DKIM Working Group   M. Kucherawy
Internet-Draft   Cloudmark
Intended status: Informational   September 29, 2010
Expires: April 2, 2011

RFC4871 Implementation Report
draft-ietf-dkim-implementation-report-01

Abstract

This document contains an implementation report for the IESG covering DKIM in support of the advancement of that specification along the Standards Track.

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 April 2, 2011.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

[DKIM], published in May 2007, has reached a level of maturity sufficient to consider its advancement along the standards track. Enclosed is a summary of collected interoperability data provided from sources that are aggregating such information as well as from a more formal DKIM interoperability event that took place in October 2007.

2. Definitions

Various terms specific to email are used in this document. Their definitions and further discussion can be found in [EMAIL-ARCH].

3. DKIM Interoperability Event

In October 2007, Alt-N Technologies of Dallas, Texas hosted an interoperability and testing event at their headquarters. Twenty organizations sent engineers and their various DKIM implementations to connect to a private internal network and exchange test messages and tabulate observed results.

3.1. Participants

The interoperability event included participants from all of the following organizations: Alt-N Technologies, AOL, AT&T Laboratories, Bizanga Ltd., Brandenburg InternetWorking, Brandmail Solutions, ColdSpark, Constant Contact, Inc., DKIMproxy, Domain Assurance Council, Google Inc., ICONIX Inc., Internet Initiative Japan (IIJ), Ironport Systems, Message Systems, Port25 Solutions, Postfix, Sendmail, Inc., StrongMail Systems, and Yahoo! Inc. Most of the participants traveled to Dallas and participated in person, but a few operated remotely.

Nearly all of the implementations were based on disjoint code development projects. A few were based on a common open source base project.

3.2. Testing Methodology

Participants were encouraged before the event to craft a set of test messages meant to exercise their own implementations as well as those of the other participants, both in terms of successful verifications as well as some expected to fail. Some test cases were developed with the intent of confounding verifiers that may not have implemented the [ABNF] of [DKIM] correctly.

The participants set up Mail Transfer Agents (MTAs) equipped with their own DKIM signing and verifying modules, and their own tools to generate mail to be signed along with tools to analyze the results post-verification. They then sent their own batteries of test messages, looking for both expected and unexpected failures in response. Some implementations included "auto-responders" that would reply with verification results, while others simply collected the results that would then be shared manually.

3.3. Observations

All of the implementations implemented all of the required portions of [DKIM] in terms of both signature and key features. Most of the implementations implemented all of the optional features of both signatures and keys. There were no notable or common exceptions.

The interoperability testing was largely successful. As might be expected, there were many verification false negatives or false positives that were the result of bugs in corner cases of some of the implementations presented for testing. In such cases the developers were able to identify the issue as resulting from their own mis-reading of the specification and not an error in the specification itself.

Several of the failures did occur as a result of specification ambiguities. The participants discussed each of these in turn and were able to come to consensus on how they believed the specification should be changed to resolve them.

3.4. Results

The handful of interoperability issues described above that referred to weaknesses or ambiguities in [DKIM] resulted in several errata being opened via the RFC Editor web site. These are being addressed in an RFC4871bis draft effort that is now starting from within the DKIM working group.

4. Collected DKIM Interoperability and Use Data

Several implementations are collecting private data about DKIM use, signature survivability, which properties of the base specification are observed in public use, etc. This section includes collection methods and summary reports provided by those implementations.

4.1. The OpenDKIM Project

The OpenDKIM Project is an open source project providing a DKIM support library, an email filter for use with MTAs, and a set of tools to assist with deployment of DKIM.

4.1.1. Details

Recent releases have included an optional feature to record statistics about messages with and without DKIM signatures. Sites enabling this feature can choose to share the data with the project's development team as part of this interoperability report work. The data can be anonymized to conceal the sending domain and client IP addresses, though these data are passed through a one-way hash to enable collation of data from common sources.

4.1.2. Results

At the time of writing of this document, the results of this effort are as follows:

Reporting Hosts: six individual MTAs representing four distinct ADMDs

Total Messages: 416393

Signatures: 304801 messages (73.2%) were not signed; 109367 (26.2%) had one signature; 2198 (0.5%) had two signatures; the remainder (less than 0.01%) had more.

Signing Algorithms: 48.7% of signatures used "rsa-sha1", while the balance used "rsa-sha256".

Header Canonicalization Algorithms: 13.7% of signatures used "simple", while the balance used "relaxed"; when grouped by domains, the percentages were similar.

Body Canonicalization Algorithms: 25.4% of signatures used "simple", while the balance used "relaxed"; when grouped by domains, the percentages were similar.

Keys in Test Mode: 55.8% of keys retrieved from the DNS were tagged as being in test mode.

Keys with Specific Granularity: Four keys were retrieved that had specific names in their "g=" tags.

Keys with Syntax Errors: Less than 0.1% of keys retrieved from the DNS had syntax errors.

DomainKeys Compatibility: 1.4% of the retrieved keys appeared to be intended for use with the older DomainKeys proposal rather than DKIM

Missing Keys: 2% of signatures received referenced keys that were not found in the DNS

Optional Signature Tags: Of the optional signature tags supported by the base specification, "t=" was seen 46.4% of the time (1% of which included timestamps in the future, even after forgiving some clock drift); "x=" was seen 4.4% of the time; "l=" was seen 4.6% of the time; "z=" was seen 4.8% of the time.

Body Length Limits: Of the signatures for which "l=" was used, 6.4% of them signed none of the body, and 100% of the rest had the body extended after signing.

Signature Pass Rates: Overall, 89.9% of observed signatures were successfully verified.

Pass Rates for Non-List Mail: Where "list mail" is defined as any mail not bearing one of the header fields defined in [LIST-ID] or in [LIST-URLS], or a "Precedence: list" field, selecting only for mail that is not list mail revealed a successful verification rate of 93.6%; selecting only for list mail produced a 84.7% success rate.

Author vs. Third-Party: 73% of the signatures observed were author signatures, meaning the "d=" value in the signature matched the domain found in the From: header field. The remainder, therefore, were third-party signatures.

DNSSEC: Only one reporting site is currently checking for DNSSEC on keys retrieved from the DNS. It found no signed keys.

Common errors: The top five verification errors observed: Key not found in DNS (18.7%), key granularity mismatch (13%), DNS retrieval failure such as timeouts (2%), key revoked (1.9%), unknown key type (1.8%).

Detected Header Field Changes: A subset of the reporting sites are capable of reporting header fields known to have been changed in transit (e.g. when "z=" tags were used by the signer). In such cases, changes to the "To:" field were more common than any other by a factor of four or more.

Most Commonly Signed Fields: From: (100%), To: (95.5%), Subject: (93.7%), Date: (92.3%), MIME-Version: (88.8%), Content-Type: (80%), Message-Id: (75.9%), Received: (59.7%). All others are below 50%.

Multiple-use Signing Domains: 9512 unique signing domains were observed. Of these, 42.7% of them sent a single signed message in the sample period, 18.6% sent two and 8.6% sent three.

4.1.3. Conclusions

The results of the OpenDKIM work are updated constantly as more data feeds come online and more data are reported. Based on the data available at the time of writing, some conclusions are possible.

At least some implementations of all of the optional signature features, all of the canonicalization combinations and all of the signing algorithms are in general use. None of the features had zero use counts.

Overall signature pass rates are generally quite high. The impact of signature survivability when correlated against MLM activity is surprisingly low based on observed data. More research into this is recommended. The DKIM Working Group is already working on an Informational draft to discuss those issues.

That the "To" field is the one most often associated with verification failures suggests some MTAs handling the message are correcting cases where the field is improperly formed. A common case is failing to quote the comment portion of that field when required to do so by [MAIL]. Such corrections cause signatures to become invalid.

The counts of low-use signing domains suggest that spammers, who typically rotate domain names with high frequency, have adopted DKIM as a tool to try to get through message filters.

4.2. AOL Data

A one-day summary of observed traffic from America Online reports the following:

Ratio of DKIM-signed mail: 42%

Properly formed signatures: 1.4 billion

Malformed signatures: 3000

Unique signing domains: 50,000-90,000

Key retrieval errors: 14 million

Signature refers to nonexistent domain: 10 million

Signature refers to nonexistent key: 36 million

Signature refers to revoked key: 138,000

Originator signatures: 1.2 billion

Third-party signatures: 184 million

Verified signatures: 1.2 billion

Failed signatures (body changed): 78 million

Failed signatures (other): 34 million

Expired signatures: less than 1 million

5. Security Considerations

This document is an implementation report and thus has no security considerations.

6. Informative References

[ABNF] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 5234, January 2008.

[DKIM] Allman, E., Callas, J., Delany, M., Libbey, M., Fenton, J., and M. Thomas, "DomainKeys Identified Mail (DKIM) Signatures", RFC 4871, May 2007.

[EMAIL-ARCH] Crocker, D., "Internet Mail Architecture", RFC 5598, July 2009.

[LIST-ID] Chandhok, R. and G. Wenger, "List-Id: A Structured Field and Namespace for the Identification of Mailing Lists", RFC 2919, March 2001.

[LIST-URLS] Neufeld, G. and J. Baer, "The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields", RFC 2369, July 1998.

[MAIL] Resnick, P., "Internet Message Format", RFC 5322, October 2008.

Appendix A. Acknowledgements

The author wishes to acknowledge the following for their review and constructive criticism of this document: Tony Hansen

The working group expresses its thanks to Alt-N Technologies for graciously hosting the 2007 DKIM interoperability event.

Author's Address

Murray S. Kucherawy
Cloudmark
128 King St., 2nd Floor
San Francisco, CA 94107
US

Phone: +1 415 946 3800
Email: msk@cloudmark.com