



## Email 101 & Mail Internals



**Anurak Chuetanapinyo**

*IronPort Systems Engineer – Thailand & Indochina*

*Security Technology Business Unit*

anurak@cisco.com

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

1

## Overview

- Email 101
  - MTAs
  - Groupware
  - POP
  - IMAP
  - SMTP overview
- Mail Internals
  - SMTP details
  - Exploring the mail envelope, headers & body

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

2

## Message Transport Agent (MTA)

- The MTA is the last outbound and first inbound hop for an enterprise message
- Primary function: route messages
  - Internet domains using DNS
  - Local domains using directory lookups
- Because of location best place to enforce policy, block spam, block viruses
- Large enterprise MTAs
  - #1: open-source sendmail on Unix servers
  - Others include postfix, qmail, exchange
- Small enterprise MTA
  - #1 Microsoft Exchange on Windows NT

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

3

## Mail User Agent (MUA)

- MUA (a.k.a. "mail client") is used to compose and read mail
- Some directly connect with MTA to send mail via SMTP
  - Outlook Express, Eudora
- Some directly connect with groupware servers to send and receive
  - Outlook 2000, Notes, Groupwise
- Others connect via HTTP to send and receive
  - Yahoo Mail, Hotmail
- **All mail goes through the MTA on its way to the recipient!**



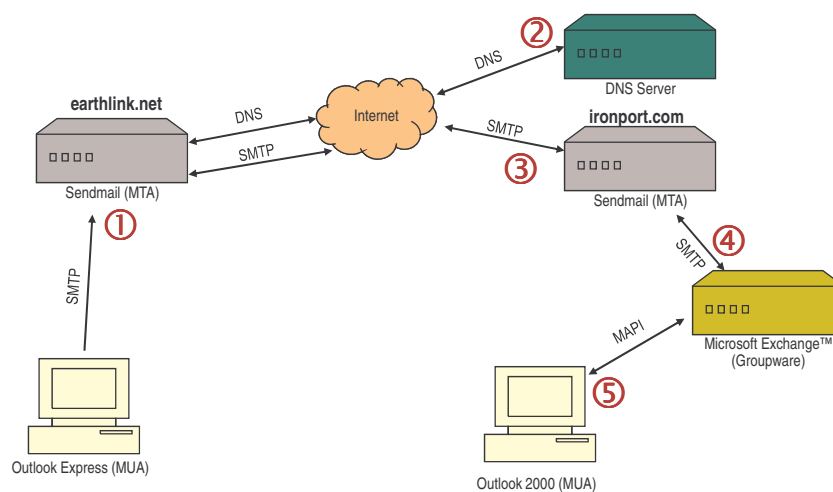
Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

4

## Groupware

- Groupware servers provide enterprise-wide messaging, calendaring, tasks, etc.
- Stores messages, syncs with client
- Provides web access
- Routes messages among workgroups

## Day in the Life of an Enterprise Message



## POP: Post Office Protocol

- Designed to support *offline* mail processing
- End user invokes an MUA to connect to the message store and download all pending mail to the user's local inbox
- After transfer, all mail processing is local to the client machine
- Intended to move mail (on demand) from the message store

## IMAP: Internet Message Access Protocol

- Method of accessing messages that are kept on a shared message store
- MUA acts on the message as if it were local, but message stays on the message store

## SMTP: Simple Mail Transfer Protocol

- The language spoken between MTAs on the Internet
- Client and server are intended to be always available on the Internet
  - DNS provides system for failover
  - DNS points to MTA for each company (MX record)
- Structured conversation includes:
  - Message Envelope
  - Message Headers
  - Message Body

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

9

## SMTP: Message Envelope

*Communicates the sender and recipient email address to the remote MTA*

- **Envelope To** (RCPT TO) used by the MTA to determine where to deliver messages
- **Envelope From** (MAIL FROM) used to determine where to return undeliverable mail

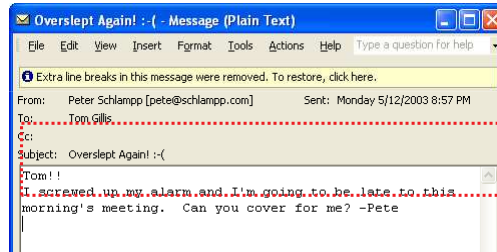
```
<< 220 smtp.ironport.com ESMTP
>> HELO mail.schlamp.com
<< 250 smtp.ironport.com
>> MAIL FROM: <pete@schlamp.com>
<< 250 sender <pete@schlamp.com> ok
>> RCPT TO: <tom@ironport.com>
<< 250 recipient <tom@ironport.com> ok
```

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

10

## SMTP: Headers

- From, To, Subject, and Date Fields in the user's MUA.
- Flags messages for special handling.



```
>> DATA
<< 354 go ahead
>> From: Peter Schlamp <pete@schlamp.com>
>> To: Tom Gillis <tom@ironport.com>
>> Subject: Overslept Again! :- (
>> Date: Mon, 12 May 2003 20:57:13 -0700
>> X-SpamScore: 100
```

## SMTP: Body

- The message body is the actual contents of the message
- Attachments are simply encoded portions of the message body

```
>> Tom!!
>> I screwed up my alarm again and I'm going
to be late to >> this morning's meeting. Can
you cover for me?
>> -Pete
>> .
<< 250 ok
>> QUIT
<< 221 smtp.ironport.com
```

## Entire SMTP Conversation

Envelope

```
<< 220 smtp.ironport.com ESMTP
>> HELO mail.schlamp.com
<< 250 smtp.ironport.com
>> MAIL FROM: <pete@schlamp.com>
<< 250 sender <pete@schlamp.com> ok
>> RCPT TO: <tom@ironport.com>
<< 250 recipient <tom@ironport.com> ok
```

Headers

```
>> DATA
<< 354 go ahead
>> From: Peter Schlamp
<pete@schlamp.com>
>> To: Tom Gillis <tom@ironport.com>
>> Subject: Overslept Again! :-(
>> Date: Mon, 12 May 2003 20:57:13 -0700
>> X-SpamScore: 100
```

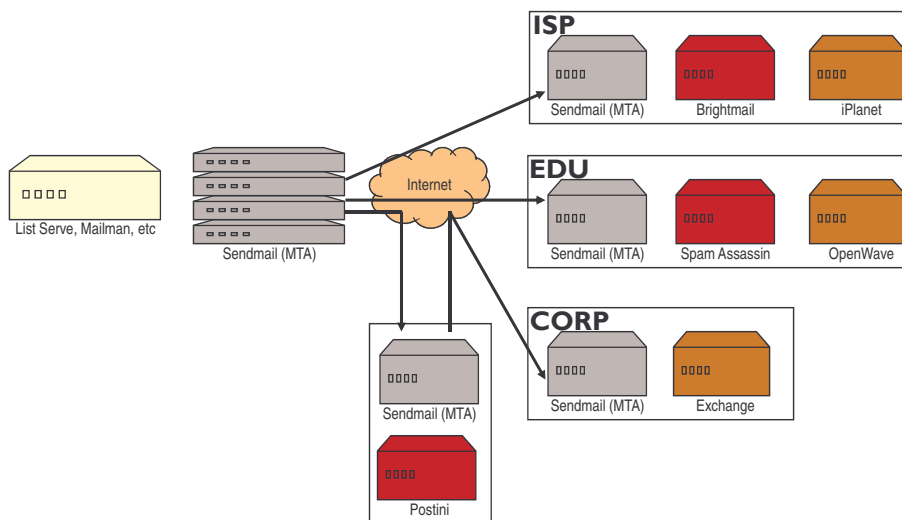
Body

```
>> Tom!!
>> I screwed up my alarm again and I'm
going to be late to
>> this morning's meeting. Can you cover
for me?
>> -Pete
>>
<< 250 ok
>> QUIT
<< 221 smtp.ironport.com
```

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

13

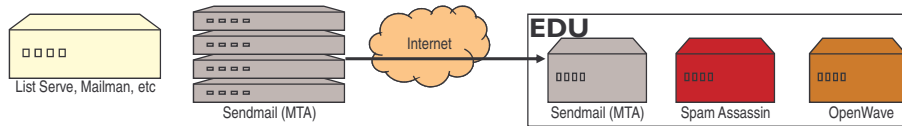
## Day in the life of a Customer Message



Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

14

## Bounce-ology



- **Conversational bounce**
  - Delivered during the SMTP conversation
  - Could be hard (bad address) or soft (mailbox full)
  - Captured in A60 bounce logs
- **Delayed bounce**
  - ISP accepts message
  - Later in time sends the message back with an error
  - Could be hard or soft
  - A60 can accept and relay to a mailbox or listmanager

## CM Glossary

- **List Manager**
  - Maintains the list of email addresses
  - Generates a campaign
  - Handles bounces and unsubscribes
- **Rendering**
  - Pulling together data to assemble actual message
- **Opt-in**
  - A user has agreed to be sent information
- **Double Opt-in**
  - A message is sent confirming the user wishes to be added to a list
- **Openrate**
  - % of messages opened – tracked with a .gif
- **Black List**
  - List of IP addresses not to accept mail from

## Mail Internals

## SMTP

- Original specification was RFC 821 (SMTP) and RFC 822 (Message Body) written in 1982.
- Updated by RFC's 2821 and 2822 published in 2001.
- Excellent source of RFC's: <http://www.faqs.org/rfcs/>
- Some also available on [private.ironport.com](http://private.ironport.com)
- DJB (author of qmail) has lots of documentation at <http://cr.yip.to/mail.html>

## SMTP Transcript

```
femur:~> telnet femur.ironport.com 8025
Trying 10.1.1.78...
Connected to femur.ironport.com.
Escape character is '^]'.
220 femur.ironport.com ESMTP
helo foo
250 femur.ironport.com
mail from:<jesse@ironport.com>
250 sender <jesse@ironport.com> ok
rcpt to:<jesse@ironport.com>
250 recipient <jesse@ironport.com> ok
data
354 go ahead
Subject: go home!
```

It's Saturday!

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

19

## Mail Message Components

- Envelope
- Message Headers
- Message Body

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

20

## Example Message

```
Return-Path: <ctran@ironport.com>
Delivered-To: ehuss@wintermute.sponsor.net
Received: (gmail 83897 invoked by alias); 12 Aug 2003 22:39:40 -0000
Delivered-To: alias-netmeridian-ehuss@netmeridian.com
Received: (gmail 83893 invoked from network); 12 Aug 2003 22:39:40 -0000
Received: from smtp2.ironport.com (63.251.108.118)
  by wintermute.sponsor.net with SMTP; 12 Aug 2003 22:39:40 -0000
Received: from mx-router.ironport.com (10.1.1.48)
  by smtp2.ironport.com with ESMTMP; 12 Aug 2003 15:39:16 -0700
Received: from diablo.ironportsystems.com (diablo.ironport.com [10.1.1.15])
  by mx-router.ironport.com (Postfix) with ESMTMP id 5DC3325DC8
  for <Ehuss@netmeridian.com>; Tue, 12 Aug 2003 15:39:16 -0700 (PDT)
X-MimeOLE: Produced By Microsoft Exchange V6.0.6249.0
MIME-Version: 1.0
Content-Type: text/plain;
  charset="iso-8859-1"
Subject: Ice Cream - in Breakroom!
Date: Tue, 12 Aug 2003 15:37:22 -0700
Message-ID: <58D2F08D49B1B94D836136DA8463E5C3706AC4@diablo.ironportsystems.com>
Thread-Topic: Ice Cream - in Breakroom!
From: "Courtney Tran" <ctran@ironport.com>
To: "All SF Employees" <sfo@ironport.com>
X-Spam-Status: No, hits=-100.0 required=4.0
  tests=USER_IN_WHITELIST
  version=2.55
X-Spam-Level:
X-Spam-Checker-Version: SpamAssassin 2.55 (1.174.2.19-2003-05-19-exp)
Status: R

There's Ice Cream in the breakroom as an afternoon WAKE-UP snack.

Thank Scott Weiss for the "budweiser" contribution!! $20 + our $10
winnings from Lotto buys a lot of ice cream! :-)
```

Courtney  
Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

21

## Mail Envelope

- Sender and Recipient
- Used by servers
- Rarely seen by end user

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

22

## Message Headers

- Received: (we can turn this off)
- >100 Received headers is a loop
- Message-ID:
- Content-type:
- Terminated by blank line
- From: and To: headers...these are displayed in the user's MUA. Servers typically ignore these.
- Reply-To:

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

23

## Message Body

- MIME, attachments
- Multipart
- Maximum size

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

24

## SMTP Conversation

- Always port 25 (unlike most protocols)
- Maximum line length, 990, 1024, 4096?
- Plain text, CRLF
- Multiple messages per connection

## HELO

- HELO hostname
- Hostname is usually ignored in favor of reverse DNS lookup
- May return 5xx to identified spammers

## EHLO (Extended SMTP)

- EHLO hostname
- Multiline response with remote server's hostname and a list of extended features, for example:

8BITMIME  
CHUNKING  
PIPELINING  
TLS

## MAIL FROM

- MAIL FROM:<arnold@ca.gov>
- Optional parameters allowed, we only recognize and handle size:
- MAIL FROM:<addr> SIZE=1024

## RCPT TO

- RCPT TO:<foo@bar.com>
- Multiple RCPT commands per MAIL
- 4xx, temporary failure, “soft bounce”
- 5xx, permanent failure, “hard bounce”
- 550, relay denied, “go away spammer.”
- “RCPT” has four syllables!

## DATA

- DJB vs. LF [cr.yip.to/docs/smtplf.html](http://cr.yip.to/docs/smtplf.html)
- Headers, blank line, body.
- Terminated with a five character CRLF.CRLF sequence, .CRLF stripped
- If you have a message with a line that starts with a “.”, the sender must escape it by inserting an additional period. The receiver does the reverse.

## Bounces

- Conversational vs. Nonconversational
- Hard vs. soft bounces
- Double bounces
- DSN format
- Dictionary harvest attack

## IronPort-specific

- Max line length ~4096
- We treat HELO 5xx as temporary
- We don't require HELO
- We accept bare LFs
- We accept nearly any address format

## Email Address Parsing

- We try harder to parse weird addresses, because customers don't want to change their infrastructure.
- <joe@foo.com>
- joe@foo.com
- "joe@foo.com"
- <"Joe Sixpack" <joe@foo.com>>
- <(joe@bar.com) otherjoe@foo.com>

## 3 Digit Reply Codes

- Less ambiguous machine processing
- Followed by arbitrary text for humans
- Each digit is progressively less meaningful

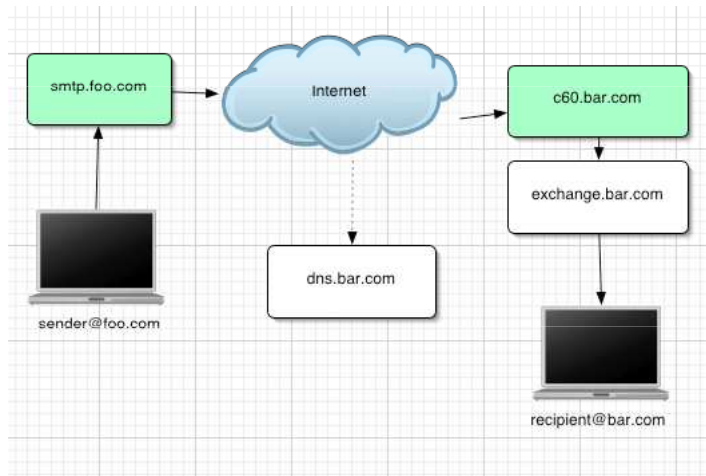
## Reply Code, First Digit

- First digit denotes whether the response is good, bad, or incomplete
  - 2 success (250 OK)
  - 3 okay so far (354 Start mail input)
  - 4 temporary failure (452 mailbox full)
  - 5 permanent failure (550 user unknown)

## Reply Code, Second and Third Digits

- Second digit categorizes the result described by the first digit:
  - 0 syntax
  - 1 connection
  - 5 mail system
- Third digit adds finer detail

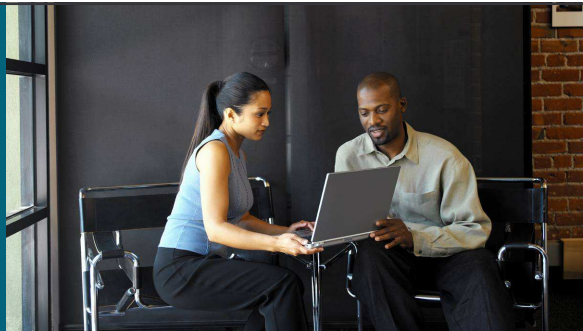
## Person to Person



Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

37

## Q and A



### Email 101

- MTAs
- Groupware
- POP
- IMAP
- SMTP overview

### Mail Internals

- SMTP details
- Exploring the mail envelope, headers & body

Presentation\_ID © 2006 Cisco Systems, Inc. All rights reserved. Cisco Confidential

38

