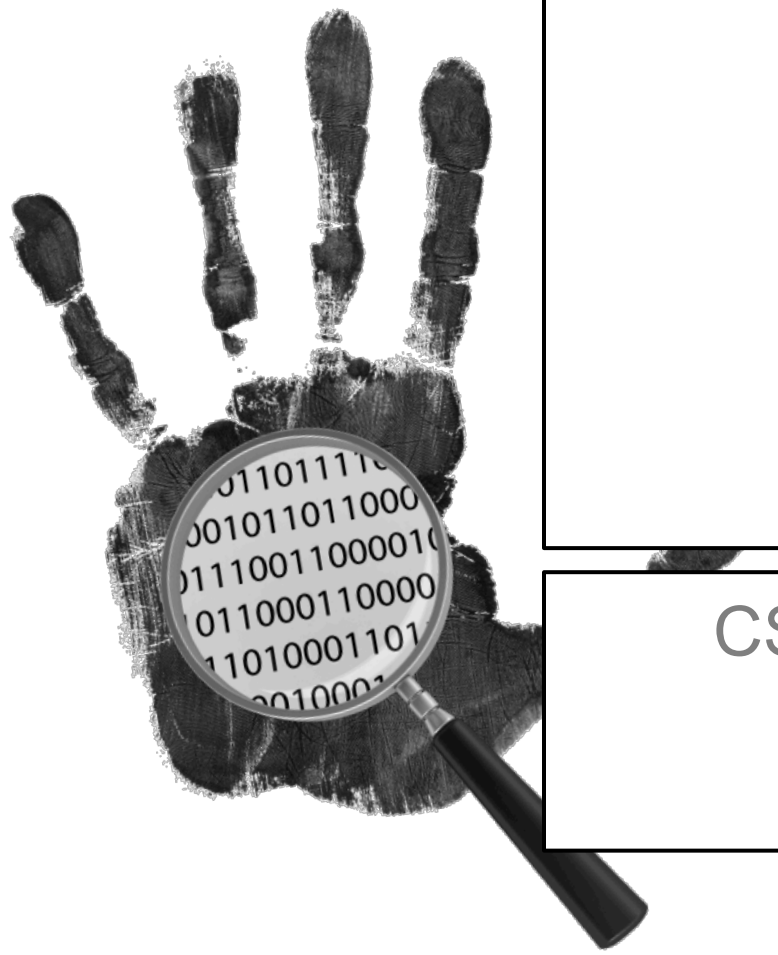
TÉCNICO LISBOA

Email Forensics

Part II.B. Techniques and Tools: Network Forensics



CSF: Forensics Cyber-Security

Fall 2015

Nuno Santos

Summary

- Introduction to network forensics
- Email forensics

Remember were we are

- Our journey in this course:
  - Part I: Foundations of digital forensics
  - Part II: Techniques and tools
    - A. Computer forensics
    - **B. Network forensics**
    - C. Forensic data analysis

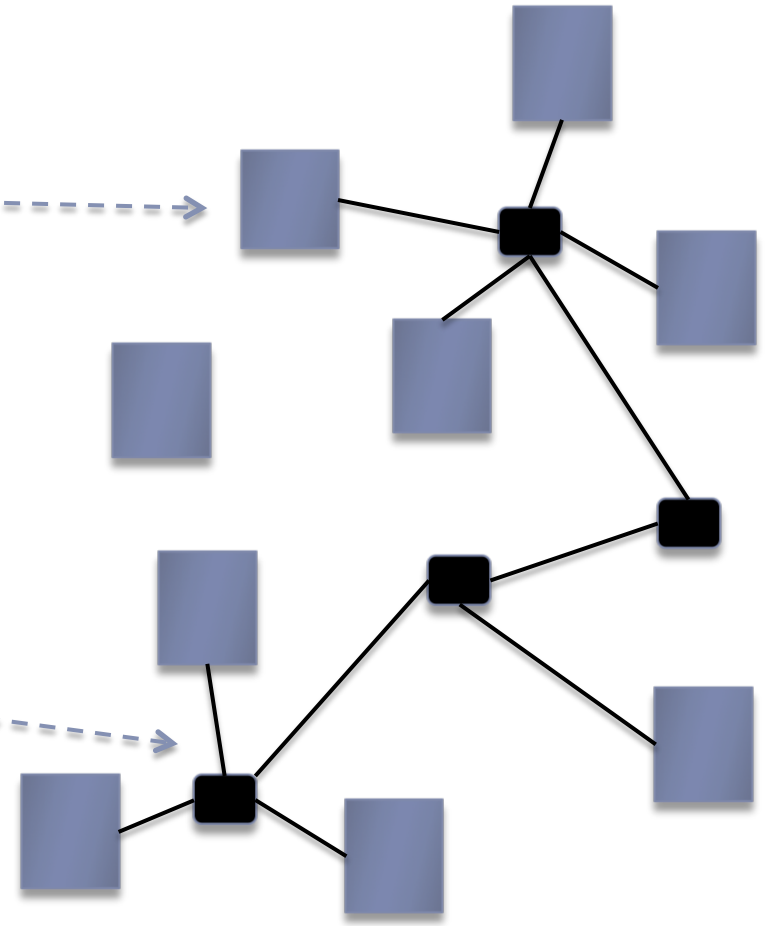
Starting today

TÉCNICO LISBOA

Introduction to network forensics

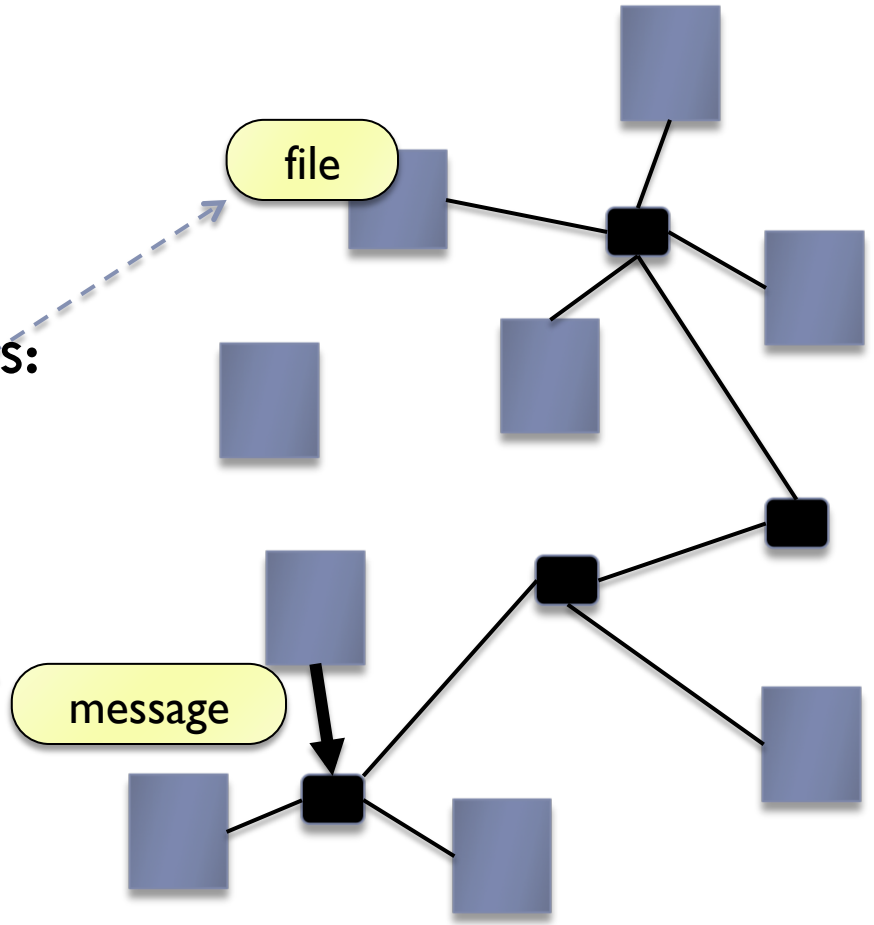
Model for reasoning about evidence sources

- Data is stored and processed in **computers**
- Data can be exchanged between computers through **networks**



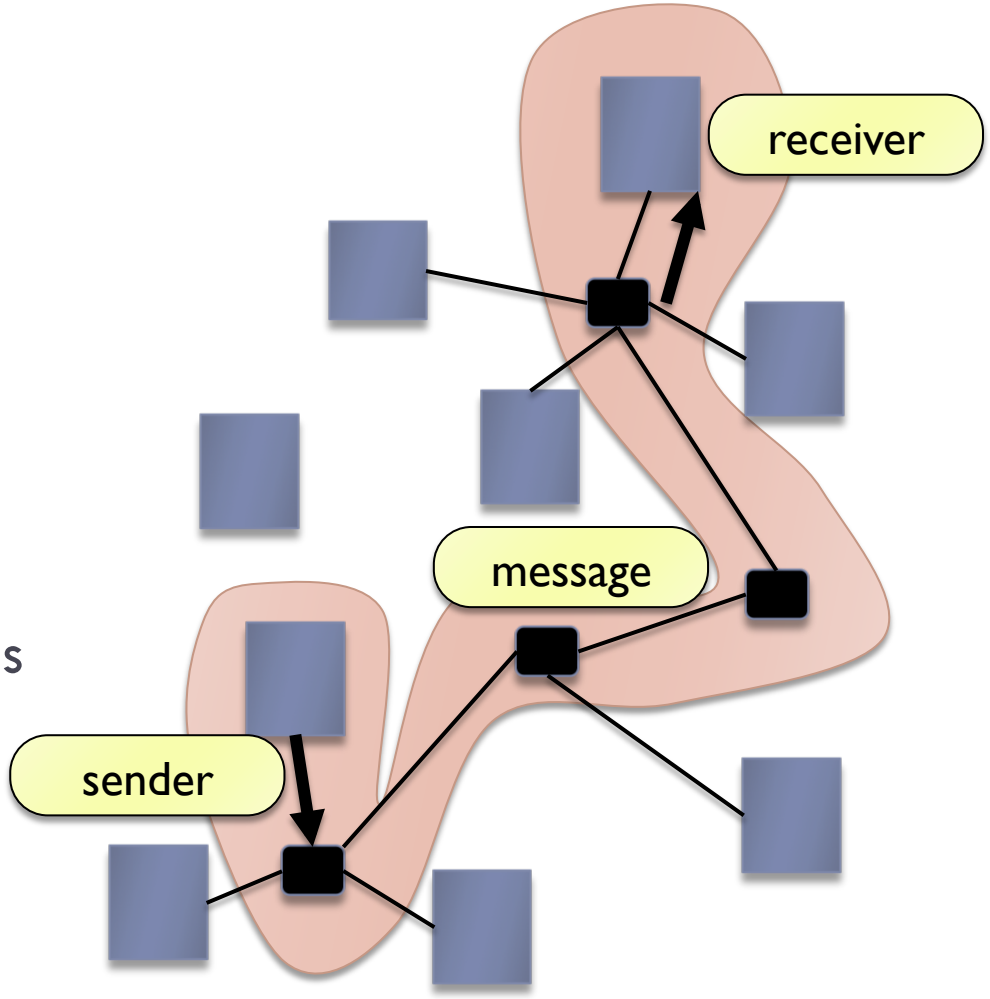
A simple way to reason about evidence sources

- **Data** are groups of 1's and 0's
- Typical data abstractions:
  - In computers: the **file**
  - In networks: the **message**
- Can be stored in **persistent** or **volatile** memory

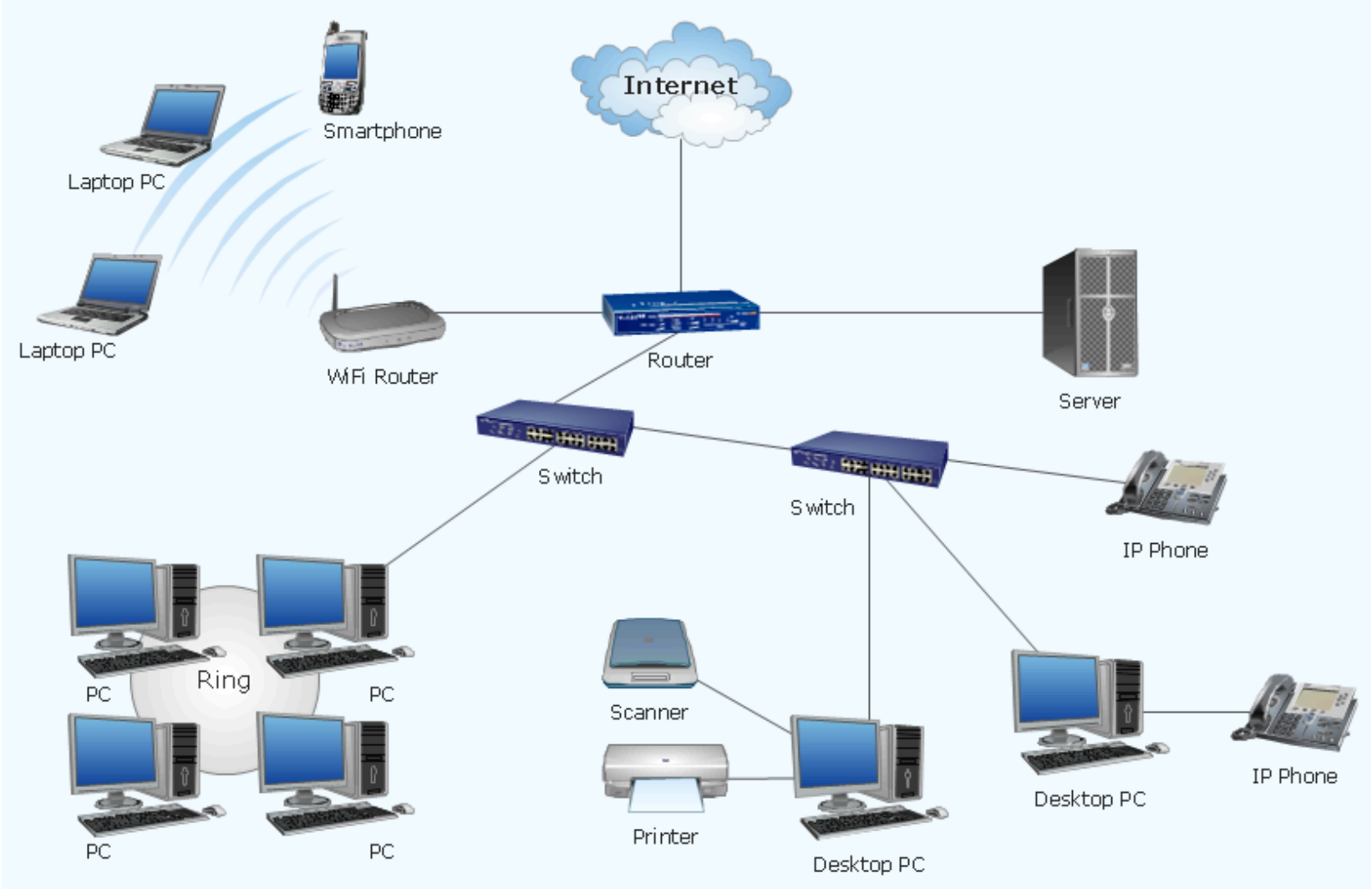


In networks, we care about messages

- In particular, we look at:
  - The content of messages
  - And traces left by messages

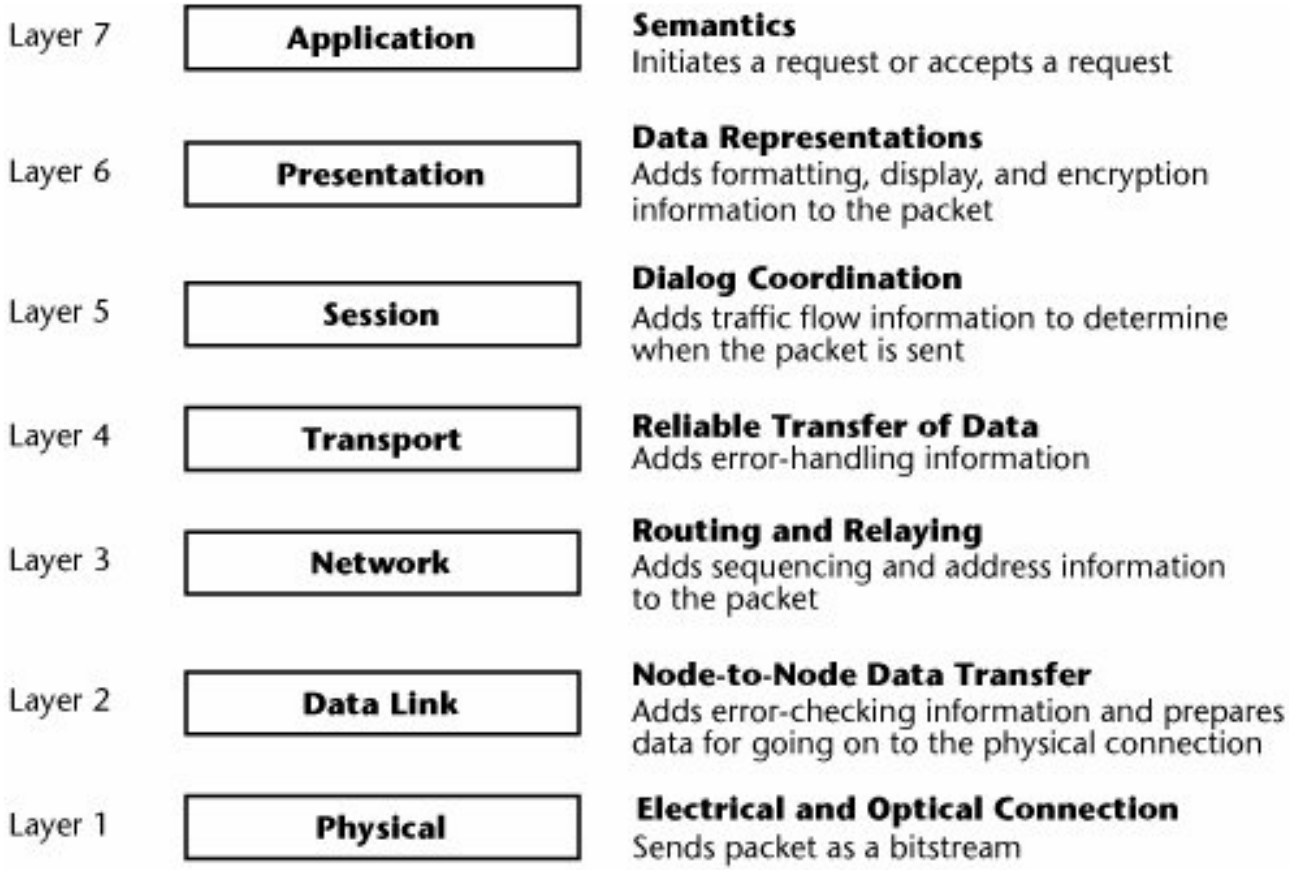


Network architecture example



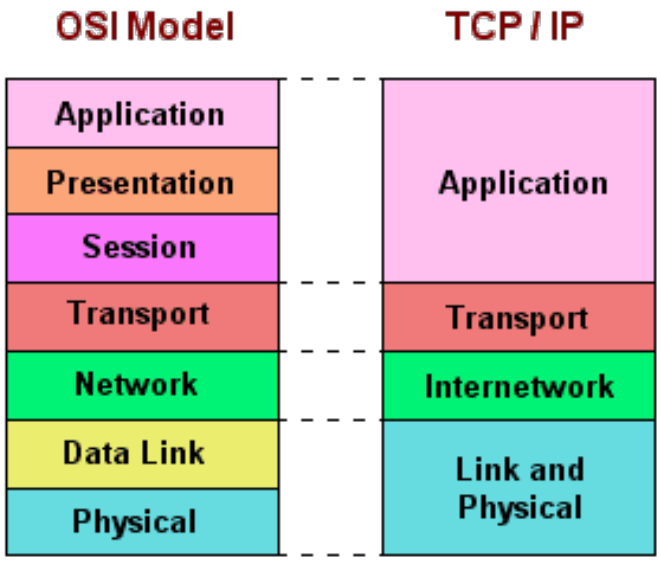
Network abstraction layers

- OSI reference model: is a reference tool for understanding data communications between networked systems

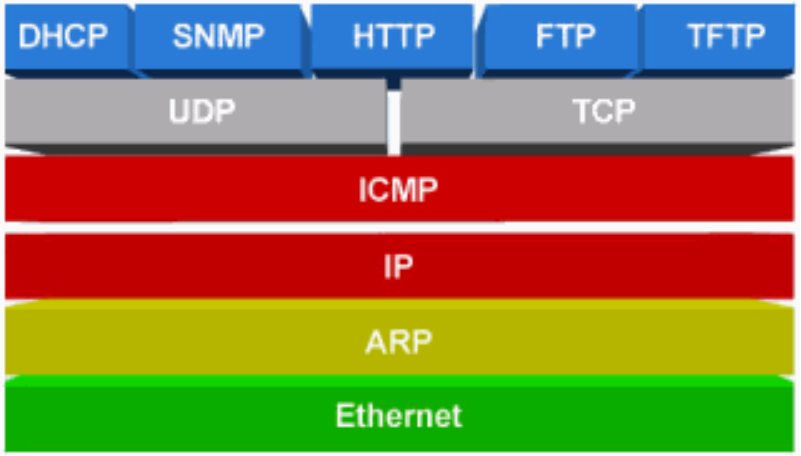


OSI vs TCP/IP models

- Correspondence between OSI and TCP/IP models



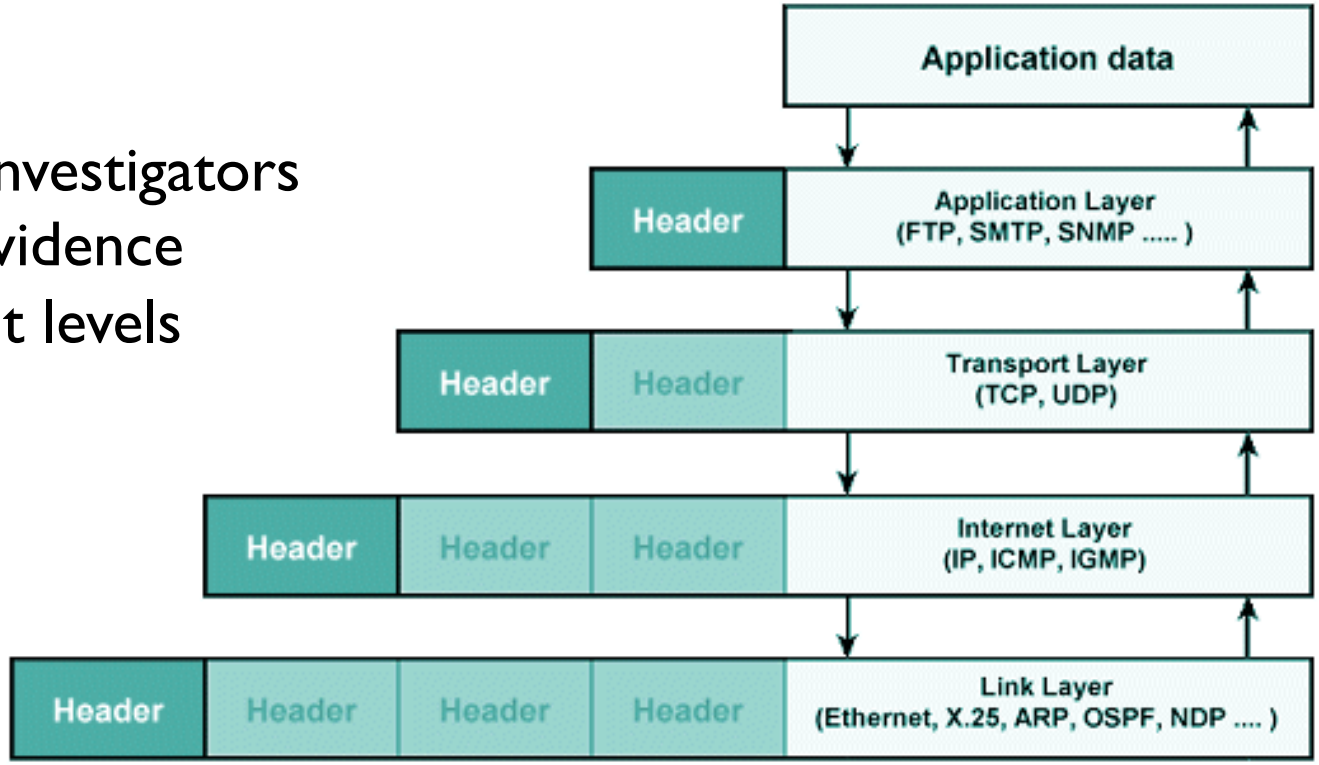
- And example protocols:



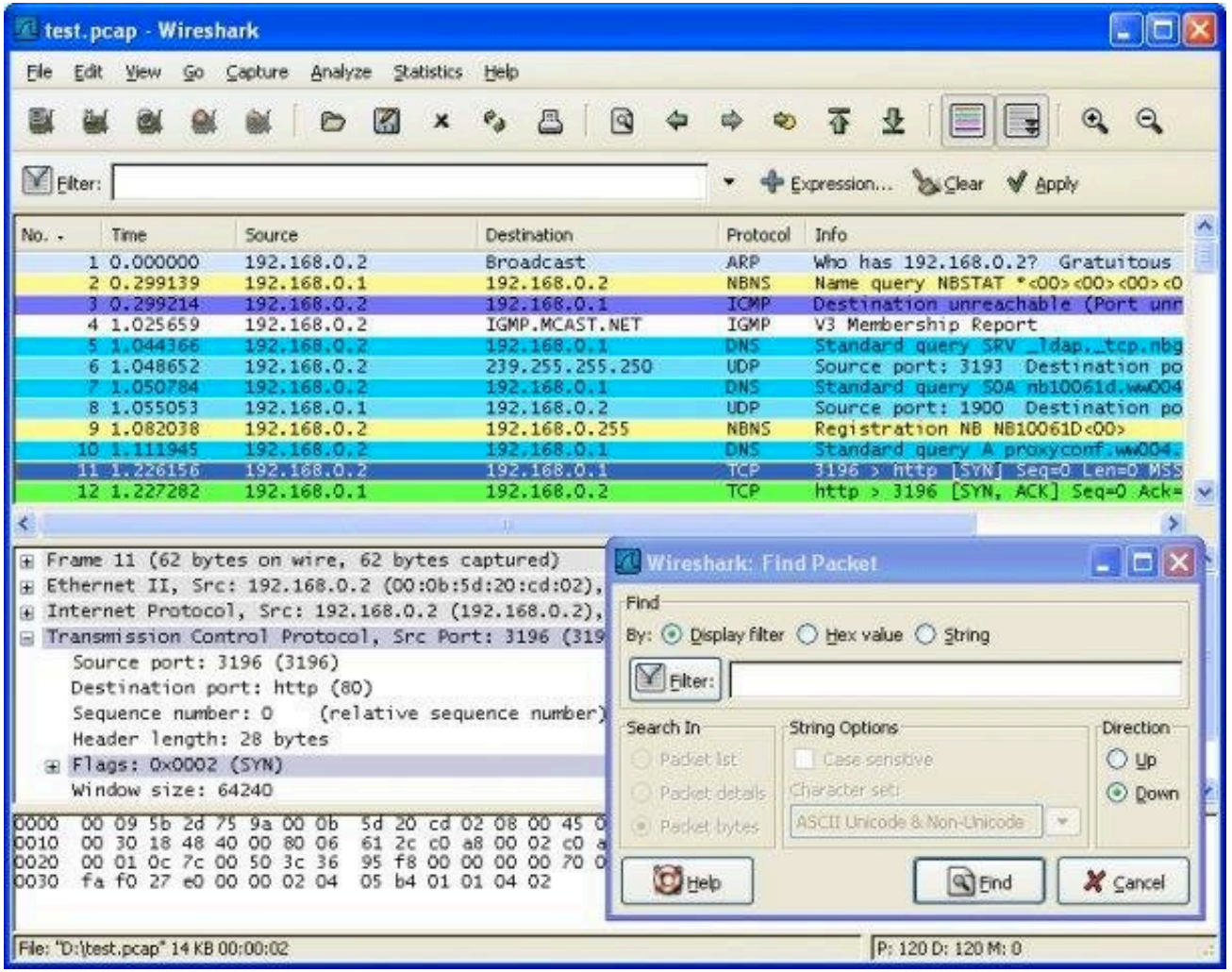
Message representation

- Packets are encoded per network stack layer:

Forensic investigators may get evidence at different levels



Need adequate tools for forensic analysis



Roadmap for network forensic classes

- Application layer
  - Email

Today

- Transport and network layer
- Data link layer

TÉCNICO LISBOA

Email forensics

Motivation for email investigations

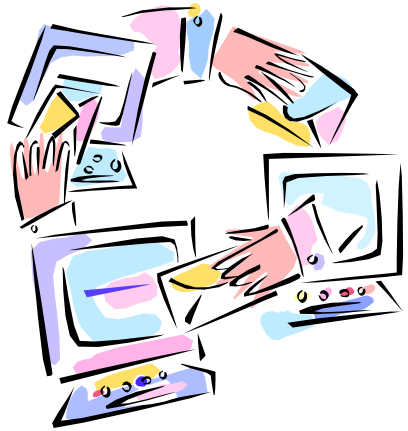
- Email has become a primary means of communication
- Email can easily be forged
- Email can be abused
  - Spam
  - Aid in committing a crime …
  - Threatening email, …

Importance of email as evidence

- E-mail can be pivotal evidence in a case
- Due to its informal nature, it does not always represent corporate policy
- Many cases provide examples of the use of e-mail as evidence
  - Enron
  - Knox vs. State of Indiana
  - Harley vs. McCoach
  - Nardinelli et al. vs. Chevron
  - Adelyn Lee vs. Oracle Corporation

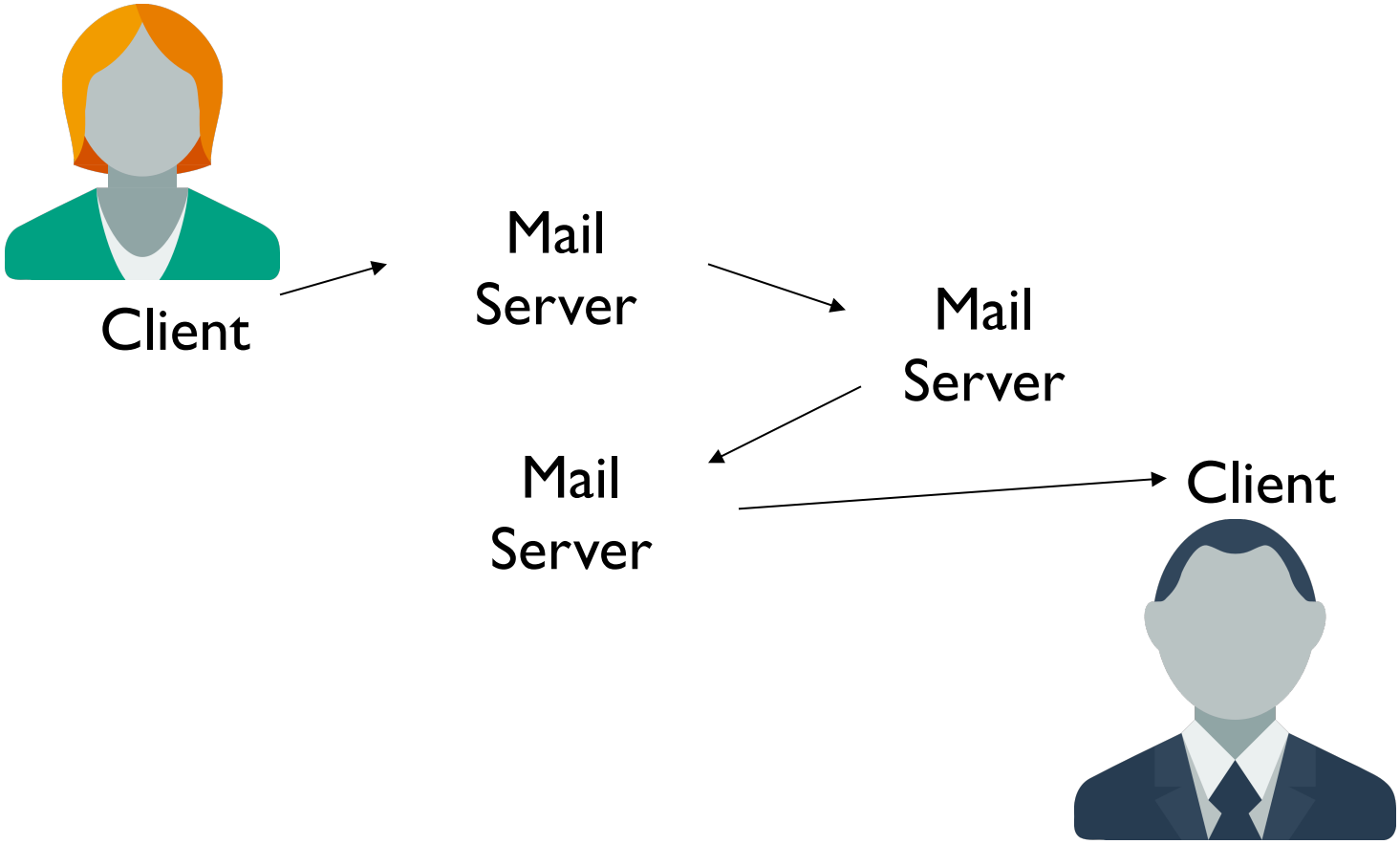
Working with email

- E-mail evidence typically used to corroborate or refute other testimony or evidence
  - Can be used by prosecutors or defense parties
- Two standard methods to send and receive e-mail:
  - Client/server applications
  - Webmail



Email fundamentals

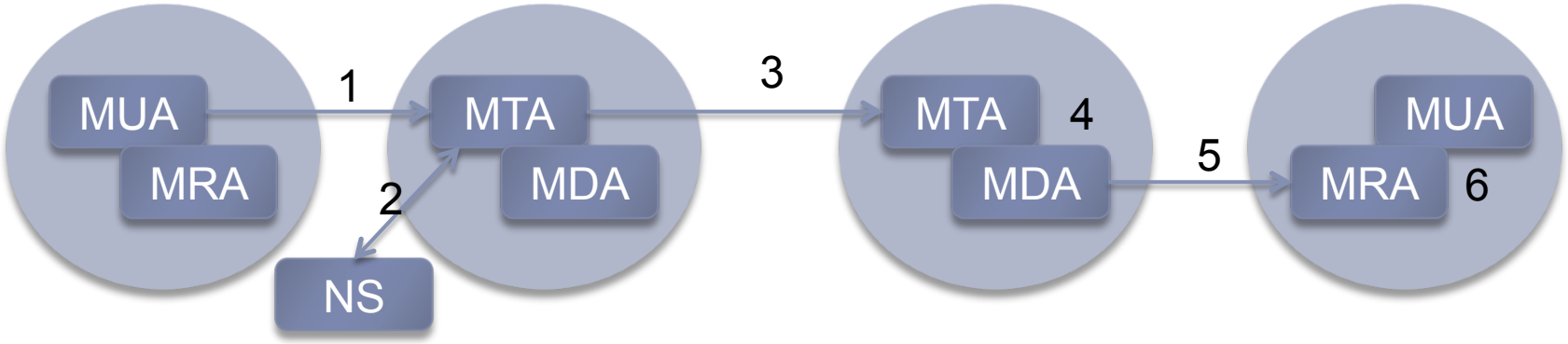
- Typical path of an email message:



Email investigations overview

- Email evidence is in the **email** itself (header)
- Email evidence is **left behind** as the email travels from sender to recipient
  - Contained in the various logs
  - Maintained by system admins
- Law enforcement can use subpoenas to collect emails headers and logs

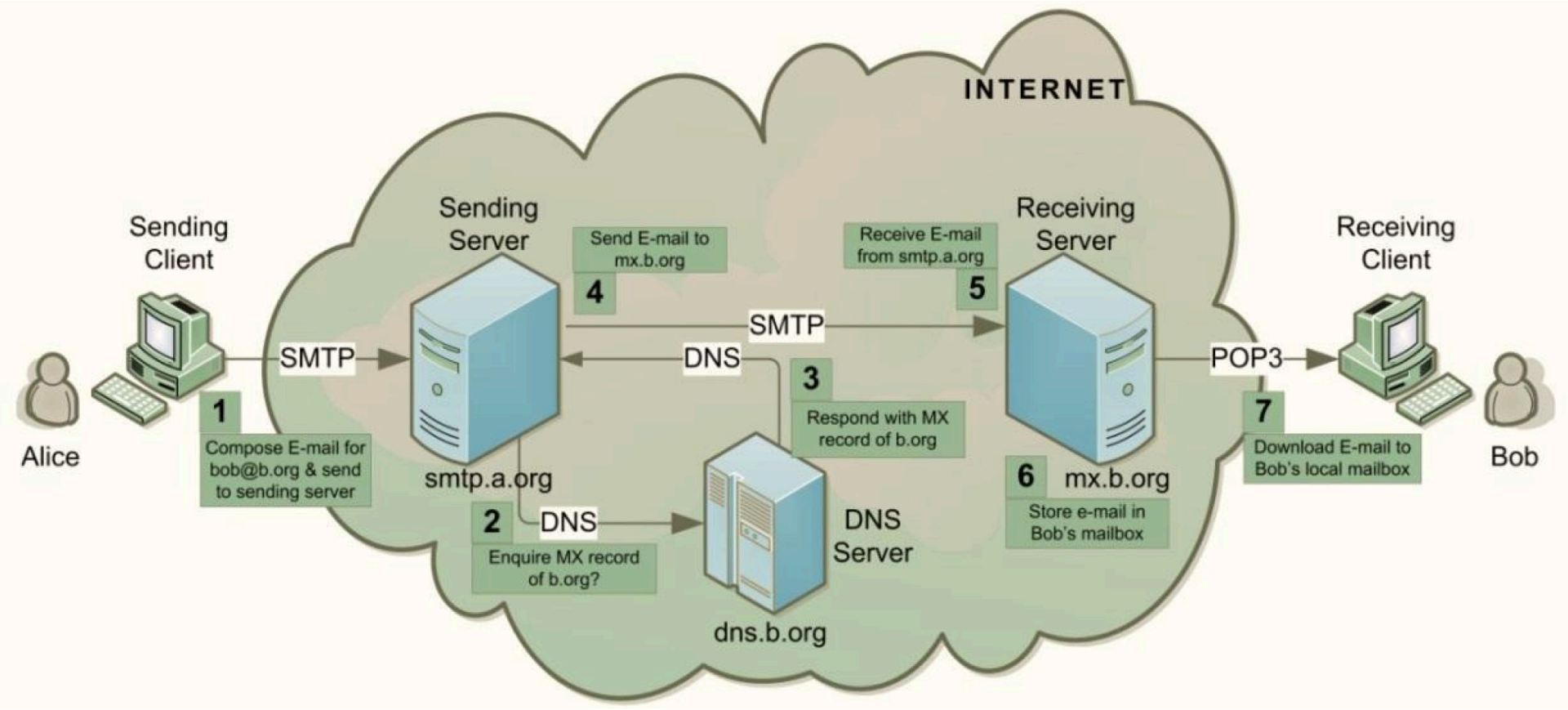
Typical actors in an email flow



- MUA – Mail User Agent
  - E.g. thunderbird, outlook
- MTA – Mail Transfer Agent
  - E.g sendmail, qmail
- MDA – Mail Delivery Agent
  - E.g procmail
- MRA – Mail Retrieval Agent
  - POP/IMAP client
- NS – Name Server
  - DNS server

1. MUA implements smtp client to smtp server
2. MTA solves address using MX record in NS
3. MTA contacts MTA though SMTP
4. Receiving MTA delivers the email to MDA
5. MRA uses IMAP/POP/MAPI to retrieve from MDA
6. MUA presents mail to user

Email communication between sender & receiver



Steps in the email communication

1. Alice composes an email message on her computer for Bob and sends it to her sending server smtp.a.org using SMTP protocol
2. Sending server performs a lookup for the mail exchange record of receiving server b.org through DNS protocol on DNS server mx.b.org for the domain b.org
3. The DNS server responds with the highest priority mail exchange server mx.b.org for the domain b.org
4. Sending server establishes SMTP connection with receiving server and delivers the email to Bob's mailbox on the receiving server
5. The receiving server receives the incoming email message
6. The receiving server stores the email message on Bob's mailbox
7. Bob downloads the message from his mailbox on receiving server to local mailbox on his client computer using POP3 or IMAP protocols (Bob can optionally use a Webmail program)

Client protocols

Post Office Service	Protocol	Characteristics
Stores only incoming messages	POP	Investigation must be at the workstation.
Stores all messages	IMAP MS' MAPI Lotus Notes	Copies of incoming and outgoing messages might be stored on the workstation or on the server or on both.
Web-based send and receive	HTTP	Incoming and outgoing messages are stored on the server, but there might be archived or copied messages on the workstation

SMTP headers

- Reviewing e-mail headers can offer clues to true origins of the mail and the program used to send it

- Common e-mail header fields include:
  - Bcc
  - Cc
  - Content-Type
  - Date
  - From
  - Message-ID
  - Received
  - Subject
  - To
  - X-Priority

SMTP headers example

- Example of a message header for an email sent from MrJones@emailprovider.com to MrSmith@gmail.com

```
Delivered-To: MrSmith@gmail.com
Received: by 10.36.81.3 with SMTP id e3cs239nzb;Tue, 29 Mar 2005 15:11:47
-0800 (PST)
Return-Path: MrJones@emailprovider.com
Received: from mail.emailprovider.com (mail.emailprovider.com
[111.111.11.111]) by mx.gmail.com with SMTP id h19si826631rnb; Tue, 29
Mar 2005 15:11:47 -0800 (PST)
Message-ID: <20050329231145.62086.mail@mail.emailprovider.com>
Received: from [11.11.111.111] by mail.emailprovider.com via HTTP; Tue,
29 Mar 2005 15:11:45 PST
Date: Tue, 29 Mar 2005 15:11:45 -0800 (PST)
From: Mr Jones
Subject: Hello
To: Mr Smith
```

The Received header

- Received is the most essential field of the email header: it creates a list of all the email servers through which the message traveled in order to reach the receiver

- The best way to read are from bottom to top
  - The bottom "Received" shows the IP address of the sender's mail server
  - The top "Received" shows the IP address of receiver mail server
  - The middle "Received" shows the IP address of the mail server through which email passes from sender to receiver

The Received headers in the example

- From mail.emailprovider.com to mx.gmail.com

```
Received: from mail.emailprovider.com (mail.emailprovider.com
[111.111.11.111]) by mx.gmail.com with SMTP id h19si826631rnb; Tue, 29
Mar 2005 15:11:47 -0800 (PST)


Received: from [11.11.111.111] by mail.emailprovider.com via HTTP; Tue,
29 Mar 2005 15:11:45 PST
```

SMTP protocol

- Neither IMAP or POP are involved relaying messages between servers
- Simple Mail Transfer Protocol: SMTP

```
S: 220 smtp.example.com ESMTP Postfix
C: HELO relay.example.org
S: 250 Hello relay.example.org, I am glad to meet you
C: MAIL FROM:<bob@example.org>
S: 250 Ok
C: RCPT TO:<alice@example.com>
S: 250 Ok
C: DATA
S: 354 End data with <CR><LF>.<CR><LF>
C: From: "Bob Example" <bob@example.org>
C: To: "Alice Example" <alice@example.com>
C: Cc: theboss@example.com
C: Date: Tue, 15 January 2008 16:02:43 -0500
C: Subject: Test message
C:
C: Hello Alice.
C: This is a test message.
C: Your friend,
C: Bob
C: .
S: 250 Ok: queued as 12345
C: QUIT
S: 221 Bye
{The server closes the connection}
```

Sending spoofed emails

- SMTP is simple, but can be spoofed easily
- How to spoof email easily:

```
C: telnet server8.engr.scu.edu 25
S: 220 server8.engr.scu.edu ESMTP Sendmail 8.12.10/8.12.10; Tue, 23 Dec 2003 16:32:07
   -0800 (PST)
C: helo 129.210.16.8
S: 250 server8.engr.scu.edu Hello dhcp-19-198.engr.scu.edu [129.210.19.198], pleased to
   meet you
C: mail from: jholliday@engr.scu.edu
S: 250 2.1.0 jholliday@engr.scu.edu... Sender ok
C: rcpt to: tschwarz
S: 250 2.1.5 tschwarz... Recipient ok
C: data
S: 354 Enter mail, end with "." on a line by itself
C: This is a spoofed message.
C: .
S: 250 2.0.0 hBO0W76P002752 Message accepted for delivery
C: quit
S: 221 2.0.0 server8.engr.scu.edu closing connection
```

Spotting spoofed messages

- Contents usually gives a hint:

1. Each SMTP server application adds a different set of headers or structures them in a different way
   - A good investigator knows these formats
2. Use internet services in order to verify header data
   - However, some companies can outsource email or use internal IP addresses
3. Look for breaks / discrepancies in the "Received" lines

Look for inconsistencies in the Received field

```
From jholliday@engr.scu.edu Tue Dec 23 16:44:55 2003
Return-Path: <jholliday@engr.scu.edu>
Received: from server8.engr.scu.edu (root@server8.engr.scu.edu [129.210.16.8])
by server4.engr.scu.edu (8.12.10/8.12.10) with ESMTP id hBO0itpv008140
for <tschwarz@engr.scu.edu>; Tue, 23 Dec 2003 16:44:55 -0800
From: JoAnne Holliday <jholliday@engr.scu.edu>
Received: from 129.210.16.8 (dhcp-19-198.engr.scu.edu [129.210.19.198])
by server8.engr.scu.edu (8.12.10/8.12.10) with SMTP id hBO0W76P002752
for tschwarz; Tue, 23 Dec 2003 16:41:55 -0800 (PST)
Date: Tue, 23 Dec 2003 16:32:07 -0800 (PST)
Message-Id: <200312240041.hBO0W76P002752@server8.engr.scu.edu>
X-Spam-Checker-Version: SpamAssassin 2.60-rc3 (1.202-2003-08-29-exp) on
server4.engr.scu.edu
X-Spam-Level:
X-Spam-Status: No, hits=0.0 required=5.0 tests=none autolearn=ham version=2.60-r
```

This looks very convincing...

Only hint: received line gives the name of my machine, defaulting to dhcp-19-198

The DHCP server logs might tell you what machine this is, given the time. But you need to know the clock drift at the various machines

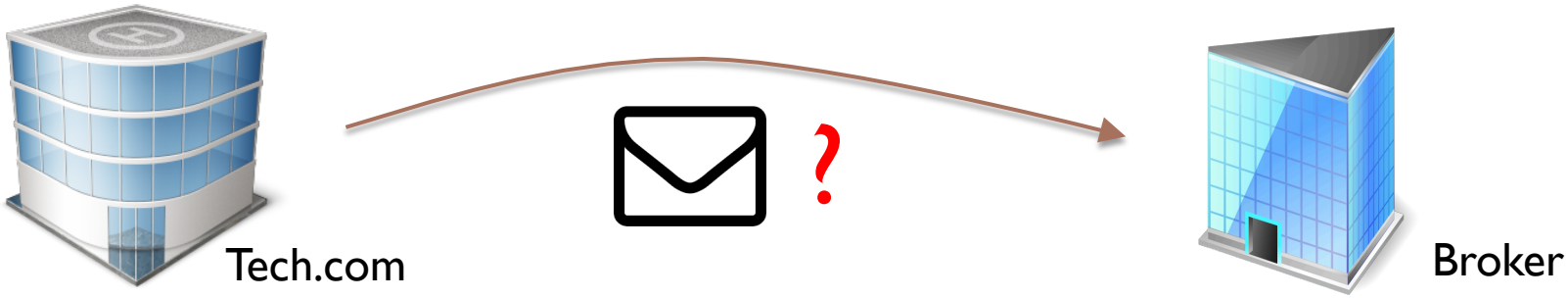
Hints for investigation of fake emails

- Verify all IP addresses
  - Keeping in mind that some addresses might be internal addresses
- Make a time-line of events
  - Change times to universal standard time
  - Look for strange behavior
  - Keep clock drift in mind
- Check server logs

Server logs

- Email logs usually identify email messages by:
  - Account received
  - IP address from which they were sent.
  - Time and date (beware of clock drift)
  - IP addresses
- Many servers keep copies of emails
  - Logs are typically purged after certain # of entries / time
- Very useful for solving cases

Email forensics: Case study



- An email attached to a $20 million dollar lawsuit purported to be from the CEO of "Tech.com" to a venture capital broker. The message outlined guaranteed "warrants" on the next round of funding for the broker.

- "Tech.com filed counterclaim and claimed the email was a forgery. Their law firm engaged us to determine the validity of the message.

Email forensics: Case study

- We imaged all of the CEO's computers at his office and his home. Recalled the email server backup tapes from off-site storage.

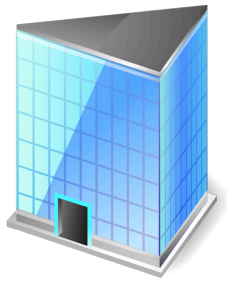
- Searched all hard drives and email server backups for "questioned" message. Search revealed no trace of the message on any of the hard drives or mail spools.

Tech.com

- When the timestamps and message ids were compared with the server logs we found that the "questioned" message could not have gone through either "Tech.com's" webmail or mail server at the time indicated by the date/ time stamp on the message.

Email forensics: Case study

- Based on our analysis defendants filed motion to image and examine broker's computers
- Federal Judge issued subpoena and we arrived at broker's business, but he refused to allow his system to imaged
- Broker's lawyer went into State Court, on a companion case, and got Judge to issue an order for a new Court appointed examiner
- The examination revealed direct proof of the alteration of a valid message's header to create the "questioned" email
- What follows are some of the tools and techniques used to document the activity



Broker

Tracking timestamp inconsistencies

The allegedly received email:

Return-Path: CEO Good_Guy@tech.com
Received: from mail.tech.com (mail.tech.com [201.10.20.152])
by hedgefund.fund.com (8.11.0/8.11.0) ESMTP id
e73MfZ331592; Thu, 3 Aug 2000 15:45:31 -0400
Received: from webmail.tech.com (webmail.tech.com
[10.27.30.190]) by mail.tech.com (Switch-2.0.1/Switch-
2.0.1) ESMTP id e73MfW903843; Thu, 3 Aug 2000
14:41:32 -0500
Received: from tech.com (ostrich.tech.com [10.27.20.190])
by webmail.tech.com (8.8.8+Sun/8.8.8) with ESMTP
id RAA01318; Thu, 3 Aug 2000 14:41:31 -0500
content-class: urn:content-classes:message
Subject: Warrants on $25 Million Funding
Date: Thu, 3 Aug 2000 14:43:47 -0500
MIME-Version: 1.0
Content-Type: application/ms-tnef;
name="winmail.dat"
Content-Transfer-Encoding: binary
Message-ID: <3989e793.87BDEEE2@tech.com>
X-MS-Has-Attach:
X-MS-TNEF-Correlator: <3989e793.87BDEEE2@tech.com>
Thread-Topic: Warrants on $25 Million Funding
Thread-Index: AcHatCZUSkaLe0ajEdaelQACpYcy8A==
From: "CEO Good_Guy@tech.com" <ceo_good_guy@tech.com >
To: "Bad_Guy_Broker" <bad_guy@fund.com>

The Received fields in more detail

```
Received:  from mail.tech.com (mail.tech.com [201.10.20.152])
           by hedgefund.fund.com (8.11.0/8.11.0) ESMTP id
              e73MfZ331592; Thu, 3 Aug 2000 15:45:31 -0400
Received:  from webmail.tech.com (webmail.tech.com
           [10.27.30.190]) by mail.tech.com (Switch-2.0.1/Switch-
           2.0.1) ESMTP id e73MfW903843; Thu, 3 Aug 2000
           14:41:32 -0500
Received:  from tech.com (ostrich.tech.com [10.27.20.190])
           by webmail.tech.com (8.8.8+Sun/8.8.8)  with ESMTP
           id RAA01318; Thu, 3 Aug 2000 14:41:31 -0500
```

- ESMTP id:
  - A unique identification assigned by each intermediate relay or gateway server. This id is usually in a hexadecimal string that is reset each day. Resulting in an id that can be resolved to a time window on a particular server.

Claimed path from the email's Received fields

Received: from mail.tech.com (mail.tech.com [201.10.20.152])
by hedgefund.fund.com (8.11.0/8.11.0) ESMTP id
e73MfZ331592; Thu, 3 Aug 2000 15:45:31 -0400
Received: from webmail.tech.com (webmail.tech.com
[10.27.30.190]) by mail.tech.com (Switch-2.0.1/Switch-
2.0.1) ESMTP id e73MfW903843; Thu, 3 Aug 2000
14:41:32 -0500
Received: from tech.com (ostrich.tech.com [10.27.20.190])
by webmail.tech.com (8.8.8+Sun/8.8.8) with ESMTP
id RAA01318; Thu, 3 Aug 2000 14:41:31 -0500

Compare against server logs: webmail@tech.com

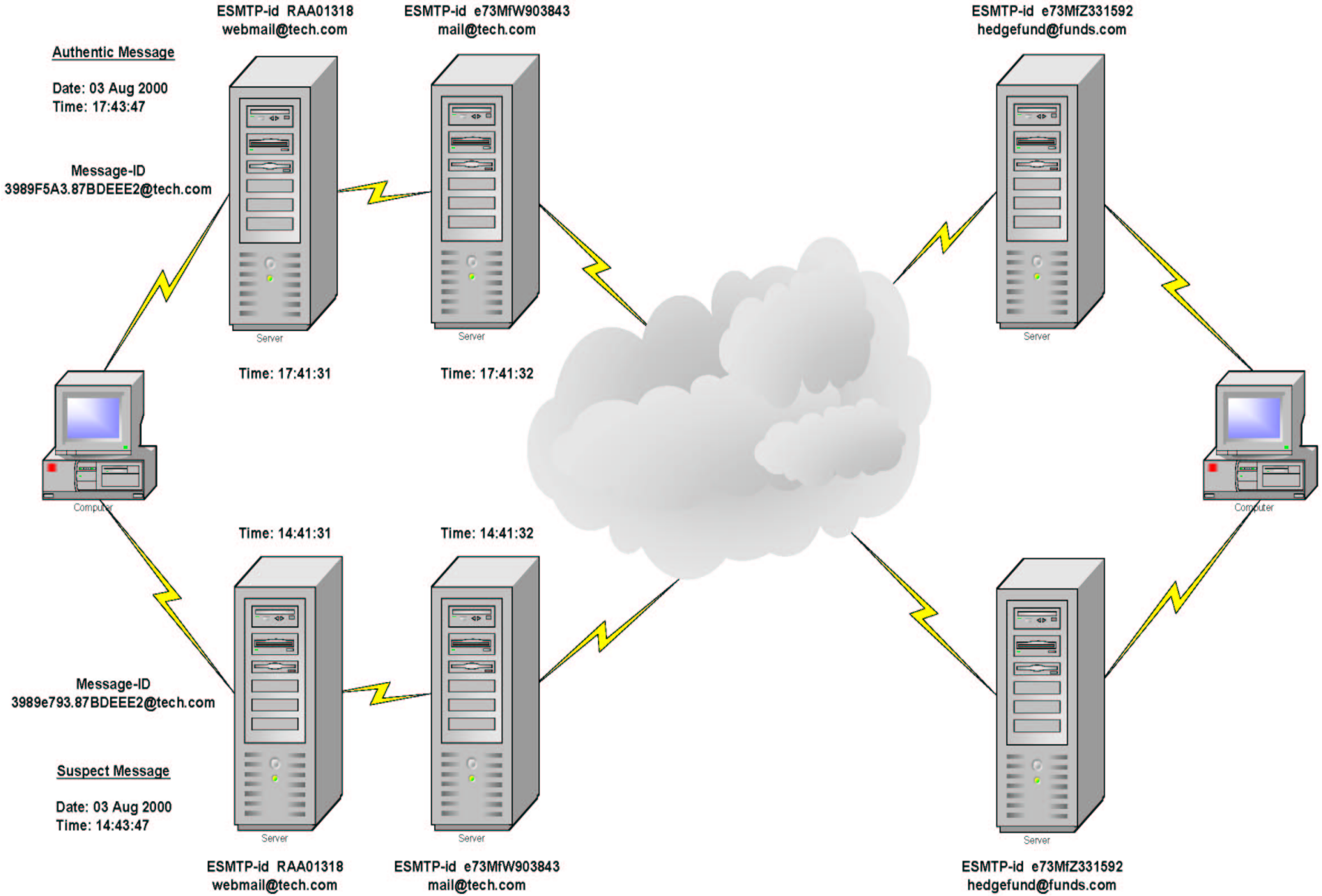
Received: from mail.tech.com (mail.tech.com [201.10.20.152])
by hedgefund.fund.com (8.11.0/8.11.0) ESMTP id
e73MfZ331592; Thu, 3 Aug 2000 15:45:31 -0400
Received: from webmail.tech.com (webmail.tech.com
[10.27.30.190]) by mail.tech.com (Switch-2.0.1/Switch-
2.0.1) ESMTP id e73MfW903843; Thu, 3 Aug 2000
14:41:32 -0500
Received: from tech.com (ostrich.tech.com [10.27.20.190])
by webmail.tech.com (8.8.8+Sun/8.8.8) with ESMTP
id RAA01318; Thu, 3 Aug 2000 14:41:31 -0500

- Analysis of the webmail server logs revealed several issues regarding the validity of the suspect message
  - Matching trace header timestamps and ESMTP ids revealed that RAA01318 was issued at 17:41:31 to the authentic message
  - Comparing the 14:41:31 timestamp of the suspect message with the log revealed the server was assigning ESMTP ids beginning with "OAA" not "RRA" as represented in the header

Compare against server logs: webmail@tech.com

Received: from mail.tech.com (mail.tech.com [201.10.20.152])
by hedgefund.fund.com (8.11.0/8.11.0) ESMTP id
e73MfZ331592; Thu, 3 Aug 2000 15:45:31 -0400
Received: from webmail.tech.com (webmail.tech.com
[10.27.30.190]) by mail.tech.com (Switch-2.0.1/Switch-
2.0.1) ESMTP id e73MfW903843; Thu, 3 Aug 2000
14:41:32 -0500
Received: from tech.com (ostrich.tech.com [10.27.20.190])
by webmail.tech.com (8.8.8+Sun/8.8.8) with ESMTP
id RAA01318; Thu, 3 Aug 2000 14:41:31 -0500

- Analysis of the mail server logs confirmed that the suspect message was not authentic
  - Matching trace header timestamps and ESMTP ids revealed that the authentic Message-ID was logged at 17:41:32 and assigned ESMTP id e73MfW903843 then it was sent to the hedgefund@fund.com server and it was assigned a new ESMTP id e73MfZ331592
  - Comparing the 14:41:32 timestamp of the suspect message with the log revealed the were no messages for over an hour during that time frame


ESMTP-id RAA01318
webmail@tech.com
ESMTP-id e73MfW903843
mail@tech.com
ESMTP-id e73MfZ331592
hedgefund@funds.com
Authentic Message
Date: 03 Aug 2000
Time: 17:43:47
Message-ID
3989F5A3.87BDEEE2@tech.com
Server
Server
Server
Computer
Computer
Time: 17:41:31
Time: 17:41:32
Time: 14:41:31
Time: 14:41:32
Message-ID
3989e793.87BDEEE2@tech.com
Suspect Message
Date: 03 Aug 2000
Time: 14:43:47
Server
Server
Server
ESMTP-id RAA01318
webmail@tech.com
ESMTP-id e73MfW903843
mail@tech.com
ESMTP-id e73MfZ331592
hedgefund@funds.com

Working with mail servers

- Some initial things to consider:
  - Which users are serviced?
  - E-mail retention policies of the company
  - Accessibility of the e-mail server
- Examining UNIX email logs: an example
  - /Etc/Sendmail.cf
    - Configuration information for Sendmail
  - /Etc/Syslog.conf
    - Specifies how and which events Sendmail logs
  - /Var/Log/Maillog
    - SMTP and POP3 communications
  - Check UNIX man pages for more information

Working with resident email files

- Some users store email is stored locally
  - Great benefit for forensic analysts because the e-mail is readily available when the computer is seized
- Begin by identifying e-mail clients on system
- You can also search by file extensions of common e-mail clients

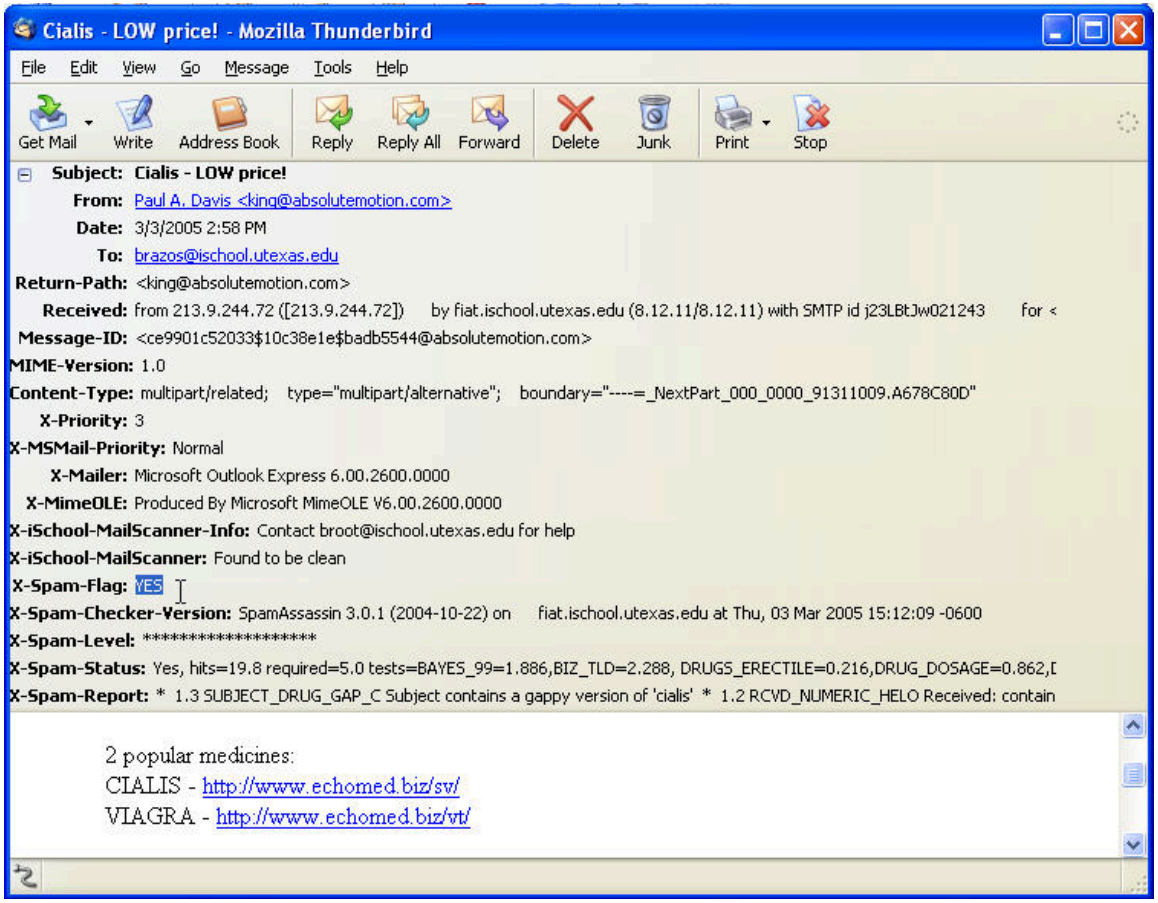
Local email storage files

- Email clients have own file formats for storing email

E-Mail Client	Extension	Type of File
AOL	.abi	AOL6 organizer file
	.aim	Instant Message launch
	.arl	Organizer file
	.bag	Instant Messenger file
Outlook Express	.dbx	OE mail database
	.dgr	OE fax page
	.email	OE mail message
	.eml	OE electronic mail
Outlook	.pab	Personal address book
	.pst	Personal folder
	.wab	Windows address book

Accessing headers from email clients

- Different tools have different ways to read headers:



To enable headers

- Eudora:
  - Use the Blah Blah Blah button
- Hotmail:
  - Options → Preferences → Message Headers.
- Juno:
  - Options → Show Headers
- MS Outlook:
  - Select message and go to options.
- Yahoo!:
  - Mail Options → General Preferences → Show all headers.

Headers on a WebMail client

This message is not flagged. [Flag Message - Mark as Unread]

From Thom Thomas Tue Jul 15 18:34:03 2003

X-Apparently-To: badboy83210@yahoo.com via 216.136.130.41; 15 Jul 2003 18:34:04 -0700 (PDT)
Return-Path: <takin00@hotmail.com>
Received: from 64.4.27.104 (EHLO hotmail.com) (64.4.27.104) by mta114.mail.scd.yahoo.com with SMTP; 15 Jul 2003 18:34:04 -0700 (PDT)
Received: from mail pickup service by hotmail.com with Microsoft SMTPSVC; Tue, 15 Jul 2003 18:34:04 -0700
Received: from 130.218.62.189 by by8fd.bay8.hotmail.msn.com with HTTP; Wed, 16 Jul 2003 01:34:03 GMT
X-Originating-IP: [130.218.62.189]
X-Originating-Email: [takin00@hotmail.com]
From: "Thom Thomas" <takin00@hotmail.com> | This is spam | Add to Address Book
To: badboy83210@yahoo.com
Bcc:
Subject: here are the headers
Date: Tue, 15 Jul 2003 21:34:03 -0400
Mime-Version: 1.0
Content-Type: text/plain; format=flowed
Message-ID: <BAY8-F104NtDEJmGzrL000148b4@hotmail.com>
X-OriginalArrivalTime: 16 Jul 2003 01:34:04.0105 (UTC) FILETIME=[57485390:01C34B3A]
Content-Length: 223

Forensic tools and services

Email forensic tools

- AccessData's FTK
- EnCase
- FINALeMAIL
- Sawmill-GroupWise
- DBXtract
- MailBag
- Assistant
- Paraben

Online services

- Geolocation of IP address
  - https://www.iplocation.net

Antiforensics: Open relays

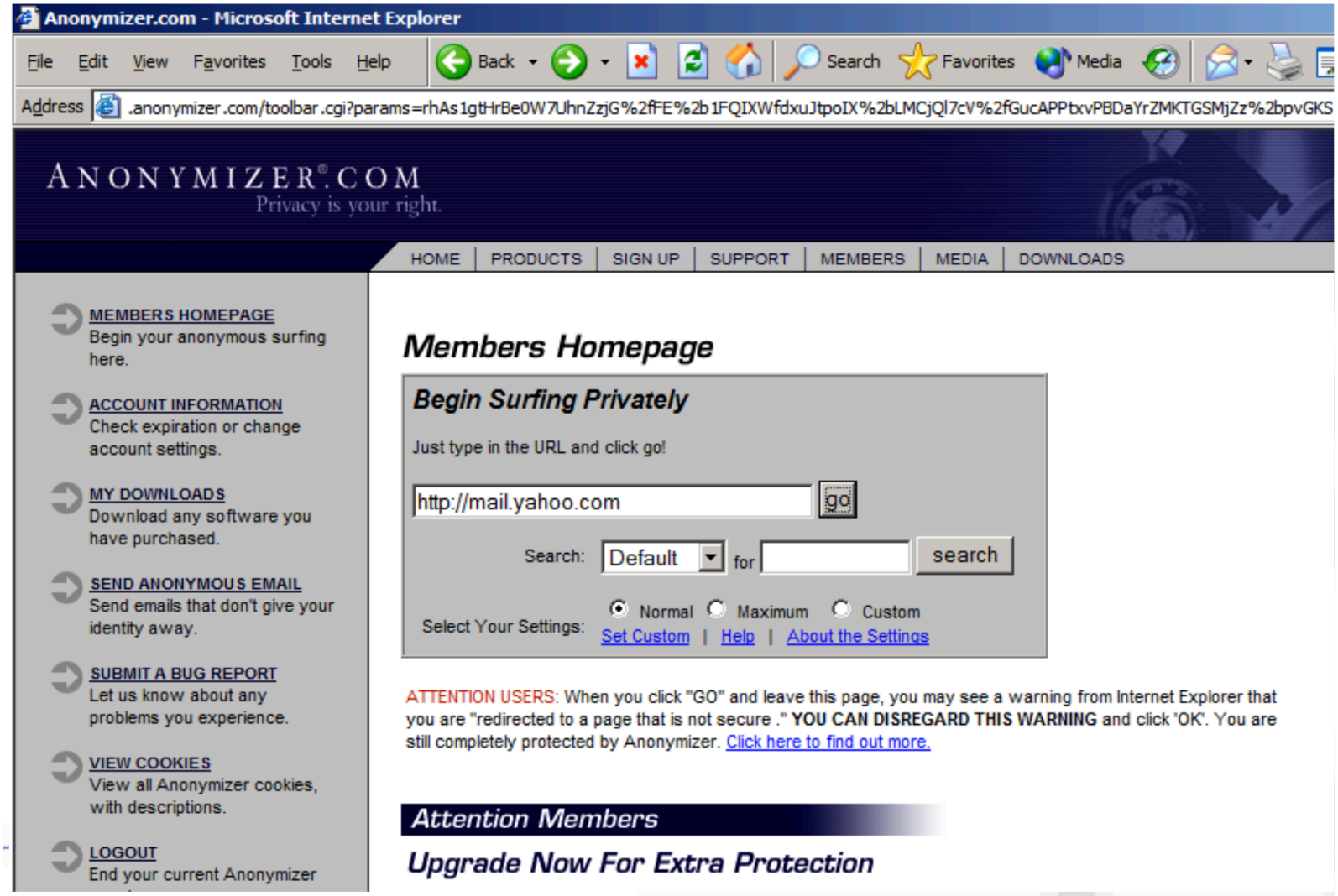
- Open relays
  - SMTP server configured in such a way that it allows anyone on the Internet to send e-mail through it, not just mail destined to or originating from known users
- Spoofers use open relays to attempt to hide the person and IP of the system that sent the email
- Where to look for evidence:
  - Email header will contain the originating address
  - Open relay log files will also contain the originating address

Antiforensics: False received from header

- Leads the investigator to the wrong server by adding a seemingly valid Received from header
- To avoid detection, the spoofer's real address will be recorded somewhere in the Received from headers, but the investigator will not know which one

- Where to look for evidence:
  - Email received from headers will contain the actual IP address of the originating system, you just won't know which header is correct
  - Trace backwards by looking at the log files of the servers the mail claims to have passed through: once you get to a server that has no record of the email, the previous system is the originating IP

Antiforensics: Anonymizer



Antiforensics: Anonymizer

- Where to look for evidence:
  - The email headers and web mail log files will point back to the anonymizer
  - You will need to look at the anonymizer's log files to determine what IP address accessed the web email account at the specific time the email was sent
  - If the anonymizer is a paying service then you can also request subscriber information for the account that was using the anonymizer to send the web based email.

Conclusions

- Network forensics cares about tracking the exchanging of messages in a networked system
- Email is a fundamental networked application that provides a very important source of digital evidence
- The primary focus of email forensics is the analysis of email headers and server logs

References

- Primary bibliography
  - [Casey05], Chapter 21, 23.2.2

Next class

- Web and online anonymity