



WHITE PAPER



## Demystifying the Anti-Spam Buzz:

*Features vs. Fluff in the Search for an Enterprise  
Anti-Spam Solution*

*March 2007*

### Summary

The options available to mitigate the problems of e-mail spam in the enterprise are varied. Just as varied is the variety of definitions of spam today. These factors often lead to some confusion about what a particular anti-spam (or anti-fraud or anti-virus) solution may or may not do.

This paper demystifies the anti-spam market and its various choices and buzzwords to help you cut through the hype and focus on the basics: determining what features you need, whether a solution you are considering includes them, and to what degree. It also defines the important key elements that an anti-spam solution should have to ensure that your investment will be a solid one for years to come.

## Spam is Spam

Ads, fraud schemes, urgent requests for help, porn, virus attachments, phishing attempts ... these are all examples of e-mail "spam" (junk or threatening e-mail) that affect organizations and end-users today. The market for solutions to stop this spam can seem overwhelming, and common pricing strategies imply that there is a difference between types of spam, or that different solutions are necessary to combat different kinds of junk e-mail.

But these distinctions are misleading. The truth is, the techniques spammers use are common across all varieties of spam. For example, spammers of all types frequently use virus and worm-infected machines to send spam. Likewise, spammers seeking to steal money or identities (phishers) use the same techniques as spammers seeking to sell the latest wonder drug.

**We treat all spam equally. If it's junk entering through your mail servers, Roaring Penguin's CanIt anti-spam solutions will stop it.**

## Spam-Fighter's Cocktail

Three primary types of spam fighting solutions are available today:

- **Client-side solutions** reside on individual users' machines. Distributed support requirements and license fees make them unsustainable in large environments, they lock an enterprise into a particular client operating system, and mobile and webmail users may not be able to access them.
- **Outsourced (third-party) solutions** require an enterprise's e-mail to be routed through third-party servers for filtering. Their primary drawback is the loss of control over an enterprise's e-mail and its filtering techniques.
- **Server-based solutions** come in the form of software and/or server appliances, and act as a gateway between the Internet and the e-mail server. This paper deals with server-based anti-spam solutions.

**This paper focuses on server-based solutions.**

### ***Scalable, Server-Side Architecture***

Server-based anti-spam solutions are preferable for enterprise environments because they can fully protect your network and users from spam and viruses with minimum administrative burden. But, with spam and virus attacks increasing mail volumes exponentially, server-based solutions must be scalable. Look for solutions with streamlined architectures, technology and cost flexibility.

For example, ensure that you can configure the solution to filter out viruses before spending resources filtering for spam. Also ensure that the solution can be set to filter particular types and sizes of e-mail before performing content-based filtering. These techniques contribute to a streamlined e-mail filtering architecture. Finally, look for a solution that is priced per-user and not per-server. Per-server pricing puts you at a disadvantage because as mail volumes grow, you have to spend more to keep the solution effective by scaling it across more servers.

**Scalable cost: Look for a solution that is priced per-user, not per-server.**

## ***Integration with Existing Environment***

You likely already have various systems in place related to e-mail. Adding spam-filtering should not require you to change your existing systems. Look for a solution that will work within your existing mail environment (i.e. mail server, O/S, mail transfer agent, etc.). Also, be sure to ask whether the solution will integrate with your existing user authentication infrastructure (e.g. LDAP, custom databases, Sendmail database, etc.) and anti-virus software.

**Integration:**  
With mail servers, authentication directory and anti-virus software.

## ***Equal-Opportunity Pricing***

The anti-spam solution you select should not require that you pay more for filtering different kinds of spam. Spam is spam, and spammers should never be allowed to dictate what you spend on a spam-fighting solution. The best approach in evaluating and selecting a solution is to begin by establishing the capabilities and features that you really need, then find a comprehensive solution that provides these.

**One-price filtering:**  
Look for a comprehensive solution that filters all types of spam.

## ***Never Lose a Valid E-Mail***

An anti-spam solution should never put you at risk of losing legitimate e-mail messages. This means that you need a solution with a quarantine function that holds suspect messages in a safe area for inspection by a human who can make the final judgment.

**Quarantine:**  
Hold suspect messages for final, human judgement.

Look for a solution with this flexibility, but be sure that your solution of choice enables you to customize spam thresholds and spam management techniques. For example, you may want to automatically quarantine or delete certain types of e-mails while only tagging messages that receive mid-range spam scores. Your solution should, at a minimum, have that level of flexibility.

## ***Centralized Solution with End-User Controls***

A solution with centralized administration and distributed end-user controls is ideal to provide the administrative benefits – such as global rules, default settings and centralized administration – while at the same time allowing you to please everyone.

**Global rules:**  
Maintain control by pushing global rule sets out to end-users.

By pushing global rules out to end users and enabling them to participate in choosing the filtering solution that best suits them, the administrative burden and complexity of defining spam in a multi-user environment is mitigated.

If you want a solution with some level of end-user control, look for one that allows you to customize the end-user interface, including its look-and-feel and the controls you provide to end-users. Only you know what your particular users and environment need and can handle. A customizable interface will help minimize disruption to end-users and will eliminate unnecessary complexity.

Also, define up-front what you mean by “end-user control”, because available solutions provide many different levels of user control. At minimum you should have the ability to give your end-users a choice among spam-filtering levels—and those levels should be determined by administrators, not by the product. You should also have the option to provide users with more advanced features beyond merely whitelisting and blacklisting.

**Customizable interface:**  
Give end-users exactly what they need while maintaining global policies.

The ability for users to set spam thresholds and obtain details about individual spam incidents may be important for certain users in your environment. If so, be sure that the solution you choose gives you the power to provide such options. Keep in mind that even if you don't start with a solution that allows for end-user controls, you may need the option in the future. For example, central administration of an organization's spam may become too great for a single IT manager or department to handle. Your organization's policies about e-mail may change, or end-users may become vocal about having greater control over their e-mail.

### **Web Interface**

Administrative and end-user interfaces of an anti-spam solution should not be dependent on any particular mail client or operating system. Look for a solution with web-based interfaces that enables administration of the tool regardless of mail client, operating system or user location. The tool's full features—not just a subset of them—should be available even to users and administrators located off-site or traveling.

**Web interface:**  
Full features available from any mail client or web browser.

### **Source Code Provided**

How you filter your organization's e-mail is your business. With the solution's source code in hand, you can implement filtering rules and methods, no matter how complex, arcane or unique they may be. Owning the source code also enables you to perform complete system security audits and make adjustments if necessary, and is the ultimate long-term insurance for your investment.

**Source code:**  
For complete control and security audits.

### **Multi-Factor Spam Analysis**

Fighting spam is an arms race. As spamware becomes more sophisticated and spamming becomes more profitable, no single tactic is sufficient to do battle with increasingly aggressive and tricky spammers. While blacklisting offending DNS domains was once sufficient to battle spam, blacklisting is today just one small aspect of effective spam management. Commercial spammers, phishers, perpetrators of e-mail frauds and virus writers all use a number of methods to obfuscate their work. Spammers devote their time to inventing new techniques—both technological and sociological—to trick end-users into responding. Such tactics include, for example, masquerading as a trusted brand by spoofing a well-known e-mail domain. Any effective anti-spam solution in today's environment must employ a variety of integrated fighting techniques to be effective—and you should not be expected to pay more for different filters. If a solution claims to stop spam, it should be expected to stop all types of spam: even those not yet invented.

**Multi-factor analysis:**  
A cocktail of spam-fighting tactics to future-proof your investment.

In the "Buzzword Compliance" section beginning below, we've detailed the requirements for a future-proof solution, as well as others that you should understand and consider – or avoid.

## Buzzword Compliance

Today's leading anti-spam solutions all lay claim to various features and capabilities that have interesting names ... but what do these features really do? How do they do it? And do you really need them?

The table on the following pages catalogs these anti-spam buzzwords, simply explains what each means, and describes in practical terms why you may (or may not) need each. It also indicates which features Roaring Penguin's CanIt solutions offer, and to what degree.

Use this table to help you sift through the "fluff" and truly understand the features that a solution offers, what they mean, and the value each will have in your particular environment.

**Buzzword IQ:**  
Even before learning what the buzzwords mean, have a sense of the main features and qualities that you need in an anti-spam solution.

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Adaptive spam analysis	Also known as "statistical" or "Bayesian" filtering, this is a type of spam filtering that learns what is or is not spam over time, based on feedback or inputs from administrators and/or users. By cataloging words and/or phrases that appear in messages rejected as spam or accepted as legitimate, adaptive filters become increasingly adept at determining the legitimacy of new messages based on the occurrence of those words and phrases.	To protect your anti-spam investment. Adaptive filters ensure that the anti-spam solution remains effective over time, even as spammers' tactics change.	<b>Yes.</b> And, in CanIt-PRO, the Bayesian filters operate on a per-user basis to learn what each individual user considers to be spam.
Advanced administrative tools	Administrator tools that let you script or automate administrative tasks.	To ease the management of multiple domains or a complex e-mail infrastructure.	<b>Yes.</b> CanIt-PRO includes command line tools for scripting purposes.
Anti-fraud	Also referred to as "anti-phishing". Phishing is the use of e-mail to spoof a well-known brand (such as a bank) in order to fool recipients into divulging personal data, such as credit card numbers, passwords and the like. Phishing e-mails are usually linked to a fraudulent web site. Anti-phishing refers to detecting and averting this type of e-mail spam.	If you want to protect your end-users from e-mail fraud schemes. It is estimated that at least 5% of all victims of phishing attacks do divulge their personal information. Filtering out phishing attempts also reduces the volume of junk e-mail in your organization.	<b>Yes.</b> CanIt's filters detect the techniques used by spam phishers. (See "mismatch rules" below as an example.) CanIt's content filters can also be easily set to detect known phishing attempts or to tag e-mails referencing commonly spoofed brands, as probably fraud attempts.
Anti-phishing	See "anti-fraud," above.		<b>Yes.</b>
Automatic Updates	Your anti-spam solution should receive automatic updates to stay on top of changing spam methods	As spammers evolve their techniques, your solution must be updated or it will fall behind and its accuracy will decrease.	<b>Yes.</b> The Roaring Penguin Training Network updates your appliance daily with Bayesian data, ensuring it keeps on top of the latest spam trends.
Bayesian filtering	See "Adaptive spam filtering" above.		<b>Yes.</b>
Blacklisting	A filtering technique whereby a list of disallowed senders or domains is created. E-mail from a sender or domain on a blacklist is always filtered as spam.	To filter out certain spammers and spamming domains. However, the increasing use of "zombie" spam machines (see below) makes blacklisting alone an insufficient solution.	<b>Yes.</b> CanIt solutions support blacklisting, and can accept realtime blacklist (RBL) directories. We recommend only careful and limited use of blacklisting.

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Content compliance	Refers to standard content filtering, but is performed on internal and/or outgoing mail.	To ensure that e-mail circulating within your domain and leaving your domain is compliant with organizational mail policies.	<b>Yes.</b> CanIt solutions can optionally filter internal and/or outbound e-mail.
Customizable interface	A feature of an anti-spam solution that enables administrators to determine what features of the product will be available via the administrative and/or end-user interfaces.	To minimize the training and support the solution requires, by customizing it to suit your particular users.	<b>Yes.</b> Administrators have full control over CanIt's interfaces. With CanIt-PRO, end-users see only those features that administrators allow them to see.
Customizable filters	A feature of an anti-spam solution that enables administrators and/or end users to alter the way e-mail is filtered. This may include building whitelists and blacklists, changing "spam thresholds" (see below), changing how file extensions and attachments are handled (see below), architecting the solution to filter for particular types of content, and more.	To tweak default filters, turn certain filters off, or upload your own existing rules.	<b>Yes.</b> CanIt's filters are completely customizable, and you can upload existing rule sets to CanIt's databases.
File extension and attachment handling	A feature that allows administrators to determine how spam filters will handle e-mail messages that contain particular types of attachments.	Because most e-mail borne viruses are distributed by way of particular file extensions (e.g. .exe, .zip, .pif), you will want to detect and control these. Controlling if and when particular types of attachments are allowed through your domain can also help with load (e.g. large files such as .mpgs).	<b>Yes.</b> CanIt's administrative interface includes full file extension and attachment handling.
Global rule sets	Rules within an anti-spam solution that all mail is subject to. Global rule sets override individual user customizations.	To enforce enterprise-wide e-mail policies and to ensure the anti-spam solution works within your environment.	<b>Yes.</b> CanIt administrators create global rules. CanIt-PRO end-users inherit global rules in addition to being able to make certain customizations.
Greylisting	Also known as "hit-and-run detection," greylisting causes a temporary delivery failure for mail from an unknown sender. Legitimate mail servers will queue and retry the delivery, whereas some spamware does not. Greylisting accepts the mail for filtering if/when it is retried. The process is transparent to both sender and recipient – neither one receives a failure message.	To eliminate a growing percentage of spam before the e-mail is even filtered. (At publication date, greylisting catches about 30% of all spam). It therefore reduces burden on your e-mail servers, databases and filters – and reduces the size of spam quarantines. Greylisting is also an effective means of detecting "zombie" spam attacks (see below).	<b>Yes.</b> Roaring Penguin invented greylisting.

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Hit-and-run detection	Also known as "greylisting" (see above).		<b>Yes.</b>
Hold-and-release	A feature by which the spam filter can be set to release suspected spam after a designated period of time.	To eliminate the possibility of Spam Trap contents growing large enough to cause storage problems. It can also be effective in training users to manage their Spam Traps regularly.	<b>Optional.</b> CanIt can automatically manage spam after a certain period of time in the Spam Trap.
Mismatch rules	Spam filter rules that check whether an e-mail claiming to be sent from a particular domain was indeed sent from that domain.	To detect a large amount of commercial spam and phishing attacks, which tend to spoof certain webmail domains such as eBay, hotmail or yahoo.	<b>Yes.</b> CanIt enables administrators to command the filters to ensure an IP relay match on any message sent from a particular domain.
Notification	A feature that notifies administrators or end-users when there is spam in the Spam Trap.	If you want a solution that gives each end-user his or her own Spam Trap, notification is an effective way to train or remind users to regularly manage their Spam Traps.	<b>Optional.</b> With CanIt-PRO, administrators can allow users to sign up for notification.
Outbound / internal mail filtering	The filtering of e-mail that exits your domain (i.e. e-mail sent by clients on your network). Internal mail filtering filters e-mail sent between your users. Outbound filtering filters e-mail sent from your users to locations outside your own domain(s).	Outbound mail filtering is not for everyone, but as the instances of "zombie" (see below) spam attacks increases, the need for it is growing. Outbound mail filtering is typically used in the following situations: <ul style="list-style-type: none"> <li>• Workplaces in highly regulated industries that need to protect their domain reputations.</li> <li>• Workplaces that need to ensure compliance to their internal e-mail policies.</li> <li>• Environments in which mail clients are commonly targeted and turned into zombie spam machines.</li> <li>• Any organization or enterprise whose domain has been blacklisted due to the amount of spam and/or viruses received from that domain.</li> </ul>	<b>Optional.</b> CanIt-PRO can be configured to filter incoming, internal and outbound e-mail. CanIt-PRO can filter one, two or all of these types of e-mail.
Policy compliance	See "outbound/internal mail filtering" above.		<b>Yes.</b>

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Quarantine	The ability of an anti-spam solution to hold suspected spam in a protected server area for further analysis or human intervention.	<p>Quarantines serve valuable purposes:</p> <ul style="list-style-type: none"> <li>* Reduce administrative burden by preventing the need for administrators to search through back-ups or archives to locate legitimate e-mails that may have been inadvertently deleted by generic spam filters.</li> <li>* Give end-users the opportunity to review/manage quarantined e-mail addressed to them.</li> <li>* Allow you to learn about the nature of spam at your organization and create custom filtering rules that meet your environment's particular needs.</li> </ul>	<p><b>Optional.</b> CanIt solutions enable administrators to handle suspected spam in any number of ways, including by quarantining them to the Spam Trap.</p> <p>By default, only a small portion of each quarantined message is kept on your server, thereby minimizing storage requirements.</p>
Reporting	A feature that enables spam activity and/or details of spam incidents to be reported in graphical form.	To generate easy-to-read reports about spam filtering activity.	<b>Supported.</b> CanIt data is easily exported to graphing software and spreadsheets.
Server-based	Refers to a type of spam filter that resides on the organization's e-mail server (rather than on client machines or on third-party servers).	To protect your entire network and all users by filtering mail at the gateway. A solution that resides within your network gives you ultimate control over your own e-mail. It also allows for centralized functionality – such as the creation of global rule-sets and centralized updates.	<b>Yes.</b> CanIt software is designed for download to your mail server. The pre-configured CanIt Appliance is a gateway that sits between the Internet and your mail server.
Spam thresholds	Refers to the ability to set a spam filter's sensitivity level when determining what is and is not spam.	To set and adjust filter sensitivity. If you want a solution that gives individual end-users control over their filters, then thresholds must be customizable per-user.	<b>Yes.</b> Spam thresholds in CanIt solutions are easily set and modified via a simple web interface, and access to thresholds can be provided to end-users if desired.
Spam trap	See "quarantine" above.		<b>Yes.</b>
Spamware detection	Detecting e-mail sent using spamware. Some telltale signs include temp-failed e-mail that is not resent (see "Greylisting" above), particular header content, false HELO commands, and more.	To eliminate a large percentage of certain spam.	<b>Yes.</b> CanIt can detect the characteristics of spamware.
Streaming technology	A mail stream is a mail filter. It has a particular set of filtering characteristics unique from all other streams on the system.	To allow per-user control (e.g. per-department or per-individual).	<b>Yes.</b> Roaring Penguin invented mail streams for its CanIt-PRO product.

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Unpacking compressed files	A filtering technique that opens compressed files (such as .zip and .tar files) attached to e-mails in order to determine whether the compressed file is hiding a malicious attachment (such as a virus) inside it.	To catch malicious e-mail attachments hidden in compressed files. However, software that unpacks compressed files must be carefully written. Compressed files can contain very large binary files that, once unpacked, essentially bombard disk space and eat up processing power – creating a denial-of-service attack.	<b>Optional.</b> CanIt is bundled with Clam AntiVirus, which carefully unpacks compressed files. You can turn this feature on or off.
Looking into compressed files	A filtering technique that looks inside compressed files and filters them as spam if they are seen to contain potentially malicious attachments.	To catch malicious e-mail attachments hidden in compressed files. If you already have anti-virus software that filters compressed files, then you're likely covered. If you don't, then this is a feature that your anti-spam solution should include.	<b>Yes.</b> CanIt's filters can look inside compressed files to detect hidden malicious attachments without actually unpacking such files.
Voting links	Literally, links appended to the bottom of each e-mail that enable the user to click on them to indicate "this is spam" or "this is not spam".	To provide users with a convenient way to train their Bayesian filters.	<b>Optional.</b> CanIt-PRO enables administrators to automatically append Bayesian voting links to e-mail messages, allowing end-users to quickly train the Bayesian filters.
Whitelisting	A filtering technique whereby a list of allowed senders or domains is created. E-mail from a sender or domain on a whitelist is always allowed through.	To reduce the load on your spam filters and streamline your filtering architecture. This feature should be used in conjunction with "mismatch rules" (see above).	<b>Optional.</b> CanIt enables the creation of optional whitelists via a simple graphical user interface.
Word-pair analysis	A type of "adaptive spam filtering" (see above) in which the spam filters analyze a message's spam probability based on the occurrence of <u>particular pairs of words</u> (not just single words).	Word-pair analysis dramatically increases the accuracy of adaptive filters.	<b>Yes.</b>

## White Paper: Demystifying the Anti-Spam Buzz

Buzzwords (A-Z)	What is it?	Why do I need it?	Does CanIt have it?
Web interface	A graphical user interface (GUI) that enables access to spam filtering features (administrative and/or end-user access) via a web browser.	<p>A web-based interface is beneficial because:</p> <ul style="list-style-type: none"> <li>• It can be administered from anywhere there is web access</li> <li>• It can work with any e-mail client or client operating system</li> <li>• It will work for users if they are using webmail rather than their mail clients</li> <li>• Users can manage their Spam Traps even while out of the office, on holiday or working remotely</li> </ul>	<b>Yes.</b> CanIt administrative interfaces, and CanIt-PRO's end-user interface, are web-based.
Zombie detection	The ability of a spam filter to detect whether an e-mail is being sent from a "zombie" spam machine (i.e. an unwittingly compromised computer used as a launch pad for spam attacks).	Zombies are usually co-opted machines residing inside legitimate domains, so special detection techniques are useful. Zombie detection is preferable to blacklisting legitimate domains and it also dramatically reduces the amount of spam that your system must filter.	<b>Yes.</b> See "greylisting" above.

### Contact Information:

Roaring Penguin Software  
17 Grenfell Cresc., Suite 209C  
Ottawa, Ontario  
K2G 0G3 Canada

E-mail: [info@roaringpenguin.com](mailto:info@roaringpenguin.com)

Tel: +1 (613) 231-6955  
Toll-Free: 1-800-210-0984

Web: [www.roaringpenguin.com](http://www.roaringpenguin.com)