

Controlling Spam

March 2006

Contents

What is Spam?	2
Why do Spammers Spam?	2
The Costs of Spam	3
How Do I Control Spam?	5
The MailMarshal Anti-Spam Solution	7
MailMarshal's Spam Detection Technology	8
Managing Spam	13
Conclusion	15

Spam is one of the issues facing email administrators today. The flood of unwanted and annoying email sent by spammers costs corporate organizations by lowering productivity and consuming IT resources. This paper outlines the spam problem, makes some suggestions about how to manage spam, and details the anti-spam technology available in MailMarshal. It contends that good email usage policies and leading filtering technology are the key elements in an overall anti-spam strategy.

WHITEPAPER - Controlling Spam

What is Spam?

Spam is unwanted advertising email invading your inbox. It promotes things like adult websites, amazing mortgage deals, and get-rich-quick schemes. Almost everyone with an email account has been “spammed” and most people agree it’s annoying and time-wasting to deal with. Spam is one of those hard to define but “you know it when you see it” phenomena. However there are some essential characteristics of all spam:

- It is unsolicited i.e. sent without the recipient’s permission
- It promotes products or services for sale.

Given this, spam is often referred to in a formal sense as unsolicited commercial email (UCE). In everyday use spam can also refer to any unwanted email such as chain letters from known senders or commercial email from retailers you have dealt with previously. In this paper we shall use the term spam to mean UCE.

Why do Spammers Spam?

Simple; there’s potential money in it.

Spammers send out their messages in the millions in the hope that a few people reply. Income is gained from actual products sold, or a percentage commission from products sold. Response rates and profit margins are typically low – but so are the costs. Tools and mailing lists for sending spam are cheap and easy to use.

The millions of messages spammers can send in a day can add up to significant revenues with even the most modest response rates. For example, with a \$1 per unit profit margin, and only a 0.1% response rate, a spammer could make \$10,000 by sending 10 million email messages.

All a spammer needs to send spam is the following:

- An email address list. These can be bought cheaply on a CD, or harvested (see below)
- Spamming software. This is inexpensive and easily obtained over the Internet
- An email server. To hide their identities, spammers often piggyback on top of an unsuspecting third party’s mail servers and relay spam through them. These servers are known as open relay servers. You can test whether your email server is an open relay by visiting

www.abuse.net/relay.html.

- A financial opportunity. There seems to be no shortage of opportunities for spammers. Common spam topics include product sales, memberships to adult websites and investment advice to name a few. It’s quite amazing what spammers will try to sell you; here is an example of an actual spam subject line: “Lose Unwanted Weight By Taking A Shower” – enough said.

WHITEPAPER - Controlling Spam

A spammer has several underhanded methods of collecting email addresses. Here are some of the most prevalent:

8 ways spammers can get your email address

1. From user registrations at unscrupulous sites
2. From user newsgroup postings
3. From user chat sessions
4. From spambots that crawl the web for any @ sign
5. From email lists the spammer buys
6. From mailing lists to which users subscribe
7. By randomly generating name combinations for your domain
8. By harvesting all the email addresses on your company's server

The Costs of Spam

So what is the big deal? As you will see below, spam is a real problem for organizations around the globe. The volume of spam is growing, it costs corporate organizations by lowering productivity and consuming IT resources, and it represents a potential legal liability.

Incidence of spam

Spam has rapidly emerged as a major issue for organizations and individuals alike, particularly in the United States. A recent Ferris Research survey indicated:

- 67% of email administrators said the increasing volume of spam is an overwhelming, major or notable problem
- 49% of end users said they received 10 or more spam messages per day.

Estimates of the incidence of spam vary depending on who you are. According to Ferris Research:

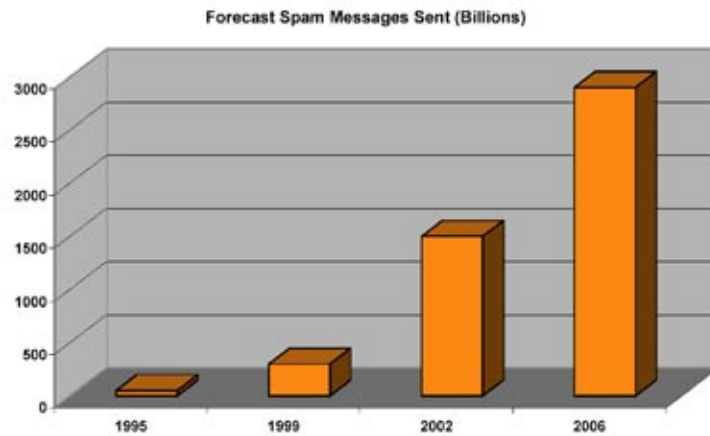
- United States based ISPs typically experience 30% of spam in their email traffic
- United States corporate organizations experience 15-20% spam
- Corporate organizations outside the US typically have a lower incidence of spam.

Gartner predicts that by 2004, unless an enterprise takes defensive action, more than 50% of corporate email message traffic will be spam.

One thing is also clear – no one expects spam volume to decrease in the near future. Research firm IDC forecasts the volume of spam will rise exponentially over the next five years:

WHITEPAPER - Controlling Spam

Forecast Spam Messages Sent (Billions)



Source: IDC; Email Usage Forecast 2002-2006

Figure 1. The volume of spam is set to double by 2006

Costs of spam

Spam is not only annoying – it costs. Ferris Research estimates that:

- Spam cost US corporate organizations \$8.9 billion in 2002
- The monthly per-user spam cost is \$10, based on a 10,000 user company
- Spam in Europe is less rampant, and the overall costs are lower at \$2.5 billion in 2002.

The costs of spam for enterprises can be broken down into three components:

- Loss of productivity. The costs associated with users dealing with spam messages. Ferris

Research estimates that each user spends 4.4 seconds dealing with each spam message.

- Consumption of IT resources. The costs incurred by IT as spam consumes bandwidth, storage space, and administrator's time.
- Help desk burden. The extra costs when annoyed users call the help desk to complain either about the volume of spam or its offensive nature.

WHITEPAPER - Controlling Spam

Legal liability

Spam also raises potential legal liability issues when it contains sexual or otherwise questionable content. This type of email is easily forwarded to people inside and outside the organization.

Email is a business tool. Anything sent from a corporate email address is effectively written on electronic company letterhead. As a result, any views, quotes, or discussions made via company email can be representative of the company and legally binding.

The casual use of profanity in business email (as in any other documented communication) has obvious implications for a business's reputation. Such emails have had more concrete repercussions as well. There have been several lawsuits involving sexual harassment in the workplace, based on lewd comments sent by email. In many cases the organization has been held responsible for not controlling their email content so as to avoid offensive exposure to employees.

How Do I Control Spam?

We believe there are two essential elements to controlling spam:

- Employ an appropriate corporate email usage policy and educate end users
- Deploy appropriate anti-spam technology

Corporate email acceptable use policy and user training

This is the first important step for any enterprise. If you have a corporate email system it is highly recommended that you also have an email acceptable use policy. A good email policy is essential for making employees aware of corporate rules and standards, giving guidelines on how to handle certain types of email and situations, and to inform employees that their email may be monitored.

While most organizations have email policies, many policies tend to lack detail on how employees can deal with inappropriate email. Regarding spam control, a good email use policy and education program would include at least the following:

- Newsletters and online sites. The conditions under which employees can sign up for newsletters, forums, newsgroups and chat rooms – these should be business related and within company guidelines.
- How to deal with spam and other inappropriate email. A good idea is to have an internal email address where spam or other inappropriate email can be forwarded and monitored by email administrators.
- How to avoid being listed. Careful use of corporate email addresses is important. Another of the basic rules is not to reply to spam asking to be removed from the list – this will only confirm a valid email address to a spammer.

Once you have an email usage policy, the next key step is to ensure all employees are educated in its contents and are regularly trained in correct email usage and handling.

WHITEPAPER - Controlling Spam

Remember, the more company email addresses are made available in public areas the more likely they will be picked up and used by spammers. Corporate email addresses should be used cautiously on the Internet. Company policy and user education is the key!

For further general information about email usage policies and how to design one, try visiting

www.email-policy.com.

Deploying anti-spam technology

The other essential element in controlling spam is the use of anti-spam technology. There are many different technical approaches to blocking spam; message content analysis, and blacklists of known spam and spammers are two examples. Many analysts agree that a combination of approaches is now essential to effectively combat spam.

“The most promising approach to reducing spam is the use of anti-spam technology. A combination of several technical approaches is required for effective blocking, with rules-based or heuristic approaches the most important.” - Ferris Research

One of the best ways to control spam in an enterprise is to employ a content filtering device between your firewall and your email server. Protecting your email system at the gateway (the point where email enters your organization) will block spam prior to it ever entering your email servers. The advantages are numerous:

- You can block spam before it reaches your email servers and therefore your users
- Management and software updates can be controlled centrally by an email administrator
- There is no need to deploy desktop software or train users on how to use it.

MailMarshal SMTP is a content filtering gateway which provides the means by which organizations can control spam. It is installed between your firewall and your email server.



Figure 2. Installing MailMarshal between your firewall and your mail server

The rest of this white paper outlines MailMarshal's anti-spam solution in more detail.

WHITEPAPER - Controlling Spam

The MailMarshal Anti-Spam Solution

Marshal has developed a comprehensive range of solutions for controlling content in and out of networks for Email and Web browsing.

MailMarshal provides a comprehensive tool to control spam based on an extensive array of functionality. The two key concepts are detection and management. MailMarshal uses technologies that enable high spam detection rates and few false positives, with exceptionally easy administration. And it does this within the context of an integrated email content management package. MailMarshal is more than an anti-spam system – it provides organizations the means to control all email content, including spam, viruses, text, and attachments within a rules-based framework. MailMarshal offers organizations an effective and flexible means to control spam, with a very rapid return on investment.

Multi-layered spam detection and management

Spam is constantly evolving as spammers employ ever more sophisticated techniques to get their message across (so to speak). No single piece of technology is likely to stop all spam. It is important therefore for enterprises to adopt a multi-faceted solution to the spam problem. MailMarshal utilizes an extensive array of anti-spam technologies:

MailMarshal's arsenal of anti-spam technologies

1. SpamCensor – MailMarshal's unique spam detection technology
2. TextCensor – advanced text analysis
3. Message header analysis
4. Email host blacklists
5. Domain or user blacklists
6. Domain or user whitelists
7. Anti-relay
8. Anti-spoofing of local email
9. DNS Blacklists, e.g. MAPS RBL
10. Reverse DNS lookups
11. Reporting

Importantly, once spam is detected, MailMarshal provides a range of management options. As a gateway solution, email can be quarantined, forwarded to administrators, or marked as possible 'spam'. All quarantined email can be managed centrally from a console. MailMarshal's standard functions also provide the ability for end users to be able to review and manage spam addressed to them via simple inbox rules.



WHITEPAPER - Controlling Spam

MailMarshal's anti-spam detection and management is illustrated below.

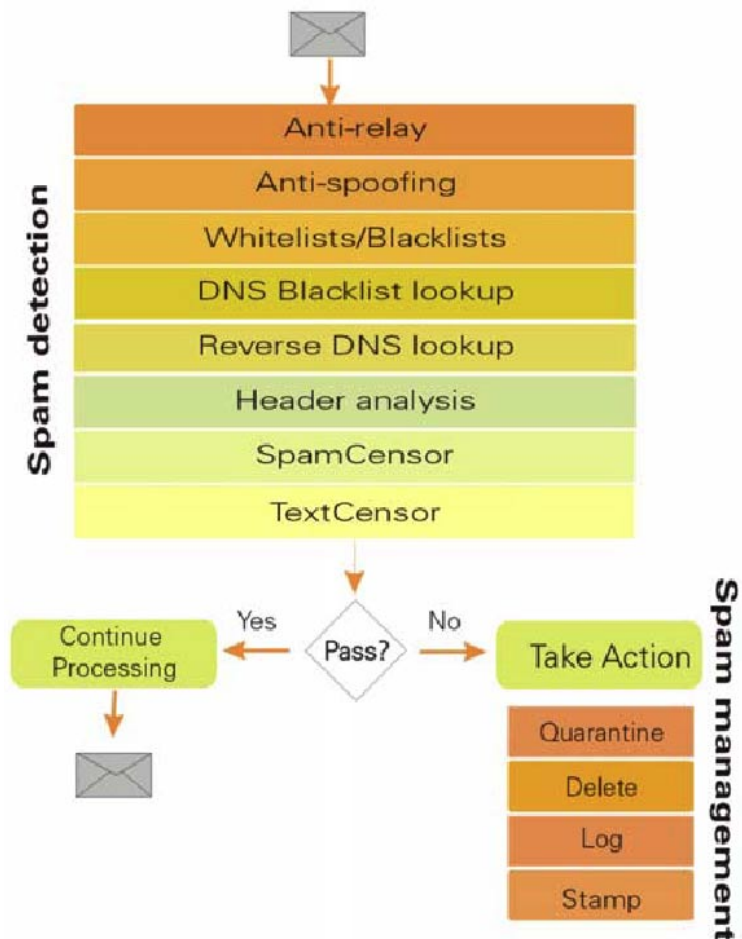


Figure 3. MailMarshal has a full range of spam detection and management capabilities

MailMarshal's Spam Detection Technology

The SpamCensor – MailMarshal's unique spam filter

New in MailMarshal 5.5 is the SpamCensor - a unique spam detection system that draws upon a range of analysis techniques to determine whether an email message is spam. The SpamCensor is an advanced filter that utilizes the following:

- Detailed header analysis. This technique closely examines email message headers for the telltale footprints of spammers. The SpamCensor looks for hundreds of typical indicators of spam, including such things as missing To: or From: addresses, invalid dates, the presence of numbers in addresses, or spaces in weird places.

WHITEPAPER - Controlling Spam

- Advanced analysis of message content. The SpamCensor filter performs advanced searches of message content. It searches for thousands of preset spam patterns such as “get rich quick”, multiple “\$” characters in a row, or g a p p y text.
- DNS lookups of blacklists. The SpamCensor can also query DNS Blacklists to determine whether the sender is blacklisted as a spammer. (See below for more information on DNS Blacklists).
- Message composition. The message size and composition is checked. Spam is typically small, and often only has an HTML part - this information is used alongside other indicators.

As The SpamCensor runs, the results from each of the thousands of tests are used to build an overall spam ‘picture’. A ‘score’ is determined of whether the message is likely to be spam. Once the score exceeds a certain trigger level, MailMarshal will treat the message as spam and take appropriate action. This approach results in high spam detection rates with very few false positives.



Figure 4. The SpamCensor combines multiple tests to arrive at a spam “score”

One of the best things about the SpamCensor is its ease of use. All the administrator needs to do is to turn it on. The SpamCensor filter has been optimized by anti-spam experts at Marshal, who have trained it using thousands of actual spam messages.

Spam is dynamic. Thousands of new spam messages are created daily, and spammers are constantly devising techniques to avoid spam filtering. As such, Marshal have engineered the SpamCensor to be extensible and flexible. The Marshal anti-spam team is constantly adjusting, and improving the SpamCensor’s filters. To remain up-to-date, MailMarshal performs automatic downloads of new SpamCensor configuration files.

The TextCensor - advanced lexical analysis

MailMarshal’s TextCensor technology is a powerful anti-spam device in its own right. The TextCensor provides a way for administrators to search email messages for words and phrases associated with spam.

The TextCensor engine has advanced search capability, including:

- Simple words or phrases. For searches of individual words or simple phrases e.g. “yours free!”
- Boolean operators (AND, OR, NOT) e.g. “Viagra AND online”
- Proximity operators NEAR, FOLLOWEDBY, INSTANCES e.g. “credit NEAR free”
- Weightings. Gives a weighting to each TextCensor phrase e.g. “FREE!” might be given a higher weighting than “buy now”

WHITEPAPER - Controlling Spam

- Limit to different parts of message. The scripts can be limited to searches the header, body, subject lines, or attachments. This is important e.g. searching subject lines for spam only can be very fruitful.

These features are combined in TextCensor scripts, which when deployed, can provide targeted and highly accurate searches of message content, with a minimum of false positives. The other main advantage of the TextCensor is its fast – much more efficient than regular expressions.

TextCensor scripts can be used in numerous ways to stop spam – they provide you with the flexibility to look for almost anything. Here are some examples:

- Specific incidents of spam. Identify specific or problem spam by adding a unique identifying line to the script. This is most useful for repetitive spam such as chain mail or fraudulent offers (for example Nigerian scams)
- Generic spam. The unvarying thing about spam is that it is always trying to sell you something! And as such, it tends to have unique characteristics that can be detected by a good script. MailMarshal's pre-supplied generic spam scripts alone have been effective at blocking over 60% of spam.
- Linked images. Another useful indicator of spam is the presence of linked images. Normal email messages usually carry images as attachments to the message. But spam frequently includes the image as an external link - the email client needs to download the image from the web. Sometimes this just confirms to the spammer that you have received the message. A more recent trend is to include the entire content of the message (including text) as an image. This negates normal lexical scanning. To identify linked images, a TextCensor script can be used to detect both the presence of an external link to an image and the absence of normal English.
- Foreign Language Character Sets. Sometimes spam arrives in a foreign language, which for organizations that do not receive any foreign email, is easy to deal with. Scripts can be created to spot foreign language character sets in the body of the message.

Message header analysis

MailMarshal has rules-based header analysis, allowing email message headers to be closely scrutinized. The technology uses a powerful regular expression engine to perform its searches.

Spammers use a number of tricks to mask their identities, or to fool you into thinking they've sent a valid email. These tricks often leave telltale footprints in the header. MailMarshal's header analysis technology can therefore be usefully employed for detecting spam. Subject lines are excellent places to perform searches because spammers want to get your attention; for example:

- A subject followed by lots of white space and a number
- A subject word interspaced with asterisks e.g. subject: b*a*d*w*o*r*d

WHITEPAPER - Controlling Spam

- A subject word with recurring spaces e.g. subject: b a d w o r d

MailMarshal can also rewrite header lines, including the subject line. This can be useful in managing detected spam with either special rules in MailMarshal or the end email client.

Email host blacklists

A host blocking blacklist is simply a list of email hosts or message sources from which you do not want to receive any email. This list can include email server host names, IP addresses, or IP address ranges. It is maintained locally by the email administrator, usually after investigation of a spam incident. It is an essential component for controlling specific problem mail hosts.

Domain or user blacklists

In addition to blocking hosts, MailMarshal also provides the ability to block domains such as baddomain.com or individual users such as harry.spammer@nuisance.com. This can be useful for blocking spam coming from a consistent domain, or simply people you would like to add to your deny blacklist. MailMarshal allows you to employ blacklists at two levels:

- In Receiver rules, where the message can be denied prior to being downloaded
- In Standard rules, where the message has been received, and action, such as quarantining the message to a folder, can be taken.

Whitelists

Often it is desirable for email from certain domains, or to or from certain users, to be always considered "OK" – i.e. not to have any anti-spam checks applied. In MailMarshal, administrators can easily create such whitelists through the standard rule wizard interface where domains, groups or individuals can be excluded from anti-spam rules.

Anti-relay

Relaying is the passing of messages to another server for delivery. An open relay server is an email server configured to accept email from any user to any destination address. Open relays allow anyone (including bulk and spam senders) to use the name and resources of that server. Best practices require relaying to be tightly controlled.

MailMarshal allows administrators to control relaying; it has full anti-relay functionality, including:

- Relay control. By default, MailMarshal prohibits relaying from external sources. However it also has the flexibility to allow you to specify exceptions for trusted email sources.
- Known relay exploit detection. MailMarshal can detect known relay exploits where Recipient fields are modified in an attempt to fool email servers into relaying the message.



WHITEPAPER - Controlling Spam

Anti-spoofing

In general, email spoofing refers to the forgery of an email header so that the message appears to have originated from someone or somewhere other than the actual source. Spammers often use spoofing in an attempt to get recipients to open, and possibly even respond to, their solicitations. MailMarshal has a specific anti-spoofing rule condition that seeks to identify messages that purport to be from a local domain but may in fact have been spoofed – i.e. they may not have originated from within the domain of a local From: address.

External DNS blacklist databases (e.g. MAPS RBL)

An established way of blocking spam is to check the sender against a blacklist of known spammers. There are a quite a number of blacklists available on the internet (over 100 at last count) of varying quality and availability. These lists are usually maintained by non-profit organizations, although some charge for certain services. One of the best known pay-sites is the Mail Abuse Prevention System Realtime Blackhole List (MAPS RBL). As the services use DNS as the method of querying their servers, they are also often referred to as DNS blacklists. Some of the well known services are:

Common DNS Blacklists

MAPS <http://www.mail-abuse.org>

SPAMCOP <http://spamcop.net>

ORBD <http://www.orbd.org>

DORKSLAYERS <http://www.dorkslayers.com>

SPEWS <http://www.spews.org>

OSIRUSOFT <http://relays.osirusoft.com>

SPAMHAUS <http://www.spamhaus.org>

The effectiveness of this method of blocking spam is entirely dependent on the quality of your chosen list, and how often it is updated. The services sometimes attract criticism because occasionally legitimate email hosts can find themselves unwittingly on the list, and it becomes difficult to send email to them! Even so, DNS blacklists are good at blocking spam and should form part of an overall anti-spam strategy.

MailMarshal has integrated DNS blacklist support, and ships with some already configured.

Administrators can apply one or more lists, either globally or in a rules-based policy.



WHITEPAPER - Controlling Spam

Reverse DNS lookups

A reverse DNS lookup is used to resolve a message sender's IP address to a valid host name. It is called reverse DNS because a normal DNS lookup is used the opposite way – i.e. to resolve a host name to an IP address.

Performing a reverse DNS check on a sender's IP address is a traditional method of trying to combat spam. Properly configured DNS should include reverse DNS (or PTR) records for email servers, and spam has tended to originate from, or be relayed through, servers without reverse DNS. Having said that, reverse DNS checks aren't as effective today as they have been in the past. There are a number of reasons for this including:

- Many legitimate email servers are not configured properly with reverse DNS records through sheer oversight or ineptitude
- The technically proficient "professional" spammers know about the reverse DNS trick and configure their servers correctly.

Therefore, reverse DNS should be used with caution, as a component, not a complete solution, for fighting spam. MailMarshal includes reverse DNS lookup functionality, and has a logging-only option so that administrators see in advance what effect it may have on their email flow.

Reporting

MailMarshal features extensive logging and reporting options. Whenever a spam rule is triggered, a custom "spam" entry can be written to a database, enabling a range of reports to be run. Using the reports, administrators can analyze mail traffic for spam. MailMarshal can answer questions such as: how much spam do we receive?; what proportion of spam to normal email do we get?; which email addresses in our organization attract the most spam?; and where does the majority of our spam come from? Armed with this information, administrators can form better strategies to combat spam, including adjusting MailMarshal to block specific spam hosts and or incidents of spam.

Managing Spam

It's great to have all this anti-spam technology, but the real power in MailMarshal comes from its flexibility. It's much more than anti-spam – it is a total email content filtering system.

Putting it together – a rules based approach

Flexibility is one of MailMarshal's strengths. You can harness the power of MailMarshal's preconfigured spam technology, such as the SpamCensor and get great results. However, MailMarshal also provides the ability to customize rules for every site. Administrators can combine rule elements together to create a policy that is greater than the sum of its parts. For example, spam messages are often not very large – often less than 50KB in size.



WHITEPAPER - Controlling Spam

This fact can be used in conjunction with TextCensor scripts and whitelists to create an accurate rule, as in the following example from the MailMarshal rule wizard:

When a message arrives

Where message is incoming

Except where addressed either to or from 'Excluded Users'

Except where addressed from 'Friendly ListServers'

Where message size is less than '50 KB'

And where message triggers text censor script(s) 'Spam'

Move the message to 'Junk'

Easy yet powerful spam management

Once a message has been determined as spam, administrators must decide what action to take with it. Gateway actions enable messages to be deleted or quarantined - this prevents spam from flooding your internal network, many organizations find that they can reduce mail volume by 40% or more by controlling spam at the gateway. MailMarshal provides a wide array of possible actions for the maximum flexibility.

- Moving the message to a quarantine folder
- Copying the message to a folder
- Blind copying to an addressee
- Running an external application
- Sending a notification message
- Stripping any attachments
- Writing a custom log message to the database for later reporting
- Adding a message stamp to the message
- Rewriting the message headers
- Parking the message for later delivery
- Routing the message to another host
- Passing the message to another rule for processing
- Deleting the message



WHITEPAPER - Controlling Spam

Spam can be passed to different quarantine folders that can be created for users, groups, or domains. Messages can be forwarded to administrators for review and additional actions such as, forward, pass-through, delete, copy, or re-route.

Messages can be monitored and controlled using the Management Console. Administrators can have any number of management consoles and they can be configured to control only specified used groups if desired.

Enabling end users to manage spam

MailMarshal can also be configured to allow end users to have extra control over 'their' spam, thereby saving administrative overhead. There are a number of ways this can be achieved, including:

- **Marking spam.** Identified spam can be marked so that it can be forwarded through to the end user where their email client can be configured to automatically move the marked spam message into a specific spam folder for review by the end user if required.
- **Sending Notifications.** Spam can be quarantined as normal, but a small notification containing such information as when sent, who from, and subject line can be sent to the end user. This enables the end user to quickly review the nature of the message. The notification message itself can be diverted into a specified spam review folder via a rule in the email client. If the end user wants the message then they can trigger its release by simply replying to the notification. This invokes MailMarshal's message release function – automatically releasing the message from the server. Due to their small size, notifications can save bandwidth on the network and also prevent the offensive nature of some spam messages residing on client machines.

Conclusion

Spam is a real issue for email administrators, and all predictions point towards it getting worse over the next few years. This white paper contends that the spam problem must be dealt with on several fronts, with a good email acceptable use policy, user education and content filtering technology as part of an overall anti-spam strategy. MailMarshal provides email administrators with the technology for controlling spam. It combines leading new technology with traditional anti-spam approaches. Above all it provides anti-spam capability in a highly flexible and proven email content filtering product. MailMarshal has:

- The range and depth of technology for maximum accuracy in spam detection
- A rules-based approach for maximum usability
- A range of management options for maximum flexibility in the enterprise.

Marshal is committed to providing the best possible anti-spam solutions. Our research & development team is working continuously to improve MailMarshal - adding greater accuracy of detection and more intuitive and flexible management functionality. For more information on MailMarshal please contact your Marshal reseller or sales representative, or visit www.marshall.com.





Marshal's Worldwide and EMEA HQ
Marshal Limited,
Renaissance 2200,
Basing View,
Basingstoke,
Hampshire RG21 4EQ
United Kingdom

Phone: +44 (0) 1256 848080
Fax: +44 (0) 1256 848060

Email: emea.sales@marshal.com

Americas
Marshal, Inc.
5909 Peachtree-Dunwoody Rd
Suite 770
Atlanta
GA 30328
USA

Phone: +1 404-564-5800
Fax: +1 404-564-5801

Email: americas.sales@marshal.com
info@marshal.com | www.marshal.com

Asia-Pacific
Marshal Software (NZ) Ltd
Suite 1, Level 1, Building C
Millennium Centre
600 Great South Road
Greenlane, Auckland
New Zealand

Phone: +64 9 984 5700
Fax: +64 9 984 5720

Email: apac.sales@marshal.com