Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 4, 2015

V. Dukhovni
Two Sigma
August 3, 2014

Opportunistic Security: some protection most of the time
draft-dukhovni-opportunistic-security-02

Abstract

This memo defines the term "opportunistic security". In contrast to the established approach of employing protection against both passive and active attacks or else (frequently) no protection at all, opportunistic security strives to deliver at least some protection most of the time. The primary goal is therefore broad interoperability, with security policy tailored to the capabilities of peer systems.

Table of Contents

1. Introduction

Historically, Internet security protocols have prioritized comprehensive protection against both passive and active attacks for peers capable and motivated to absorb the associated costs. Since protection against active attacks relies on authentication, which at Internet scale is not universally available, while communications traffic was sometimes strongly protected, more typically it was not protected at all. The fact that most traffic is unprotected facilitates nation-state pervasive monitoring (PM [RFC7258]) by making it cost-effective (or at least not cost-prohibitive). Indiscriminate collection of communications traffic would be substantially less attractive if security protocols were designed to operate at a range of protection levels; with encrypted transmission accessible to most if not all peers, and protection against active attacks still available where required by policy or opportunistically negotiated.

Encryption is easy, but key management is difficult. Key management at Internet scale remains an incompletely solved problem. The PKIX ([RFC5280]) key management model, which is based on broadly trusted public certification authorities (CAs), introduces costs that not all peers are willing to bear. PKIX is not sufficient to secure communications when the peer reference identity ([RFC6125]) is obtained indirectly over an insecure channel or communicating parties don't agree on a mutually trusted CA. DNSSEC ([RFC4033]) is not at this time sufficiently widely adopted to make DANE ([RFC6698]) a viable alternative at scale. Trust on first use (TOFU) key management models (as with saved SSH fingerprints and various certificate pinning approaches) don't protect initial contact and require user intervention when key continuity fails.

Without Internet-scale key management, authentication required for protection against active attacks is often not possible. When protocols only offer the options of authenticated secure channels or else cleartext, most traffic is sent in the clear. Therefore, in order to make encryption more ubiquitous, authentication needs to be optional. When authenticated communication is not possible,

unauthenticated encryption is still substantially stronger than cleartext. Opportunistic security encourages peers to employ as much security as possible, without falling back to unnecessarily weak options. In particular, opportunistic security encourages unauthenticated encryption when authentication is not an option.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The following definitions are derived from the Internet Security Glossary [RFC4949], where applicable.

Perfect Forward Secrecy (PFS): For a key management protocol, the property that compromise of long-term keys does not compromise session/traffic/content keys that are derived from or distributed using the long-term keys.

Man-in-the-Middle attack (MiTM): A form of active wiretapping attack in which an attacker intercepts and may selectively modify communicated data to masquerade as one of the entities involved in a communication. Masquerading enables the MiTM to violate the confidentiality and/or the integrity of communicated data passing through it.

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

Unauthenticated Encryption: Encryption using a key management technique that enables unauthenticated communication between parties. The communication may be 1-way or 2-way unauthenticated. If 1-way, the initiator (client) or the target (server) may be anonymous.

3. Opportunistic Security Design Philosophy

Interoperate to maximize deployment: The primary goal of designs that feature opportunistic security is to be able to communicate

with any reasonably configured peer. If many peers are only capable of cleartext, then it is acceptable to fall back to cleartext when encryption is not possible. If authentication is only possible for some peers, then it is acceptable to authenticate only those peers and not the rest. Interoperability must be possible without a need for the administrators of communicating systems to coordinate security settings. Applications employing opportunistic security need to be deployable at Internet scale, with each peer independently configured to meet its own security needs (within the practical bounds of the application protocol). Opportunistic security must not get in the way of the peers communicating if neither end is misconfigured.

Maximize security peer by peer: Subject to the above Internet-scale interoperability goal, opportunistic security strives to maximize security based on the capabilities of the peer (or peers). For some opportunistic security protocols the maximal protection possible may be just unauthenticated encryption to address passive monitoring. For others, protection against active MiTM attacks may be an option. Opportunistic security protocols may at times refuse to operate with peers for which higher security is expected, but for some reason not achieved. The conditions under which connections fail should generally be limited to operational errors at one or the other peer or an active attack, so that well-maintained systems rarely encounter problems in normal use of opportunistic security.

Encrypt by default: An opportunistic security protocol MUST interoperably achieve at least unauthenticated encryption between peer systems that don't explicitly disable this capability. To facilitate incremental deployment, opportuistic security protocols may tolerate cleartext connections or sessions with peers that don't support encryption. Over time, as peer software is updated to support opportunistic security, only legacy systems or a minority of systems where encryption is disabled should be communicating in cleartext. Whenever possible, opportunistic security should employ Perfect Forward Secrecy (PFS) to make recovery of previously sent keys and plaintext computationally expensive even after disclosure of long-term keys.

No misrepresentation of security: Unauthenticated communication or use of authentication that is vulnerable to MiTM attacks is not represented as strong security. Where protection against active attacks is required, unauthenticated opportunistic security is not a substitute.

In summary, opportunistic security is an umbrella term that encompasses protocol designs that remove barriers to the widespread use of encryption in the Internet. The actual protection provided by opportunistic security depends on the capabilities of the communicating peers; opportunistic security MUST attempt to at least encrypt network traffic, while allowing fallback to cleartext with peers that do not appear to be encryption capable.

It is important to note that opportunistic security is not limited to unauthenticated encryption. When possible, opportunistic security SHOULD provide stronger security on a peer-by-peer basis. For example, some peers may be authenticated via DANE, TOFU or other means. Though authentication failure MAY be a reason to abort a connection to a peer that is expected to be authenticated, it MUST NOT instead lead to communication in cleartext when encryption is an option. Some Message Transfer Agents (MTAs, [RFC5598] Section 4.3.2) employing STARTTLS ([RFC3207]) have been observed to abort TLS ([RFC5246]) transmission when the receiving MTA fails authentication, only to immediately deliver the same message over a cleartext connection. This design blunder MUST be avoided.

4. Security Considerations

Though opportunistic security potentially supports transmission in cleartext, unauthenticated encryption, or other protection levels short of the strongest potentially applicable, the effective security for users is increased, not reduced. Provided strong security is not required by policy or securely negotiated, nothing is lost by allowing weaker protection levels, indeed opportunistic security is strictly stronger than the alternative of providing no security services when maximal security is not applicable.

5. Acknowledgements

I would like to thank Steve Kent. Some of the text in this document is based on his earlier draft.

6. References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[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.

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

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

[RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, May 2008.

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

[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.

[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