IRONPORT

IronPort Email Authentication

WHITE PAPER

TABLE OF CONTENTS

Executive Summary

The problems of spam, viruses, phishing and most email denial-of-service attacks can all be traced back to a single common cause – lack of authentication in the email protocol SMTP.

This lack of authentication means that a receiving mail server cannot reliably verify that a particular message is in fact from the sender it purports to be from, making it harder to identify friend from foe.

The industry has recognized this shortcoming, and a great deal of effort has been put into developing a new standard that will "overlay" SMTP and provide the sender authentication that is so desperately needed. This paper will present a brief history of how this problem evolved, explore the pluses and minuses of the leading standards proposals, and highlight some recommendations.

DEFINITIONS

Email nomenclature can be a bit confusing, so it is useful to start with some definitions. An email message has an addressing scheme similar to a postal message:

HELO/EHLO: The initial contact command between a sending and a receiving mail server, indicating an SMTP conversation.

Envelope sender[1]: The address of the sending mail server; not exposed to the end-user, used for managing bounces. For example, a commercial email from Citibank may well have an envelope sender from a service provider such as DoubleClick. Envelope sender is similar to the return address on a postal envelope.

"From:" address: This is a header in the message body that is displayed to the end-user. In many real-world-use cases, a "From:" address will not match the address of the server that delivered the message. In the postal example, the "From:" address is similar to the "From:" address on a formal business letter.

"Sender," "reply to," "resent sender" and "resent from" headers: Additional headers that can be inserted into a message body as a message and forwarded across the Internet. All are intended to specify the legitimate originator of a message in the event of a retry or a multi-hop message routing. None can be easily verified without one of the new authentication protocols.

HISTORY

Modern email history began in 1982 with the creation of RFC 821, "Simple Mail Transfer Protocol," and RFC 822, "Standard for the format of ARPA Internet Text Messages." This technology was designed to allow research colleagues at ARPANET to collaborate across unreliable data links in a totally trusted network. The resulting SMTP email protocol allows any host injecting mail into the mail system to identify itself as any arbitrary domain name during the SMTP conversation. Specifically, a sending mail server can assume the identity of any random sender for all sender attributes, including:

- HELO/EHLO domain
- Envelope sender
- "From", "sender," "reply to," "resent sender" and "resent from" headers

Technical solutions to the authentication problem began in 2003. The current proposal known as "SPF Classic" grew out of earlier proposals known as RMX (Reverse MX) and DMP (Designated Mailer Protocol) technologies.

IRONPORT

1 Envelope sender is also known as MAIL FROM or return path.

Microsoft's Caller ID technology merged with SPF to form Sender ID in 2004. In parallel, Yahoo! was developing a technology called DomainKeys, and Cisco was developing a similar technology called Internet Identified Mail. The two technologies were merged in July 2005 to form DomainKeys Identified Mail (DKIM). The current Request for Comment (RFC) drafts for Sender ID and DKIM have been submitted to the IETF, but none are yet approved standards. Links to these RFCs can be found in Appendix A.

THE AUTHENTICATION PROBLEM

Although the lack of authentication went largely unexploited for 20 years, the last few years have seen massive abuse of this weakness. Almost 80 percent of all email is spam, with the vast majority spoofing the sender's identity for all sender attributes. Spoofing of the sender's domain allows phishing email to defraud consumers, damage corporate brands and create bounce-based distributed denial-of-service attacks, and it makes spam more difficult to identify.

The email in Figure 1 shows a phish email that purports to be from Citibank. The email was actually sent from a compromised host in Taiwan[2].

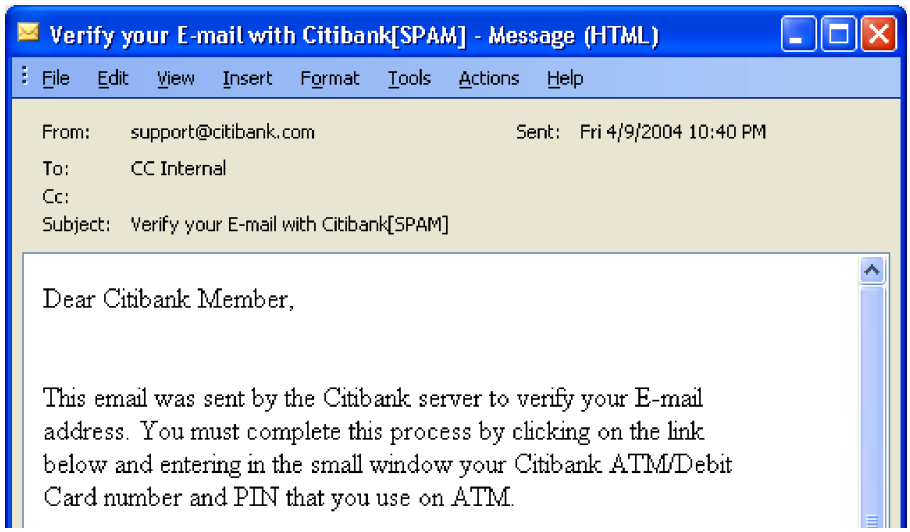
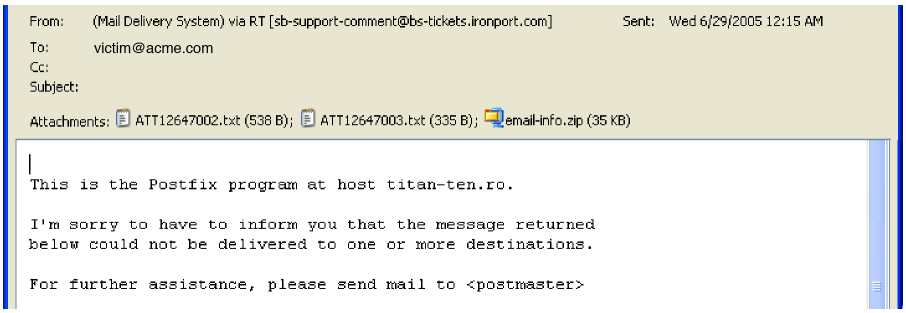


Figure 1: Email Spoofing the Citibank "From:" Header

2 The only unspoofable attribute of email technology is the IP address of the message transfer agent (MTA) sending the email (source IP address). IronPort Systems' SenderBase (www.senderbase.org) provides reputation, geographic location and other useful information for every IP address on the Internet sending email. See SenderBase for a report on the IP address that sent this Citibank phish. Without Sender ID or DKIM, Citibank cannot indicate that this compromised host in Taiwan is not authorized to send on their behalf.

The example in Figure 2 shows a misdirected bounce resulting from a spoofed envelope sender address.

Figure 2: Bounce Illustrating Results of a Spoofed Envelope Sender



In this case, a message with a spoofed envelope sender was sent to a recipient at the "titan-ten.ro" domain. The recipient address at titan-ten.ro was invalid, so the Postfix email program at titan-ten.ro bounced the email to the spoofed envelope sender, *victim@acme.com*. This technique is widely used to reflect bounces to an unsuspecting target. In this specific example, the original email was the MyTob virus and the bounce message included the actual virus in the attachment "email-info.zip."

SENDER ID AND DOMAINKEYS IDENTIFIED MAIL

Email authentication allows the recipient to determine whether an email is really from the purported sending domain. Two types of email authentication are primarily used: path-based and crypto-based.

> ***Sender ID*** *is a path-based authentication technology that authenticates the sending domain, based on the network path the email took. Network path is defined by source IP address.*
>
> ***DomainKeys Identified Mail (DKIM)*** *is a crypto-based authentication technology that authenticates the sending domain, based on a cryptographic signature contained within the email.*

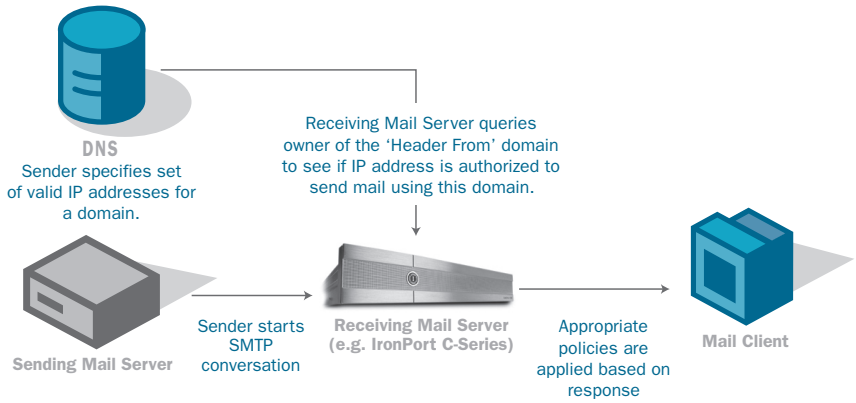
Other authentication technologies such as Client SMTP Validation (CSV), Bounce Address Tag Validation (BATV), MTA Mark and SES are not addressed in this white paper.

Sender ID

Sender ID technology allows email senders to associate their domains with allowed network paths (source IP addresses) for their email. A sender who wishes his or her email to be authenticated will publish a list of allowed network paths in the form of a DNS text record. A receiving mail server authenticates an email by extracting the sending domain from the email, querying the sender's DNS and extracting the allowed network paths (source

IP addresses) from the DNS text record. The receiving mail server then compares the IP address of the "last hop" sending mail server with these allowed network paths retrieved from the alleged sending domain. Email from allowed network paths is authenticated, while email from IP addresses that are not on the allowed network path fails authentication. Currently there are three Sender ID record types: "SPF1," "SPF2.0/mfrom" and "SPF2.0/pra." The fact that there are three separate and distinct types of SPF records can be a bit confusing, as all three are often referred to as simply "SPF" records, but they actually serve very different purposes.

Figure 3: Sender ID



Sender ID: Maps IP addresses to 'Header From' Domains

SPF classic and SPF2/mfrom are used to solve the problem of misdirected bounces. They do this by authenticating the HELO/EHLO hostname or envelope sender. Validation of the envelope sender allows recipients to avoid sending misdirected bounces because the true sender of the mail can be validated before a bounce is generated. Referring to the earlier example, if a spammer sends out messages with an envelope sender of *victim@acme.com*, the receiving mail server will look up the SPF records of acme.com and see that the IP address of the spammer's server attempting to deliver the mail is not on the allowed network path published by acme.com. Thus the message can simply be dropped. Another benefit of envelope sender authentication is that it happens before the message body is delivered. The receiving mail server will pause the SMTP conversation while it authenticates the envelope sender address. If the message is rejected, then the receiving mail server will have never accepted the body of the message, saving system resources and bandwidth since the body is typically the largest part of the message.

Envelope sender authentication has significant value because a spammer will often send out millions of messages with a single forged envelope sender — such as the example, *victim@acme.com*. This action results in millions of bounces delivered to acme.com from legitimate servers across the Internet, a massive distributed denial-of-service attack. Widespread adoption of SPF verification would eliminate misdirected bounce problems. Innocent compa-

nies can also end up on blacklists because misdirected bounces they sent hit spam traps or generated end-user complaints (some blacklists view a misdirected bounce as spam). SPF verification by receivers would identify forged return addresses and prevent this blacklisting.

SPF 1 and SPF2/mfrom focus on validating the envelope sender, or HELO domains, creating a solution to misdirected bounces. But these sender attributes are used only by the mail servers and are not exposed to end-users. So SPF1 and SPF2/mfrom are not useful in controlling phishing attacks, where the sender is forging the "F:" header that appears in the end-user client, such as the Citibank forgery in Figure 1.

SPF2/pra is designed to counter the phishing problem. It uses a mechanism similar to SPF2/mfrom to verify the "From:" address that is displayed to the end-user. A sender publishes a list of allowed network paths – a list of IP addresses of servers that can send mail on behalf of the domain owner and allow the domain owner to show up in the end-user's mail client. Authenticating the "From:" address has the added complexity of determining which "From:" address to use. A basic message will include a "From:" address of a sender. However, mail forwarding services and mailing lists will often insert another header such as "Resent from." SPF2 includes an algorithm called PRA that attempts to identify the most recent sender by parsing through any forwarding headers. The PRA then looks to authenticate the last hop server. For example, if a message was sent from ebay.com to a relay server at harvard.edu and finally delivered to acme.com, the mail server at acme.com would run the PRA to determine that the forwarding sender was harvard.edu. And if harvard.edu is also running the PRA it would have authenticated the message coming from ebay.com, creating a chain of trust. If the forwarding server at harvard.edu didn't run the PRA, then any spammer could send forged mail through the forwarding server and the receiving mail server would not be able to detect the forgery.

Technical Challenges with SPF/Sender ID

The complexities associated with forwarding mail and with third-party mailing lists pose a significant challenge to any path-based authentication scheme such as SPF and Sender ID. The core issue is that in order to positively authenticate either the "envelope sender" to thwart bounce problems or to authenticate the "From:" address to thwart phishing, the sender must publish a complete list of every allowed network path. Every possible mail forwarding service must be mapped and published by the sender – a task that is clearly not feasible. The authors of the specification have suggested that mail forwarding services need to run the PRA themselves and rewrite the mail they forward with the "Resent from" header. While this solution would be useful, there are a large number of forwarding services out there and getting them all to change in the near term is also not feasible. So SPF and Sender ID suffer from this significant challenge that is known in the industry as "the forwarding problem." The forwarding problem doesn't mean that

these important protocols have no value. As will be explained in the following section, these authentication schemes are very powerful when combined with sender reputation data.

DomainKeys Identified Mail (DKIM)

In contrast to the path-based authentication of SPF/Sender ID, DKIM uses a cryptographic stamp to authenticate message senders. In the DKIM scheme, a sender makes a hash of every outgoing message and encrypts that hash using a private key in a PKI pair. The public key from the pair is then published in a DNS text record. A receiving mail server authenticates the message by extracting the sending domain from the email, retrieving the public key from the DNS text record and validating the signature against the message's contents. Email with a valid signature is authenticated and email with an invalid signature fails authentication.

Figure 4: DKIM Signing

Figure 5: DKIM Validation

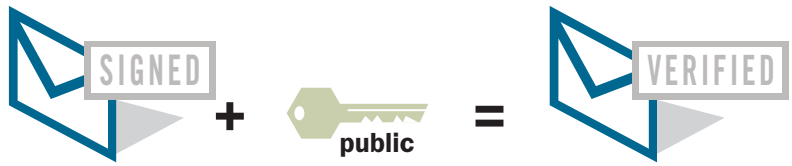
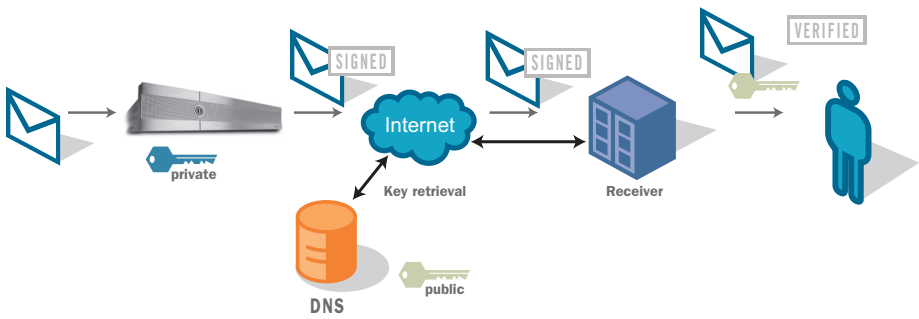


Figure 6: DKIM In Action



Although the protocol is flexible, DKIM almost always validates the domain in the "From:" header. Since this header is always presented in email clients to end-users, this technology is an excellent solution to the phishing problem. This technology is not currently used to limit misdirected bounces and requires the entire message to be received with full signature validation before any validation data can be extracted.

The advantage of DKIM is that it is a robust solution to the mail forwarding problem. A message can be forwarded repeatedly and the digital signature travels right along with it. If a spammer attempts to modify or manipulate a message in transit, it will break the decryption process and fail authentication. The downside of this solution is that if a forwarding service modifies the message body at all, it will also fail decryption, but this is a matter of ensuring that forwarding mail servers remain RFC compliant. One notable challenge is mailing lists. A mailing list will often modify a message by adding unsubscribe links or sponsorships. Thus mail forwarded through a mailing list will not pass authentication. A simple solution is to not sign mail bound for well-known mailing lists.

Perhaps the greatest challenge to DKIM acceptance is the performance impact associated with scanning, encrypting and decrypting messages. Estimates range for performance impacts of anywhere from 20 percent to 50 percent for DKIM signing on open source MTAs. Furthermore, DKIM can be a fairly complex system to set up. The good news is that commercial vendors, such as IronPort® Systems, are building high-performance DKIM solutions that are easy to set up and use.

Sender ID and DKIM are often compared side by side. But in many ways they are different answers to different problems. Sender ID has the advantage of being easy to implement and therefore has higher adoption rates. It struggles with complexities associated with the various sender identity attributes in SMTP and the use cases around forwarding and mailing lists. DKIM is harder to implement, yet once implemented is more effective in its use case. DKIM in its current form is only useful for stopping phishing – Sender ID attempts to solve phishing and bounce problems. A summary of these approaches is provided in Table 1.

Table 1. Advantages and Disadvantages of Sender ID, SPF and DKIM

	SENDER ID	SPF	DKIM
Validation Method	PRA domain against source IP via DNS (PRA is most commonly "From:" header)	Envelope sender against source IP via DNS	Cryptographic signature in email header using DNS public key
Strengths	Addresses phishing problem Very easy to deploy	Eliminates misdirected bounces Checking is performed before message data is received	Addresses phishing problem More robust, unaffected by multiple SMTP hops
Weaknesses	Does not address misdirected bounce problem Validates ONLY last hop	Does not address misdirected bounce problem Validates only last hop	More difficult to implement, both sending and receiving
Challenges	Mail forwarding can cause validation failure	Mail forwarding will cause validation failure	Mailing lists can cause validation failure. Any modification to message in transit will cause validation failure

ADOPTION STATUS

Sender ID was first to market, is easier to implement and is more prevalent than DKIM. IronPort recently published a study showing that one-third of all email has a Sender ID record[3]. Most major ISPs and many large corporations publish Sender ID records, with Microsoft being the leading proponent. In June 2005, Hotmail and MSN introduced a Sender ID pass/fail indication in their email client. Key findings from the IronPort study include:

- 35 percent of all Internet email is now authenticated using SIDF
- 75 percent of all Fortune 100 companies use SIDF for marketing related email
- Nine of the top ten most phished domains use SIDF

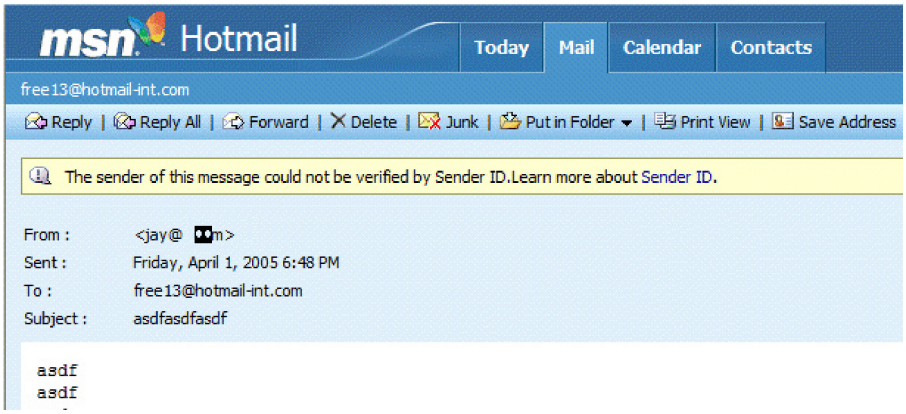


Figure 7: Hotmail Client Indicating Sender ID Verification Failure

DomainKeys Identified Mail adoption is more difficult to measure, but recent studies indicate that 10 percent of all Internet email now uses DKIM. Yahoo!, Earthlink and Google's Gmail have been the leading adopters, with each indicating DKIM pass/fail in their email clients. Yahoo! receives over one billion messages per day signed with DKIM (Figure 8). The IronPort study also indicates:

- 45 percent of all Fortune 100 companies use DKIM marketing related email
- Adoption of DKIM has increased from an negligible value in just nine months
- Five of the top ten most phished domains use DKIM

IRONPORT

3 "Email Authentication: Sender ID Framework (SIDF) and DomainKeys Identified Mail (DKIM) Scorecard." This study is available to download at: www.ironport.com/authentication

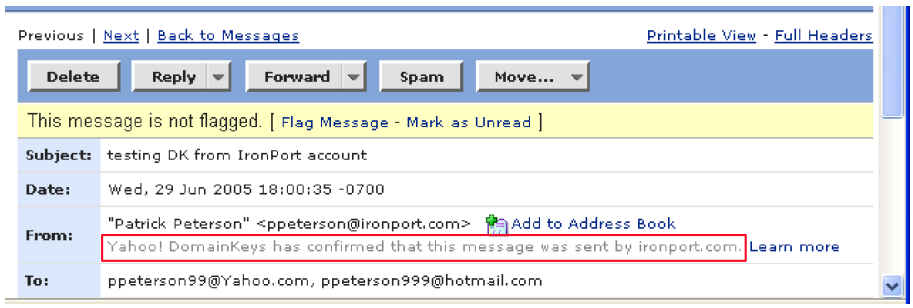


Figure 8: Yahoo! Client Indicates DomainKeys Validation

WHY AUTHENTICATE?

Sender authentication is a long-term solution to the email security problem. It is not a quick fix – it will take years for the global email infrastructure to be fully upgraded. However, there is significant value in publishing authentication records today. For a receiver, most of the benefit of authentication is associated with a negative authentication – being able to reject or discard messages that do not authenticate properly. With path-based systems, there are still too many legitimate use cases (such as mail forwarding) that would cause a "false negative" authentication. Thus most receivers won't discard a negative Sender ID authentication.

However, in order to drive adoption, major ISPs are exposing positive authentication to end-users – a message that authenticates correctly will have a trust flag displayed to the end-user. Over time, more large ISPs are expected to follow this practice, which will begin to address the core issue of consumer trust. When consumers learn that they can trust an authenticated email, the absence of authentication will become suspicious. Establishing trust involves a shift in consumer behavior, but it is not outrageous to think that in 24 months consumers will expect to see positive authentication in email from corporate mail systems; therefore, corporations are well advised to begin authenticating now.

The other reason to adopt an authentication system now is that, while a negative authentication is not always enough data for a receiver to drop a message, it is an important data point. When authentication is combined with sender reputation, an intelligent mail system can make more accurate decisions.

IronPort Systems invented the concept of reputation more than three years ago, and in that time every major email security vendor has followed to one degree or another. A reputation system looks at a wide range of data to assess the behavior of a given mail sender. IronPort's SenderBase® Network collects more than 120 data points from more than 100,000 networks to characterize the trustworthiness of any given sender. This data is rolled into a score of -10 to +10, and is made available to IronPort appliances receiving mail. The appliances then have the ability to "push back," or rate limit senders, based on how suspicious they appear.

Combining this data with authentication yields powerful results. If an SPF record doesn't authenticate, the forwarding problem makes it difficult to say with certainty that the message is invalid and should be dropped. But if an SPF record doesn't authenticate and the sending mail server has a very low reputation score, the message can be dropped with confidence. Likewise, an SPF record that doesn't authenticate but the sending mail server has a very positive reputation score (such as a Fortune 500 company) the mail will be accepted. When combined with an intelligent reputation system, path-based authentication can be very effective.

THE SOLUTION TO BOUNCE ATTACKS

Bounce Address Tag Validation (BATV) is an initiative in the Internet community to address bounce attacks. BATV provides email administrators the tools required to protect themselves from bounce attacks, with minimal overhead and no ongoing maintenance.

BATV digitally signs the envelope's return address (SMTP's Mail From:) with a private key.

Normally, the envelope's return address is:

MAIL FROM: support@bigbank.com

BATV converts the address to:

MAIL FROM: pvrs=support=3201EA1CF@bigbank.com

When bounce messages are returned to an email gateway, the existence of the correct signature determines legitimate bounces from bounces that didn't originate from the organization's domain.

IronPort's Bounce Verification technology represents the first appliance-based solution compliant with BATV. IronPort also provides additional functionality to support deleting, quarantining or marking the subject line of fraudulent bounces.

What is unique about IronPort Bounce Verification™ is that, unlike other email authentication technologies, it does not require industry adoption to be effective. The uni-directional nature of IronPort Bounce Verification provides immediate benefit to those who deploy the technology.

IRONPORT SYSTEMS' ADOPTION RECOMMENDATIONS

With spam making up 80 percent of all email, massive credit card theft from phishing attacks and nonstop virus attacks, email is clearly broken. Email authentication technologies offer a critical solution to fixing email. IronPort has been active in development and testing of email authentication technologies from the beginning. The IronPort security gateway appliances currently

support a range of technologies, including a sophisticated feature set for signing outbound email with DomainKeys and IronPort Bounce Verification technology.

Technical challenges need to be overcome, but a separate challenge is the adoption problem. With the exception of BATV, authentication solutions require senders to publish records (Sender ID) and sign their email (DKIM). They also require receivers to validate the email. Lack of adoption by either senders or receivers hinders the value of the system to everyone – a classic "chicken and egg" problem.

Publishing records and signing email is essential today. This approach is particularly beneficial to widely phished financial services companies and companies using email marketing, enabling them to reduce phishing and brand abuse. All senders should inventory their domains and mail servers and publish Sender ID records as soon as possible. Companies whose existing solution does not support DKIM signing should begin investigating options to add this capability to their infrastructure (see Figure 9).

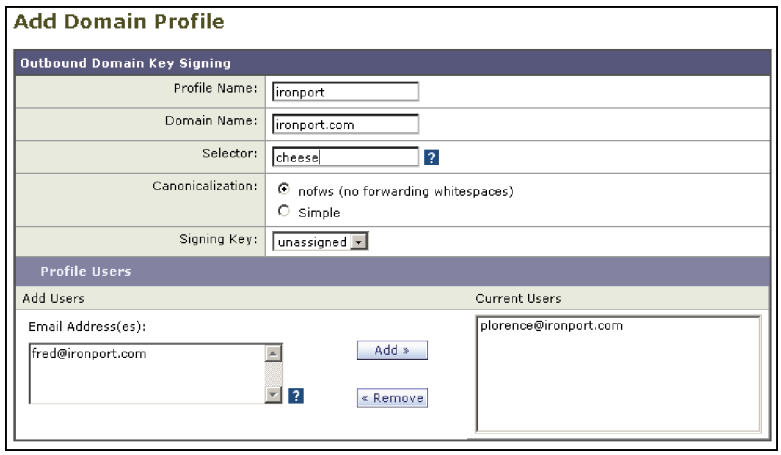


Figure 9: IronPort's "point and click" DKIM Signing Technology

On the receive side, positive authentication is useful today; however, more adoption and technology enhancements will be required before negative authentication can be used as a strong indication of invalid mail. While there is clearly a strong correlation between illegitimate messages and authentication failures, the technology is not yet reliable enough to be a stand-alone factor in message disposition decisions.

For these reasons, IronPort recommends caution in deploying email authentication on the receive side. To address bounce attacks, deploy BATV-related technologies to filter bounces at the perimeter. When it comes to phishing and other spoofing attacks, validation cannot be used as a sole criterion to reject mail. Sophisticated senders should implement solutions that use validation results as one ingredient in a disposition decision. Receivers should be cautious in passing Sender ID and DKIM results, based on mail headers to end-users, as the authenticated header might not be displayed in the user's email client.

IRONPORT

IronPort has been using Sender ID records as a factor in SenderBase Reputation Scores since 2004. IronPort's sophisticated, multifaceted variable response to threats allows the administrator to extract as much protection from the technologies as possible, mainly by factoring authentication into reputation analysis. IronPort expects this protection will increase over the next 12 months, as adoption grows.

APPENDIX

Appendix A: Current Email Authentication Draft Specifications

Sender ID resources:

- http://www.microsoft.com/mscorp/safety/technologies/senderid/resources.mspx

DKIM resources:

- http://mipassoc.org/dkim/ietf-dkim.htm

BATV resources:

- http://mipassoc.org/batv/index.html

Appendix B: Email Authentication Implementation Guide

Publish SPF/Sender ID records as soon as possible. The steps required to publish these records include:

1. Determine domain(s) to authenticate
   - Inventory IP addresses of all servers sending email.
   - IP addresses controlled by your organization can easily be extracted via SenderBase (www.senderbase.org).
   - Identify any IP addresses of international locations.
   - Identify any external sources of email, such as email marketing.
   - Identify any other sources, such as third parties allowed to spoof your domain, clients allowed to send directly to the Internet, etc.
2. Create SPF records using SPF record-creation wizards and the IP addresses identified in step 1 above. If a receiver receives an email that is not from one of the domain's stated IP addresses, it may be a result of spoofing or mail forwarding. For this reason, IronPort suggests starting by publishing records, recommending a neutral action if the SPF validation fails. This strategy can be modified to a soft fail and then fail as the deployment matures.

3. Publish SPF records in DNS. Each record is a simple text record published in the domain's zone.

Begin a project to sign outbound email. This step requires deploying gateway security technology capable of signing outbound messages. The technology will use the private key, and the public key must be added to the domain's DNS record. As with SPF/Sender ID, start by treating failures neutrally and gradually increase the severity of an authentication failure.

IRONPORT

IronPort Systems
950 Elm Avenue, San Bruno, California 94066
TEL 650.989.6500 FAX 650.989.6543
EMAIL info@ironport.com WEB www.ironport.com

IronPort Systems, a Cisco business unit, is a leading provider of anti-spam, anti-virus and anti-spyware appliances for organizations ranging from small businesses to the Global 2000. IronPort appliances utilize SenderBase, the world's largest email and Web threat detection network and database. IronPort products are innovative and easy-to-use—providing breakthrough performance and playing a mission-critical role in a company's network infrastructure.

IronPort is now part of Cisco.

CISCO