CS155b: E-Commerce

Lecture 23: April 17, 2003 E-Mail Abuse: Spam and Viruses

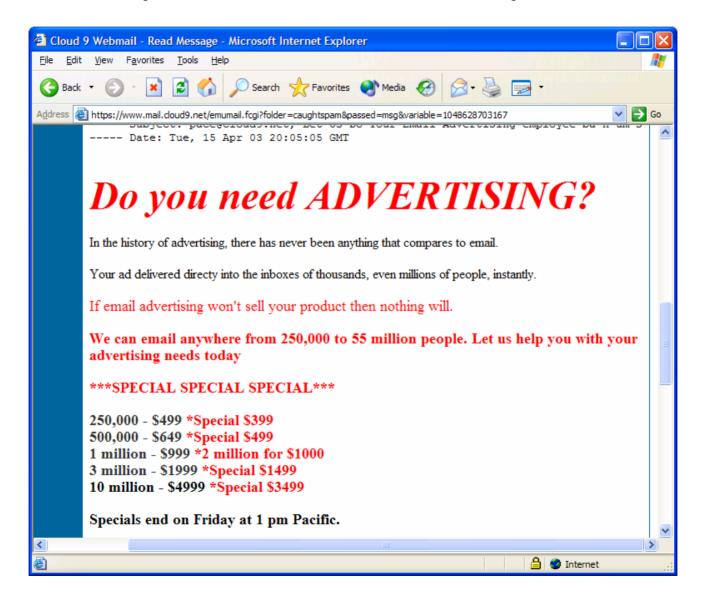
Acknowledgements: V. Ramachandran (Yale) and C. Dwork (Microsoft)

What is Spam?

Source: Mail Abuse Prevention System, LLC

- Spam is unsolicited bulk e-mail (primarily used for advertising).
- · An electronic message is spam IF:
 - (1) the recipient's personal identity and context are irrelevant because the message is equally applicable to many other potential recipients; AND
 - (2) the recipient has not verifiably granted deliberate, explicit, and still-revocable permission for it to be sent; AND
 - (3) the transmission and reception of the message appears to the recipient to give a disproportionate benefit to the sender.

Spam About Spam



Why is Spam such a problem?

- · Simple answer: People don't like it!
- · Cost:
 - Postal mail and telephone calls cost money.
 - Sending e-mail does not (in general).
- · Speed:
 - Messages created and sent to many users instantaneously, without human effort.
 - (Almost) Instant notification of success or failure to reach destination.

Consequences of Spam

- Large amounts of network traffic (?)
 - Network congestion
 - Mail servers can be overloaded with network requests; could slow mail delivery
- Wasted Time and Storage
 - Downloading headers & checking mail takes longer
 - More unwanted mail to delete
 - E-mail must be stored at servers
 - Microsoft: 65-85% of storage costs go to Spam

How is E-mail Sent?

Source: RFC 821 (SMTP)

```
SMTP
User |<-->|
 ----+ | Sender- |Commands/Replies| Receiver-|
+----+ | SMTP |<----+
| File |<-->|
                 l and Mail
                                  |<-->| File |
|System|
                                          |System|
         Sender-SMTP
                              Receiver-SMTP
                  Model for SMTP Use
                      Figure 1
```

Example Mail Exchange

```
[vijayr@cyndra ~]$ telnet netra 25
Trying 128.36.229.21...
Connected to netra.cs.yale.edu (128.36.229.21).
Escape character is '^]'.
220 netra.cs.yale.edu ESMTP Postfix
HELO cyndra
250 netra.cs.yale.edu
MAIL FROM: vijayr@cs.yale.edu
250 Ok
RCPT TO: vijayr@whiqclio.princeton.edu
250 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
This is a test
250 Ok: queued as EE0A5D728E
QUIT
221 Bye
Connection closed by foreign host.
```

Tracking Spam

- SMTP runs on top of TCP.
 - Packets are acknowledged.
 - Source of packets is known in any successful mail session.
- SMTP servers add the IP address and hostname of every mail server or host involved in the sending process to the e-mail's message header.
- But, dynamic IP addresses and large ISPs can make it difficult to identify senders.

E-Mail Headers

```
Return-<iw3bvad9nfk@kali.com.cn>
     Path:
 X-Original-ram@cloud9.net
       To:
 Delivered-ram@cloud9.net
 Received: from localhost (localhost [127.0.0.1]) by russian-caravan.cloud9.net (Postfix) with
           ESMTP id EF454AADC; Tue, 15 Apr 2003 14:09:07 -0400 (EDT)
 Received: from russian-caravan.cloud9.net (localhost [127.0.0.1]) by localhost (VaMailArmor-
           2.0.1.7) id 09388-51E384ED; Tue, 15 Apr 2003 14:09:07 -0400
 Received: from host217-40-121-145.in-addr.btopenworld.com (host217-40-121-145.in-
           addr.btopenworld.com [217.40.121.145]) by russian-caravan.cloud9.net (Postfix)
           with SMTP id B6BEAAA23; Tue, 15 Apr 2003 14:08:06 -0400 (EDT)
 Received: from wzr4k.wb23acf.com [110.70.78.125] by host217-40-121-145.in-
           addr.btopenworld.com id u4162Pp3anwF for <pace@cloud9.net>: Tue, 15 Apr
           2003 20:05:05 +0100
  Message- <8$$7g2$-0lnu1u$-a4-s93-5pw5x@zat9bhegt.y0t>
        Id:
     From: "Hubert Rivers" <iw3bvad9nfk@kali.com.cn>
       To: pace@cloud9.net
       Cc: <photo@cloud9.net>, <promo@cloud9.net>, <ram@cloud9.net>,
           <reynolds@cloud9.net>, <rl@cloud9.net>, <robertl@cloud9.net>
   Subject: pace@cloud9.net, Let Us Do Your Email Advertising employee ba h um s
     Date: Tue, 15 Apr 03 20:05:05 GMT
 X-Priority: 1
  X-Msmail-High
   Priority:
  X-Mailer: MIME-tools 5.503 (Entity 5.501)
     MIME- 1.0
   Version:
   Content-multipart/alternative; boundary="4DFAC9BD.DC. 5ED6.9"
     Type:
X-Antivirus: checked by Vexira MailArmor (version: 2.0.1.7; VAE: 6.19.0.3; VDF: 6.19.0.6; host:
           russian-caravan.cloud9.net)
```

Spoofing E-mail Headers

- Most e-mail programs use (and most people see) only the standard "To," "Cc," "From," "Subject," and "Date" headers.
- All of these are provided as part of the mail data by the mail sender's client.
- · Any of this information can be falsified.
- The only headers you can always believe are message-path headers from trusted SMTP servers.

Open Mail Relays

- An open mail relay is an SMTP server that will send mail when the sender and recipient are not in the server's domain.
- These servers can be used to obfuscate the mail-sending path of messages.
- Mail-sending cost can be offloaded to servers not under spammers' control.
- Most servers are now configured to reject relays, and many servers will not accept mail from known open mail relays.

Relay Rejection

```
[vijayr@cyndra ~]$ telnet mail.cloud9.net 25
Trying 168.100.1.4...
Connected to russian-caravan.cloud9.net (168.100.1.4).
Escape character is '^]'.
220 russian-caravan.cloud9.net ESMTP Postfix
MAIL FROM:user@cloud9.net
250 Ok
RCPT TO:vijayr@cs.yale.edu
554 <vijayr@cs.yale.edu>: Relay access denied
QUIT
221 Bye
Connection closed by foreign host.
```

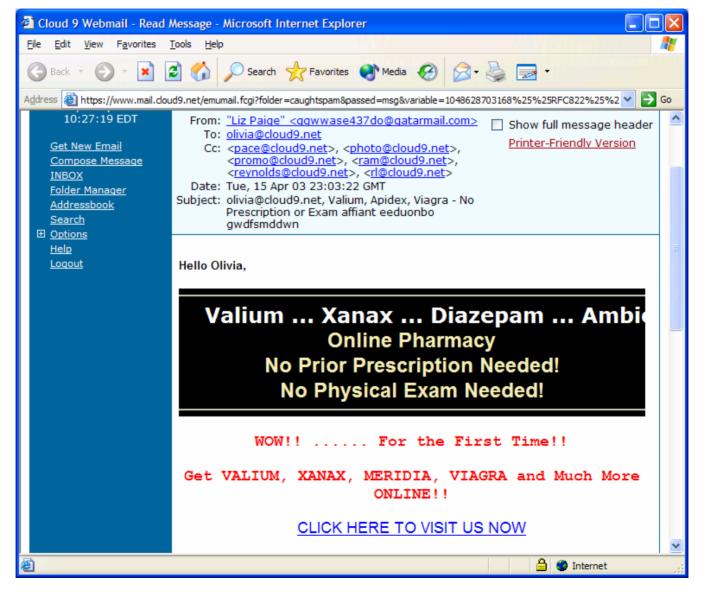


- SpamAssassin is a spam-fighting tool.
- Primary development efforts exist for the open-source, UNIX-compatible version. The source code and select Linux binaries are available for free download (for noncommercial use).
- Commercial and Windows-compatible products are available that use the technology.
- SpamAssassin is installed on many ISP mail servers and is used by the CS dept. at Yale.

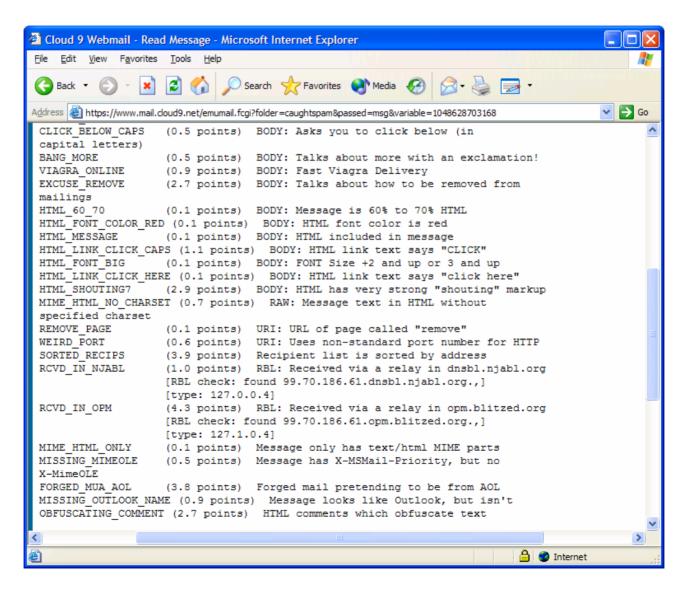
SpamAssassin: Overview

- Filtering is done at the mail server.
 (But, the technology can also be used to create plug-ins for mail clients.)
- Messages receive a score.
 - Message content and headers are parsed.
 - The more occurrences of Spam-like items in the message, the higher the score.
- Messages with scores above a threshold are automatically moved from the user's INBOX.
- Tolerance for Spam is user-configurable.

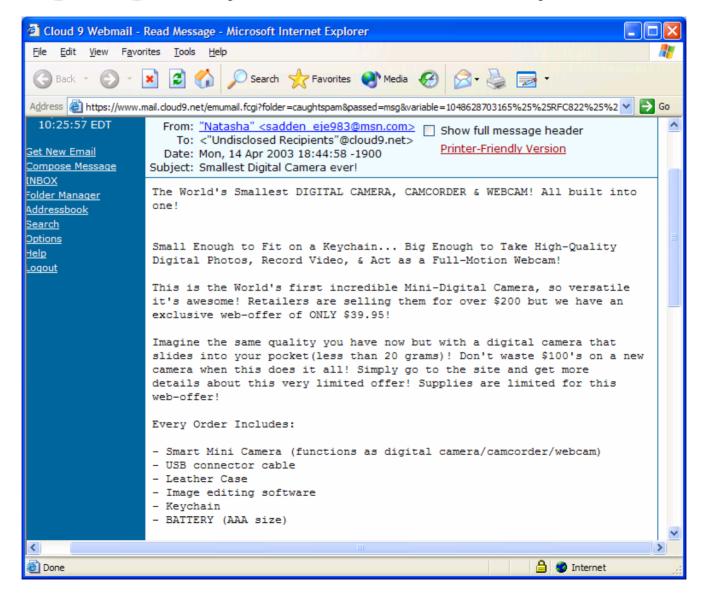
Judging Spam: Example #1



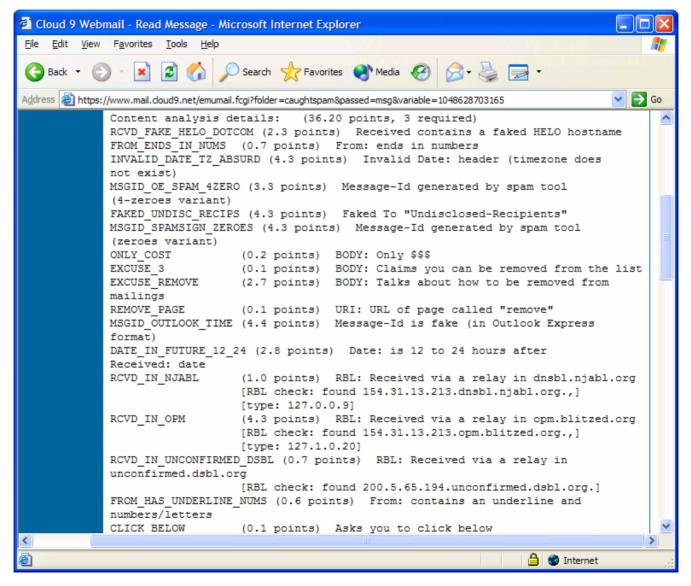
Judging Spam: Results #1



Judging Spam: Example #2



Judging Spam: Results #2



SpamAssassin: Techniques

Source: SpamAssassin.org (developers' website)

The spam-identification tactics used include:

- header analysis: spammers use a number of tricks to mask their identities, fool you into thinking they've sent a valid mail, or fool you into thinking you must have subscribed at some stage. SpamAssassin tries to spot these.
- text analysis: again, spam mails often have a characteristic style (to put it politely), and some characteristic disclaimers and CYA text. SpamAssassin can spot these, too.
- blacklists: SpamAssassin supports many useful existing blacklists, such as <u>mail-abuse.org</u>, <u>ordb.org</u> or others.
- Razor: Vipul's Razor is a collaborative spam-tracking database, which works by taking a signature of spam messages. Since spam typically operates by sending an identical message to hundreds of people, Razor short-circuits this by allowing the first person to receive a spam to add it to the database -- at which point everyone else will automatically block it.

Once identified, the mail can then be optionally tagged as spam for later filtering using the user's own mail user-agent application.

Tricks to Avoid Filters

- Use MIME-/UU-encoding for messages.
 - E-mail messages can be in complex formats; this allows messages to contain multiple parts and attachments.
 - To preserve warping of content, message parts and attachments can be transformed using a standard encoding method.
 - E-mail clients are supposed to decode message parts when presented to the reader.
 - Basic filters often do not process encoded text!
- Insert HTML comments between words.

Examples of Tricks

Source: spam-stopper.net

Reply-To: <yobaby5132h16@yahoo.com>

Message-ID: <031c06e62c2b\$8445d5b2\$5da01aa2@qjwmpp>

From: <yobaby5132h16@yahoo.com>

To: Lower bills Subject: ** Approved.

Date: Tue, 24 Sep 2002 11:24:41 +0600

MiME-Version: 1.0

Content-Type: multipart/mixed;

boundary="----=_NextPart_000_00A3_83C84A5C.B4868C82"

X-Priority: 3 (Normal) X-MSMail-Priority: Normal

X-Mailer: Internet Mail Service (5.5.2650.21)

Importance: Normal

-----=_NextPart_000_00A3_83C84A5C.B4868C82 Content-Type: text/html; charset="iso-8859-1"

Content-Transfer-Encoding: base64

PGh0bWw+DQo8Ym9keT4NCjxmb250IGNvbG9yPSJmZmZmZmYiPnNreTwZm9u dD4NCjxwPllvdXlgaG9tZSByZWZpbmFuY2UgbG9hbiBpcyBhcHByb3ZIZCE8 Ynl+PC9wPjxicj4NCjxwPlRvlGdldCB5b3VylGFwcHJvdmVklGFtb3VudCA8 YSBocmVmPSJodHRwOi8vd3d3LjJnZXRmcmVlcXVvdGVzLmNvbS8iPmdvDQpo ZXJlPC9hPi48L3A+DQo8Ynl+PGJyPjxicj48Ynl+PGJyPjxicj48Ynl+PGJyPjxicj48Ynl+PGJyPjxicj48Ynl+PGJyPjxicj48Ynl+DQo8cD5UbyBiZSBleGNsdWRIZCBmcm9tlGZ1cnRoZXlgbm90aWNlcyA8YSBo cmVmPSJodHRwOi8vd3d3LjJnZXRmcmVlcXVvdGVzLmNvbS9yZW1vdmUuaHRt bCl+Z28NCmhlcmU8L2E+LjwcD4NCjxmb250IGNvbG9yPSJmZmZmZmYiPjFnYXRl DQo8L2h0bWw+DQo4MzM0Z1RpbzqtbDk=

As se<!--5-->en on NB<!--D-->C, CBS, and CN<!--H-->N, and even Opr<!--D-->ah! The health

health<b

been report<!--resale-->ed on by the Ne<!--test-->w Engl<!--->and Jour<!--F-->nal of Medi<!--F-->cine.
br> For<!--resale-->get aging and d<!--->ieting forever! And it's Gua<!--S-->ranteed!
br>

* Red<!--lo-->uce body fat and build lean muscle WIT<!--resale-->HOUT EXERCISE!

* Enha<!--resale-->ce se<!--la-->xual perf<!--hehe-->ormance

- * Rem<!--resale-->ove wrinkles and cellulite
br> * Lower blood pres<!--resale-->sure and improve choles<!--->terol profile
br> * Imp<!--resale-->rove sleep, vision and me<!--->mory
br>
- * Resto<!--resale-->re hair color and gro<!--->wth
* Stren<!--resale-->gthen the immune sys<!--->tem
* Incre<!--resale-->ase ener<!--->gy and card<!--->iac output
* Turn bac<!--resale-->k your body's biol<!--->ogical time cl<!--->ock 10-20 years
in 6 months of usage !!!
br>
's a href="http://www.chinaniconline.com/ultimatehgh/">FOR FRE<!--o->E INFO<!--you-->RMATION AND G<!--love-->ET FREE 1 MON<!--resale-->TH SUPPLY OF HG<!--->H CLICK HERE
br>
BR>
br>
BR>
br>
BR>
br>
BR>
br>
br>
com/ultimatehgh/remove.php?userid=resale@globals pider.net">Click Here

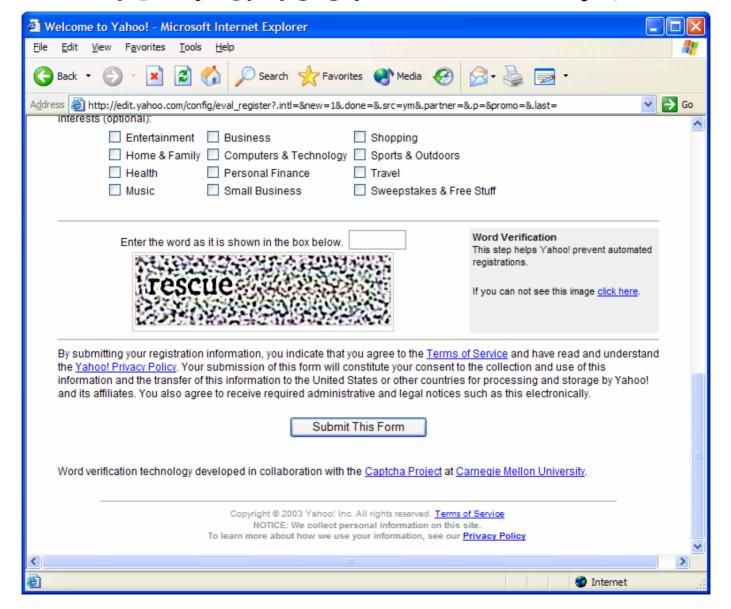
Proposals to Eliminate Spam

- · Charge a micro-payment for e-mail.
- · Computational method: force senders to "prove" that they spend some minimum amount of time per sender per message.

Prove You are a Human

- CAPTCHA: Completely Automated Public Turing test for telling Computers and Humans Apart
- Require people to pass CAPTCHAs to sign up for free e-mail accounts.
 - Perform some easy-for-human but difficult-for-computer computation
 - Identify words, or find objects in pictures, e.g.
- ? The future: build into the e-mail sending process some way to prove e-mail senders are humans or authorized automated agents

The Yahoo! CAPTCHA



Viruses

A computer virus is a piece of code, often malicious, that is intended to transmit itself between computers and replicate itself and/or execute instructions without the user's knowledge or intent.

Examples: Michelangelo, I-Love-You, Melissa, Slammer, Code Red

How Does One Get Infected?

Simple answer:

Run malicious code on your computer.

Simple reaction:

Then I won't.

Problem:

What if you are tricked into doing it? Or don't know it's happening?

Types of Viruses

- Trojan Horses: disguised to do one thing, but do another when run
- Boot Sector Viruses: reside in system sectors; run in the background while resident in memory; copy themselves to other disks
- File Infectors: modify portions of executable files on disk so that virus code is unknowingly executed
- Macro Viruses: take advantage of the programmability of documents; run when infected files are accessed
- Worms: replicate across networks, possibly through proprietary software protocols
- E-mail Viruses: transmitted through e-mail, often through attachments

Can you get infected simply by reading an e-mail or viewing a web page?

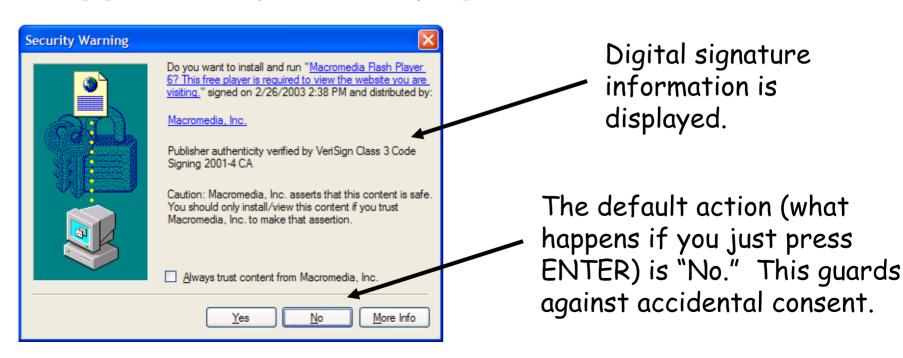
Can you get infected simply by reading an e-mail or viewing a web page?

YES. But your security settings have to allow it, e.g., if you permit scripts to run in HTML e-mail that could contain malicious code.

Plain text cannot contain a virus.

Consent to Run Code

Most browsers that have the capability to execute malicious, remote code will ask you for consent before running anything triggered by a web page.



Can you get infected by viewing a picture attachment to an e-mail?

Can you get infected by viewing a picture attachment to an e-mail?

NO. But you can be fooled by receiving an attachment that *looks like* a picture but is really something else.

Always check the type of a file.

Can I get infected if I own a Mac?

Can I get infected if I own a Mac?

YES. You might not be affected by the same viruses because the code might not run, but there are some Mac worms and e-mail viruses, and Mac files can be carriers of Windows macro viruses.

Beware of Attachments

- Back in the days of MS-DOS, code lived in three types of files: COM, EXE, BAT.
 Problem: If you have a virus WP.COM and a program WP.EXE, typing "WP" causes the virus to run because of precedence rules.
- As programs become more feature-rich and systems become more complex, executable code becomes part of more file types.

Files That Can Contain Code

How many extensions do you recognize?

.com	.exe	.bat	.scr	.pif
.vbs	.js	.vbx	.ocx	.dll
.doc	.xls	.ppt	.eml	.pl
.class	.htm(I)	.hta	.asp(x)	

Example: Melissa

- Microsoft Word macro virus
- On document load, the AutoExec macro runs, containing code that:
 - uses Microsoft Office / Windows features to access the address book and e-mail others infected files; AND
 - infects the default template for Word documents, so that any new Word file on the machine contains the infected **AutoExec** macro.

Example: Code Red

- Microsoft IIS worm
- Uses a "buffer overflow" bug in web server software to transmit and run itself.
- Replicates wildly by sending requests across the Internet from infected machines, causing congestion.
- · Changes web pages on infected machines.
- Launches a DDoS attack on <u>www.whitehouse.gov</u>.

Other Nasty Virus Tricks

- Modify system files.
- Force system to run virus at start-up.
- Intercept and modify requests to the operating system and provide false information (e.g., as done by "stealth" viruses).
- · Change local security settings.
- Run as an Internet server in the background, creating a "back door."

Viruses and Business

- Consider Slammer, the SQL-server worm. SQL server is a Microsoft database product. Hosts running it are often connected to the Internet so that systems can easily share data.
- Slammer infected 90% of vulnerable computers in 10 minutes and reached its peak traffic rate of 55M scans/sec after three minutes (CNET.com).

The Cost of Disinfection

Source: CNET.com News

Productivity Losses:

Klez:

\$9 billion

LoveLetter:

\$8.8 billion

Code Red:

\$2.6 billion

SQL Slammer:

\$0.95-1.2 billion