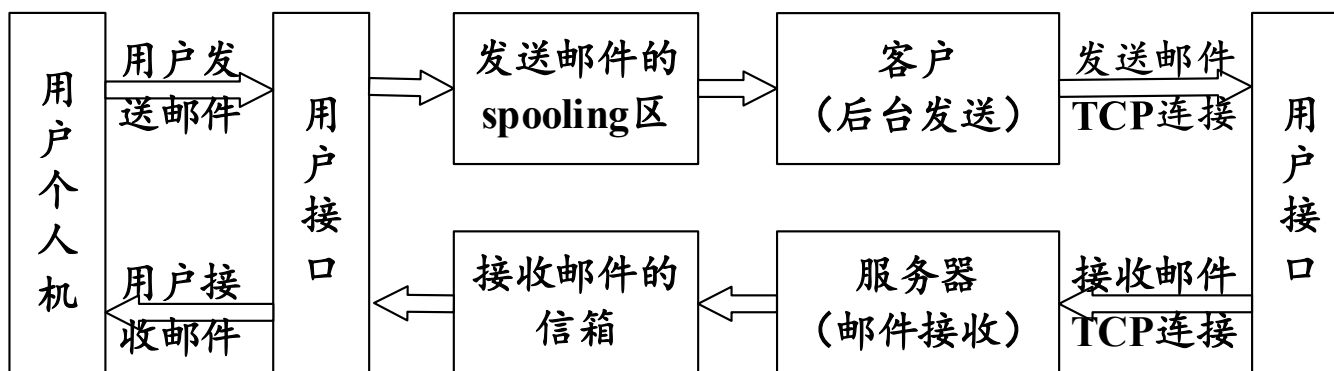
- 邮件服务器需要使用发送和读取两个不同的协议
 - 功能：发送和接收邮件，同时还要向发信人报告邮件传送的情况（已交付、被拒绝、丢失等）。
- 一个邮件服务器既可以作为客户，也可以作为服务器。
 - 当邮件服务器 A 向另一个邮件服务器 B 发送邮件时，邮件服务器 A 就作为 SMTP 客户，而 B 是 SMTP 服务器。
 - 当邮件服务器 A 从另一个邮件服务器 B 接收邮件时，邮件服务器 A 就作为 SMTP 服务器，而 B 是 SMTP 客户。

电子邮件如何传输



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



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。

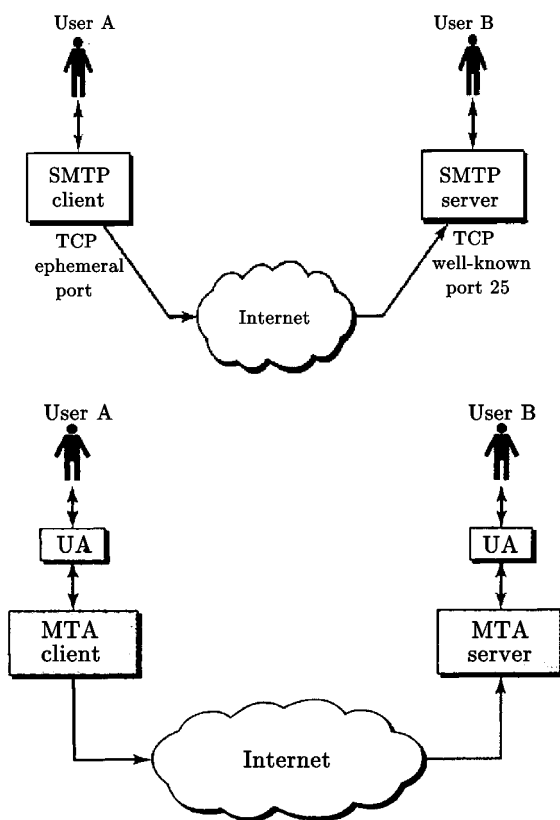
- 默认端口号：25 (明文), 465 (SSL加密)

- 特点

- 基于TCP的流模式

- 只发送ASCII码文本

- 原始SMTP不支持鉴权，现已增强



SMTP通信的三个阶段

- 连接建立

- 发送主机是客户端和接收主机是服务器端。
- 直接发送，不使用中间的邮件服务器。

- 邮件传送

- 邮件接收后，收件人不需要马上阅读

- 连接释放

- 邮件发送完毕后，SMTP 应释放 TCP 连接。

服务器：220 service ready

客户端：HELO: abc.com

服务器：250 OK

服务器：...

客户端：QUIT

服务器：221 service closed

命令和响应

- SMTP用命令和响应在MTA客户端和服务器的传输
- 命令或答复按格式，以回车结束

```
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
```

类型	格式	示例
命令	关键字 参数	HELO example.edu MAIL FROM: <whuang@xmu.edu.cn>
响应	三位数代码 消息	250 OK 221 Bye

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

SMTP消息实例

```
S: 220 xmu.edu.cn Anti-spam GT for Coremail System (xmu.edu.cn[20221101])
C: EHLO YITADELL
S: 250-mail | PIPELINING | AUTH LOGIN PLAIN | AUTH=LOGIN PLAIN | coremail
1Uxr2***7kG0xki17xGrUDI***FY2U3Uj***28x1UUUUU7Ic2I***UFxE7***Ca0xDrUUUUj |
STARTTLS | SMTPUTF8 | 8BITMIME
C: AUTH LOGIN
S: 334 dXNl***hbWU6
C: User: d2h1Y***QHhtdS5lZHUuY24=
S: 334 UGFzc3d***Q6
C: Pass: d***dU***zg1
S: 235 Authentication successful
C: MAIL FROM: <whuang@xmu.edu.cn>
S: 250 Mail OK
C: RCPT TO: <****@xmu.edu.cn>
S: 250 Mail OK
C: RCPT TO: <whuang@xmu.edu.cn>
S: 250 Mail OK
C: DATA
S: 354 End data with <CR><LF>.<CR><LF>
C: DATA fragment, 470 bytes

S: 250 Mail OK queued as BAAi0gBnUDwZ18hkfKI7BA--.54010S2
C: QUIT
S: 221 Bye
```

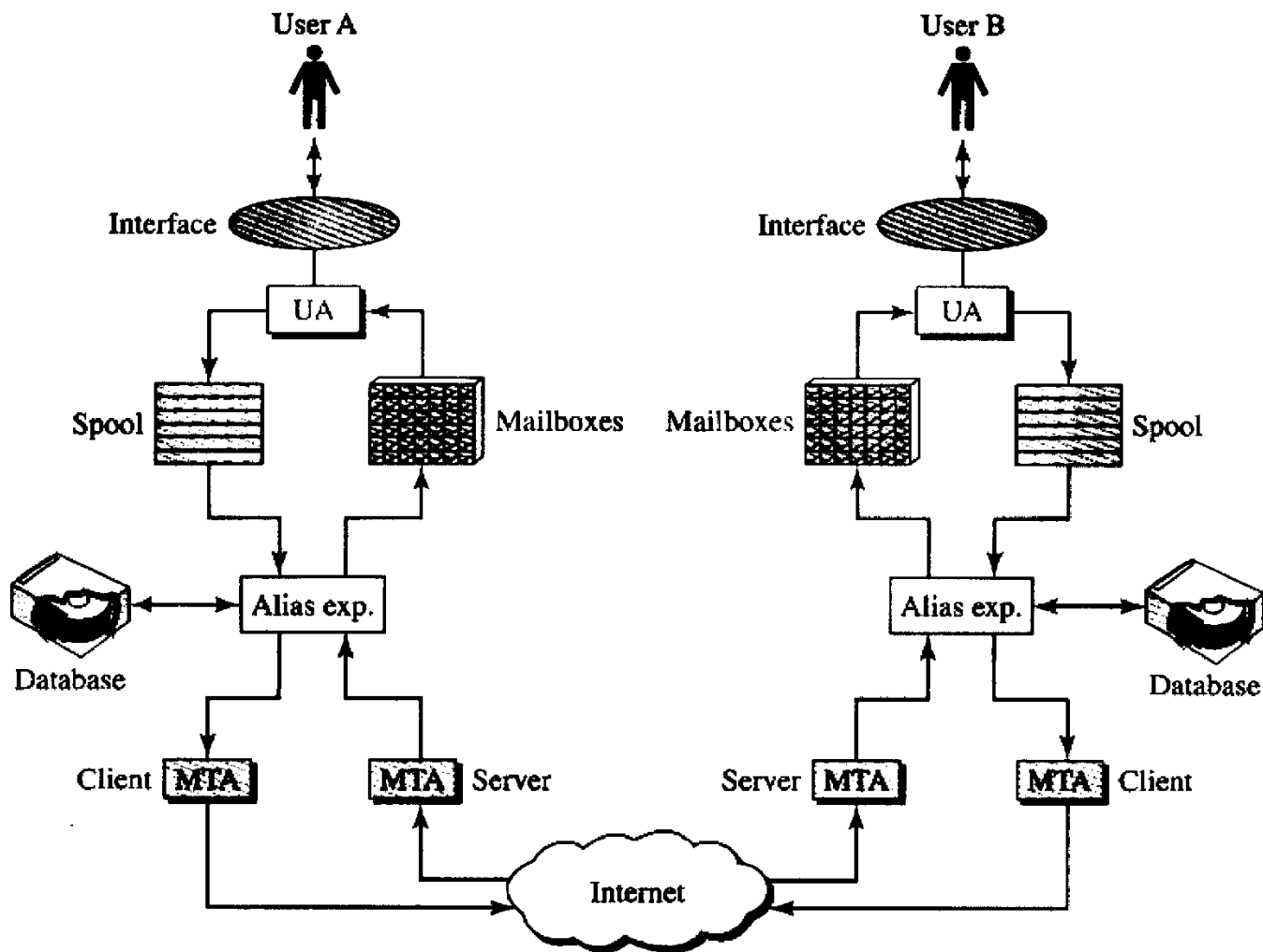
说明：

星号“*”处原为1位英文字母或数字，此处出于隐私考虑隐藏其真实值。

说明：

此处470 Bytes内容为前述消息。

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。
 - MIME不是邮件协议，不能替代SMTP。
- 在原始SMTP头部分添加五个标头以定义传输参数

MIME头	示例	说明
Mime-Version	Mime-Version: 1.0	标识MIME的版本
Content-Type	Content-Type: multipart/mixed	说明邮件的性质
Content-Transfer-Encoding	Content-Transfer-Encoding: base64	主体编码
Content-Id		邮件的唯一标识符
Content-Description		内容描述

MIME邮件格式

成分	名称	实例
1	传统消息头	Received: from YITADELL (unknown [180.233.66.123]) by mail2 (Coremail) with SMTP id BAAi0gBHOUW02shk_8E7BA--.54323S2; Tue, 01 Aug 2023 18:13:09 +0800 (CST) Date: Tue, 1 Aug 2023 18:13:10 +0800 From: "whuang@xmu.edu.cn" <whuang@xmu.edu.cn> To: "whuang@xmu.edu.cn" <test@abc.com> Subject: FYI
2	MIME在消息头的扩展	Mime-Version: 1.0 Message-ID: <202308011813100755976@xmu.edu.cn>
3	MIME头1	Content-Type: multipart/mixed; boundary="-----=_001_NextPart608765212147_-----" This is a multi-part message in MIME format.
4	MIME头2	-----=_001_NextPart608765212147_----- Content-Type: multipart/alternative; boundary="-----=_002_NextPart524780714027_-----"

MIME邮件格式

成分	名称	实例
5	MIME头3 (正文文字)	<pre>-----=_002_NextPart524780714027_----- Content-Type: text/plain; charset="us-ascii" Content-Transfer-Encoding: base64 VGhpcyBpcyBhIG1haWwuDQo=</pre>
6	MIME头4 (正文HTML)	<pre>-----=_002_NextPart524780714027_----- Content-Type: text/html; charset="us-ascii" Content-Transfer-Encoding: quoted-printable <html><head><meta http-equiv=3D"content-type" content=3D"text/html; charse= t=3Dus-ascii"><style>body { line-height: 1.5; }body { font-size: 14px; fon= t-family: "Microsoft YaHei UI"; color: rgb(0, 0, 0); line- height: 1.5; }</= style></head><body>=0A<div>This is a mail.</div>=0A</body></h= tml> -----=_002_NextPart524780714027_-----</pre>

MIME邮件格式

成分	名称	实例
7	MIME头5 (附件1)	<pre>-----=_001_NextPart608765212147_----- Content-Type: application/octet-stream; name="readme.txt" Content-Transfer-Encoding: base64 Content-Disposition: attachment; filename="readme.txt" VGhpcyBpcyBhIFJFQURNRSBmaWx1Lg==</pre>
8	MIME头6 (附件2)	<pre>-----=_001_NextPart608765212147_----- Content-Type: application/octet-stream; name="logo.jpg" Content-Transfer-Encoding: base64 Content-Disposition: attachment; filename="logo.jpg" /9j/4AAQSkZJRgABAQEAYABgAAD/2wBDAAyEBQYFBAYGBQYHBwYICgAKCgkJChQODwwQFxQYGBcU FhYaHSUfGhSjHBYWICwgIyYnKSopGR8tMC0oMCUoKSj/2wBDAQcHBwoIChMKChMoGhYaKCgoKCgo KCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCj/wAARCAIAAgADASIA ... WIKKKKACiigAooooAKKKKACiigAooooAKKKKACiigAooooAKKKKACiigAooooAKKKKAP/9k= -----=_001_NextPart608765212147_-----</pre>

MIME头的参数值

- **MIME-Version**的值为1.0
- **Content-Type**的参数是[type]/[subtype]; parameter
- **Content-Transfer-Encoding**的参数是[mechanism]

类型	子类型	描述
text	plain	未格式化的文本
multipart	mixed	正文包含不同数据类型的有序部分
	parallel	同上，但无顺序
	digest	类似mixed，但默认为message/rfc822
	alternative	同一消息的不同版本
message	rfc822	正文是一个封装的消息
	partial	正文是更大消息的片段
	external-body	正文是对另一条消息的引用
image	jpeg	JPEG格式的图像
	gif	GIF格式的图像
video	mpeg	MPEG格式的视频
audio	basic	8KHz语音单通道编码
application	postscript	Adobe PostScript
	octet-stream	通用二进制数据（八位字节）

类型	描述
7bit	NVT ASCII字符和短行
8bit	非ASCII字符和短行
binary	不限行长的非ASCII字符
base64	Base64编码
quoted-printable	非ASCII字符编码为等号，后跟ASCII码

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。
- 3个字符用4个字符传，适用于ASCII码不多的情况。

• 过程

- 将数据分成24位一组。
- 每组分为4字块，每块6位。
- 每块6位解释为一个可打印字符。

源码	M						a						(空)											
位组	0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	1	0	0	-	-	-	-	-	-
码字	T						W						E						=					

6位值	0	...	25	26	...	51	52	...	61	62	63	填充
对应码	A	...	Z	a	...	z	0	...	9	+	/	=

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

邮局协议3 (POP3)

- POP3 (Post Office Protocol version 3)
 - POP3是邮件客户端从服务器接收邮件的协议
 - 允许列表、收取和删除邮件
 - 默认端口号：110 (明文)，995 (SSL)
- POP3协议使用命令与响应进行交互

类型	格式	示例
命令	关键字 参数	RETR 70 QUIT
响应	状态代码 消息	+OK 166114 octets -ERR

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

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

成分	名称	实例
1	欢迎辞	+OK Welcome to coremail Mail Pop3 Server (163coms[...
2	鉴权	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)]
3	总邮件数	STAT +OK 7 231904 LIST +OK 7 231904 1 9854 2 140937 3 10224 4 17139 ... UIDL +OK 7 231904 1 1tbiMAkaCFEAER4pHgAAsC 2 1tbiNQuc...

POP3 会话实例

成分	名称	实例
4	获取邮件	RETR 7 +OK 1698 octets
5	返回邮件内容	Received: from [192.168.7.128] (unknown [119.233.1... f0uMueDze6FWdWBSy3zL/K1nBCoAwR8FFrAhNS4gN8AhsIy++3... (此处具体邮件内容为邮件的MIME格式)
6	退出	QUIT +OK core mail

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
(特色化示范性软件学院) 博士, 副教授
School of Informatics Wei Huang