Global Information Assurance Certification Paper

"SPAM": RECOURSE AND EDUCATION

By

Rodney Caudle

Submitted in partial fulfillment of the requirements for the

GIAC Security Essentials Certification (GSEC) v1.4b

SANS (http://www.sans.org)

November 10th, 2003

TABLE OF CONTENTS

ACKNOWLEDGMENTS

The author wishes to thank his colleague Arthur Stephens for providing proof reading and valuable feedback on this paper.

GLOSSARY

IMAP. Internet Message Access Protocol; a standard protocol for accessing e-mail from your local server. IMAP is a client/server protocol in which e-mail is received and held for you by your Internet server.

POP3. Post Office Protocol 3; a less sophisticated protocol where your e-mail is saved for you in your mail box on the server. When you read your mail, all of it is immediately downloaded to your computer and no longer maintained on the server.

SMTP. Simple Mail Transport Protocol; a TCP/IP protocol used in sending and receiving e-mail. However, since it's limited in its ability to queue messages at the receiving end, it's usually used with one of two other protocols, POP3 or IMAP.

Spam. Unsolicited bulk e-mail (or junk e-mail'), which can be either commercial (such as an advertisement) or noncommercial (such as a joke or chain letter). (Supreme Court of the State of Washington, USA)

ABSTRACT

E-mail is the golden application of the Internet. Hundreds of millions of people use e-mail everyday to facilitate business and personal communications. Encroaching on this widespread use of e-mail is unsolicited bulk e-mail (UBE), better known as "spam". This paper explores the widespread impact of spam by taking a three step approach at looking into this disturbing issue. First, some background on the issue in question is given in the first chapter describing the definition of UBE, or spam, with some numbers illustrating the severity of the problem quantitatively. Second, the technical aspects of e-mail are explained focusing on the disparity between protocols that provides for misleading or inappropriately addressed e-mail to be sent and received. Third, the risks associated with e-mail are explored focusing on sending and receiving UBE. Finally, technical and legislative recourse is explored to provide mitigation and appropriate defensive measures against the growing threat of spam.

Chapter 1

INTRODUCTION

E-mail is the golden application of the Internet. Every day hundreds of millions of e-mail messages are sent across the Internet. Some of the e-mails are legitimate messages to loved ones, business associates or customers. However, the rest of the e-mails consist of advertisements for lunchmeat, steaks, cameras, spice racks or services. Sometimes the ratio of unwanted e-mail (or spam) to valid e-mail can be as high as six out of every ten. Spam is a growing point of concern and contention effecting businesses and individuals on the Internet. Large quantities of spam, numbering in the tens or hundreds of millions, are sent every day. Every person who has an e-mail account has received spam; most likely several each day. However, even though spam is a common phenomenon, there is still a lot of confusion surrounding spam. Every year there are new bills introduced into Congress and the Senate which attempt to address the growing concern of an inundated public. Yet how can one adequately address an issue as elusive as spam? What exactly is spam and what distinguishes a legitimate email from spam? What are the inherent risks that spam pose to businesses (and individuals) and their relationships? Why is spam such a problem and can we stem the tide before it's too late? Effectively, every business in operation today conducts business via e-mail. This paper explores the phenomenon of spam, the underlying technologies, and the risks which spam poses.

Spam is defined in varying ways depending on the perspective from which you view email. Even among advocates of the anti-spam community the definition is vague and difficult to define. As monkeys.com states, "The anti-spam community on the Internet has long grappled with a problem of terminology. Just as one U.S. Supreme Court Justice once said about pornography 'I can't define it, but I know it when I see it', we in the anti-spam community have generally preferred to leave our definition(s) of the term 'spam' somewhat loose and ambiguous,...". However, the following is a collection of commonly understood definitions for the term "spam".

> *SpamHaus and mail-abuse.org: 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.*
>
> *Dictionary.com: Unsolicited e-mail, often of a commercial nature, sent indiscriminately to multiple mailing lists, individuals, or newsgroups; junk e-mail.*
>
> *Russell Nelson: Email spam is Unsolicited Bulk Email (UBE).*
>
> Unsolicited *means that you lack affirmative consent from the recipient. If you found an address on a web page, on a mailing list, or on Usenet, you don't have consent. If you got an address in gift, sale or trade, you don't have consent. If someone gave you an address for a particular purpose (for example, a commercial transaction, information about your products, or after-sales support) you only have consent to use it for that particular purpose. Use for any other purpose requires a new consent.*
>
> Bulk *means that you sent a substantively similar message to more than 200 addresses a day. A message that differs from recipient to recipient only by details (e.g., the recipient's name, account number, blocks of random words, characters, numbers, or non-rendered text) is the same message. A message that uses different wording to express the same idea is the same message. If you sent the same message to 200 different people day after day, it's spam.*
>
> *Monkeys.com: Internet spam is one or more unsolicited messages, sent or posted as part of a larger collection of messages, all having substantially identical content.*

Probably the most well known definition of spam, and the one used by this paper, was provided by the Supreme Court of the State of Washington in the famous "State v. Heckel (Jason), d/b/a Natural Instincts" case decided in June of 2001. The court defined spam as "... unsolicited bulk e-mail (or 'junk e-mail'), which can be either commercial (such as an advertisement) or noncommercial (such as a joke or chain letter)."

This definition of spam seems hauntingly similar to physical mail which is delivered to houses every day. This mail comes addressed to "occupant" or "resident" and bears advertisements for companies within a certain radius of the recipient. However, there are important differences between this 'junk mail' and the spam which comes to e-mail accounts. First, the burden of cost is on the sender not the receiver. In the case of physical mail, the person (or company) sending the advertisements pay for the privilege of sending the letter while the recipient does not pay for receiving the letter. This is not the case for spam as both the sender and the recipient must pay for bandwidth charges. However, the

receiver must pay additional charges for anti-virus scanning (at the least) to ensure that the e-mail is safe to receive. In the case of companies this cost can be measured by e-email, by MB, or by a flat monthly fee. Secondly, for the most part the percentage of junk mail to good mail is heavily weighted toward legitimate mail. However, for e-mail the odds tip to the other side with spam accounting for sometimes as much as sixty percent of e-mail. Table 1 contains figures from the last six months from my employer which show an increasing trend for spam. Every month spam has increased somewhere between six to twenty four percent. The monthly percentage of spam has increased by seventeen point sixty-five percent. At this rate of increase (three percent a month) within a year over ninety percent of all e-mail received would be spam.

Month	Total Email	Good Email	Spam	% By Volume	Total Email (MB)	Good Email (MB)	Spam Size (MB)	% By Size
December-02	660,703	352,113	308,590	46.71%	19,005.44	17,315.84	1,689.60	8.89%
January-03	742,511	383,479	359,032	48.35%	20,480.00	18,647.04	1,832.96	8.95%
February-03	742,437	360,872	381,565	51.39%	20,500.48	18,657.28	1,843.20	8.99%
March-03	896,659	411,210	485,449	54.14%	22,312.96	20,039.68	2,273.28	10.19%
April-03	994,977	420,222	574,755	57.77%	25,569.28	23,162.88	2,406.40	9.41%
May-03	1,138,460	405,744	732,716	64.36%	25,231.36	22,128.64	3,102.72	12.30%
June-03	1,096,718	396,786	699,932	63.82%	24,872.61	21,779.42	3,093.19	12.44%
	6,272,465	**2,730,426**	**3,542,039**	**56.47%**	**157,972.13**	**141,730.78**	**16,241.35**	**10.28%**

Timeframe	Delta Total Email	Delta Good	Delta Spam	Delta Email (MB)	Delta Good (MB)	Delta Spam (MB)
Dec-Jan	11.66%	8.53%	15.11%	1,475	1,331	143
Jan-Feb	-0.01%	-6.07%	6.09%	20	10	10
Feb-Mar	18.82%	13.04%	23.96%	1,812	1,382	430
Mar-Apr	10.40%	2.17%	16.85%	3,256	3,123	133
Apr-May	13.45%	-3.51%	24.16%	-338	-1,034	696
May-Jun	-3.73%	-2.23%	-4.58%	-359	-349	-10

Table 1: Six Month Spam Trend

Gray Cary Ware & Freidenrich, LLP

This disturbing trend will result in numerous issues for companies and individuals who try to communicate through e-mail. Public e-mail accounts such as Yahoo!, Hotmail, or MSN are inundated with spam. It is not uncommon to find an individual who has an account which they do not use because spam was overwhelming the usefulness of the account. What would happen to the Internet if the "killer application" became unusable because of an overflow of spam?

Chapter 2

UNDERSTANDING THE TECHNOLOGY

To understand why spam is effective, a comprehensive review of the protocols exploited by spam must be presented. Just like in a personal computer, where multiple components work together as a system, e-mail consists of multiple protocols, each protocol performing a single function but together defining a system. The sending protocol, Simple Mail Transport Protocol (SMTP), is used to pass an e-mail message between servers from the sender's domain to the recipient's domain. Defined by Request For Comments (RFC) 821, SMTP is the basis for delivering e-mail between domains. The actual message is formatted according to RFC 822 which refers to e-mail as "Internet Text Message". Once the e-mail is resting in the recipient's e-mail server, Internet Message Access Protocol (IMAP) or Post Office Protocol 3 (POP3) is used by the e-mail client to pickup and read the message. Alternative e-mail clients common to the UNIX environment, such as Mutt or Pine, will read directly from the file system instead of over the network, but require the program to reside on or have access to the local file systems.

When considering e-mail over the Internet, SMTP is a "trusting" protocol. This means there is effectively no authentication in the use of SMTP when receiving e-mail from the Internet. While authentication is technically built into SMTP, this would cause issues with communications to a domain from anonymous Internet sources. This would allow for communication with existing business contacts but would not allow initiations of new communications. For the purposes of this paper we will consider only the case of anonymous sources of e-mail and will disregard the instances where SMTP authentication is required.

The first step in sending an e-mail message is to determine the recipient's e-mail gateway, called a Mail Transfer Agent (MTA). This is accomplished by querying the Domain Name Service (DNS) Server for the domain in question. The DNS server will return a list of Mail Exchange (MX) records with a

weight, i.e. order of priority, associated to each. The lowest weighted MX record is the primary MTA to use.

```
$ dig MX sample.com

;; QUESTION SECTION:
;sample.com.                    IN      MX

;; ANSWER SECTION:
sample.com.    86400   IN    MX      10 smtp.sample.com.
sample.com.    86400   IN    MX      20 smtp2.sample.com.
```

Figure 1: DNS Query

Once the recipient's MTA is determined a connection must be established to the server. Since SMTP is a clear text protocol, a simple telnet session to port twenty five is sufficient to connect to the MTA. The server should return a code of 220 which means the server is ready to receive the incoming e-mail. The server may also return a code of 250 OK in addition to the code 220. The exact syntax of the text after the return code is dependent on the vendor of the MTA but a return code of 250 means OK or command accepted. The examples here were performed using a Qmail MTA.

```
$ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
250 Ok
```

Figure 2: Telnet Connection to Server

To send the e-mail message one needs to use the correct syntax as defined in RFC 821 (link), Simple Mail Transfer Protocol (1982). The first line to send is a HELO command to identify the sending MTA to the receiving MTA. However, no actual verification of this command is used and almost anything is considered valid input. The MTA will acknowledge with a success code of 250 or a failure code of another value.

```
$ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
HELO anywhere.com
250 Hello... this is the Sample.Com SMTP Server
```

Figure 3: HELO Command

Next the MTA expects an identification of who is sending the e-mail message. This is completed with the MAIL FROM: command. Again there is no verification of this field so anything can be typed here. The MTA will acknowledge with a success code of 250 or another value for a failure.

```
$ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
HELO anywhere.com
250 Hello... this is the Sample.Com SMTP Server
MAIL FROM: <JohnDoe@anywhere.com>
250 Ok
```

Figure 4: MAIL FROM Command

The MTA now expects a list of recipients for this e-mail message. The RFC defines the RCPT TO command for this requirement. For multiple recipients, one should send one RCPT TO command per recipient. The SMTP server should return a success code of 250 for each recipient if they exist or 550 or another error code if the request is invalid. A return code of 550 is indicative of a "no one here by that name" response from the MTA. If the MTA is configured to forward for this domain or is not the end point of delivery for this e-mail, these commands will probably succeed even if the user doesn't exist. However, eventually the e-mail message will reach the final destination MTA which will return a code of 550 for an unknown address.

```
 $ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
HELO anywhere.com
250 Hello... this is the Sample.Com SMTP Server
MAIL FROM: <JohnDoe@anywhere.com>
250 Ok
RCPT TO: <user1@sample.com>
250 Ok
RCPT TO: <noone@sample.com>
550 No such user here
```

Figure 5: RCPT TO Command

Once the recipient list is completed the MTA expects the DATA command. This command will return a code of 354 instructing the client to start sending the e-mail message body. This message body has a special format and will be discussed shortly. The DATA field ends when the client sends a sequence of `<CRLF>.<CRLF>` to the MTA. Once the termination code is received the MTA will

return a code of 250 Ok. Now that the transaction is complete the next command to be sent to the MTA is the QUIT command. This disconnects the session and instructs the MTA that you are finished. The MTA will acknowledge with a return code of 221.

```
 $ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
HELO anywhere.com
250 Hello... this is the Sample.Com SMTP Server
MAIL FROM: <JohnDoe@anywhere.com>
250 Ok
RCPT TO: <user1@sample.com>
250 Ok
RCPT TO: <noone@sample.com>
550 No such user here
DATA
354 Send data. End with CRLF.CRLF
From: <sender list>
To: <recipient list>
Subject: <enter subject of e-mail message>
Text of the email message...
This is another line of text.
.
250 Ok
QUIT
221 Closing connection
```

Figure 6: DATA Command

With the email transaction completed a thorough examination of the content sent during the DATA portion of the e-mail will yield additional information (Figure 6). This is important in defining the characteristics of an e-mail as there are essentially two parts to an e-mail message. Similar to a physical letter mailed to a relative or business associate the e-mail has an envelope and a message. Up to this point, all of the information entered, the envelope of the e-mail, will last only until the transaction with the MTA is disconnected and, just like a person receiving the letter, the envelope is discarded once opened. The message body will continue unchanged throughout the path of the e-mail from MTA to MTA. The format of the message body is defined by RFC 822 (link), The Standard for the Format of ARPA Internet Text Messages (1982). This standard defines a number of fields that can be used in the header of an e-mail message to identify important items such as "To", "From" and "Subject" as in our example. These fields are read by the e-mail client, such as Microsoft Outlook, and displayed for the user to view. The following is a list of fields important in defining an e-mail message:

Name	Description

1.	Date	Date and time when the message was sent
2.	From	Originator of this e-mail
3.	To	List of recipients of the e-mail
4.	Subject	Subject of the e-mail message.
5.	Reply-To	Address to direct replies to this e-mail message to.
6.	Message-ID	Unique identifier of the e-mail message as defined by the first SMTP server encountered.
7.	Bcc	List of recipients not displayed by the client application… Blind Carbon Copy
8.	Cc	Additional list of recipients of the e-mail
9	Received	RFC 822 required field; a copy is filled in by each MTA along the path

Table 2: List of important fields for RFC 822

The most important aspect of this RFC-822 encoded message body embedded inside the RFC 821 SMTP envelope is actually the lack of continuity inherent between the two portions of the e-mail. The fields defined by the message (RFC 822) have no correlation or effect on the fields used by the envelope (RFC 821) for delivery of the e-mail. While this may not seem significant to the casual observer, the impact of this is felt by every recipient of e-mail today. This lack of continuity can essentially render an appropriately formatted e-mail message virtually untraceable as well as allowing for falsification of identity of origination point. Consider the following example:

```
$ telnet smtp.sample.com 25

Trying 10.0.0.25...
Connected to smtp.sample.com.
Escape character is '^]'.
220 Hello... this is the Sample.Com SMTP Server ESMTP
HELO anywhere.com
250 Hello... this is the Sample.Com SMTP Server
MAIL FROM: <JohnDoe@anywhere.com>
250 Ok
RCPT TO: <user1@sample.com>
250 Ok
DATA
354 Send data. End with CRLF.CRLF
From: “Meg Ryan” <meg.ryan@anywhere.com>
To: “One Hot List” <>
Reply-To: “One Hot List” <>
Subject: You’ve got to see this
Message-ID: <4321.765.098-msg@anywhere.com>
Mime-Version: 1.0
Content-Type: text/plain; charset=us-ascii;
    format=flowed Content-Transfer-Encoding: 7bit

Hello Baby,

Do you want to see my pictures?

Go to:  http://www.sexypics.com/
.
```

```
250 Ok
QUIT
221 Closing connection
```

Figure 7: Sample Spam Message

Obviously, Meg Ryan did not send this email and the message definitely wasn't sent to anyone specific. However, the e-mail message was delivered to user1@sample.com. In addition, no record of the originator's e-mail address (<JohnDoe@anywhere.com>) or the recipient (<user1@sample.com>) can be found in the received e-mail message. This is because the MTA looks only at the RFC 821 envelope fields to deliver the e-mail message. Once the message has traversed the final MTA, these fields are discarded as unnecessary and the RFC 822 encoded message is stored (or converted) by the e-mail server in preparation for the user's client to pickup the message. Table 3 contains a complete breakdown of the fields for this sample e-mail message.

	RFC 821: Simple Mail Transport Protocol (1982)	
	Name	**Value**
1	HELO	anywhere.com
2	MAIL FROM	<JohnDoe@anywhere.com>
3	RCPT TO	<user1@anywhere.com>
4	DATA	See Below ↓
	RFC 822: Standard for the Format of ARPA Internet Text Messages (1982)	
4.1	Date	Not Used
4.2	From	"Meg Ryan" <meg.ryan@anywhere.com>
4.3	To	"One Hot List" <>
4.4	Subject	You've got to see this
4.5	Reply-To	"One Hot List" <>
4.6	Message-ID	<4321.765.098-msg@anywhere.com>
4.7	Bcc	Not Used
4.8	Cc	Not Used
4.9	Received	Added by each MTA that participates in the delivery of the message

Table 3: List of fields for example e-mail message

Providing information necessary to trace the path of an e-mail is critical to diagnosing the origin of an e-mail message. The Received field in RFC 822 is a required field that is to be filled in by each MTA

prior to delivering the message. The information included in the Received field is a listing of all MTA's that the message traversed to reach its destination. This can provide some of the critical information which is otherwise discarded when the message is delivered. Figure 8 is a listing from a sample e-mail message header.

```
Received: from DOMAIN (SOURCEIP) by RECIPIENTSERVER with SMTP
(Microsoft Exchange Internet Mail Service Version 5.5.2650.21)  id
L6BT75LL for RECIPIENT; DATE
```

Figure 8: SMTP relay stamp

For each MTA the message passes through the Received field which will contain at least the sending MTA and the receiving MTA. Occasionally additional information is included in this field as a way of preserving the information from the SMTP envelope as the message travels. Table 4 gives a description of the information included in the sample received field in Figure 8. The information available in the received field is dependant on the vendor's implementation of RFC 822 and may include different information.

	Name	Value
1	DOMAIN	Domain stated in the HELO command of the SMTP connection
2	SOURCEIP	Originator of the SMTP connection
3	RECIPIENT	Some SMTP servers, such as POSTFIX, add the RCPT TO field to this stamp as well
4	DATE	Timestamp (according to the SMTP server) of the SMTP connection
5	RECIPIENTSERVER	SMTP Server receiving the e-mail

Table 4: SMTP relay stamp fields

Figure 9 shows the header of an actual e-mail received from another domain. The path of the e-mail as it traverses MTAs is clearly defined by the Received fields: [R1] → [R2] → [R3] → [R4] → [R5] → [R6] → [R7]. The header information has been cleansed of identifying names and IP addresses for the purpose of this example.

```
[R7]Received: from SERVER6 (IP6) by SERVER7 with SMTP (Microsoft
Exchange Internet Mail Service Version 5.5.2650.21) id L6BT0LDT;
Wed, 2 Jul 2003 09:34:28 -0700
[R6]Received: (qmail 29238 invoked by uid 900); 2 Jul 2003 16:34:22
-0000
[R5]Received: from unknown (HELO SERVER5) (IP5) by 900 with SMTP; 2
Jul 2003 16:34:22 -0000
[R4]Received: from SERVER4 (SERVER4 [IP4]) by SERVER5(Postfix) with
ESMTP id 964B015D96 for < JohnDoe@sample.com >; Wed,  2 Jul 2003
16:34:21 +0000 (UCT)
[R3]Received: by SERVER4 (MessageSwitch) id 1057163661131205_31410;
Wed,  2 Jul 2003 16:34:21 +0000 (UCT)
[R2]Received: from SERVER2 (unknown [IP2]) by SERVER3 (Postfix) with
ESMTP id 081F015D92 for < JohnDoe@sample.com >; Wed,  2 Jul 2003
16:34:21 +0000 (UCT)
[R1] Received: from SERVER1 (unknown [IP1]) by SERVER2 (Postfix)
with ESMTP id B7FC21CF15AD8 for <JohnDoe@sample.com>; Wed,  2 Jul
2003 12:37:47 +0000 (/etc/localtime)
Date: Wed, 2 Jul 2003 12:37:47 UT
To: JohnDoe@sample.com
From: eNews and Views <eNewsandViews@eletters.eweek.com>
Reply-To: eNewsandViews@eletters.eweek.com
Subject: Mozilla 1.4's Key Improvements Out of Sight
Message-Id: <20030702123747.B7FC21CF15AD8@SERVER2>
```

Figure 9: sample e-mail header

This is an excellent example of the varying information available in the received field for different vendors. However, this information is "trusted" as accurate and assumes that the first MTA has been given accurate information and not adjusted the e-mail in any way. This is a naïve assumption because falsifying the information in the Received field is a trivial attack and can lead to a misleading or obstructed delivery path. Assume for instance that one has control of a server and has written a program to send e-mail by opening a telnet session to localhost [127.0.0.1] and typing information similar to the examples above. The insertion of a couple of Received lines into the header would provide enough misleading information to bring into question the exact path or source of the e-mail message.

In summary, the SMTP protocol, as defined by RFC 821 provides the basis for delivery of RFC 822 encoded message bodies between domains. However, the information in the RFC 821 envelope used by SMTP is not validated against the RFC 822 encoded message body. The Received field from RFC 822 makes an attempt at retaining some of the RFC 821 envelope and provides for a way to trace the origin of an e-mail. The ease of filling out an e-mail provides a communication protocol that is easy to use and easy to exploit.

Chapter 3

RISK ANALYSIS: "SPAM"

Risk is defined as "the danger or probability of loss" (dictionary.com) and can be quantitatively defined as the threat of exploit multiplied by the ease of exploitation multiplied by the value of the resource. As we've demonstrated in chapter 2, creating a falsified e-mail message is trivial and can easily be automated using programming languages such as PHP, Perl, BASH, Visual Basic, and more. This is important to understand because the increasing use of the Internet for providing critical business transactions is also increasing the importance of e-mail and associated electronic communications. However, these business communications are built on a basis of trust established by sending and receiving e-mail (or other communications) through the Internet. The impact of e-mail on today's business environment is monumental. The increasing speed of communicating across long distances or countries through e-mail has empowered even the smallest business to reach a worldwide audience easily. Because of the importance of e-mail to businesses today, the risks from spam are important to understand and address appropriately to ensure the integrity of the established trust between parties.

The risks from spam fall into two categories according to whether one is the recipient or (supposed) sending party of the spam e-mail. The perception of the recipient is important to take into account when understanding the impact of spam. Spam can appear to come from noncommercial entities such as charities or a close friend. In addition spam may appear to, or actually come from, a business acquaintance or business entity. As a recipient of spam, the risks and associated impact are easier to address. Table 5 contains a complete list of the risks and impact as a recipient of spam.

	Risk	**Description**	**Mitigation**
1	Virus (attachment)	Receiving a contaminated e-mail and opening or causing the payload to execute on an unprotected system causing further outbreak. This is the same risk associated with e-mail in general.	• Desktop Antivirus Software with appropriate policies for updates • Inbound e-mail antivirus scanning software with appropriate policies for updates
2	Virus (html link)	Receiving a HTML e-mail with a virus embedded within the e-mail causing the payload to execute when the e-mail is viewed.	• Desktop antivirus software with appropriate policies for updates

			• Antivirus and/or URL filtering at the firewall/gateway level for HTTP requests
3	Offensive Content	Receiving an e-mail message which contains vulgar, lewd or another form of offensive content	• Heuristic, keyword or some other filtering technology at the inbound e-mail relay
4	Large Quantities of Unsolicited Bulk E-mail	Receiving large quantities of UBE sometimes reaching 65% or greater of total e-mail received resulting in a difficult and sometimes impossible to utilize e-mail address	• Heuristic, keyword or some other filtering technology at the inbound e-mail relay

Table 5: Risks as recipient of spam

There may be additional risks associated with spam, but these are the largest issues facing businesses and individuals today. For the recipient of spam the impact of receiving large quantities daily can be daunting and sometime insurmountable. Depending on the business role of the recipient, the spam-to-e-mail ration can be as high as sixty five or seventy out of every one hundred e-mails. Sales persons and individuals who require a lot of social interaction for success are most often the recipients with the highest spam ratios.

The second category of risks is for the sender of spam. These risks are not as clearly defined and depend on the content of the messages being distributed, whether the sender was a willing participant and the manner in which the spam is distributed. Spammers will often hijack a computer and use it to send hundreds of thousands of spam messages daily. High volume e-mail relays specially configured to process e-mail can send upwards of a million messages an hour. A few years ago these e-mail relays were combined with an "Anonymizer" service which effectively cleansed the transactions passing through the service of any identifying information which might lead to the source of the transmission. With the increase in accountability for the originating point of transmissions these services have become unprofitable and for the most part have disappeared. Table 6 outlines some of the possible risks associated with sending spam. Depending on a company's business model these risks may or may not be applicable. Careful evaluation is necessary to determine applicability.

	Risk	**Description**	**Mitigation**
1	Business Reputation	A company sending large quantities of spam is often considered disreputable. If the company is unaware of its actions then distrust can often be added to the company's reputation. The damaged	• Screening of outbound SMTP connections • Restriction of outbound SMTP to authorized servers only

		reputation, if widely publicized, can have a significant impact on a company's bottom line.	
2	Legal Ramifications	Depending on the content of the spam this class of risk can be extremely costly for a company. Pornographic spam and fraudulent offers can result in a suspended business license.	• Screening of outbound SMTP connections and content
3	Blacklisting	This is a popular tactic for fighting spam. An IP address from which suspected spam originates is added to a blackhole list. This list is distributed to subscribers who disallow connections from any IP on the list.	• Monitoring common blackhole lists and prompting for removal from any lists • Customer training on alternative technologies instead of blackhole lists

Table 6: Risks as sender of spam

The largest risk a company runs, as an origination point for spam, is temporary or permanent damage to their business reputation and relationships. Once the business relationships are impacted, rebuilding the trust essential to facilitate effective interaction is difficult at best and sometimes not possible. A review of these risks and ways to mitigate the impact to business are presented in the final chapter.

Chapter 4

RECOURSE AND MITIGATION

There are many approaches to combating spam. These include technological solutions like keyword and pattern matching, heuristic filters and blackhole lists. In addition, there are and have been several proposed laws designed to enforce standards in spam and to punish non-adherence. The Direct Marketing Association (DMA) is lobbying against these proposed laws because of the restrictions imposed on advertising over e-mail. Because of the popularity and cost effectiveness of e-mail marketing, the proposed laws will most likely be ineffective in the fight against spam. Many people believe that technology is not effective in addressing the growing effects of spam. In some cases this opinion would seem justified. However, there are some technological solutions which, when implemented correctly, can yield great success in dealing with spam.

Another aspect of dealing with spam is the need to ensure that legitimate business-related e-mail is not quarantined or denied. Incorrectly configured technological approaches most often result in high ratios of mistakenly quarantined e-mails, also known as false positives. The exact false-positive ratio that is acceptable varies from business to business depending on the field and size. In the case of the legal industry a very low false positive ratio is essential to ensure the integrity of business relationships. An example of a low false positive ratio is one false positive in 250,000 e-mails or approximately 0.0004 percent.

BlackHole Lists

Overview

The first, and one of the oldest, technologies is the blackhole list. These blackhole lists are collections of IP addresses. Depending on the list, there are varying reasons for the IP address appearing on the list. A common reason was because the IP was previously an open relay. An open relay is a SMTP Mail Transfer Agent which does not authenticate access either by incoming IP address or user accounts. These open relays can be used by anyone to send e-mail anywhere. Another common

reason to be included in a blackhole list is because of attacks originating from an IP address. There are less common reasons such as upsetting the administrator of the list and having an entire list of networks added to the list. These "religious wars" have reduced the effectiveness of publicly available lists as more innocent ISPs have been impacted by upsetting an administrator. A blackhole list can result in a high false positive ratio depending on the number of business relationships impacted by the block.

Recommendation

In general black hole lists are not as beneficial unless a thorough evaluation has been performed on the IP addresses received. If an e-mail relay service is used, then blackhole lists are not effective because of the need to restrict inbound e-mail to the provider's IP addresses only. A properly tuned blackhole list can provide an effective starting point in the fight against spam but requires a lot of time to maintain as the list of spammers move constantly. If a business is considering this method of recourse it might be well to consider purchasing a blackhole list subscription. However a thorough analysis of the IP addresses contained in the list and how the IP addresses pertain to one's business can significantly reduce the false positive ratio of this technology.

Keyword and Phrase Filters (Basic)

Overview

The second technological approach is almost as old as blackhole lists and involves matching clear text keywords to block e-mail. These filters can be finely tuned providing weighted measures, varying algorithms and multiple words to result in a successful block. The short fall of this approach is the need to maintain the keyword list. All except the most complex keyword filters are easily fooled by adding spaces, non-standard characters, non-printable ASCII text, or using images of blank spaces between letters. The result is a constant battle which consumes resources and provides mediocre results at best. This technology better suited to enforcement of an offensive language policy rather than combating spam.

Recommendation

This technology has a high resource requirement with mediocre results and a high rate of false positives. If considering this form of recourse, ensure that adequate staffing is assigned to culling through the quarantined messages looking for false positives and inappropriately quarantined messages.

Reverse DNS Lookup Filter

Overview

The Reverse DNS Lookup as a filter has become a popular technique for combating spam. The idea behind this class of filters is to enforce the idea that if a company is legitimate then they will take the time to enforce standards of interaction on the Internet. One of these standards is the idea of matching the forward and reverse DNS lookups for the mail relay (MTA) interacting with the Internet. If these don't match then there's a good chance the connection is originating from a dialup account or home broadband system. A few years ago this principle held a lot of credibility within the community as there were few offices or legitimate sources of email originating from small networks or dialup accounts. However, with the advent of telecommuting and the growth of broadband Internet access this assumption is somewhat sketchy and can lead to a large false positive rate. Small businesses that are connecting to the Internet are often doing so through a broadband connection. With the leasing of small sections of IP addresses, the ISP is usually denying the ability of managing the reverse lookups for this IP range. The use of this technology may leave your company without exposure to one of the most rapid growth sectors today.

Recommendation

This technology was effective a few years ago, but as the Internet has grown up with the increase in broadband Internet access, the effectiveness of this technology has decreased. Depending on the business model in use, this technology may leave a company without the ability to effectively communicate with a growing market whose primary connection is through the use of high speed broadband access. The capital expenditures of this growing segment of the small business market can result in significant numbers of potential sales for businesses aware of this area. Proceed tentatively

with this technology and perform analysis based on IP addresses from logfiles to determine if this technology is applicable before implementation.

Keyword and Phrase Filters (Heuristic)

Overview

The fourth technological approach is part of a new breed of defensive measures. These measures are called "heuristic" because they use statistical mathematical formulas to determine the probability that a single e-mail message is spam. One such technique was first popularized by Paul Graham in his paper "A Plan for Spam" in August 2002. He detailed a formula for determining the probability that a piece of e-mail is spam based on some simplified equations that he included in his paper. The features of his filter require two buckets to work effectively: Good E-Mail and Bad E-Mail. Given a large enough sampling of e-mail in each bucket, the filter yields excellent results without the high false positive rates seen with simple keyword filters or blackhole lists. In addition, this type of filter has a tendency to change overtime as more samples are added, thus making the filters more effective and able to adjust to new techniques in spam. In addition, this type of filter can be tuned to an individual's preferences allowing personalized filters to exist. There are other statistical methods available and different filters target different portions of the e-mail from the header section to the body only.

Recommendation

Heuristic filters are a big improvement from the manually maintained keyword lists. Because they can be configured to maintain themselves while providing a low false positive rate, these are the best filters available today for catching spam. The drawbacks are the large sample base needed and the inability to accurately handle binary image files. If one is planning to implement these technologies, they need to allocate enough time to design the filter to operate as self-sustaining a mode as possible. In addition, configuring the ability to automatically add new samples as often as necessary will ensure the accuracy of the filter. With appropriate planning these filters can yield amazingly effective results with extremely low false positive rates.

Image File Filters and Scanners

Overview

Image File Filters and Image File Scanners, as a class of filtering, is just beginning to reach the stage where the results can be visualized. The idea behind this class of filter is to be able to take a raw image, like a JPEG, and determine if it contains offensive words, nudity or any other undesirable content. Currently, this technology of image recognition is not very accurate. However, this class of filtering technology yields amazing potential and several companies are pouring resources into the development and refinement of this technology.

Recommendation

This technology is fairly young and somewhat inaccurate in the current set of implementations. However, these filters have incredible potential and can yield interesting results. The next few years will be critical to the development of this class of filtering technology. Companies might consider implementing this class of filters in an effort to assist with the growth and refinement as the technology matures.

URL Filters

Overview

The URL Filter class is an implementation of a well-known application to a new purpose. An HTML e-mail can contain links to content not embedded in the e-mail itself. These links may not trigger any alarms on spam filters to deny the e-mail, but the content linked within an e-mail could be wildly explicit. This class scans those links, much like an actual web request, and determines if the links yield known offensive content. If so, then the e-mail is classified as spam or weighted more severely than otherwise intended.

Recommendation

This is a proven class of filtering technology that is currently employed by companies every day. This implementation of this technology as a class of e-mail filters will continue to develop and become more effective. However, the drawback of the URL Filtering technology (like antivirus scanners) is that, while incredibly effective; they rely on known signatures or URLs. This means that someone gets

hit before this class of filters becomes effective in stopping that signature. The good news is these filters are pulling from all customers so there is a good chance that the initial burden will lie somewhere else. Leveraging an outsourced service for this technology is useful for a business to mitigate the risk of the first exposure and gain additional traffic of visited URLs from neighboring users.

Opt-In Access Filters (WhiteList)

Overview

WhiteList filtering has become more widespread by larger corporations today as they continue to handle the brunt of the spam flood. The idea is to deny e-mail access from any IP address until a registration process has been completed in order to match an incoming e-mail address to the permitted IP address. The effectiveness stems from the gamble that automated mailers used by spammers will not take the time to fill in the form before sending their e-mail. However, automating an HTML form is a simple scripting challenge and the spammers will adapt over time and begin to circumvent this white listing filter. With appropriate automation, this technology can be rendered ineffective even with resources dedicated to validating the entered IP addresses. A script can send twenty e-mails faster than a person can click a "Deny Access" button. In addition, this technology takes a risk of alienating the portion of the market who does not understand the technology.

Recommendation

White List Filtering technology may appear effective initially, but will tend to have little effect over time. This technology should be used as a stop-gap while a more effective filtering solution, like the heuristic filters described above, are being developed and implemented. If this technology is a necessity, one should be prepared to dedicate resources to cleaning the table structures to deny the automated entries that are destined to occur.

Tarpitting

Overview

The seventh class of filters consist of the idea of tar pitting, or slowing down, recurrent connections to an e-mail relay or MTA. This idea states that if an additional connection is attempted within a quiet

period, said connection will be paused an incremental amount of time before allowing the connection to proceed. This will effectively slow down the ability of the sending relay to deliver e-mail, not only to the single recipient, but in general as more connections are tied up in the tar pit. Tarpitting is a clever idea which operates on the assumption that legitimate e-mail does not come from the same source consistently, but multiple spam messages will come in a flood.

Recommendations

Depending on the business, tarpitting may provide some reprieve over time if enough border points adopt similar technology. There are no circumventions to this method since e-mail would need to be delivered from multiple connection points simultaneously which ensures that the cost of bypassing this technology is not beneficial to the spammer. Spammers often prefer to use a single relay to flood spam in a constant pipe to the Internet. Tarpitting stems the pipe for all users making it much less likely for spam to continue to grow as more and more tar pits appear as a standard. The downside to this class of technology is that this approach is not wide spread yet and if a business uses a hosting service to filter e-mail the tar pit will cause more problems than benefits. In addition, while tarpitting is simple the implementation does require some skill to apply the appropriate code patches and build the mail relay or MTA appropriately.

Legislation

Overview

SPAM legislation provided by the government enforces the restriction of spam activity to legal distribution of e-mail only. There are thirty four states in the United States which currently have enacted anti-spam legislation. Furthermore, there is mounting public pressure for Congress to enact nation-wide legislation in conjunction with the state laws currently in effect. Appendix A lists a collection of pending anti-spam legislation in the U.S. and European Union. However, even though individual states have legislation opposing spam, the general problem has only become worse. As stated above, the Direct Marketers Association (DMA) is lobbying Congress to legalize spam by setting standards on e-mail sent within the U.S. As of today, based on the increasing trend of spam, the DMA can be said to be ineffective or unable to provide self-regulation and control their members.

How can it be said that the DMA is the source of the problem and not rogue spammers not associated with the DMA?

There are several issues facing effective legislation of anti-spam efforts. First, the Internet is tentatively related to the physical boundaries defined by international entities today. While these boundaries can be enforced to some degree within an E-Commerce world because of the need to physically deliver products to a place or destination, e-mail is not confined by these boundaries and the impact of restricting and enforcing such restrictions are much less promising. For instance, some companies actively block ranges of IP addresses known to be associated with illegal activities originating in Asia and others do not. There is no question that traffic from these types of contacts would be considered harmful by all parties, but there are only so many links across the oceans. ISPs are not able to block these addresses because some of them may be used for legitimate business interaction. This is a similar problem to what is faced by corporations attempting to use blackhole lists or other technologies based on IP addresses. Legislation that does not take into account the restrictiveness of the jurisdiction of their mandates where the Internet is concerned is guaranteed to be ineffective in the fight against spam. The playing field is changing and the old physical boundaries are blurring into a global-space utilized by everyone. Spammers know this and will move offshore or pay for services from another country willing to accept the traffic in exchange for cash.

The second issue facing Congress is the growing trend of virus-like spam distribution. A trend that started recently is for a spammer to launch a virus-like attack on a directly connected computer like a home DSL connection. To date, most home users have been held harmless for the unintended actions of their computer if they were unknowingly infected by a virus. In this case, there would be no possible recourse unless the perspective of the judicial system changes. The responsibility of the security of a broadband users' home connection is beyond the scope of this paper, but would be an interesting topic of discussion and further study.

The third issue facing law makers is the tentativeness of a "successful venture" where these legislations are concerned. By defining what constitutes legitimate marketing e-mail from spam, e-mail legislation is opening up a flood of additional inquiries and a possible increase in marketing e-mail instead of

spam. Once acceptable conduct is introduced into the picture, businesses or individuals will be able to legally flood companies and persons with advertisements without the risk of retaliation. If the volume of e-mails continues to grow, the amount of spam seen by the end-user will increase accordingly unless advances in the technological filters described can provide additional performance. Even with increased performance of current filters the amount of spam could become unmanageable and render e-mail unusable as a medium of communication.

Given the issues facing Congress, their hesitancy in enacting anti-spam legislation into law is understandable. With the current approach to the issues cited above, one has to question the real need for and possibility of enforcement. If an anti-spam legislation is enacted, will it have any effectiveness in stemming the flood of spam e-mails? Given the global nature of the Internet and the limited jurisdiction of the U.S. justice system, the possibility of encountering an unenforceable policy condition in another country is likely. What happens to the average broadband consumer who becomes infected and is a jumping point for anonymous spam? This approach of sending spam renders the original connection untraceable. Without significant expenditures to provide education and protection for every person connecting to the Internet, these isolated incidents will yield very large distracting targets for enforcement of the legislation. Once the acceptable parameters for spam are defined, how does one ensure that the amount of spam decreases instead of increases? Can a penalty or fine for spam be collected realistically given today's environment, or will this law become unenforceable?

Recommendations

There are many issues facing Congress when considering anti-spam legislation. Perhaps the perspective of the current approach should be adjusted to align more closely with the nature of the Internet and the global perspective that it represents. First, the desired effect for the general public is to reduce the constant flood of e-mail coming into their inbox. Second, on the other side of the argument, the DMA would like to be able to retain an effective and cheap way of advertising to masses of people worldwide. Third, the underlying protocol is fundamentally flawed in such a way that non-repudiation is not guaranteed, thus decreasing the chance or likelihood of enforcing any legislation. Because of these issues the current focus of criminalizing spam is bound to fail to accomplish its goals.

Providing incentives for more effective communication techniques could possibly yield better results than assigning punitive measures to spam. For example, if a proposal to adopt digital signatures at the company level were instituted this would provide a way for all e-mail coming from a company entity to be digitally signed to ensure non-repudiation. The recipient entity could filter based upon digital signature of inbound e-mail providing more efficient processing of e-mail from semi-trusted sources or partners. In addition, if you had a collection of persons who did not regularly receive e-mail from certain sources, these could easily be filtered by their signature. This approach differs from the use of the "FROM" address because, with the digital signature, the address cannot be falsified without access to their private keys and the necessary infrastructure to process appropriately.

An alternative method of handling spam (and phone calls) is to mandate uninvited callers to "make a binding offer to pay an interrupt fee to the recipient" as a way of validating the contact. The system includes an accept list that is managed by the owner of the e-mail account. Uninvited callers must make a binding offer to pay an interrupt fee to the recipient in exchange for receiving the e-mail. While this system seems to result in the same system as the "White List or Opt-In Access Filters" described above, the main difference is the fee attached to sending e-mail. Repetitively sending e-mail to the company can cost a substantial amount of money depending on the interrupt fee negotiated. Without the recurring fee, the system is flawed with the ease of scripting needed to bypass these filters. In addition, the recipient can opt-out of accepting the fee associated with the call. This allows for the ability to freely exchange e-mails between companies or friends without cost. While this may seem to be an unusual suggestion, it could be on the right track. This solution provides the necessary transaction environment for enforcing, tracking and assessing the cost of an action or infringement against a recipient of e-mail.

For legislation to succeed in providing a realistic and effective spam deterrent, the proper perspective must be attained. In addition, a global perspective needs to be considered in order to assess the effectiveness of proposed nation-wide anti-spam legislation. A consideration for the enforceability of a proposal needs to be assessed. And lastly, the effectiveness of the proposal needs to be measured and reported to ascertain a successful initiative. Without these three key elements, an attempt at deterring spam will only provide a rock, not a dam, in the flood of e-mail spam.

Chapter 5

DEFENSE IN DEPTH

Protecting against the many maladies that affect e-mail is a requirement of today's business environment. A comprehensive approach to handling e-mail is required to provide complete protection for the end-user. In order to analyze the protection needed one must understand the requirements of a comprehensive system. The requirements should address the three core initiatives of **C**onfidentiality, **I**ntegrity, and **A**vailability.

Confidentiality		
	Encryption	Encrypting e-mail during delivery and while at rest will ensure that only the appropriate parties are able to read the messages. If a non-encrypted e-mail were received from the company it could be discarded as invalid or falsified.
Integrity		
	Digital Signature	Digitally signing each e-mail sent will allow verified authenticity and non-repudiation of e-mail origination. A non-signed e-mail could be filtered as invalid or falsified. The digital signature can be applied at the individual or corporate entity level.
Availability		
	Aggressive Filtering	Antivirus and spam filtering is needed to ensure that valid e-mails are received and inappropriate content is filtered for additional verification or discarded without impacting the delivery of valid e-mail messages.
	Redundancy	Redundant e-mail gateways provide the capability to continue to receive e-mails during an outage or loss of service situation.
	Queuing	Queuing of e-mails provides the ability to continue to receive e-mails in the case of a complete loss of service or extended outage situation.
	Advanced routing and redirection	Additional actions to standard policy and content filters provide extra levels of availability and functionality to ensure that communications are delivered appropriately.

This paper focuses on the **A**vailability and **I**ntegrity of e-mail because of the threats posed by the increasing flood of unsolicited bulk e-mail (or spam). Multiple layers of filtering and redundancy are needed to ensure adequate protection and isolation of inappropriate material. Figure 10 displays a

graphical representation of the architecture for a comprehensive system using the technologies detailed earlier in this paper.

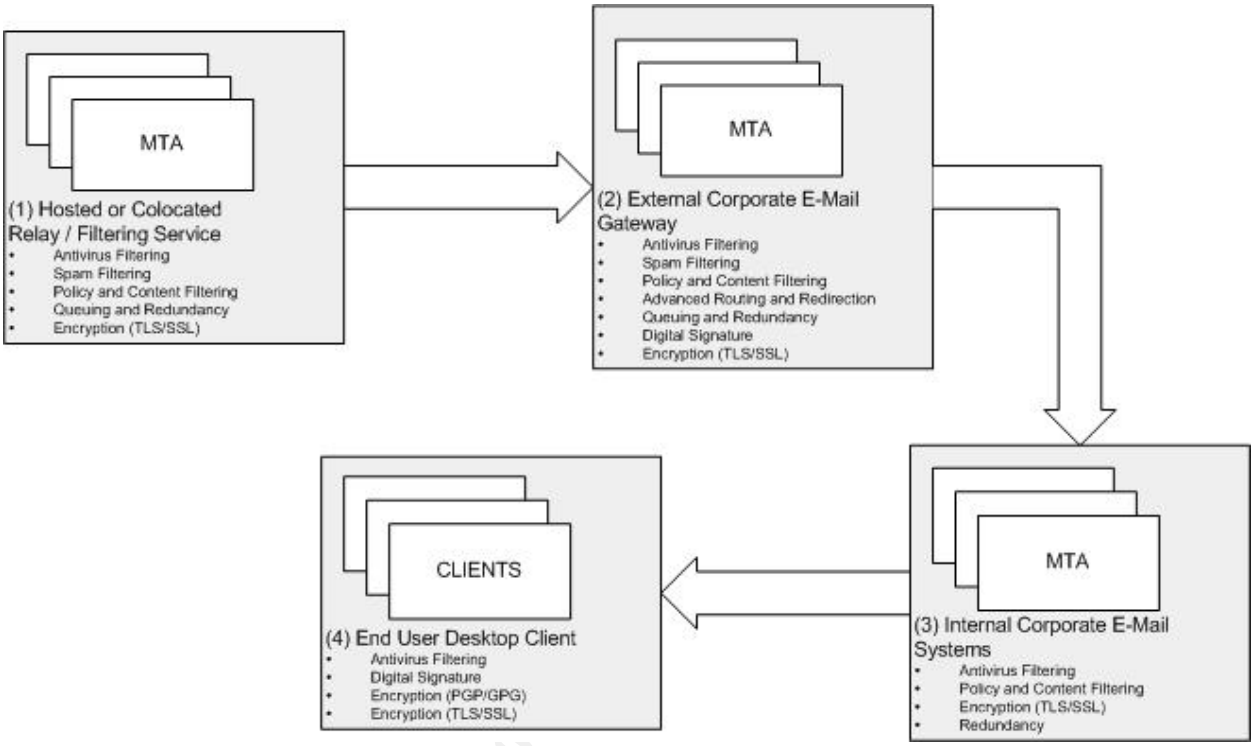

Figure 10: Comprehensive E-Mail Perimeter Protection

This completes the overview of the defensive architecture to provide a protected e-mail environment for the end-user. A detailed inspection of the applicability of each of the technologies listed in Chapter 4 as well as additional technologies described in this chapter will assist in isolating appropriate choices for each step of the architecture.

Step 1: Hosted or Collocated MTAs

This step is the gateway for all e-mail transactions entering or departing the company and should provide an initial (and final) filter prior to presentation to the external entities. The functionality represented here covers **C**onfidentiality, **I**ntegrity and **A**vailability.

Confidentiality		
	Encryption (TLS/SSL)	Encrypting e-mail during delivery and while at rest will ensure that only the appropriate parties are able to read the messages. TLS/SSL is

		encryption for transmission between MTAs and end-user stations. TLS/SSL does not encrypt the e-mail at rest or while stored on a hard drive, etc.
Integrity		
	Antivirus Filtering	Multiple antivirus engines scanning inbound/outbound e-mail to ensure close to 100% virus free e-mail exchange
	Spam Filtering	Multiple filtering techniques provide the ability to filter above 90% of all spam received with a false positive rate of less than 1:250,000
	Policy and Content Filtering	The ability to block inbound/outbound e-mail based on key elements such as Subject, To, From, Attachment type, and more
Availability		
	E-Mail Queuing	queuing of inbound /outbound e-mail for future delivery in the event of a power outage, loss of service or catastrophic event which can last for several days
	Redundancy	multiple, redundant and load balanced servers provide a seamless interface with external entities

Confidentiality of communications is guaranteed by utilizing the TLS/SSL libraries for passing critical e-mail messages. This encryption wrapper for the SMTP protocol is an easy fit onto an existing e-mail system, but must be enforced from end-to-end to completely guarantee confidentiality of communications. Transport Layer Security (TLS) utilizes encryption standards such as shared-key or certificate based encryption to transport data securely over untrusted network segments. Similar to IPSEC, an incorrect shared key phrase or certificate will result in failure to establish communications. Secure Sockets Layer (SSL) is identical to the SSL used for securing web communications using HTTP/S. When a MTA (or MUA) connects to send/receive messages the certificate is exchanged and communications start. This certificate exchange ensures encrypted communication over the complete delivery path. However, each segment, from MTA to MTA, is separate and the certificates change accordingly. Either TLS or SSL provides adequate protection for protecting SMTP communications across untrusted network segments. However, the use of a third party certificate authority to provide remediation of communication certificates can provide a simplified implementation of these technologies.

Integrity of communications is guaranteed using a combination of antivirus and spam filtering technologies as describe previously. Technologies that affect the isolation of external entities such as

Blackhole Lists, Reverse DNS Lookup Filters, and Tar pitting can only be used at this step in the architecture. In addition to these technologies, the rest of the filtering technologies such as keyword and phrase filters, image file filters, URL scanning filters, and opt-in filters can be implemented as measures against spam at this step. A large portion of inappropriate and malicious content should be removed from e-mail messages while passing through this step of the architecture.

Availability of communications is guaranteed with a combination of redundancy and message queuing. This provides mitigation of long term and short term outages to critical e-mail infrastructure components. Most of the time recovery from an outage will be automatic and happen without user intervention.

Step 2: External Corporate E-Mail Gateway

The second step (2) in the architecture consists of providing a mail transfer agent (MTA) which acts as a gateway between the corporate network and the external entities. The functionality represented here covers **C**onfidentiality, **I**ntegrity and **A**vailability.

Confidentiality		
	Encryption (TLS/SSL)	Encrypting e-mail during delivery and while at rest will ensure that only the appropriate parties are able to read the messages. TLS/SSL is encryption for transmission between MTAs and end-user stations. TLS/SSL does not encrypt the e-mail at rest or while stored on a hard drive, etc.
Integrity		
	Digital Signature	Digitally signing each e-mail sent will allow verified authenticity and non-repudiation of e-mail origination. A non-signed e-mail could be filtered as invalid or falsified. The digital signature can be applied at the individual or corporate entity level.
	Antivirus Filtering	Multiple antivirus engines scanning inbound/outbound e-mail to ensure close to 100% virus free e-mail exchange
	Spam Filtering	Multiple filtering techniques provide the ability to filter above 90% of all spam received with a false positive rate of less than 1:250,000
	Policy and Content Filtering	The ability to block inbound/outbound e-mail based on key elements such as Subject, To, From, Attachment type, and more
	Advanced Routing and Redirection	The ability to provide additional functionality and actions to the policy and content filtering. Additional actions are provided for e-mails which match criteria such as redirect, blind-copy, drop, reject, and route.

Availability		
	E-Mail Queuing	The ability to queue inbound /outbound e-mail for future delivery in the event of a power outage, loss of service or catastrophic event which can last for several days
	Redundancy	Multiple and/or redundant and load balanced servers provide a seamless interface with the hosted/collocated service and internal systems.

With the exception of Advanced Routing and Redirection and Digital Signature capabilities, the requirements for guaranteeing **I**ntegrity are identical to step (1). However, because of the placement in the chain, technologies such as black hole filters, tar pitting and reverse DNS lookup filters cannot be used without impacting performance of the system. Technologies such as keyword and phrase filters, image file filters, opt-in white lists and URL filters can provide adequate filtering for a second layer of spam filtering. Because of the location of this step in the architecture a higher false positive ratio can be supported, but should be kept to a minimum to facilitate manageability of the system. Digital Signatures can be applied at this level for a corporate entity to validate outbound e-mail. This provides non-repudiation of e-mail messages for the corporation in question.

Confidentiality is guaranteed by proving TLS/SSL encryption as described in the previous section. Either technology will suffice since this is a marginally trusted area and contains communications with the internal systems.

Step 3: Internal Corporate E-Mail Systems

The third step (3) in the architecture is the final MTA or endpoint for the e-mail message. The systems at this level commonly hold the e-mail message until the user picks up the message. The functionality represented here covers **C**onfidentiality, **I**ntegrity and **A**vailability.

Confidentiality		
	Encryption (TLS/SSL)	Encrypting e-mail during delivery and while at rest will ensure that only the appropriate parties are able to read the messages. TLS/SSL is encryption for transmission between MTAs and end-user stations. TLS/SSL does not encrypt the e-mail at rest or while stored on a hard drive, etc.
Integrity		
	Antivirus Filtering	Multiple antivirus engines scanning inbound/outbound e-mail to ensure close to 100% virus free e-mail exchange

	Policy and Content Filtering	The ability to block inbound/outbound e-mail based on key elements such as Subject, To, From, Attachment type, and more
Availability		
	Redundancy	Multiple and/or redundant and load balanced servers provide a seamless interface with the end-user.

Confidentiality is provided by the use of TLS/SSL to encrypt communications between end-users and the servers in step 2.

Integrity is provided by simplifying the architecture for this point. Antivirus filtering and policy and content filtering provide adequate coverage for end-users.

Availability is provided with redundant internal servers, according to policy, which allows increased uptime for end-users.

Step 4: End-User Desktop Client

The fourth, and final, step (4) consists of the mail user agent (MUA) which provides an interface between the end-user and the e-mail architecture. The functionality represented here covers **C**onfidentiality and **I**ntegrity.

Confidentiality		
	Encryption (TLS/SSL)	Encrypting e-mail during delivery will ensure that only the appropriate parties are able to read the messages. TLS/SSL is encryption for transmission between MTAs and end-user stations. TLS/SSL does not encrypt the e-mail at rest or while stored on a hard drive, etc.
	Encryption (PGP/GPG)	Encrypting e-mail during rest provides protection of e-mail content during rest and transmission. Decryption of messages only happens when read. This protection is possible using technologies such as PGP or GPG and requires a PKI infrastructure or similar functionality to facilitate the key exchange between entities and individuals.
Integrity		
	Digital Signature	Digitally signing each e-mail sent will allow verified authenticity and non-repudiation of e-mail origination. A non-signed e-mail could be filtered as invalid or falsified. The digital signature can be applied at the individual or corporate entity level.

Confidentiality at this level provides a new encryption type which protects the e-mail at rest and during transmission. Encrypting the message using technology such as PGP or GPG an end-user can ensure that potential eavesdropping results in unreadable messages. This level of encryption requires a wide-spread infrastructure to facilitate key distribution. In addition, protection of the private keys used for decrypting a message needs to be in place, as a lost or stolen key can cause a significant loss of confidentiality for the system. Also, the end-user can use TLS/SSL to encrypt communications with the internal server to further protect their communications.

Integrity is provided by personal digital signatures ensuring non-repudiation of e-mails at an individual level. Again, a robust infrastructure is needed to distribute keys for this functionality to be effective.

Conclusion

Ensuring adequate protection and filtering technology for email is an important part of a proactive security posture for companies today. Ensuring **C**onfidentiality, **I**ntegrity and **A**vailability for e-mail infrastructure faced with the increasing flood of unsolicited bulk e-mail or spam requires the integrated use of many separate but effective technologies. Technologies are more effective than ever before and have the potential to continue to be effective in reducing the flood of spam. As emerging technological areas such as Image File Filtering and Scanning continue to mature, their effectiveness will become more potent. Technology has the ability to effectively address the onslaught of spam with the proper perspective regardless of origination.

While legislation may provide non-technical approaches to addressing spam, their effectiveness is hindered by a lack of jurisdiction for offenders outside the boundaries of the U.S. This lack of a global perspective essentially ensures a safe haven of operations for spammers from legal actions brought by U.S. citizens and businesses. In fact, by providing fixed parameters which define spam legislation may increase the amount of spam sent. Once a definition is considered "good", spammers will be free to spew forth appropriately labeled e-mail messages without worry of legal action. Appendix A provides a list of some of the proposed legislation in the U.S. Senate and U.S. House of Representatives. Some carry potential to reap some effectiveness but how effective these will be remains to be seen.

Appendix A: Pending Spam Legislation

Bill	What is Proposed	Sponsors	Status / Notes
United States Senate			
Can-Spam Act *Controlling the Assault of Non-Solicited Pornography and Marketing Act*	• Legitimate return address and advertising labels • Prohibits use of confusing and deceptive subject lines • Consideration for a Do-Not-Spam registry	Conrad Burns (Republican Montana) Ron Wyden (Democrat Oregon)	• First introduced in April 2003. • Approved unanimously on June 19th, 2003 by the Commerce Committee.
Spam Act *Stop Pornography and Abusive Marketing Act.*	• Federal registry for non-spam • Requires labels for adult content messages • Allows individuals (or companies) to sue spammers	Charles Schumer (Democrat New York) Lindsey Graham (Republican North Carolina)	• Introduced in June 2003 • Pending vote • Endorsed by Christian Coalition
Criminal Spam Act	• Criminal penalty of up to five years for breaking into computers to send spam and using	Orrin G. Hatch (Republican Utah) Patrick J. Leahy (Democrat Vermont)	• Introduced in June 2003 • Pending vote • Loosely defines

	fake identities for accounts		spam as an e-mail designed to promote a product or service.
Ban on Deceptive Unsolicited Bulk E-mail Act	• Prohibits false information on subject lines and harvesting of e-mails from websites • Requires opt-out mechanism	Bill Nelson (Democrat Florida)	• Introduced in May 2003 • Pending vote
Computer Owner's Bill of Rights	• Nationwide "do-not-email" registry by the FTC • Empowers FTC to impose civil penalties on spammers who send to addresses on the registry	Mark Dayton (Democrat Minnesota)	• Introduced in March 2003 • Pending vote
United States House of Representatives			
Anti-Spam Act	• Must include legitimate return addresses and advertising labels • Requires an opt-out mechanism	Heather A. Wilson (Republican New Mexico) Gene Green (Democrat Texas)	• Introduced in June 2003 • Pending vote • Over three dozen sponsors currently in the house

	• Enforces a reply to requests within a ten-day period • Prohibits deceptive or misleading subject lines • Addresses cannot be shared to third parties • Proposes misdemeanor criminal penalties for spammers	John D. Dingell (Democrat Michigan) John Conyers Jr. (Democrat Michigan) Anna Eshoo (Democrat California)	• Gaining momentum
Rid Spam Act *Reduction in Distribution of Spam Act*	• Requires legitimate return addresses, advertising labels and opt-out • Opt-out would expire after three years • Information can be shared if consumers given notice • Proposed misdemeanor criminal penalties	Richard M. Burr (Republican North Carolina) James Sensenbrenner (Republican Wisconsin) Billy Tauzin (Republican Louisiana)	• Introduced May 2003 • Pending vote • Loosely defines spam as an e-mail designed to promote a product or service.
Reduce Spam Act	• Offers bounties	Zoe Lofgren	• Introduced in May

Restrict and Eliminate the Delivery of Unsolicited Commercial Electronic Mail or Spam Act	equal to twenty percent for individuals providing information to identify and prosecute spammers • Prohibits deceptive or misleading subject lines • Individuals can sue spammers • Requires labeling for advertising and pornography	(Democrat California)	2003 • Pending vote
Wireless Telephone Spam Protection Act	• Prohibits use of wireless SMS to send spam	Rush D. Holt (Democrat New Jersey)	• Introduced in January 2003 • Pending vote • Closest bill for protecting against wireless spam

Appendix B: Technology to Architecture Mapping

Confidentiality

Technology	**Step 1**	**Step 2**	**Step 3**	**Step 4**
Encryption (SSL/TLS)	X	X	X	X
Encryption (PGP/GPG)	---	---	---	X

Integrity

Technology	**Step 1**	**Step 2**	**Step 3**	**Step 4**
Digital Signatures	---	X	---	X
Antivirus Filtering	X	X	X	X
Policy and Content Filtering	X	X	X	---
Advanced Routing and Redirection	X	X	---	---
Black hole Lists	X	---	---	---
Keyword and Phrase Filters (Basic)	X	X	---	---
Keyword and Phrase Filters (Heuristic)	X	X	---	---
Reverse DNS Lookup Filters	X	---	---	---
Image File Filters and Scanners	X	X	---	---
URL Scanning and Filters	X	X	---	---
Opt-In Access Filters (White list)	X	X	---	---
Tar pitting	X	---	---	---

Availability

Technology	**Step 1**	**Step 2**	**Step 3**	**Step 4**
E-Mail Queuing	X	X	---	---
Redundancy	X	X	X	---

BIBLIOGRAPHY

Fahlman, Scott E. *Selling interrupt right: A way to control unwanted e-mail and telephone calls*, IBM Systems Journal, Vol 41, NO 4, 2002.

www.monkeys.com: *Spam Defined.* http://www.monkeys.com/spam-defined/definition.shtml

www.spamlaws.com: *Hotmail v. Van$ Money Pie.* http://www.spamlaws.com/cases/vanmoneypie.html

www.spamlaws.com: *Ferguson v. Friendfinders, Inc.* http://www.spamlaws.com/cases/ferguson.html

www.wa.gov: *Junk Email: Washington's Law* http://www.wa.gov/ago/junkemail/index.shtml

www.crynwr.com: *The Definition of Spam* http://www.crynwr.com/spam/definition.html

www.mail-abuse.org: *Mail Abuse Prevention System: Definition of "spam"* http://www.mail-abuse.org/standard.html

www.dictionary.com: *spam* http://dictionary.reference.com/search?q=spam

www.spamhaus.org: *The definition of spam* http://www.spamhaus.org/definition.html

Last Updated: July 23rd, 2017

Upcoming Training

SANS Boston 2017	Boston, MA	Aug 07, 2017 - Aug 12, 2017	Live Event
SANS Prague 2017	Prague, Czech Republic	Aug 07, 2017 - Aug 12, 2017	Live Event
Community SANS Omaha SEC401*	Omaha, NE	Aug 14, 2017 - Aug 19, 2017	Community SANS
SANS New York City 2017	New York City, NY	Aug 14, 2017 - Aug 19, 2017	Live Event
SANS Salt Lake City 2017	Salt Lake City, UT	Aug 14, 2017 - Aug 19, 2017	Live Event
SANS Adelaide 2017	Adelaide, Australia	Aug 21, 2017 - Aug 26, 2017	Live Event
Community SANS Trenton SEC401	Trenton, NJ	Aug 21, 2017 - Aug 26, 2017	Community SANS
Virginia Beach 2017 - SEC401: Security Essentials Bootcamp Style	Virginia Beach, VA	Aug 21, 2017 - Aug 26, 2017	vLive
SANS Chicago 2017	Chicago, IL	Aug 21, 2017 - Aug 26, 2017	Live Event
SANS Virginia Beach 2017	Virginia Beach, VA	Aug 21, 2017 - Sep 01, 2017	Live Event
Community SANS Pasadena SEC401 @ NASA	Pasadena, CA	Aug 23, 2017 - Aug 30, 2017	Community SANS
Mentor Session - SEC401	Minneapolis, MN	Aug 29, 2017 - Oct 10, 2017	Mentor
SANS San Francisco Fall 2017	San Francisco, CA	Sep 05, 2017 - Sep 10, 2017	Live Event
SANS Tampa - Clearwater 2017	Clearwater, FL	Sep 05, 2017 - Sep 10, 2017	Live Event
Mentor Session - SEC401	Edmonton, AB	Sep 06, 2017 - Oct 18, 2017	Mentor
SANS Network Security 2017	Las Vegas, NV	Sep 10, 2017 - Sep 17, 2017	Live Event
Community SANS Albany SEC401	Albany, NY	Sep 11, 2017 - Sep 16, 2017	Community SANS
Mentor Session - SEC401	Ventura, CA	Sep 11, 2017 - Oct 12, 2017	Mentor
Community SANS Columbia SEC401	Columbia, MD	Sep 18, 2017 - Sep 23, 2017	Community SANS
Community SANS Dallas SEC401	Dallas, TX	Sep 18, 2017 - Sep 23, 2017	Community SANS
Community SANS New York SEC401	New York, NY	Sep 25, 2017 - Sep 30, 2017	Community SANS
Rocky Mountain Fall 2017	Denver, CO	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS London September 2017	London, United Kingdom	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Baltimore Fall 2017	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Copenhagen 2017	Copenhagen, Denmark	Sep 25, 2017 - Sep 30, 2017	Live Event
Community SANS Boise SEC401	Boise, ID	Sep 25, 2017 - Sep 30, 2017	Community SANS
Baltimore Fall 2017 - SEC401: Security Essentials Bootcamp Style	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	vLive
SANS DFIR Prague 2017	Prague, Czech Republic	Oct 02, 2017 - Oct 08, 2017	Live Event
Community SANS Charleston SEC401	Charleston, SC	Oct 02, 2017 - Oct 07, 2017	Community SANS
Community SANS Sacramento SEC401	Sacramento, CA	Oct 02, 2017 - Oct 07, 2017	Community SANS
Mentor Session - SEC401	Arlington, VA	Oct 04, 2017 - Nov 15, 2017	Mentor