CLOUDMARK®

Clustered Email Environments
Aggressive Throttling on Multiple Time Windows

November 2014

Introduction

In the email world, spam and other malicious message types are an ongoing threat to customer satisfaction, provider capacity, and system security. While some semi-legitimate senders with fixed IP addresses send unwanted messages, much more traffic originates from large IP network ranges acquired by bad actors.

This paper considers techniques employed at a large service provider in order to minimize the damage done by massive spam campaigns initiated from vast numbers of IP addresses controlled by spammers.

Characteristics of the Attack

In many parts of the world it is extremely easy to gain access to large IP network ranges. As a consequence, a spammer in control of such a network has an enormous number of IP addresses available for sending malicious messages. Spammers are well aware that sophisticated providers will try to track the reputation of individual sending IP addresses in order to blacklist bad actors. The idea behind using a large number of sending addresses is to either overwhelm such reputation systems or to slip in "under the radar" with distributed campaigns. In either case, insufficient precision in reputation tracked over a window of time is being exploited to get the spam through.

In what follows, we describe two attack patterns and the manner in which they intend to gain advantage over the service provider's defenses and the ability of the Cloudmark Security Platform for Email to address them.

"Hailstorm" Attacks

In the "hailstorm" attack, the spammer initiates a buildup of many simultaneous connections from IP addresses under his control. In a coordinated fashion, all sending clients are triggered, dumping massive quantities of malicious messages into the provider's network over a very short amount of time. We have seen instances of hundreds of thousands messages arriving within a few minutes.

Such an attack typically succeeds for several reasons. First, per-IP address throttling limits may not trigger because spammers are very quick to adapt to such limits. Second, content filters require some amount of time to identify a stream of messages as spammy. Once again, spammers can stop their spam run before this happens.

As long as the spam run stops before the content filter indicates spam, the likelihood that the IP address would be blacklisted locally is decreased.

"Snowshoe" Attacks

During the "snowshoe" attack, only a small number of messages are sent from a very large number of IP addresses. At no point would per-IP address throttling limits be triggered. As in the "hailstorm" attack above, content filters will lag by a small amount of time. Should the spam run stop before the per-IP spam sending thresholds are tripped, the receiving system typically wouldn't blacklist the sending IP address.

This attack type may be aggravated if the service provider has a large multi-node cluster, but the nodes do not share reputation information. Given the low sending volumes from any particular IP address, the same message type may be routed through different nodes in the cluster, never allowing a single node to accumulate the necessary reputation statistics needed to block the message.

What's the Problem?

Both of the above attack patterns depend on the fact that individual sending IP addresses are not blacklisted in time, or in fact, might never blacklisted at all. In the ideal case for the spammer, this allows him to reuse IP addresses in the future. In the worst case, should the provider blacklist some addresses, there will be plenty more addresses in reserve. This is compounded massively in an IPv6 environment where the address space is so large that it becomes impossible to track the reputation of individual IP addresses.

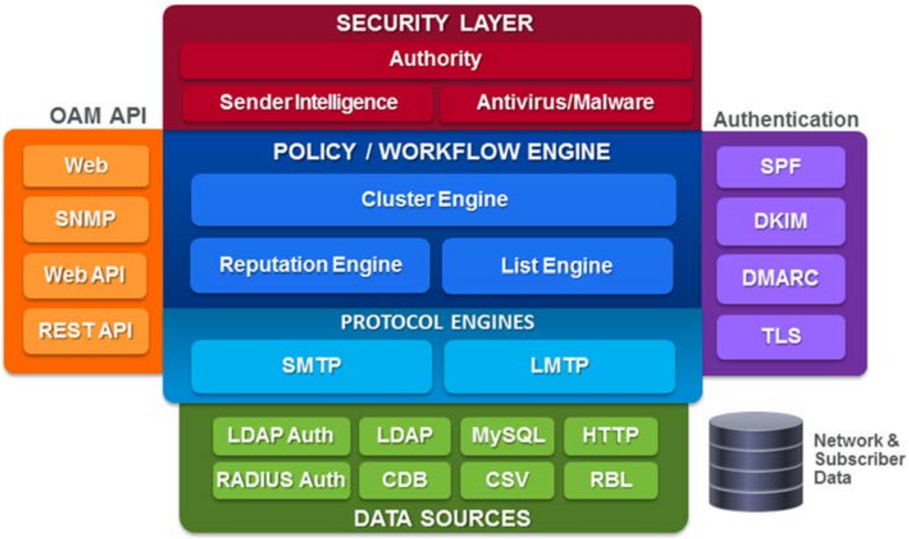
Working together with the provider, we came to the conclusion that the overarching problem was a lack of precision of reputation tracked on the part of individual IP addresses. Something more aggressive was required.

The Solution

Seeing that the tracking of reputation on a per-IP address basis was insufficient, we worked on taking a more aggressive stance. In principle, we decided to delineate the world into known senders and suspect senders. Known senders are treated separately based on historical sending patterns. All other sending IP addresses are considered potentially suspect should be aggressively throttled.

In considering the treatment of suspect senders, three observations were made. First, look at IP address ranges rather than individual IPs. Second, because an MTA node doesn't operate in isolation, it should share reputation statistics with all other nodes in the cluster. Third, in order to avoid spikes of spam activity, abnormal message delivery patterns should be forcibly distributed over time.

The actions taken in implementing the above observations in policy are described in the sections that follow. Making the changes was easy because the Cloudmark Security Platform for Email running within the provider network natively supports the functionality required to enforce these policy concepts.



Cloudmark Security Platform for Email

Reputation Across IP Address Ranges

Given that per-IP Address reputation statistics are insufficient, we began tracking reputation of IP addresses rolled up to IP address ranges. This means that the activity of a single IP address will count towards the total within the range, making it much more likely that the total will reach predefined throttling limits. It needs to be stated explicitly that whereas individual IP addresses may be blacklisted, we are instead proposing strict throttling limits to minimize the damage from spam sent from potentially bad IP address ranges. We are not advocating outright blocking of the range except in cases where a person makes the final decision based on additional information.

Two questions must be answered to make this approach successful. First, how does one identify a bad address range? We argue that a much more tractable problem is to first identify known senders and then assume that everyone else is potential suspect sender. Known senders may include known good senders, but may also include known bad senders who may be blocked outright. At the very least, known high-volume senders need to be identified. We don't want to lump their reputation statistics into IP address range aggregations. Lists of known senders can be built manually or an external feed may be used. For instance, the Cloudmark Sender Intelligence reputation feed provides information on known senders, including both volume and rating information.

The second question that comes up concerns the existence of legitimate senders within the aggregated IP address range. Wouldn't they be blocked as well? While a valid concern, the strategy is not to produce a blacklist. Rather, we're advocating heavy throttling, not outright blocking. This means that the occasional message from an unknown, but legitimate, sender is still very likely to get through.

Once known senders have been dealt with separately, the remaining IP addresses are aggregated into ranges to track reputation. As an example, by default, we used /24 aggregation in IPv4 and /32 aggregation in IPv6, while applying throttling limits commensurate with the size of the provider's installation. At the provider being discussed, the limits could be as low as 5,000 messages per aggregation per hour, guaranteeing that spammers cannot deliver an arbitrarily large number of messages during that time. Layering on top of outright sending limits, policy can also be implemented to bandwidth constrain SMTP traffic from these networks.

By taking this step, all suspect sending IP addresses contribute to the reputation statistics over a larger range, increasing the likelihood that a spam campaign will be quickly blocked by providing additional time for the content filtering engine to receive threat updates, maximizing its efficacy.

Cluster-Wide Reputation

In a large distributed cluster, it can be problematic if individual nodes maintain their own reputation statistics. Especially for low-volume sending IP addresses, there is no guarantee that any individual node will have the precision to take appropriate action when subsequent connections from the same IP address may be arbitrarily routed to an alternate system.

In this case, the provider has more than 50 MTA nodes in the cluster. By enabling cluster-wide reputation tracking in Cloudmark Security Platform for Email, each node gains directly from the observations made by other nodes.

By taking this step, clustered reputation statistics converge much more quickly than if locally observed reputation data was stored in isolation on every individual node. In this particular case, convergence was 50 times faster.

Reputation Over Time

The previous aggregation-based strategies proved very successful in reducing the absolute numbers of bad messages delivered into the provider's network. However, the provider still experienced spiky transmission patterns that seemed to let through more spam than expected.

As an example, when we tracked message volumes on a daily basis, we found large delivery spikes of spam at each day's boundary with an elevated likelihood that such messages would be delivered. As previously mentioned, content filters do the heavy lifting of protecting against spam messages that have evaded all other layers of defense, but they need time to gain precision.

By smoothing message delivery over the course of the day, we significantly reduced overall delivery of spam by giving the content filter time to gain enough precision to stop the attack. In our particular case, we configured the Cloudmark Security Platform for Email to track cluster-wide reputation statistics on a per-IP address range basis over 5 minute, 1 hour, and 24 hour time windows. Maximum sending limits were then applied over each time window.

The actual limit sizing is installation dependent, but consider the following example. If you want to enforce a limit of 10,000 messages in a 24-hour period, the 1-hour limit might be 1,000, while the 5 minute limit might be 250. Notice that we don't just divide the 24-hour limit by 24 in order to get the 1-hour limit. The 1-hour limit is proportionally higher than the 24-hour limit. This is to accommodate what might be a legitimate burst of traffic.

By taking this step, the probability that the system gains precision on a new spam campaign is greatly improved.

Conclusion

For unknown sending IP addresses, tracking of reputation statistics on a per-IP basis tends not to provide sufficient precision to identify bad actors and block spam. We used the native capabilities and flexible policy framework of the Cloudmark Security Platform for Email to efficiently gain necessary precision to track the aggregate behavior of a range of IP addresses. After excluding known senders, particularly big senders that can skew statistics, we tracked the reputation of aggregated IP address ranges and aggressively throttled message throughput in excess of limits commensurate with an installation's size.

By following the strategies described in this paper, we were able to reduce the amount of spam delivered into the provider's network by over 95% relative to previous attacks of the same nature

Americas Headquarters
Cloudmark, Inc.
San Francisco, USA

Europe
Cloudmark Europe Ltd.
London, UK

Paris
Cloudmark Labs
Paris, France

Japan
Cloudmark Japan
Tokyo, Japan

CLOUDMARK®

www.cloudmark.com