

How Spamhaus Cost-Effectively Eliminates Spam

An Osterman Research White Paper

Published March 2008



Why You Should Read This White Paper

If you use email, you receive spam. While the problem is onerous for individual users, it is dramatically worse for companies, network operators, Internet service providers and other that process large amounts of email. The problems caused by spam include annoyance, slower message delivery times and the like. However, the most serious problem caused by spam is higher costs: loss of employee productivity; additional investments required in spam-filtering hardware, software or services; crashed servers that take more IT time to address; additional investments in IT staff to resolve spam-related problems; and a host of other problems.

In essence, the goal for email system administrators boils down to three fundamental issues:

- Eliminating as much spam as possible without losing legitimate email
- Keeping mail flowing without spam filtering impeding mail flow
- Doing both of these as cost effectively as possible

Despite the fact that billions have been invested to eliminate spam, the problem is getting worse: spam volumes are increasing at a rapid pace and the proportion of email represented by spam continues to increase. Spammers continue to devise new and innovative methods for circumventing anti-spam defenses and they often succeed.

What organizations need, therefore, is a cost-effective, robust and easy-to-use approach to solving the spam problem and to maintain current, real-time defenses to combat the newest threats from spammers.

This white paper, sponsored by Spamhaus and Messaging Architects, discusses the various techniques that can be used to thwart spam, and it discusses Spamhaus' offerings – in use

by thousands of organizations worldwide – that are designed to block the vast majority of spam before it can enter and impact a corporate network. For example, one Spamhaus customer was able to eliminate \$400,000 in third-party service costs through the use of Spamhaus and it halved the number of email and filtering servers it operated. This goal – significantly reducing the cost of spam to organizations of all sizes – is the focus of this white paper.

Significantly reducing the cost of spam to all organizations of all sizes is the focus of this white paper.

Spam is Getting Worse

It almost goes without saying that the spam problem is getting worse. Depending upon the time of day, the day of the week and time of year, spam represents anywhere from 80% to 95+% of all email sent across the Internet. However, the severity of the problem is actually masked by these statistics: the absolute volume of spam is growing at a rapid pace. For example, between May and November 2006, spam volumes doubled.

THE COST AND PAIN OF SPAM

Spam carries with it a wide variety of problems, including its impact on individual users, companies, network operators, Internet service providers, hosted messaging service providers, companies that are purportedly sending phishing attempts and others. Among the problems that spam causes are:

- **Increased storage requirements**

Email that is suspected to be spam is placed into quarantines that users can inspect in order to check for false positives, or legitimate email that has mistakenly been identified

as spam. If we assume that the average corporate user receives 75 spam messages per day, that each spam is 50 kilobytes in size and spam is retained in the quarantine for 30 days, the quarantine will contain roughly 110 megabytes of content per user. In an organization of 5,000 users, that represents 550 gigabytes of unwanted content that must be managed by IT.

Between May and November 2006...spam volumes doubled.

- **Bandwidth constraints**

Transmitting spam across a corporate network consumes network bandwidth, resulting in slower email delivery, slower access to Web sites, the need to add bandwidth as spam volumes increase, and so forth. Further, newer types of spam messages, such as image spam, are up to ten times larger than conventional spam, and so place even greater demands on both bandwidth and storage.

- **Severe drains on network and server resources**

Spammers will often attempt to gather new email addresses through directory harvest attacks, in which email servers are flooded with a massive number of emails in an attempt to determine which addresses are valid. Because the SMTP protocol is designed to return information on which email addresses are not valid, those addresses that do not bounce back are assumed to be legitimate, generating new addresses for spam campaigns. These attempts can consume a large proportion of an email server's resources and can result in a server crash.

- **Loss of employee productivity**

Even with good spam-filtering technology in place, some spam still gets through. This means that employees will spend time looking at some spam messages, particularly those with seemingly valid subject lines, and then deleting those messages. This results

in loss of employee productivity and, in some cases, employees actually purchasing items they see advertised in these messages.

- **Financial losses**

Perhaps among the most serious consequences of spam is the fact that some people will be fooled by these messages, such as phishing attempts, and provide spammers with sensitive information.

- **The bottom line: spam drives up costs**

The cost of managing spam is roughly proportional to the amount of spam that is received. The more spam that enters an organization, the greater the number of email servers, spam filtering servers or appliances, support resources, anti-spam services, network bandwidth and other infrastructure elements that are required, driving up the cost of managing a network and messaging system.

Best Practices in Spam Remediation

Given the current severity of the spam problem, the fact that spam is getting worse and its impact on virtually anyone who sends or receives email, there are a variety of best practices that any organization should implement to thwart the problem.

WHITELISTS

A whitelist is a simple list of addresses of known legitimate senders of email. These can be maintained at both the corporate directory level and by individual users, and can speed delivery of email by passing through content from those on the whitelist without having to scan the content for its 'spamminess'.

BLACKLISTS / BLOCKLISTS

Blacklists and blocklists are just the opposite – these are lists of known or suspected spammers' IP addresses or domains whose content can be used to manage incoming email. If a suspected spammer's content enters a network or messages are received from a suspect IP address or domain, the content can simply be blocked.

Using a Domain Name System Block List (DNSBL), such as Spamhaus, can provide a company, network operator or other email processor with useful, real time information about incoming email. These are passive tools in that they simply provide information about

The advantages of a DNSBL are that it is inexpensive in both outright cost and in CPU cycles.

each incoming email, allowing the recipient of the email to use this information in a manner consistent with its own policies. For example, if an incoming email sender's IP address is listed on the DNSBL, it can be rejected outright, accepted or tagged as suspect and then passed along to a secondary filtering system. The advantages of a DNSBL are that it is inexpensive in both outright cost and in CPU cycles, and it provides information to the

sender so that legitimate sources listed on the DNSBL can take corrective action for future messages.

HEURISTICS AND RELATED TECHNIQUES

Heuristic filtering works by running each message and its various parts through a gauntlet of pre-defined rules and then scoring each message based on its content, the proximity of words to one another and other criteria. Messages that reach a particular threshold score are assumed to be spam and are then placed into a quarantine. Messages that score lower than this level are sent through to their intended recipients.

A somewhat related technique is Bayesian filtering, in which filters are 'trained' by examining a corpus of legitimate email and spam. Based on the characteristics of these two types of email, the filter can then determine with a high level of probability the likelihood of future messages being legitimate.

REPUTATION ANALYSIS

Reputation analysis is a newer technique that examines the reputation of sending sources to determine the likelihood that a message is legitimate. Using real-time or near real-time email traffic statistics and data on the amount of spam sent by each IP address, the reputation for a particular IP address can be determined. The theory behind reputation analysis is that if an email is received from an IP address that previously has sent large quantities of spam, the new email will be more likely than not to be spam. Content received from those IP addresses can then be throttled back to receipt of perhaps 100 messages per hour. This will allow legitimate email to be received in the event that the reputation score of the IP address was incorrect, but it will slow spam coming from that IP address to a trickle.

The fundamental advantage of reputation analysis is that it can stop a large proportion of email from entering a network, relieving spam-filtering systems from the burden of processing most of the spam that they otherwise would be required to manage.

Organizations should implement a multi-layered defense that use RBLs as a simple and low cost solution to block as much spam as possible before it is accepted by mail servers.

BEST PRACTICES: THE BOTTOM LINE – WHY USE RBLs?

Organizations should implement a multi-layered defense that uses RBLs as a first stage filter. This simple and low cost solution blocks significant volumes of spam (85-95%) before it is accepted by mail servers. This approach will dramatically reduce resource requirements, making mail servers more responsive because they are not using CPU cycles to process the large volume of spam that would have been accepted at the gateway. This approach will also reduce latency in overall message delivery.

How Spamhaus Helps Organizations Defeat Spam

Spamhaus is an international organization whose mission is to track the Internet's Spam Gangs, to provide dependable real-time anti-spam protection for Internet networks, to work with Law Enforcement Agencies to identify and pursue spammers worldwide, and to lobby governments for effective anti-spam legislation. Founded in 1998, Spamhaus is based in Geneva, Switzerland and London, UK and is run by a dedicated team of 25 investigators and forensics specialists located in 10 countries.

Spamhaus maintains three lists that contain the IP addresses of known or suspected spammers:

- **Spamhaus Block List (SBL)**
Contains IP addresses that are controlled by known spammers.
- **Exploits Block List (XBL)**
Contains IP addresses of virus-compromised computers that are sending spam.
- **Policy Block List (PBL)**
Contains IP addresses that should not be delivering unauthenticated SMTP email.

A TWO-STAGE PROCESS IS USED

Spamhaus uses a two-stage approach to eliminate the vast majority of spam:

- **Stage 1**
The first stage is to install the Spamhaus ZEN blocklist on incoming mail relay(s). ZEN, which is a combination of Spamhaus' SBL, XBL and PBL blocklists, will identify and reject 75% of a normal mail relay's incoming mail traffic. Stage 1 effectively rejects 80-85% of junk email, keeping this traffic off the network and freeing internal resources from having to process this content.
- **Stage 2**
Over 60% of spam contains the URLs of spammer Web sites whose Web server IP addresses are listed on the Spamhaus SBL. Therefore, the second stage is to scan the 25% of mail that makes it past first stage IP filtering, looking for URLs in the message body and testing their host IPs against the SBL. This is accomplished by installing an application capable of scanning message bodies for URLs and checking them against the SBL.

Why Spamhaus? Spamhaus is the #1 rated, most widely used blocklist..the most accurate and effective in the industry.

Testing of the dual-stage Spamhaus approach has shown that 99.6% of spam can be stopped with no false positives.

HOW EFFECTIVE IS SPAMHAUS?

Spamhaus lists the IP addresses of spammers' Web servers and DNS servers, in addition to spam sources in the SBL for this purpose. Spammers may find fresh sources not yet on Spamhaus' DNSBLs, but in most cases they need to advertise a Web site hosted somewhere.

Remaining spam, which should now be reduced to less than 7% of an organization's total incoming email traffic, is managed easily by a spam filter's other components, including SURBL, with the result that the total spam catch rate should now average 99.6%, or 299 in every 300 spam messages.

WHY SPAMHAUS?

Spamhaus is the #1 rated, most widely used blacklist. Independent rating services (stats.dnsbl.com) rate Spamhaus as the most accurate and most effective in identifying spam, and also generates, by far, the lowest false positives in the industry.

Spamhaus RBLs are used by many ISPs, corporations and universities and is considered to be a 'Best Practice' approach in fighting the costs and effects of spam.

SPAMHAUS USED TO BE FREE, BUT...

Spamhaus currently protects more than 1.2 billion user mailboxes worldwide, making it the largest single provider of spam-source data. Subscribers to the Spamhaus Datafeed Service receive a continuous feed from the three block lists. When an SMTP connection is made to one of its customers' servers, these block lists are checked before the email is accepted. If the sender is on the list, the SMTP connection is simply not accepted, allowing customers to block spammers before their content can enter the network.

Spamhaus is offered in two versions:

- No charge (for fewer than 100,000 queries per day)
- Fee-based (for more than 100,000 queries per day)

While Spamhaus used to be offered as a free service, it no longer is for larger customers, as noted above. Given that the mission of Spamhaus is to help the user community in the fight against spam, the change in licensing for use of Spamhaus' technology is, for all intents and purposes, a request to this community to help fund the Project.

Many organizations continue to use Spamhaus without a license, many likely not realizing that the service is extremely inexpensive, as shown in the following figure:

Spamhaus Annual Pricing

Users	ISP/ Enterprise
501 to 1,000	\$1,500
1,001 to 5,000	\$1,850
5,001 to 10,000	\$2,250
10,001 to 20,000	\$3,200
20,001 to 50,000	\$5,700
50,000 to 100,000	\$10,000
Unlimited	\$16,800

WHAT IS THE ROI?

Given the very low prices for Spamhaus and the significant benefits it offers, the return on investment for Spamhaus is extraordinarily high, as noted in the following examples.

REAL-WORLD EXAMPLES

What follows are examples of how Spamhaus has provided significant financial benefits for various types of organizations:

- **Large ISP**
 - Tier 1 ISP that manages 70 million domains
 - Spamhaus stops more than 80% of spam in Phase 1
 - Saves the ISP more than \$500,000 annually by reducing the number of filtering servers and other IT required
 - ROI in excess of 7,800%
- **Small ISP**
 - Processes five million emails per week, Spamhaus blocks 45 million spam emails per week
 - Using Spamhaus, the number of servers, storage and IT administration resources needed to manage email cut in half
 - **COST SAVINGS**
 - Eliminate two email servers at \$6,000 each: \$12,000
 - Reduced 550Gb of storage: \$1,000
 - Reduced IT admin time (0.1 FTE): \$4,000
 - TOTAL COST SAVINGS: \$17,000**
 - SPAMHAUS COST: \$3,200**
 - ROI: 531%**
- **Two universities in the Big Ten**
 - More than 100,000 users each
 - Spamhaus has generated just one false positive in two years
 - Used to operate 10 email servers and filter servers, but Spamhaus has allowed the number of email and filter servers to be reduced to six each
 - Administration and storage have been cut by 50%
 - **COST SAVINGS**
 - Eliminate four email servers at \$6,000 each: \$24,000

Reduced one terabyte of storage: \$2,000

Reduced IT admin time (0.2 FTE): \$8,000

TOTAL COST SAVINGS: \$34,000

SPAMHAUS COST: \$8,400

ROI: 405%

- Per the email infrastructure manager at a Big Ten university, “without Spamhaus we’re dead, would need twice the servers, our email experience long delays, and we would receive many user complaints”.

- **Mid-sized corporation**

- 40,000 email users on the Sendmail platform

- COST SAVINGS

Eliminate three email servers at \$6,000 each: \$18,000

Reduced four terabytes of storage: \$4,000

Reduced IT admin time (0.1 FTE): \$5,000

TOTAL COST SAVINGS: \$27,000

SPAMHAUS COST: \$5,700

ROI: 473%

- As one senior manager noted, “Prior to using Spamhaus, [we] had six email servers and 20-40 load averages on these servers. Since turning on Spamhaus, we’re down to three email servers with load averages of 3-4.”

- **Large corporation**

- 120,000 email users running Exchange

- Three-stage, layered messaging filters

- Using Spamhaus in the first stage as an initial perimeter filter allowed them to eliminate the second stage (Brightmail) filter entirely and save \$400,000 per year

- COST SAVINGS

Eliminate second stage filter (Brightmail): \$400,000

TOTAL COST SAVINGS: \$400,000

SPAMHAUS COST: \$16,800

ROI: 2,380%

- As one senior manager at this company noted, “Very simple check to eliminate [the] vast majority of junk email. Works great!”

About Spamhaus

ABOUT SPAMHAUS

The Spamhaus Project is an international non-profit organization whose mission is to track the Internet's Spam Gangs, to provide dependable real-time anti-spam protection for Internet networks, to work with Law Enforcement Agencies to identify and pursue spammers worldwide, and to lobby governments for effective anti-spam legislation. Founded in 1998, Spamhaus is based in Geneva, Switzerland and London, UK and is run by a dedicated team of 25 investigators and forensics specialists located in 10 countries.

© 2008 Osterman Research, Inc. All rights reserved.

No part of this document may be reproduced in any form by any means, nor may it be distributed without the permission of Osterman Research, Inc., nor may it be resold by any entity other than Osterman Research, Inc., without prior written authorization of Osterman Research, Inc.

All Product and brand names used in this document may be trademarks or registered trademarks of their respective owners. Any such trademarks or registered trademarks are the sole property of their respective owners.

THIS DOCUMENT IS PROVIDED "AS IS". ALL EXPRESS OR IMPLIED REPRESENTATIONS, CONDITIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OR FITNESS FOR A PARTICULAR PURPOSE, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE DETERMINED TO BE ILLEGAL.