White Paper

Outbreak Filters

Spear phishing and targeted attacks using new tools and more reconnaissance are on the rise. Cisco IronPort™ Outbreak Filters are designed to combat smarter attackers and sharpened tactics. This paper outlines a variety of these attacks, as well as what Cisco IronPort Outbreak Filters are, how they work, and their configuration options.

A Turning Point

In late 2010, spam volume showed the first signs of decreasing. Starting in October 2010, several prolific spammers were and several botnets were shut down.

These arrests, combined with the shutdown of the botnet command and control infrastructure used for sending threats, have contributed to the drastic reduction in spam volumes. While there are still systems sending massive volumes of spam, there has been a definite change in tactics. In addition to sending bulk spam blasts, attackers are now increasingly sending specially crafted targeted attack messages in attempts to avoid antispam security systems.

Staying Ahead of Targeted Attacks

Historically, spam email was easy to identify because it was full of grammar, spelling, and punctuation errors, and used formal language. Now, attackers have polished their language and content along with their overall message, making it more difficult to distinguish a real email from spam.

Targeted attacks represent less than 1% of the inbound spam messages faced by organizations globally. These attacks are categorized as:

- **Advanced Persistent Threats (APTs):** These messages are sent as part of a broader attack that attempts to break into an organization's network over an extended period of time.
- **Spear phishing and whaling:** These messages are targeted toward specific individuals in an attempt to steal money or information. An example is a targeted attack on an organization's Accounts Payable group in an attempt to install software that steals the organization's banking information.

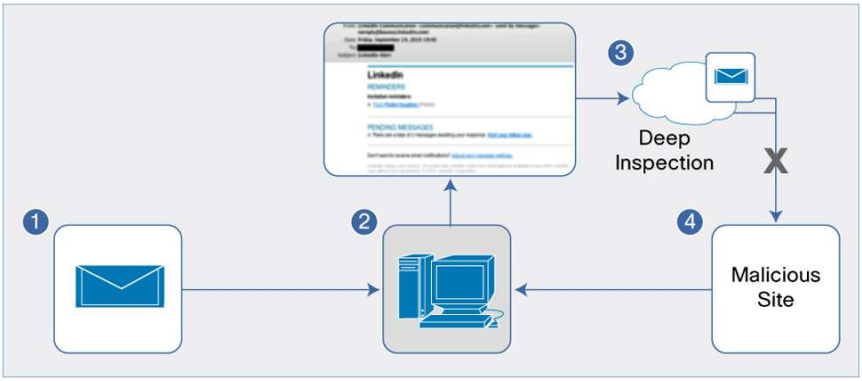
Targeted attacks use data that can easily be harvested from social networking sites such as Facebook and LinkedIn. By including city names or a target's name, or by making an email appear to be from a friend or business associate, attackers increase their chances of success. Attackers know the weakest link in the security chain is the general end user - and they know that social engineering is providing the highest level of returns.

Staying ahead of these attacks takes a coordinated effort. Cisco IronPort Outbreak Filters technology is designed to tackle this challenge, using a combination of Cisco IronPort SenderBase® information about "bad" email senders; Cisco SensorBase™ information that has been collected by Cisco security appliances; and Outbreak Intelligence, web scanning technology acquired through Cisco's purchase of ScanSafe. Cisco IronPort Outbreak Filters use refined rule sets to inspect emails, a dynamic quarantine to hold suspect emails for rescanning, and Outbreak Intelligence to scan suspect URLs.

How Do Outbreak Filters Work?

Cisco IronPort Outbreak Filters build upon Cisco IronPort Virus Outbreak Filters. Like their predecessors, the Outbreak Filters protect systems against new outbreaks of viruses and other malware delivered via attachments, but they take this technology to a new level by scanning URLs and processing them in real time - as the user opens them - to block malicious sites. In addition, the filters send data about these websites to Cisco Security Intelligence Operations (SIO) to protect all users of Cisco security products that incorporate SensorBase technologies, including Cisco's firewall, web security, and intrusion prevention products.

Figure 1. A simple flow of Cisco IronPort Outbreak Filters



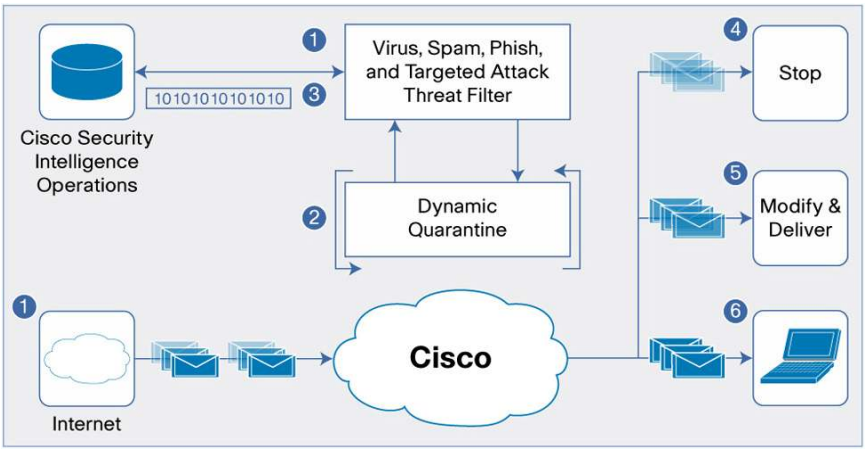
1. An incoming email is scanned by Outbreak Filters. The refined rule set identifies this as a potential phishing or targeted attack email and handles it as configured on the appliance. By default, a disclaimer is prepended to the email text identifying it as a phish and the URL contained in the email is rewritten.
2. The email with the rewritten url is delivered to the user's inbox.
3. If opened, this rewritten email link sends the user to a public proxy where the webpage content is intercepted and scanned in the cloud in real time using Outbreak Intelligence from ScanSafe.
4. If malware is detected on the page, a blocked page message is served up to the user and information about the URL is passed from the ScanSafe system back to Cisco SIO. Otherwise, the user is given a choice: surf the page through the proxy or go directly to the site.

This ScanSafe technology is the same that powers our cloud-based web security offering, which scans over seven billion web requests per day. It uses a combination of deep content analysis, structural content investigation, and virtualized script emulation to scan requested web pages for a large variety of malware (covered in another section).

Outbreak Filters use telemetry from the Cisco IronPort Web Security Appliance and from Cisco ScanSafe. However, customers do not need to have either the Web Security Appliance or ScanSafe to use Outbreak Filters. Existing Virus Outbreak Filters customers can take advantage of the new Outbreak Filters by simply upgrading to Cisco IronPort AsyncOS® 7.5 or later.

A More Detailed Look

Figure 2. A deeper look into the message flow through Outbreak Filters



1. An email comes into an organization and is scanned by Cisco IronPort AntiSpam using the most current rule sets and attack heuristics that have been published by Cisco SIO.
2. If the rule sets determine the email to be a potential attack, it is assigned a dynamic quarantine retention value, or time for which the message will be quarantined.
3. Once this quarantine timer expires or the maximum quarantine time configured has been reached, the message is released from quarantine and rescanned by the IronPort AntiSpam engine using the latest rule set.
4. This updated rule may identify the message as spam to be processed according to configured policy.
5. If the message is determined to be safe, it is delivered to the mailbox unchanged.
6. If the message is determined to be a possible threat, it is sent to the user's inbox after being processed according to the configuration - such as rewritten URL, prepended message body text, and prepended subject line.

When the modified, suspect URL is opened, the user is redirected to the cloud and the URL is verified. Once this verification is complete, the webpage is pulled and scanned in real time using Outbreak Intelligence from ScanSafe. If the webpage is malicious, it will be blocked, and a notification will be sent to the user indicating that the page contained malware. Figure 3 shows a sample notification.

Figure 3. Sample Blocked Page Notification

Detailed information about the disposition of the URL is passed back to Cisco SIO. This information is processed and added into IronPort AntiSpam updates, which will block any quarantined messages that are waiting to be released.

Note: Quarantined messages are continuously rescanned by the antispam engine.

The User Experience

The user simply receives an email in their inbox that, by default, will have a prepended subject line and message body with a warning message along with rewritten URLs. If the user clicks on the rewritten URL and opens it in a browser, they will connect to a public web proxy. If the webpage contains malicious content, they will simply receive a block page. If the page contains a suspicious file, the user will be presented with a warning screen asking them if they wish to download the file. Also, if the page appears to have no malware but still appears suspicious, the user can open the website through the proxy for protection or can go to the website directly.

In all cases, the user is prompted within the browser environment through the proxy server connection.

Components of Outbreak Filters

Outbreak Filters are comprised of Cisco IronPort AntiSpam rule sets and Cisco SIO refined rule sets, Cisco IronPort AntiSpam URL rewrites, and Cisco IronPort Email Security Appliance targeted attack heuristics, as well as Cisco ScanSafe technology to block malicious web content.

The Cisco SIO rule sets are the standard Cisco IronPort AntiSpam updates that are used for filtering incoming emails as standard spam or attack messages that should be blocked or quarantined. These rule sets include feedback from URLs that have been blocked previously by Outbreak Filters and are used to rescan messages that are exiting the outbreak quarantine.

Targeted attack heuristics combine intelligence gained from SenderBase and security appliances, as well as content derived from known attacks. This intelligence is used to evaluate various aspects of each message, such as the URLs, headers, and message body, to identify threats within the message. These heuristics work with other Outbreak Filters rules to determine if a given message is a targeted threat.

Rewriting the URLs is the job of the Cisco IronPort AntiSpam engine. The URL is rewritten to include the URL for the proxy on the public Internet, a hash used for authentication of the proxy request, and the original URL. As an example, if the URL http://www.youtube.com/watch?v=FCARADb9asE were rewritten in an email, it may come out looking like this: http://secure-web.cisco.com/auth=11vKpmpe8R7CLj6iG0cbr40VY5yOTR&url=http%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DFCARADb9asE. This will cause the request to flow through the proxy for the web content to be scanned by Cisco ScanSafe.

Deep content analysis is the first of the three Cisco ScanSafe processes that look for and block malicious web content. Content is examined that would indicate malicious intent. This is done by comparing to the known “normal” statistical model kept accurate by AI learning. For example, does an animated GIF have only one frame? Does an image file contain executable code or other anomalous content?

The structural content investigation process examines the structure of the content, again looking for signs of potential risk. For example, based on content analyzed by Cisco SIO, it could be determined that a new obfuscated executable file has a probability of over 95% of being malicious. This is done by scanning the page content using multiple proprietary scanning engines, or scanlets, such as reputation scanlets and scanlets for Flash, Java, PDF, archives, executables, file anomalies, and more. These scanlets run in parallel to ensure high performance.

The third component, especially important for dynamic content such as scripts, is virtualized script emulation. Running the script within the cloud infrastructure allows for monitoring of malicious behavior such as a hidden redirect or “drive-by” downloads attempting to edit user settings on a computer. If malicious behavior is detected, the script is not allowed to pass to the end user and is blocked.

These web components from Cisco ScanSafe have been proven to accurately identify and block zero-day malware, which makes up more than 25% of the malware blocked.

What Do Outbreak Filters Detect?

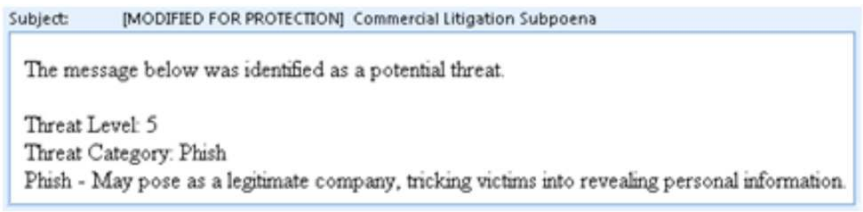
Cisco IronPort Outbreak Filters are organized into four categories; malware, phishing, scam, and virus-infected emails and detect more than 20 different types of scams. The following list covers some of the types of attacks Outbreak Filters can detect:

- Phish
- Charity
- Robbed Abroad
- Seminar
- Inheritance
- Financial URL
- Fake Deal
- Bank Transfer
- Fake Cashier’s Check
- Money Mule
- On Craigslist
- Loan
- Financial Phone
- And many more

This list dynamically changes, keeping up with the current trends that attackers use in their attempts to get users to open URLs and/or attachments.

A configuration setting in Outbreak Filters allows for the system to prepend information about the nature of the email above the body text and to put a message on the subject line (Figure 4).

Figure 4. Subject modification and body text



Customizing Notifications

If an organization’s policies require a specific message to be presented to an end user, a custom text message can be configured through a text resource of the type Disclaimer Template. These text resources can use any of

the existing notification variables or any of the following four new variables, along with custom text required by organizational policies:

- $threat_category
- $threat_type
- $threat_description
- $threat_level

Once configured, the Preview Text option can be used to see an example of what the configured disclaimer would look like to an end user.

Troubleshooting

With Cisco IronPort Outbreak Filters, troubleshooting is simple. The filters add a line to the message_logs log subscription. These logs can be accessed via the CLI, downloaded from the GUI, or exported from the appliance to an enterprise logging system.

Here is an example log entry for a message deemed to be a threat:

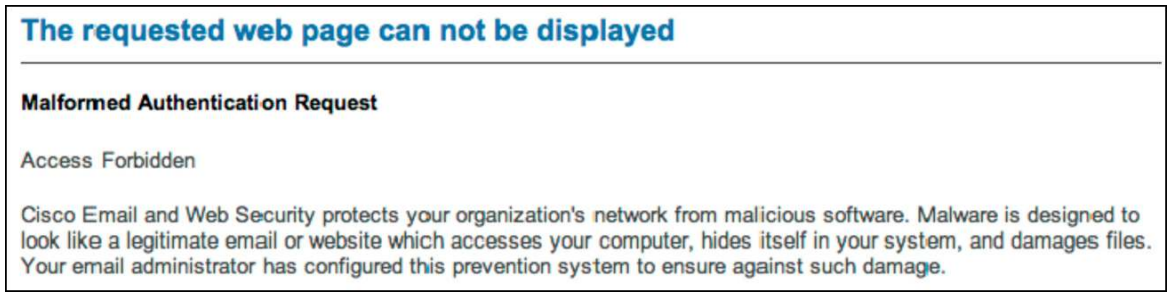
```
Wed Mar 9 17:43:51 2011 Info: New SMTP ICID 1720507 interface Management (10.0.0.42) address
67.215.21.16 reverse dns host www.wyodino.org verified yes
Wed Mar 9 17:43:51 2011 Info: ICID 1720507 ACCEPT SG SUSPECTLIST match sbrs[-2.0:0.0] SBRS
-1.0
Wed Mar 9 17:43:51 2011 Info: Start MID 487757 ICID 1720507
Wed Mar 9 17:43:51 2011 Info: MID 487757 ICID 1720507 From: <KarmelaMeir@aol.ca>
Wed Mar 9 17:43:51 2011 Info: MID 487757 ICID 1720507 RID 0 To: <jdoe@customer.com>
Wed Mar 9 17:43:51 2011 Info: MID 487757 Message-ID '<058fb89e-40612-1e4e1551531134@user-
pc>'
Wed Mar 9 17:43:51 2011 Info: MID 487757 Subject 'Payment Confirmation - eBay Item number :
390294669929'
Wed Mar 9 17:43:51 2011 Info: MID 487757 ready 1318 bytes from <KarmelaMeir@aol.ca>
Wed Mar 9 17:43:51 2011 Info: MID 487757 matched all recipients for per-recipient policy strong in the
inbound table
Wed Mar 9 17:43:51 2011 Info: ICID 1720507 close
Wed Mar 9 17:43:51 2011 Info: MID 487757 using engine: CASE spam negative
Wed Mar 9 17:43:51 2011 Info: MID 487757 interim AV verdict using Sophos CLEAN
Wed Mar 9 17:43:51 2011 Info: MID 487757 antivirus negative
Wed Mar 9 17:43:51 2011 Info: MID 487757 Threat Level=2 Category=Phish Type=Financial Url
Wed Mar 9 17:43:51 2011 Info: MID 487757 queued for delivery
Wed Mar 9 17:43:51 2011 Info: New SMTP DCID 89502 interface 10.0.0.42 address 10.0.0.119 port 25
Wed Mar 9 17:43:52 2011 Info: Delivery start DCID 89502 MID 487757 to RID [0]
Wed Mar 9 17:43:52 2011 Info: Message done DCID 89502 MID 487757 to RID [0]
Wed Mar 9 17:43:52 2011 Info: MID 487757 RID [0] Response '2.0.0 p2A1iFIH017778 Message
accepted for delivery'
Wed Mar 9 17:43:52 2011 Info: Message finished MID 487757 done
```

Cisco IronPort's Outbreak Filters use their own quarantine, which is separate from antispam or DLP policy quarantines where messages could be sitting. Messages that are deemed false positives can be released from quarantine manually with an option to send a copy of the message to Cisco for examination.

Users or administrators can easily report missed outbreaks by forwarding the emails as an attachment to outbreaks@ironport.com or phish@access.ironport.com, and forwarding false positives as an attachment to ham@access.ironport.com. Please note that Entourage and Microsoft Outlook both do not forward the emails with headers intact. Microsoft Outlook users should install the Spam and Encryption Flag plug-in software available from Cisco IronPort, which will properly forward the messages for action.

If a user tampers with a URL to try and use the proxy for other websites, access will fail. The rewritten URL has a hash that identifies the URL as one that is valid; if the URL is changed, the user will receive an error message from the proxy (Figure 5).

Figure 5. Altered URL Error Message



Cisco IronPort Outbreak Filters combine the best of Cisco IronPort and ScanSafe technology to offer superior protection against low-volume targeted attacks. For more information, please contact your local Cisco sales representative.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-684611-00 10/11