Messaging, Malware and Mobile Anti-Abuse Working Group

M³AAWG Initial Recommendations for Addressing a Potential Man-in-the-Middle Threat

July 2015

Introduction

The messaging community has made impressive progress when it comes to encouraging deployment of opportunistic (best effort) encryption of email traffic. Opportunistic encryption, as described in the TLS for Mail: M³AAWG Initial Recommendations and IETF Opportunistic Security: Some Protection Most of the Time documents, however, is *not* necessarily secure against Man-in-the-Middle (MITM) attacks.

To understand why this is true, it is necessary to consider what typically happens when opportunistic encryption cannot be adequately negotiated. In that case, MTA-to-MTA (mail-transfer-agent to mail-transfer-agent) transmissions normally fall back to sending email traffic in plain text – that is, totally unencrypted. Consequently, a provider's choice lies between tolerating best-effort encryption or else living with no encryption at all. That is not much of a choice and this paper assumes opportunistic TLS is the preferred option. That being said, even though opportunistic encryption protects the messages during transmission from sender to receiver, it is still possible for a MITM attacker with a self-signed certificate to impersonate the intended destination.

This brief document describes the MITM situation and various methods that bad actors can use to conduct MITM attacks and covers the components for deterring these attacks. It also introduces a new technology, DANE (DNS-based Authentication of Named Entities), that can assist messaging providers in validating they are communicating with an intended destination when using SSL/TLS.

Mitigating Man-in-the-Middle (MITM) Attacks

In MITM attacks, adversaries interpose themselves between the sender of a message and the intended recipient of that message:

The following methods have all been used by bad actors to come between senders and receivers. This list should not be considered exhaustive:

1. ARP spoofing[1]
2. Rogue DHCP servers[2]

[1] ARP spoofing, http://en.wikipedia.org/wiki/ARP_spoofing
[2] Rogue DHCP, http://en.wikipedia.org/wiki/Rogue_DHCP

3. Web Cache Communication Protocol (WCCP)[3]
4. Web Proxy Autodiscovery Protocol (WPAD)[4]
5. Spoofed WiFi wireless access points ("evil twin" access points)[5]
6. DNS poisoning[6]
7. BGP route injection[7]
8. Physical (inline) network traffic interception devices

This discussion does not consider interception attacks that are executed on the endpoint itself nor does it consider Man-in-the-Browser attacks and the like. As with any technology, without secure endpoints there can be no assurance of general data security.

Risks of MITM Attacks

If adversaries are able to successfully execute a MITM attack against clear-text network traffic, they can eavesdrop on that traffic, modify it and/or impersonate parties to the communication. If the traffic is encrypted in transport but the endpoints are not cryptographically protected against MITM attacks, an adversary can execute many of the same attacks against this encrypted traffic as it can against clear-text traffic. It is therefore extremely important that cryptographically protected transmissions be protected against MITM attacks as well.

In an ideal world, traffic would be protected end-to-end by users implementing PGP/GPG or S/MIME and server-to-server traffic would also be protected with SSL/TLS as well. But most users do not use PGP/GPG or S/MIME. This makes server-to-server encryption that is less susceptible to MITM attacks all the more essential. Messaging providers who want to deter MITM attacks can help protect server-to-server traffic by using methods where:

1. All mail servers identify themselves using a globally trustworthy certificate—that is, the server uses a certificate signed by a globally trusted certificate authority.

2. The name of the server corresponds to one of the domain names for which the certificate was issued (the server and certificate "match").

3. The Online Certificate Status Protocol (OCSP) and/or a Certificate Revocation List (CRL) has been checked and the certificate has not been revoked.

4. The certificate is not used before it is first valid nor after it has expired.

5. The certificate is signed using an industry-standard (SHA-2) signature[8].

---

[3] Web Cache Communication Protocol, http://en.wikipedia.org/wiki/Web_Cache_Communication_Protocol
[4] Web Proxy Autodiscovery Protocol, http://en.wikipedia.org/wiki/Web_Proxy_Autodiscovery_Protocol
[5] Evil twin (wireless networks), http://en.wikipedia.org/wiki/Evil_twin_%28wireless_networks%29
[6] DNS spoofing, http://en.wikipedia.org/wiki/DNS_spoofing
[7] Kim Zitter, "Revealed: The Internet's Biggest Security Hole," *Wired Magazine*, August 26, 2008, http://www.wired.com/2008/08/revealed-the-in/
[8] SHA-2, https://en.wikipedia.org/wiki/SHA-2

6. The certificate covers a strong (2048- or 4096-bit RSA) key pair.

7. The originating and receiving mail server support the most recent version of the TLS protocol (TLS 1.2 at the time this document was published).

8. The servers mutually agree on using a cipher suite that supports forward secrecy[9] for the purpose of key exchange (normally Ephemeral Diffie-Hellman [EDH][10] or Elliptic curve Diffie-Hellman Ephemeral [ECDHE][11]).

9. A strong symmetric cipher is negotiated (ideally AES-128 or AES-256).

If *any* of the preceding conditions are not satisfied between the sending MTA and the receiving MTA, the sending server should *not* transmit the message to the receiving MTA.

Disposition of Messages That Cannot Be Securely Conveyed

If the sending server cannot transmit a message to the intended receiving server how can the non-deliverable message be securely processed? Options might hypothetically include:

1. The message can be rejected outright and returned to the sender for processing, assuming the sending host and the receiving host reach an agreement that they CANNOT securely exchange a message while a connection is still established. Messages that cannot be securely delivered shall NOT be bounced to apparent message body senders (due to the possibility of a spoofed apparent sender).

2. Alternatively, the message can be temporarily queued and then retried one or more times, thereby helping to address transient non-deliverability.

3. After proceeding with either step (1) or (2) above, the message can be summarily dropped. This presumes the sender has an application-level delivery confirmation mechanism to detect silent non-deliveries if/when they occur.

In general, non-deliverable messages should be handled consistent with the recommendations found in Section 3.8 of the M[3]AAWG Sender Best Common Practices.[12]

Future Directions with DANE

DNS-based Authentication of Named Entities (DANE[13, 14]) is an IETF proposal for a method that allows certificates to be bound to DNS (Domain Name Server) names using DNSSEC. DANE allows a site to specify the certificate it uses and that should be seen by third parties interacting with that site. The identity

---

[9] Forward secrecy, http://en.wikipedia.org/wiki/Forward_secrecy
[10] Diffie-Hellman key exchange, https://en.wikipedia.org/wiki/Diffie-Hellman_key_exchange
[11] Elliptic curve Diffie-Hellman, https://en.wikipedia.org/wiki/Elliptic_curve_Diffie-Hellman
[12] M[3]AAWG Sender Best Common Practices, Version 3.0, Updated February 2015, https://www.m3aawg.org/sites/maawg/files/news/M3AAWG_BCP_Ver3-2015-02.pdf
[13] RFC 6698: https://tools.ietf.org/html/rfc6698
[14] RFC 7218: https://tools.ietf.org/html/rfc7218

of that certificate is specified by special records contained in the DNS. DNSSEC allows third parties to trust the DANE records that are published in the DNS. DANE will be explored more fully in a separate M³AAWG document.

Conclusion

Deploying opportunistic encryption as described in TLS for Mail: M³AAWG Initial Recommendations is an excellent way to start protecting email traffic between providers. It is not, however, designed to thwart more sophisticated Man-in-the-Middle attacks. The Messaging, Malware and Mobile Anti-Abuse Working Group recommends that industry messaging providers use the principles outlined in this paper to fight MITM attacks by paying closer attention to certificates and how they are validated; that is, by tying an identity to a cryptographic key pair. These guidelines are not intended to be viewed as comprehensive and M³AAWG is working to create additional guidance to improve protection of user messaging.

References

[1] ARP spoofing, http://en.wikipedia.org/wiki/ARP_spoofing

[2] Rogue DHCP, http://en.wikipedia.org/wiki/Rogue_DHCP

[3] Web Cache Communication Protocol, http://en.wikipedia.org/wiki/Web_Cache_Communication_Protocol

[4] Web Proxy Autodiscovery Protocol, http://en.wikipedia.org/wiki/Web_Proxy_Autodiscovery_Protocol

[5] Evil twin (wireless networks), http://en.wikipedia.org/wiki/Evil_twin_%28wireless_networks%29

[6] DNS spoofing, http://en.wikipedia.org/wiki/DNS_spoofing

[7] Kim Zitter, "Revealed: The Internet's Biggest Security Hole," *Wired Magazine*, August 26, 2008, http://www.wired.com/2008/08/revealed-the-in/

[8] SHA-2, https://en.wikipedia.org/wiki/SHA-2

[9] Forward secrecy, http://en.wikipedia.org/wiki/Forward_secrecy

[10] Diffie-Hellman key exchange, https://en.wikipedia.org/wiki/Diffie-Hellman_key_exchange

[11] Elliptic curve Diffie-Hellman, https://en.wikipedia.org/wiki/Elliptic_curve_Diffie-Hellman

[12] M³AAWG Sender Best Common Practices, Version 3.0, Updated February 2015 https://www.m3aawg.org/sites/maawg/files/news/M3AAWG_Senders_BCP_Ver3-2015-02.pdf

[13] RFC 6698: The DNS-Based Authentication of Named Entities (DNS) Transport Layer Security (TLS) Protocol: TLSA, https://tools.ietf.org/html/rfc6698

[14] RFC 7218: Adding Acronyms to Simplify Conversations about DNS-Based Authentication of Named Entities (DANE), https://tools.ietf.org/html/rfc7218

M³AAWG097