Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 16, 2015

V. Dukhovni
Two Sigma

August 15, 2014

Opportunistic Security: Some Protection Most of the Time
draft-dukhovni-opportunistic-security-03

Abstract

This memo introduces the concept "Opportunistic Crypto-Security" (OCS). OCS is a set of protocol design principles that attempt to remove barriers to the widespread use of encryption on the Internet. OCS is not a protocol. Protocols that adhere to OCS guidelines may offer additional crypto-security services, e.g., integrity and authentication, if these services are supported by all parties to a communication. The OCS design philosophy departs from the common practice of other Internet security protocols; they commonly require cryptographic protection against both passive and active attacks, or offer no protection at all. OCS protocols strive to offer encryption even if authentication is not available. This document encourages designs in which cryptographic protection against both passive and active attacks can be deployed incrementally, without creating barriers to communication.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 16, 2015.

Copyright Notice

Table of Contents

1. Introduction

The development of Opportunistic Crypto-Security (OCS) is motivated by the concerns raised in [RFC7258]. Pervasive monitoring (as defined in that RFC) is feasible because of the lack of widespread use of encryption for confidentiality. Although the IETF has developed many security protocols (e.g., TLS, IPsec, SSH, …) that employ encryption for confidentiality, most of them also require one-way or two-way authentication. Authentication is mandated by the protocols to protect against active attacks. If communicating peers are unable to meet the authentication requirements imposed by these protocols, the result may be no communication, or plaintext communication.

The ability to authenticate any potential peer on the Internet requires an authentication mechanism that encompasses all such peers. No IETF standards for authentication meet this criteria. The Public Key Infrastructure (PKI) model employed by browsers to authenticate web servers (often called the "Web PKI" [cite]) imposes cost and management burdens that have limited its use. The trust-on-first-use (TOFU) authentication approach assumes that an unauthenticated public key obtained on first contact (and retained for future use) will be good enough to secure future communication. TOFU-based protocols, e.g., SSH [cite] work well in enterprise environments, but were not designed to scale for Internet-wide use.

DNS-Based Authentication of Named Entities (DANE [RFC6698]) defines a way to distribute public keys bound to DNS names. It can provide an alterative to the Web PKI (for other than EV certificates [cite]). DANE should be used in conjunction with DNSSEC [RFC4033]. At time, DNSSEC is not sufficiently widely deployed to allow DANE to satisfy the Internet-wide, any-to-any authentication criteria noted above. Thus protocols that mandate authenticated communication cannot generally do so via DANE (at time).

OCS provides a near-term approach to removing barriers to widespread use of encryption, while offering a path to authenticated, encrypted communication in the future. The primary goal of OCS is to counter attacks, consistent with the goals established in [RFC7258]. However, OCS does not preclude offering protection against active attacks, if suitable authentication capabilities are available. OCS is not intended as a substitute for authenticated, encrypted communication when such communication is already available to peers, e.g., based on TLS, IPsec, SSH, etc.

To achieve widespread adoption, OCS must support incremental deployment. Incremental deployment implies that security capabilities will vary from peer to peer, perhaps for a very long time. Thus use of an OCS protocol by one peer may yield communication that is unauthenticated but encrypted, authenticated and encrypted, or plaintext. This last outcome will occur if not all parties to a communication support OCS (or if an active attack makes it appear that this is the case). OCS protocols will attempt to establish authenticated, encrypted communication whenever both parties are capable of such, but will fallback to unauthenticated encrypted communication if authentication is not possible. Fallback to plaintext communication will occur as noted above.

OCS protocols do not prohibit the use of local security policies. A security administrator may specify security policies that override opportunistic security. For example, a policy might require authenticated, encrypted communication, in contrast to the default OCS security policy.

The remainder of this document provides definitions of critical terms, enumerates the OCS design principles/guidelines, and provides an example of an OSC design, in the context of communication between mail relays.

2. Terminology

Perfect Forward Secrecy (PFS): As defined in [RFC4949].

Man-in-the-Middle (MiTM) attack: As defined in [RFC4949].

Trust on First Use (TOFU): In a protocol, TOFU calls for accepting and storing a public key/credential associated with an asserted identity, without authenticating that assertion. Subsequent communication that is authenticated using the cached key/credential is secure against an MiTM attack, if such an attack did not succeed during the (vulnerable) initial communication. The SSH protocol makes use of TOFU. The phrase "leap of faith" (LoF, [RFC4949]) is sometimes used as a synonym.

[note that this is still not quite correct. In an enterprise environment it is common for the enterprise to provide an out-of-band means of verifying the asserted identity, e.g., based on the hash of the public key.

One-way and Two-way Authentication <fill in>

3. Opportunistic Crypto-Security Design Principles

As noted in Section 1, OCS aims to remove barriers to the widespread use of encryption on the Internet. A secondary goal is protection against active attacks, by enabling incremental deployment of authenticated, encrypted communication. OCS seeks to achieve the best protection possible, based on the capabilities of communicating peers.

1. Determine Peer Security Capabilities: An OCS protocol first determines the capabilities of the peer with which it is attempting to communicate. Peer capabilities may be discovered by out-of-band or in-band means. (Inband determination implies negotiation between peers. Out-of-band mechanism include the use of DANE records or cached keys/credentials acquired via TOFU.) The capability phase determination may indicate that the peer supports authenticated, encrypted communication, unauthenticated encrypted communication, or only plaintext communication. (Note that use of out-of-band capability determine, e.g., DANE or TOFU, is downgrade resistant, and thus preferred over in-band negotiation techniques. The goal of this design principle is to maximize the offered security services on a pairwise, peer basis.

2. Apply Security Policy: Having determined peer security capabilities, an OCS protocol next applies any local security polices in addition to the default OCS policy (see below). Local policies may require security services in addition to encryption, e.g., authentication. A policy might restrict the set of algorithms that are employed (for encryption, authentication, integrity, etc.) The OCS default policy is simple: establish encrypted communication if possible; authenticate the peer if the capability exists; revert to plaintext if encrypted communication is not possible. Reverting to plaintext merely because authentication was not possible is inconsistent with the default policy! However, explicit, local policy overrides the default OCS policy.

3. Employ Perfect Forward Secrecy: OCS protocols SHOULD employ PFS to protect previously recorded encrypted communication from decryption even after a compromise of long-term keys.

4. No misrepresentation of security: Unauthenticated encrypted communication must not be misrepresented to users (or in logs) of non-interactive applications as equivalent to communication over an authenticated encrypted channel. This principle is consistent with the goal of not encouraging use of OCS in lieu of protocols that offer additional security services, when such protocols can be employed successfully.

4. Example: Opportunistic TLS in SMTP

Many Message Transfer Agents (MTAs, [RFC5598]) support the STARTTLS ([RFC3207]) ESMTP extension. MTAs acting as SMTP clients are generally willing to send email without TLS (and therefore without encryption), but will employ TLS (and therefore encryption) when the SMTP server announces STARTTLS support. Since the initial ESMTP negotiation is not cryptographically protected, the STARTTLS advertisement is vulnerable to MiTM downgrade attacks. Further, MTAs do not generally require peer authentication. Thus the use of STARTTLS for SMTP protects only against passive attacks.

MTAs that implement STARTTLS establish either an authenticated, encrypted session or deliver messages over a plaintext channel. Recent reports [cite?] from a number of large providers suggest that the majority of SMTP email transmission on the Internet is now encrypted, and the trend is toward increasing adoption.

The STARTTLS advertisement is vulnerable to active attacks and some MTAs that advertise STARTTLS exhibit various interoperability problems in their implementations. As a result, it is common for a pair of STARTTLS-enabled MTAs to fall back to plaintext communication when the TLS handshake fails, or when TLS fails during message transmission. This is a reasonable trade-off, consistent with OCS principles, since STARTTLS protects against only passive attacks; absent an active attack TLS failures are simply interoperability problems.

Some MTAs employing STARTTLS abandon the TLS handshake when the peer MTA fails authentication, only to immediately deliver the same message over a plaintext connection. Other MTAs have been observed to tolerate unverified self-signed certificates, but not expired certificates, again falling back to plaintext. These and similar behaviors are NOT consistent with OCS principles, since they revert to plaintext communication when authentication fails, instead of employing unauthenticated, encryption, communication.

Protection against active attacks for SMTP is described in [I-D.ietf-dane-smtp-with-dane]. That draft introduces the terms "Opportunistic TLS" and "Opportunistic DANE TLS"; this draft is consistent with the OCS design principles defined in this document.

6. Security Considerations

OCS supports communication that is authenticated and encrypted, unauthenticated and encrypted, or plaintext. The security services offered to communicating peers is not reduced by the use of OCS. This is because the default OCS policy employs the best security services available based on the capabilities of the peers, and because local security policies take precedence over the default OCS policy. OCS is an improvement over the status quo; it provides better security than the alternative of providing no security services when authentication is not possible (and not strictly required)

OCS coexists with and is preempted by local, non-OCS security polices. Non-OCS policies may inhibit use of encryption when many peers cannot offer authenticated, encrypted communication. Unless authenticated, encrypted communication is necessary, non-OCS local policies of this sort run counter to the goals established in [RFC7258].

7. Acknowledgements

I would like to thank Steve Kent. Some of the text in this document is based on his earlier draft. I would like to thank Dave Crocker, Peter Duchovni, Paul Hoffman, Scott Kitterman, Martin Thomson, Nico Williams, Paul Wouters and Stephen Farrell for their helpful suggestions and support.

8. References

8.1. Normative References

[RFC3207] Hoffman, P., "SMTP Service Extension for Secure SMTP over Transport Layer Security", RFC 3207, February 2002.

[RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, March 2005.

[RFC5116] McGrew, D., "An Interface and Algorithms for Authenticated Encryption", RFC 5116, January 2008.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

[RFC6125] Saint-Andre, P. and J. Hodges, "Representation and Verification of Domain-Based Application Service Identity within Internet Public Key Infrastructure Using X.509 (PKIX) Certificates in the Context of Transport Layer Security (TLS)", RFC 6125, March 2011.

[RFC6698] Hoffman, P. and J. Schlyter, "The DNS-Based Authentication of Named Entities (DANE) Transport Layer Security (TLS) Protocol: TLSA", RFC 6698, August 2012.

8.2. Informative References

[I-D.ietf-dane-smtp-with-dane]
Dukhovni, V. and W. Hardaker, "SMTP security via opportunistic DANE TLS", draft-ietf-dane-smtp-with-dane-11 (work in progress), August 2014.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", RFC 4949, August 2007.

[RFC5598] Crocker, D., "Internet Mail Architecture", RFC 5598, July 2009.

[RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an Attack", BCP 188, RFC 7258, May 2014.

Author's Address

Viktor Dukhovni
Two Sigma

Email: ietf-dane@dukhovni.org