



REPUTATION-BASED MAIL FLOW CONTROL

Blocking Extreme Spam and Reducing “False Positives”

Blocking unsolicited commercial email or “spam” is an increasingly important but difficult task for IT staff. Great effort has been spent developing sophisticated content filters that can identify and score messages to determine their level of “spaminess”. While these systems continue to improve with the introduction of techniques such as bayesian filtering, their efficacy will always be compromised by “spammers,” with a money motive, who will continually modify their content to get through these filters. Some senders of unsolicited commercial email have taken the process of avoiding content filters one step further — using widely available tools to scan their messages through popular anti-spam packages before they are sent so they can modify their content to make sure it gets through.

“A ‘Reputation Filter System,’ allows differing mail throughput policies to be implemented for different senders. These policies will slow mail delivery for less trustworthy senders, while allowing much higher throughput for mail delivery from more trustworthy senders.”

— MICHAEL OSTERMAN
Analyst, Osterman Research

Reputation-Based Mail Flow Control

Content filters face a natural trade-off – the more aggressive they are at blocking spam, the greater the risk of inadvertently blocking legitimate email. Content filters are also highly CPU intensive, requiring additional hardware investments and sometimes delaying the delivery of critical messages. Finally, content filters are defenseless against dictionary harvest attacks, denial of service attacks and other important risks to an enterprise.

A MULTI-LAYERED SPAM DEFENSE

The overall effectiveness of content-based filters can be enhanced when used in combination with a reputation-based mail flow control system. Reputation-based systems are the next generation of “identity-based” spam-fighting approaches like blacklists and whitelists and make decisions based on comprehensive information about the source of the message. Unlike blacklists (lists of suspected “bad” mail servers) and whitelists (lists of assumed “good” mail servers), reputation filters rely on objective data to assess the probability that a message from any given IP address is spam. This probability is based on data such as how many messages a mail server is sending, how many complaints they get, whether or not the mail server is sending to “honeypot” accounts, where the sending organization is located, how long the organization has been sending email from a given location and a number of other factors.

Using a reputation filter in combination with a content filter has numerous benefits for an organization. These benefits include:

- **Improved catch-rate (blocking more spam).** Reputation filters introduce important new sources of data tied to a sender’s IP address that can be used to block egregious spam directly or as an input into third-party anti-spam systems.
- **Reduced false-positives (incorrectly identifying fewer legitimate messages as spam).** Because reputation-filters are probability-based (not binary like black & whitelists), receiving MTAs (Message Transfer Agent) can limit the rate they receive messages from someone without blocking their messages. This creates a prohibitive cost for spammers while ensuring that critical email still gets delivered.
- **Lower hardware costs and increased message throughput.** Research from IronPort Systems indicates that a well-designed reputation filter identifies over 50% of a typical corporation’s email as either “known good” or “known bad” based on the message source. Messages from “known bad” senders can be blocked at the gateway, while messages from “known good” senders can bypass content filters. This configuration halves the number of messages that must be sent through content filters, doubling the throughput of a typical enterprise’s messaging infrastructure.
- **Reduced risk from denial of service or dictionary harvest attacks.** By rate limiting based on a sender’s source IP, the reputation filter can throttle senders with bad reputations, minimizing the damage from malicious attacks.

KEY COMPONENTS OF A REPUTATION-BASED SYSTEM

When deploying reputation-based mail flow control system, there are several things companies should evaluate. Any solution should be:

- **Non-spoofable.** The email sender’s reputation should be based on the IP addresses of the email sender. Because SMTP is a two-way conversation over TCP/IP, it is nearly impossible to “spooof” an IP address – the IP address presented must actually be controlled by the server sending the message.

Reputation-Based Mail Flow Control

- **Comprehensive.** The more data that is maintained about a sender's reputation, the better the decisions that can be made about a message from that source. At a minimum, any reputation-based system should track data on complaints, message volume, messages sent to "Honeypots", sender location and open proxy status.
- **Reliable.** A reputation filter should provide fast and reliable access to network data. Where possible, reputation data should come from one source, limiting the number of threads that must be opened to process an incoming message.
- **Variable.** The richness of the data in the reputation assessment should allow for a graduated response to spam – the more suspicious a sender appears, the more restricted their access. A company may choose to block senders with terrible reputations, while limiting the number of recipients/hour for senders with bad reputations.

SenderBase™ from IronPort Systems is a powerful new sender reputation service that provides a variety of data on the identity of any sender, allowing email administrators to establish trust levels for a given sender much as a credit rating service allows a merchant to establish credit worthiness of a customer. The data is free and can be accessed at www.senderbase.org, and can be used as the foundation of a reputation-based mail flow control system that can be used to determine access privileges granted to any sender. These privileges range from unlimited delivery rates with unlimited attachment sizes and bypassed content scanning to limited delivery rates with no attachments and full content scanning to the complete refusal of incoming connections from the sender.

“Access to SenderBase could have a major impact on the accuracy of spam filters”.

—DANIEL QUINLAN
SpamAssassin Developer

A WORLD WIDE NET OF SENSORS

One of the most telling data points collected on any sending IP address is global volume of mail sent. Spammers are incredibly effective at making their content look like legitimate email but the one parameter they cannot mask is volume. It is difficult for any one mail administrator to measure volume since spam gets spread across millions of domains. However, with a network of more than 10,000 ISPs, universities and corporations, SenderBase provides a global view of mail being sent by any IP address. This global view can be used to identify and block high-volume spam sources.

SenderBase is an open, free service being rapidly embraced by the spam-fighting community. Email administrators access SenderBase data via a web browser or via machine-readable responses to DNS queries. SenderBase DNS zones can be queried for each incoming message being received and a rich set of data is returned via compact DNS TXT records. This information is read by smart messaging gateways or spam filters to help make better decisions about how to handle an incoming message.

SenderBase data is also available to organizations writing rules to more effectively catch spam. “SenderBase provides us with an entirely new set of data, such as sender mail volume and organization data, that we can use to analyze the email as it is received” said Daniel Quinlan, a SpamAssassin developer. “Access to SenderBase could have a major impact on the accuracy of spam filters”.

Reputation-Based Mail Flow Control

LINKING DATA BY IP, DOMAIN, AND ORGANIZATION

SenderBase indexes all data at the IP, Domain and Organization level. This is a simple but incredibly powerful way to combat spam. Spammers will often have large networks with multiple domains and IP addresses. If a spam message is received from a single IP address, blocking that IP is a bit like squashing a single ant found in your kitchen. There are literally hundreds or thousands of more IP addresses and domains, akin to thousands of ants in the nest, waiting to attack your network. SenderBase provides a simple tool to root out the entire nest – block thousands of IP addresses across multiple domains all rolling up to the same organization with a single click.

EXAMINE THE WHOLE NETBLOCK

SenderBase allows the administrator to rapidly “zoom out” and examine other mail servers on the same netblock as the server in question. Practical limitations force spammers to operate out of a limited number of co-location facilities. As a result, where you find one spammer, you will often find others on the same netblock. A quick “zoom out” to examine the /24 level of any IP address will often reveal a pattern of servers with high volumes and high complaint rates. Conversely, if you are tempted to block an entire netblock, a quick examination of the entire netblock in SenderBase can help identify any innocent mailers that may inadvertently be affected, so that you can narrow your block or whitelist those senders specifically.

AGGREGATED COMPLAINT DATA AND BLACKLISTS

In addition to volume and network data, SenderBase aggregates 10 black lists and provides a consolidated view of whether or not the IP address in question appears on any of them. SenderBase, in partnership with SpamCop and other sites, also contains data on end user complaints and samples of messages sent from the IP address in question. While much of this data is available in a variety of places on the Internet, SenderBase provides a unified view that is indexed to the other SenderBase data.

IMPLEMENTING A REPUTATION-BASED MAIL FLOW CONTROL SYSTEM

The reputation data provided by SenderBase can immediately add value to virtually any email gateway system, even ones based on traditional gateway software such as open-source SendMail, qmail, or Postfix. SenderBase Reputation data can dramatically streamline the process of managing whitelists or blacklists. With a single click, SenderBase users can export the full list of IP addresses sending mail from a given domain or organization. This is a powerful tool to help quickly build a whitelist that includes all of the many mail servers a Fortune 500 customer might have. Conversely, it is a very efficient tool to block the many thousands of IP addresses and domains used by an organization to send spam. Users can select an “IP Export” in file formats that can be readily loaded into Sendmail, qmail or Postfix.

For a more powerful solution, IronPort Systems has developed a family of Messaging Gateway™ appliances that have the SenderBase Reputation data built in. System administrators get a comprehensive view of their incoming and outgoing mail flow, and can drill down to see details on any of their incoming mail senders. The SenderBase service also provides a statistical score to the IronPort appliance that assesses the trustworthiness of that sender based on a broad set of SenderBase data. The more suspicious a sender appears, the more restrictive email policies will be applied to the sender. This is best illustrated with an example. If a sender has high global volumes

Reputation-Based Mail Flow Control

of mail – say 200 Million messages per day – from a network of 5 different domains and 1,700 IP addresses that have only been sending mail for 15 days yet have a high end user complaint rate and they don't accept incoming mail, they will have a very low reputation score and the IronPort C60™ will limit the amount of mail they can deliver to the corporation. If a sender is a Fortune 500 company, they will likely have much more modest global email volumes – say 500,000 messages per day – will have a smaller number of IPs and domains with a long sending history, they will accept incoming email and have low (or zero) end user complaint rates.

One of the most unique aspects of the IronPort C60 is it's ability to offer a variable response to spam – the more suspicious a sender appears, the slower they are allowed to go. This offers a dramatic improvement over traditional white and blacklists which have a binary response – a sender is either good or bad with nothing in between. Spammers are very effective at appearing to be “in between”.

The SenderBase Reputation data is also very helpful to an IronPort administrator for diagnosing anomalies. The IronPort C60 has a powerful Mail Flow Monitor™ round-robin database that records all email traffic patterns and identifies unusual changes in these patterns. The system can be programmed with an automated response to these anomalies or it can generate an alert and allow the administrator to rapidly assess the threat using on-board SenderBase Reputation data.

CONCLUSION

The next generation of spam control solutions will need to move beyond just content filtering in order to be effective. Reputation-based information makes existing filters more effective and will form the foundation of more sophisticated mail flow control systems. When used in conjunction with a content based system, the net result is a more accurate and more secure solution for incoming mail filtering.

Email receivers can begin using reputation-based information to research any sender in the world today at www.senderbase.org. Receivers looking for more powerful controls over incoming mail can also take advantage of the advanced mail flow monitoring and flow control capabilities of products such as IronPort's Messaging Gateway appliances.



IronPort Systems, Inc.
1100 Grundy Lane, Suite 100
San Bruno, California 94066
tel 650.989.6500 **fax** 650.989.6543
email info@ironport.com
www.ironport.com

THE IRONPORT STORY

IronPort Systems is focused on one goal: to revolutionize Internet Messaging. We have developed a family of products called Messaging Gateway™ appliances that offer breakthrough performance, unprecedented ease of use, and reduced total cost of ownership. The IronPort team has done it before: our technology stems from experience at Hotmail, eGroups, ListBot, and Yahoo!