Guest Lecture - CSCI E-40

Email & Spam Filtering

Joe Pranevich

November 14, 2017

About Me

- IT/Operations professional, 19 years running leading sites and infrastructure
- Harvard Extension Teaching Assistant/Fellow since 2005
- Cloud, Datacenters, Networks, Servers

Why Talk About Spam?

Spam is the single greatest challenge to email providers and users.

Without the constant (and often invisible) level of anti-spam efforts by ISPs and client software, email would be useless.

Current estimates suggest ~60% of mail is presently spam.

Tragedy of the Commons

Spam Definition

A message is Spam only if it is both <u>Unsolicited</u> and <u>Bulk</u>.

- Unsolicited Email is normal email

(examples: first contact enquiries, job enquiries, sales enquiries)

- Bulk Email is normal email

(examples: subscriber newsletters, customer communications, discussion lists)

(Source: Spamhaus.org)

Terminology

Spam by any other name:

- "Unsolicited Bulk Email" (UBE)
- "Unsolicited Commercial Email" (UCE)
- I still say "spam"

Timeline of Email & Spam

- 1960-65 – Transfer of “email” across different users of the same system
- 1966 – Email over a network invented
- 1969 – Email over the ARPANET
- 1971 – @-sign first used for email
- 1971 – “Mailbox Protocol” (RFCs 196, 221)
- 1978 – First spam-mail sent – 600 recipients, advertising Digital Equipment Corp
- 1982 – SMTP (RFC 821)
- 1992 – First large-scale spams sent
- 2010 – Spam decreases for the first time - 90%+
- 2017 - Spam rates continue to decline - ~60%

The World's First Spam - 1978

DIGITAL WILL BE GIVING A PRODUCT PRESENTATION OF THE NEWEST MEMBERS OF THE DECSYSTEM-20 FAMILY; THE DECSYSTEM-2020, 2020T, 2060, AND 2060T. THE DECSYSTEM-20 FAMILY OF COMPUTERS HAS EVOLVED FROM THE TENEX OPERATING SYSTEM AND THE DECSYSTEM-10 <PDP-10> COMPUTER ARCHITECTURE. BOTH THE DECSYSTEM-2060T AND 2020T OFFER FULL ARPANET SUPPORT UNDER THE TOPS-20 OPERATING SYSTEM. THE DECSYSTEM-2060 IS AN UPWARD EXTENSION OF THE CURRENT DECSYSTEM 2040 AND 2050 FAMILY. THE DECSYSTEM-2020 IS A NEW LOW END MEMBER OF THE DECSYSTEM-20 FAMILY AND FULLY SOFTWARE COMPATIBLE WITH ALL OF THE OTHER DECSYSTEM-20 MODELS.

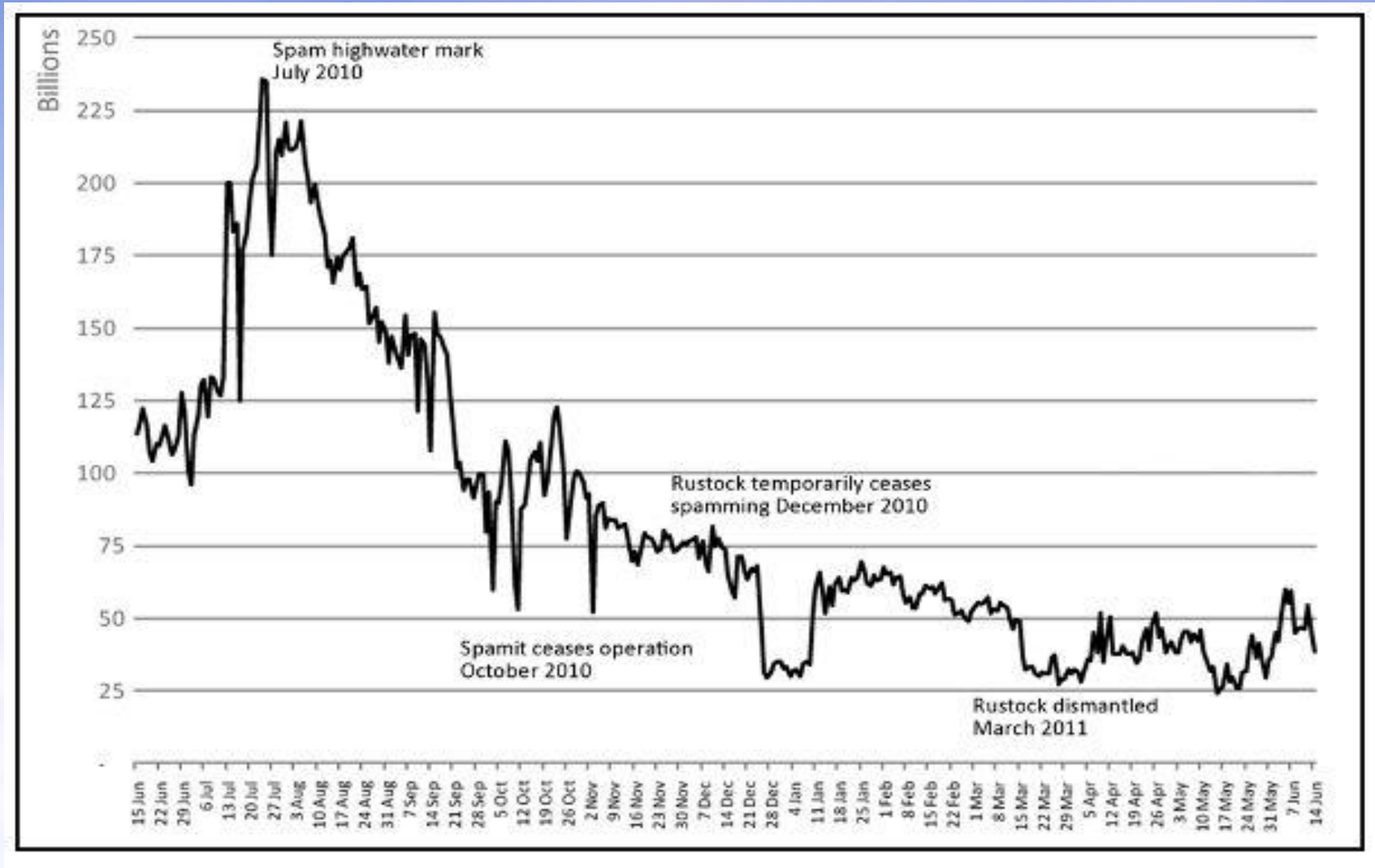
WE INVITE YOU TO COME SEE THE 2020 AND HEAR ABOUT THE DECSYSTEM-20 FAMILY AT THE TWO PRODUCT PRESENTATIONS WE WILL BE GIVING IN CALIFORNIA THIS MONTH. THE LOCATIONS WILL BE:

TUESDAY, MAY 9, 1978 - 2 PM
HYATT HOUSE (NEAR THE L.A. AIRPORT)
LOS ANGELES, CA

THURSDAY, MAY 11, 1978 - 2 PM
DUNFEY'S ROYAL COACH
SAN MATEO, CA
(4 MILES SOUTH OF S.F. AIRPORT AT BAYSHORE, RT 101 AND RT 92)

A 2020 WILL BE THERE FOR YOU TO VIEW. ALSO TERMINALS ON-LINE TO OTHER DECSYSTEM-20 SYSTEMS THROUGH THE ARPANET. IF YOU ARE UNABLE TO ATTEND, PLEASE FEEL FREE TO CONTACT THE NEAREST DEC OFFICE FOR MORE INFORMATION ABOUT THE EXCITING DECSYSTEM-20 FAMILY.

Spam Decreasing Since 2010



http://mashable.com/2011/07/04/spam-decreased-82percent/

The Challenge

The numbers are huge:

> 20-100 billion email messages per day

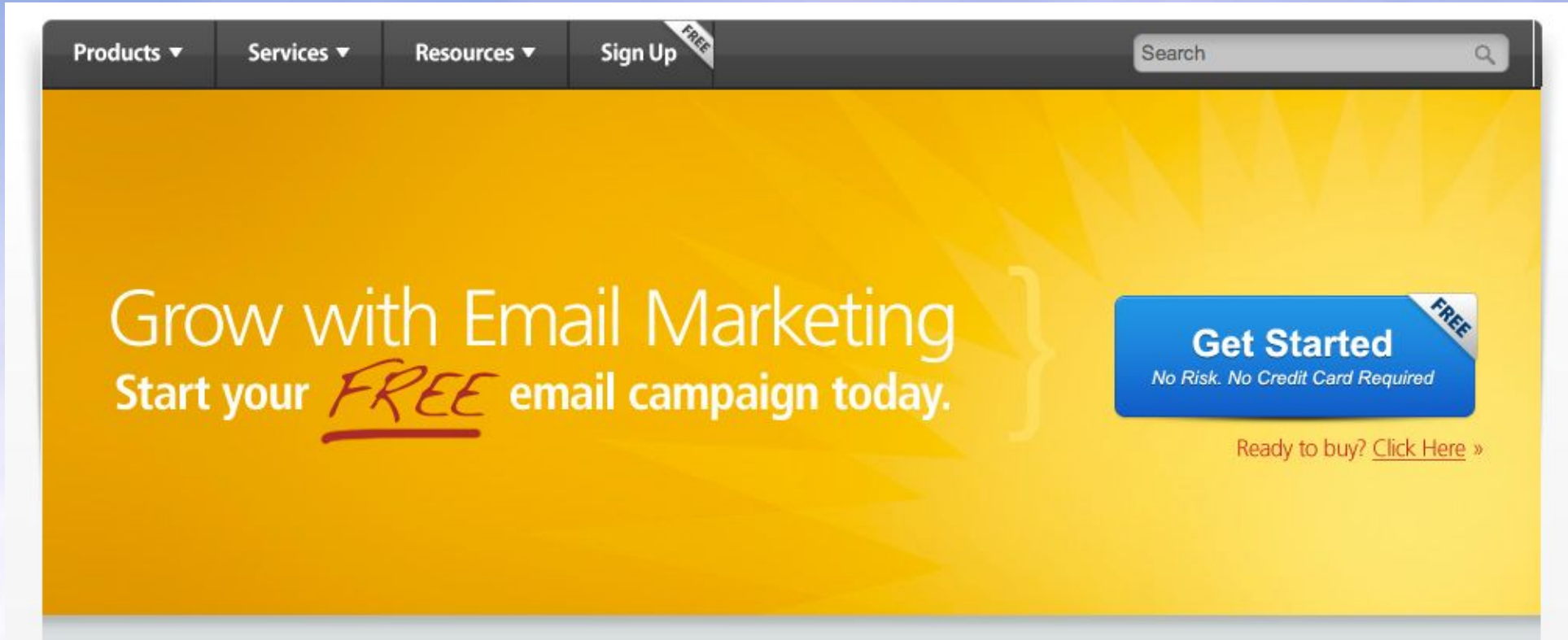
Widely different estimates, but no one doubts the numbers are huge.

Any discussion of email must necessarily devolve into a discussion of Spam

Where does it come from?

- Commercial email senders
  - Many legitimate senders, some gray areas
- Abused mail providers (ISPs & “webmail”)
  - Free and paid services (credit card fraud)
- Botnets
  - Compromised systems – virus writers found they could make an easy buck

Not The Enemy:



Botnets

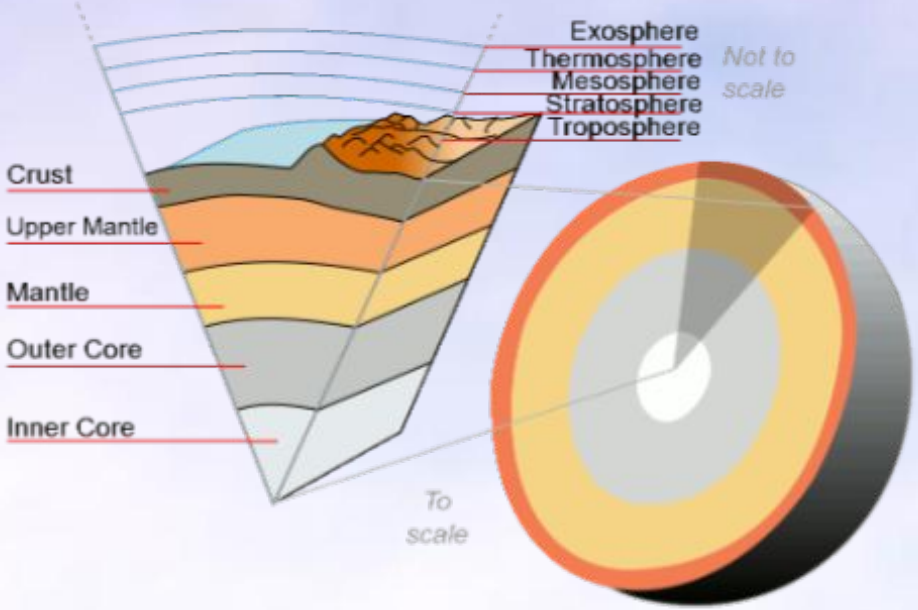
- Botnets are the "big evil" in the email world: computer viruses with a commercial purpose.
- First used for spam in 2001.
- Should ISPs block outbound port 25? What does that say about the end to end principle?

Costs of Spam

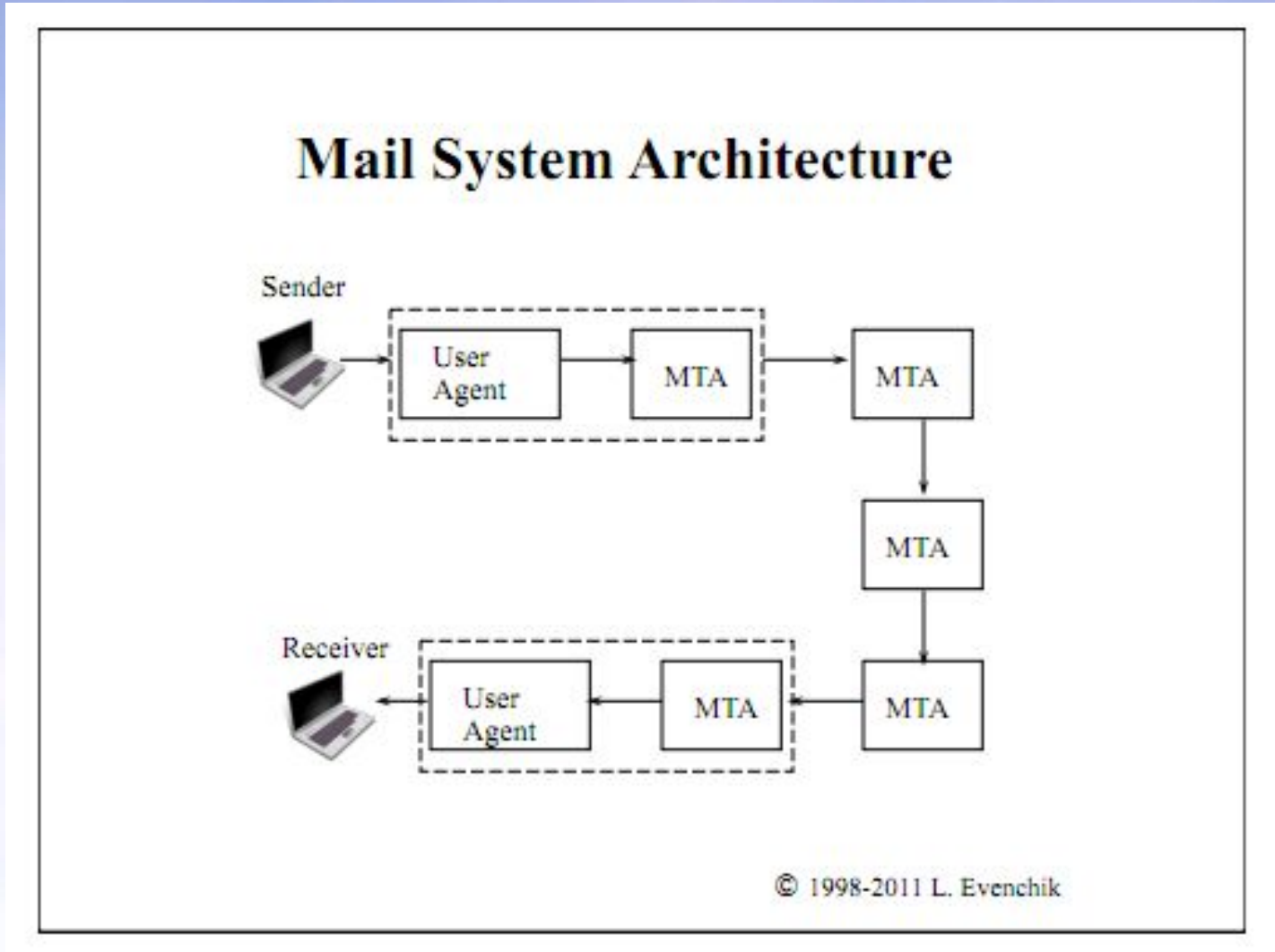
- Network bandwidth
- Disk space
- Processor Cycles
- Lost time (mine & yours)
- 2004 National Technology Readiness Survey estimated almost $22 billion annually loss due to spam. (How much today?)

Layered Approach

- No “silver bullet” to spam prevention
- Use multiple techniques to gradually weed as much out as possible, as early as possible.
- First defenses are key. The more time we spend per message, the worse off we are.



Q: Where Do You Block Spam?



A: Everywhere

1. ***At The Sender*** - If you are an ISP or email provider, but many end-user organizations do this now as well.

2. ***At The Network*** - Block known senders from communicating at all.

3. ***At The MTA*** - Using rules based, signatures, and other systems.

4. ***At The User Agent*** - Last line of defense.

Anti-Spam Techniques

Network-level

- ACLs
- Real-time Block Lists (DNSBLs)
- Greylisting
- DomainKeys

MTA-level

- Rules-based systems
- Bayesian filters

First Phase: Network

Blocking spam prior to

TCP connection establishment

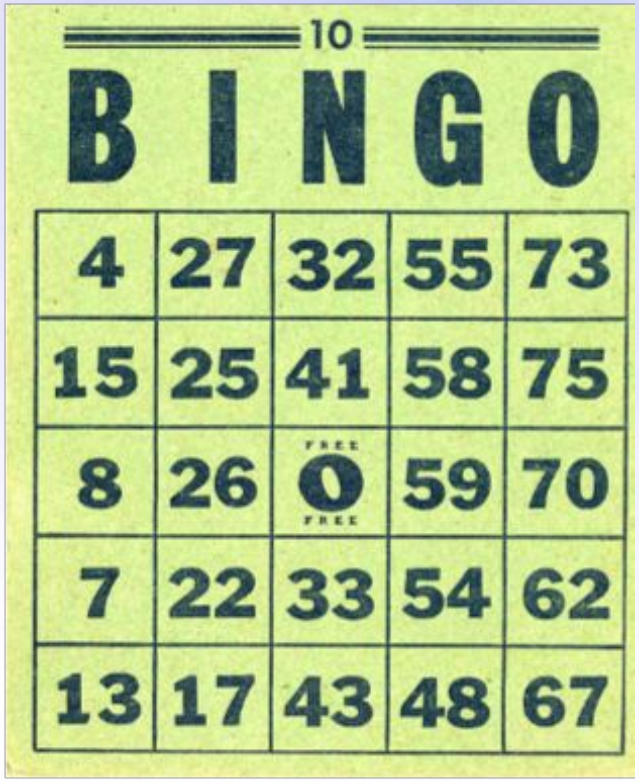
(Think: Firewall)

Traditional Block-lists (ACLs)

A block-list is a IP address or range from which you will never accept mail.

If you receive spam from an address, block it!

In practice, impossible.

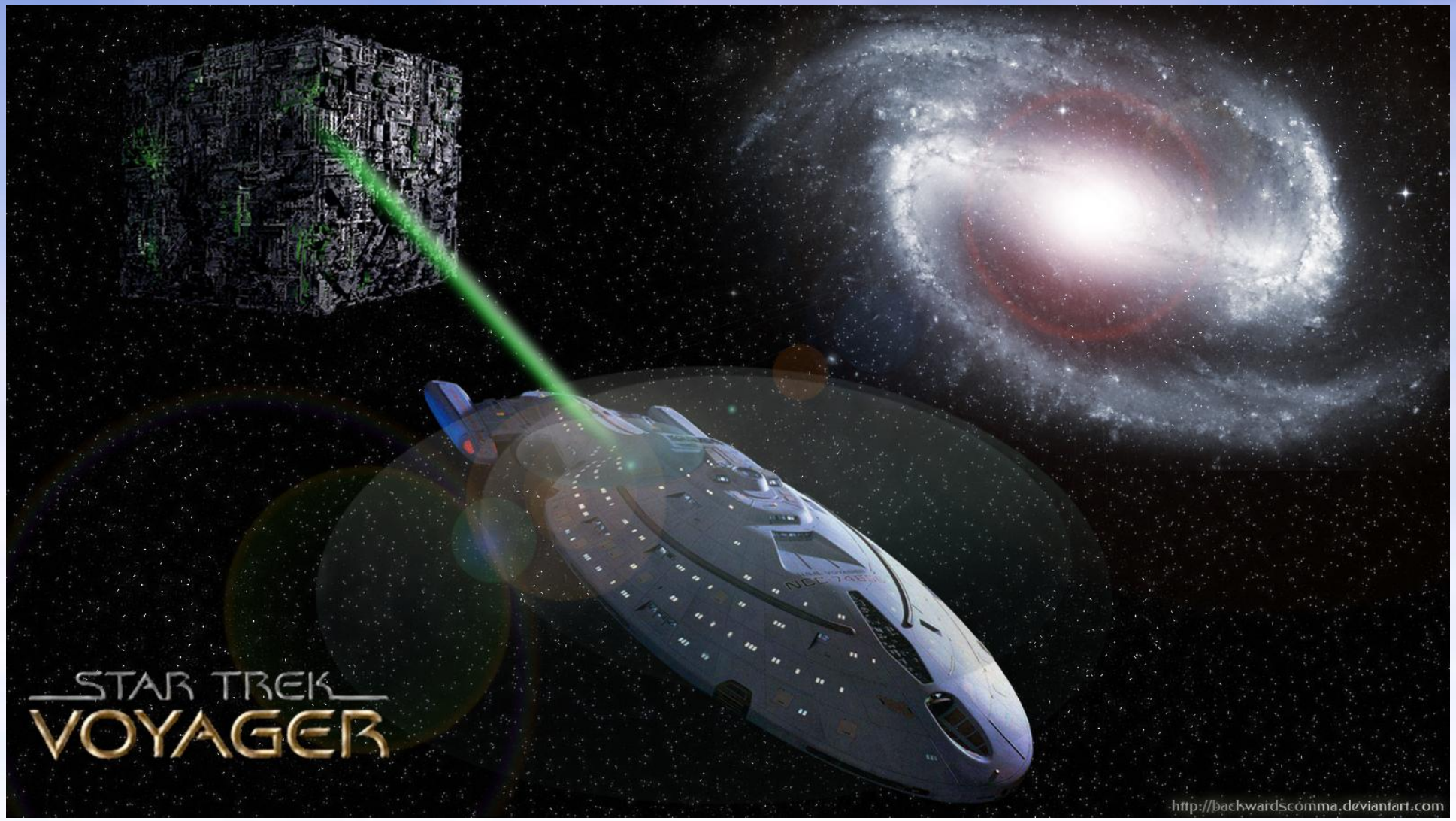


Real-time Block-lists (aka DNSBL)

- “DNS Blacklists and Whitelists” – RFC 5782 (February 2010)
- Protocol invented by Vixie Enterprises in 1997. (Part of “Mail Abuse Prevention System”)
- DNS-based protocol to serve custom IP databases, with associated text fields. Works similar to reverse DNS.
- Today, many anti-spam groups maintain these databases.
- Queried for every incoming connection, so protocol needed to be fast and cache-able.

Cooperative Approach

"Rotate the shield frequencies!"



Many DNSBLs

- There are many DNSBLs with different strategies.
- Most are free for individuals, but may be expensive for large organizations.
- Can be "stacked" to use multiple lists.

Common DNSBLs

- Spamhaus
  - SBL
  - XBL
  - PBL
- SpamCop
  - SCBL
- SORBS

THE SPAMHAUS PROJECT

Spamhaus is one of the most respected of the commercial / free lists. Three DNSBL services:

- SBL – Verified spam sources, determined using a “honey pot” system
- XBL - “Exploit block list”; compromised systems, open proxies, etc.
- PBL - “Policy block list”, DHCP servers, etc.

Honeypots?

A Few Techniques:

- Place a never-used email address on a webpage. Any mail that arrives is spam.
- Bounce message analysis from large ISPs
- Once used addresses, now all messages are spam (!!)

The Spamhaus Project – SBL

http://www.spamhaus.org/sbl/sbl.lasso?query=SBL233

Apple Yahoo! Google Maps YouTube Wikipedia News (3716) Popular

SPAMHAUS

THE SPAMHAUS PROJECT

Spamhaus | SBL | XBL | PBL | ROKSO | DROP

SBL Home | SBL FAQs | SBL Listing Policy | SBL Delisting Procedure |

SBL Advisory

Please read this page carefully. It contains the reason for the listing and who to contact about it.

Ref: SBL233

127.0.0.2/32 is listed on the Spamhaus Block List (SBL)

10-Jul-2008 18:59 GMT | SR01

Spamhaus Block List (SBL) - LOOPBACK ADDRESS

DO NOT CONTACT SPAMHAUS REGARDING SBL233 127.0.0.2/32!

127.0.0.2 is the loopback address of the SBL DNS zone "sbl.spamhaus.org" used for testing SBL configuration on mailservers. It is also listed in most other DNSBL systems as the standard testing address for those zones. This SBL listing, or the listing of 127.0.0.2 in any other DNSBL or other spam-blocking tool, does not affect your e-mail deliverability in any way -- that is simply impossible in the IPv4 internet environment.

127.0.0.2 is designated by RFC3330 as a reserved internal address (that is, a LAN or intranet). It is normally used for loopback testing so that packets are sent and received by the very same node (whatever local machine is sending them receives them, too). That means that node "127.0.0.2" can be used by your network *and* by any other internal networks simultaneously -- it is local to your network *and* to any other network which chooses to use it! It is not routable on the open internet, so your mail is most certainly not being delivered to other networks by it, therefore your mail CAN NOT be rejected due to SBL233.

Please do not contact Spamhaus regarding mail you think is rejected due to SBL233. It is not, we promise! Contact your system administrator, ISP, or mailserver consultant, show them the error messages you have received, and diagnose the problem on your system.

Thank you.

The Spamhaus Project

whois.arin.net

OrgName: Internet Assigned Numbers Authority

...

Risks In Using DNSBLs

- Messages are blocked <u>before</u> content is seen: impossible to verify accuracy.

- Most DNSBLs recommend that you “tag, not block” based on their recommendations, but in practice this is rarely applied.

Second Phase: SMTP

Blocking spam during SMTP connection phase

Greylisting

- Not a “blacklist” or a “whitelist”, but something in between
- Goal is still to block mail before we get past the initial headers

Greylisting Process

- When mail arrives, store source IP, To, and From addresses in a database. Return a SMTP status code 450 indicating “try again later”.
- Legitimate mailers (with queues) will try again later. If the mail then matches what is in the database, allow it through.
- Spammers almost never retry, so the mail will be dropped.
- Downside: much slower mail delivery.
- Can run only on a subset of addresses.

Sender Policy Framework (SPF)

- Extension to SMTP - RFC 4408, updated by RFC 6652
- Using DNS (TXT or SPF records), a mailer identifies which servers in his domain send email.
- This list is consulted by the receiver prior to accepting the message.
- This is useful for dealing with forged “ReturnPath”or other email headers or compromised machines.

Example SPF Record (DNS)

jpranevich — bash — 80×24

```
temp:~ jpranevich$ dig @8.8.8.8 -t TXT _spf.google.com

; <<>> DiG 9.7.3-P3 <<>> @8.8.8.8 -t TXT _spf.google.com
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 39035
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;_spf.google.com.               IN      TXT

;; ANSWER SECTION:
_spf.google.com.        300     IN      TXT     "v=spf1 ip4:216.239.32.0/19 ip4:
64.233.160.0/19 ip4:66.249.80.0/20 ip4:72.14.192.0/18 ip4:209.85.128.0/17 ip4:66
.102.0.0/20 ip4:74.125.0.0/16 ip4:64.18.0.0/20 ip4:207.126.144.0/20 ip4:173.194.
0.0/16 ?all"

;; Query time: 52 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Sun Dec  4 21:28:26 2011
;; MSG SIZE  rcvd: 248

temp:~ jpranevich$
```

SPF Result Header Example

Received-SPF:

pass (google.com: domain of foo@gmail.com designates 10.68.35.225 as permitted sender)

client-ip=10.68.35.225;

Authentication-Results: mr.google.com;

spf=pass (google.com: domain of foo@gmail.com designates 10.68.35.225 as permitted sender) smtp.mail=foo@gmail.com;

dkim=pass header.i=foo@gmail.com

Third Phase: Message Analysis

Blocking spam based on content before passing to User Agent

(These are expensive options!)

DKIM (Formerly DomainKeys)

IETF Draft Standard – RFC 6376

- Adopted by Yahoo!, Google, and others.
- Emails are signed with the private key of the sending domain, recipients know that emails came from where they say

Example DKIM Query

jpranevich — bash — 80×24

```
temp:~ jpranevich$ dig @8.8.8.8 -t TXT gamma._domainkey.gmail.com

; <<>> DiG 9.7.3-P3 <<>> @8.8.8.8 -t TXT gamma._domainkey.gmail.com
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 36596
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;gamma._domainkey.gmail.com.    IN      TXT

;; ANSWER SECTION:
gamma._domainkey.gmail.com. 300 IN      TXT     "k=rsa\; t=y\; p=MIGfMA0GCSqGSIb
3DQEBAQUAA4GNADCBiQKBgQDIhyR3oItOy22ZOaBrIVe9m/iME3RqOJeasANSpg2YTHTYV+Xtp4xwf5g
TjCmHQEMOs0qYu0FYiNQPQogJ2t0Mfx9zNu06rfRBDjiIU9tpx2T+NGlWZ8qhbiLo5By8apJavLyqTLa
vyPSrvsx0B3YzC63T4Age2CDqZYA+OwSMWQIDAQAB"

;; Query time: 51 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Sun Dec  4 21:39:09 2011
;; MSG SIZE  rcvd: 287

temp:~ jpranevich$
```

Rule-Based Systems

- SpamAssassin is a rule-based tool which identifies spam based on content and headers
- SpamAssassin is Perl-based and open source
- Other open source and commercial alternatives are out there.

SpamAssassin

SpamAssassin works through “rules” and “scores”:

- Usually, triggering a single rule is not enough to mark a message as spam.
- Each rule gets a score. When those scores add up to an admin-configured level, the message is marked as spam. (Deleted or moved to “Junk” folder.)

SpamAssassin Rules

Some example rules:

HTML contains far too many close tags – Score 1.041
HTML font size is large – Score 0.147
HTML font size is huge – Score .804
HTML font color similar to background - Score 0.131
HTML font face is not a word – Score 0.92
HTML includes a form which sends mail – Score 1
And many others…

Scores adjusted using AI techniques.

SpamAssassin headers usually stripped before sent to end user.

URL-in-Email Blocking

- DNSBLs (such as SpamHaus PBL & DBL) are used for URLs embedded in email messages.
- If an email contains a URL, the IP or name of the web server is checked against a list and blocked, if appropriate.
- Some ISPs are even more aggressive and block based on number of links or other factors.

Fourth Phase: Annoyed Users

Spam got through, but at least we can learn something!

Bayesian Filtering

- “Smart” filters that learn which emails a user wants based on user action, such as moving a mail to or from the “Junk” folder

- Filtering works because, historically spam mails haven’t looked like “regular” mails. The filter adapts to each user over time.

Bayesian Downsides

One user's filter may not work well for another user.

Filters may take a long time to "train" properly and so users will have a higher false-positive rate initially, discouraging them from using the filter.

"Bayesian Poisoning" is a common approach for spammers today. This is when a spammer sends mail with lots of legitimate text (such as a snippet from a book, newspaper, or website), in addition to the spam.

When the user marks that message as "spam", the Bayesian engine will interpret patterns in that legitimate text as spam, making it more likely that legitimate mails will be caught by the filter and vice-versa.

Feedback Loops - (M)ARF

- Some ISPs (AOL, Yahoo, etc.) employ feedback loops to notify other ISPs that spam mail has been sent from them.
- ARF – “Abuse Reporting Format” being popularized by AOL.
- Only works after the mail has been sent out and read (possibly by hundreds or more.)

WWGD?
(What Would Google Do?)

- Large ISPs are now resorting to "crowd sourcing" email spam filtering.
- If an email is marked by spam by many receivers, it can be removed from the inboxes of others before they see it.
- Google, Microsoft, and others have "secret sauce" around their spam prevention efforts. This is frequently based on neural networks.

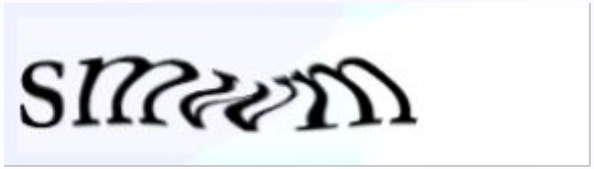
Challenges for ISPs

- Because anti-spam efforts are so aggressive, legitimate mailers often end up on the wrong side of a block.
- Sometimes, they are innocent.
- Often, spammers are using these semi-open systems to send their bulk mail, resulting in havoc for legitimate users of the system.

Some Solutions

- CAPTCHA
- Feedback Loops
- Outbound filtering

CAPTCHA

- "Completely Automated Public Turing test to tell Computers and Humans Apart".
- Should be easy for humans, but very hard for computers.
- Developed by Carnegie Mellon
- Example CAPTCHA: 
- (Source: Wikipedia)

Outbound Filtering

- Many ISPs now apply message-level techniques on emails sent from their own users
- Because of privacy policies, etc. not all organizations can do this.
- SORBS and other organizations now recommend this for all ISPs

CAN-SPAM

- “Controlling the Assault of Non-Solicited Pornography And Marketing Act of 2003”
- Requires bulk email to be labeled and to offer an “opt out”, plus other features.
- Ignored by foreign mailers. Law has many loopholes.
- A few convictions.

CAN-SPAM for ISPs

One very good provision:

(b) ISP HELD HARMLESS FOR GOOD FAITH PRIVATE ENFORCEMENT- An ISP is not liable, under any Federal or State civil or criminal law, for any action it takes in good faith to block the transmission or receipt of unsolicited commercial e-mail.

A word about opt outs...

CAN-SPAM requires opt outs, however this is *exactly what the spammers want.* By “opting out”, you are verifying to them that:

- You have a valid email address
- Your ISP lets their mail through
- You read their bulk mail

And you expect them to stop mailing you?

Future of Email

- Popularity of social networks over traditional mailing platforms.
- Other closed systems (IM?) which have been built with abuse-prevention in mind.
- Stronger measures?

Arms Race

- Who wins: The spammers or the filtering technologies?

Any Questions?

Thank You!