

Spam Filter Testing Best Practices

Content & Collaboration Strategies

Matt Cain

FOCAL POINT

Given the significance of the spam blight and the competitive nature of the spam-blocking vendor landscape, most organizations are diligently evaluating suppliers, and in many cases bringing in products for hands-on testing. In addition, many trade publications are doing on-site bake-offs to determine the effectiveness of various solutions, including on-premises software, appliances, and managed services. In some cases, the testing methodology is flawed, and the results do not represent the actual effectiveness of the product or service. The root cause of the invalid testing is that testers typically take a corpus of mail and forward it to the spam-blocking service or product. In such cases, because of the message forwarding, the vendor is unable to perform a series of sender IP validation tests, nor is it able to glean intelligence from the SMTP setup. In some cases, these real-time tests can contribute up to 20% of spam being blocked.

Furthermore, the header information that is forwarded along with the message is subject to spammer manipulation and is therefore not necessarily a productive interrogation target. For example, spammers now routinely add legitimate IP sending addresses to header information in hopes of the spam being allowed to pass through the blocking service without scrutiny. In fact, spammers now take great pains to hide the originating IP address in the header information, which affects not only whitelist performance, but also blacklist, traffic shaping, and reputation filter effectiveness. Therefore, customers need to understand spammer header tricks as well as the value of real-time spam evaluation services, and change testing methodologies accordingly to get a more accurate picture of the effectiveness of spam-blocking products.

Here, we describe various real-time blocking techniques from a sampling of vendors and conclude with best practices for accurate testing methodologies.

CONTEXT

It is not news to any organization using e-mail that spam threatens the effectiveness of e-mail systems. Left unattended, spam clogs inboxes, compromises user efficiency, and overwhelms system components such as message stores and message transfer agents (MTAs). Furthermore, spam is a conduit of all types of salacious content and fraudulent come-ons that seek to cajole users into disclosing confidential information such as credit card and Social Security numbers, bank account information, and passwords (known as phishing). Approximately 70% of most organizations' inbound SMTP traffic is spam. Therefore, it is mandatory that organizations aggressively deploy top-tier spam-blocking solutions to mitigate the risks and problems associated with spam.

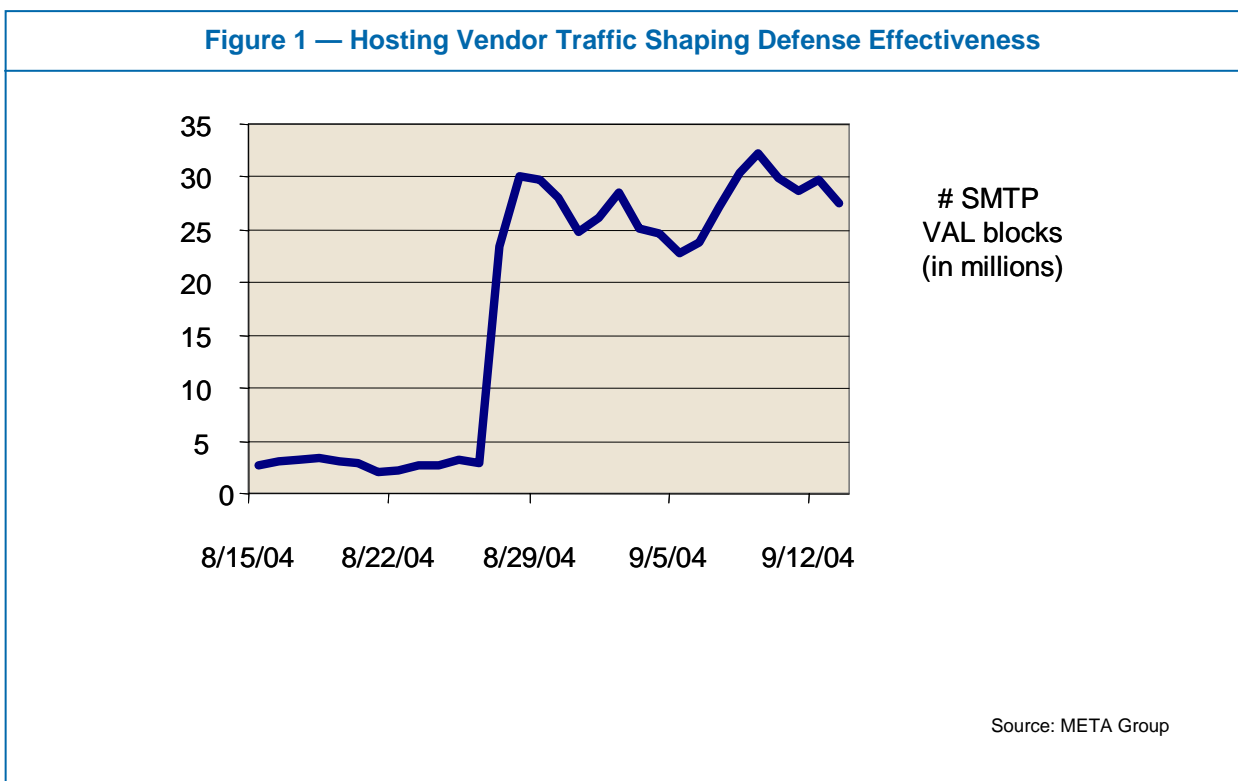
Many spam-blocking services use multiple strategies that are not invoked in a mail-forwarding testing situation:

- **IP-blocking reputation lists:** Some spam-blocking companies filter more than 100 million messages a day, and from this large volume they are able to glean intelligence about the sending patterns of a particular IP address. If they find a high correlation between a particular IP address and an unusual volume of mail or certain types of mail coming from the same address, they will refuse connections — for at least a period of time — from that address. Some companies issue a 550 SMTP error message (access denied) to the sender. In this scenario, some vendors have a technician examine the mail flow and determine whether the messages are spam and then act accordingly.

META Trend: *As ad hoc electronic communication grows in importance (e.g., e-mail, instant messaging, Web conferencing), organizations will be challenged to create a hygienic and low-cost infrastructure. Through 2006, special attention will be focused on spam blocking and policy enforcement (e.g., regulatory compliance). By 2007, rising electronic communication volumes will frustrate users coping with information overload and drive organizations to employ common filters, queuing services, and categorization engines to ease communication burdens.*

Companies also do in-depth log analysis to determine the validity of sending IP addresses. In addition, vendors have automated the reputation process, and in cases of, for example, real-time mail flood attacks, the system can shut down connections, though only after human oversight. With these IP-based reputation filter approaches, vendors estimate that they stop between 5% and 8% of all spam flowing through its network. Although there is a common belief that IP addresses are spoofable, SMTP connections require a confirmation packet from the recipient MTA sent to the sending IP address. Although not technically impossible, in practicality, it is extremely difficult to spoof an IP address.

- **TCP/IP blocking:** Another blocking method that would not be applicable in a mail-forwarding testing scenario is blocking messages at the TCP/IP level. Vendors have determined that spammers often display certain behavior during the SMTP conversation string — when the recipient and sender MTA first establish a connection. Vendors will not disclose the specific behavior for fear of tipping off spammers, but when this common behavior is identified, vendors issue a 550 SMTP error message (access denied). One vendor is blocking between 25 million and 30 million messages a day based on this method (see Figure 1). This per-message blocking service is also invoked using e-mail authentication standards such as Sender Policy Framework (SPF) and directory fail attempts.



- **Traffic shaping:** Some vendors use a third spam-blocking method called traffic shaping or IP throttling, which would not be invoked in a mail-forwarding testing situation. In this case — called graylisting — vendors again correlate message flow and type with a particular IP address. But instead of dropping the connection, vendors slow down delivery rates — issuing an SMTP 451 error message (connection temporary unavailable). SMTP relays of legitimate sending organizations will retry later to get the message through, but a spammer — which is typically paid on volume of messages sent — will quickly lose patience and move on to another recipient MTA. This method will also effectively stop dictionary harvest attacks, where the spammer attempts to collect legitimate mail addresses by bombarding the recipient MTA with a large volume of mail addressed to common names.
- **Header walk:** Other real-time techniques are emerging that help determine the validity of messages. “Header walk” services allow the interrogation of header hop data (the routing path the message took to get to the destination), enabling the discovery of the sending source IP address. After determination of the source IP address, it will perform a real-time lookup to see whether that IP address is registered to that domain.

- **Sender query:** Vendors use services that validate whether the sending address is legitimate by, for example, sending a message back to the sender, enabling the company to ascertain that the sending address is legitimate by scanning for delivery error codes.
- **Header/conversation data comparison:** Companies are also investigating a service that would allow the comparison of SMTP conversation data (e.g., sending domains) with header information to see whether they match. Currently, spammers will often spoof header information, and evidence of that spoofing will be revealed by comparing the SMTP conversation data with the header data the recipient sees.
- **Sender IP identification:** Although sender IP addresses can usually be found in the receive headers in a forwarding scenario, hygiene vendors take additional actions to find the appropriate IP address. Certain MTAs do not include the original IP address, and some open-source MTAs have a tendency to botch the proper placement of the sender IP address. Therefore, in a mail-forwarding scenario, it could be impossible to find the original sender's IP address.
- **Proxy server identification:** At a lower level (TCP/IP), vendors can often identify when a proxy server is being used to relay mail — an almost certain indication that a spammer is the source of the messages. Again, in a mail-forwarding scenario, this intelligence is lost, creating suboptimal results for the spam-blocking system under interrogation.
- **Source origin loss:** Losing the source origin of the message also creates other problems; some Bayesian filters look at the receive header to find legitimate hops. In a mail-forwarding situation, the final hop is always legitimate, thereby tricking the filter into awarding positive scores that are fed into the overall statistical analysis of the message.
- **Message modification:** Any modification to a message before it arrives at the filter can create artificial results — whether they are changes in the headers or content. For example, adding a forward descriptor into the subject line can obscure results. Forwarding scenarios may also result in changes to the MIME boundaries that separate messages into logical parts such as text and attachments, leading, again, to compromised results.
- **Stale mail:** Although not a real-time issue, testing results can be obscured by use of obsolete messages. Most vendors retire mail-blocking rules, heuristics, and signatures regularly. If an older corpus of mail is used in a test, blocking effectiveness can be compromised if the relevant rules and other data have expired.

Testing Methodologies

Many evaluations focus exclusively on spam capture rates and false-positive generation. A broader testing methodology is more appropriate, where factors such as end-user satisfaction, ease of administration, and operational control are considered. This approach enables greater leeway for so-called graymail (messages the recipient might have solicited at one point, but no longer wants to receive). Likewise, it is inappropriate to compare vendors representing the three major delivery modalities (hosted, appliances, and traditional software load) with the same criteria because each delivery mechanism has a different value proposition. Finally, the plug-and-play method of spam evaluation — where testers merely turn on the service with little or no tuning — does a disservice to the vendors. Testers need to do the appropriate tuning, including quarantine conditioning — enabling recipients to set up block/allow lists to get a more accurate picture of blocking effectiveness.

We suggest the following for more accurate testing:

- For hosting vendors such as MessageLabs, FrontBridge, and MX Logic, the most accurate testing scenario is to change the destination MTA message exchange record to the hosted vendor, which will filter the mail and forward it to the recipient domain.
- For on-premises traditional software load and appliance vendors such as CipherTrust, an IP load balancer should be placed in front of the hygiene vendors, and a real-world mail feed should be balanced equally across all the vendors being tested.

META Practice

These approaches have two main virtues: use of real-time, real-world e-mail feeds, and no changes to mail headers and other data typically altered in a mail-forwarding testing scenario. Testers should ensure that a statistically relevant volume of mail is tested for legitimate results.

Furthermore, for both approaches, we make the following suggestions for improving the overall testing methodology:

- Blocking engines should be appropriately tuned before the actual tests start. In the case of hosted vendors, end users should be allowed to configure their personal blocking preferences.
- Testing should be done on real business users. They should give feedback to testers when spam gets through and when false positives are detected. These users should come from various corporate departments such as human resources, accounting, and customer support. This ensures a real-world representation of a corporate mail stream.
- Definitions of spam should be agreed on before testing (e.g., all messages with salacious content, regardless of sender, should be considered spam).

Bottom Line

For effective mail hygiene vendor comparisons, organizations must ensure that test beds represent real-world conditions, creating a more accurate picture of actual spam-blocking capabilities. Evaluations should also focus on the broader capabilities of the vendors, not just exclusively on spam capture rates and false-positive generation.

Business Impact: E-mail is a crucial corporate communication tool; organizations must implement sophisticated e-mail hygiene practices to ensure stable, secure, and operationally sound messaging infrastructures.