Internet Engineering Task Force	P. Hallam-Baker
Internet-Draft	Comodo Group Inc.
Intended status: Standards Track	R. Stradling
Expires: April 5, 2012	Comodo CA Ltd.
	October 3, 2011

TOC

The di (DIGEST) URI Scheme
draft-hallambaker-digesturi-02

Abstract

A URI scheme for referencing static data abjects by means of a cryptographic digest mechanism is specified. The format is designed to resist content type substitution attacks and supports a choice of digest algorithms.

Status of this Memo

Copyright Notice

Table of Contents

TOC

1. Definitions

TOC

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in **RFC 2119** [RFC2119].

TOC

1.2. Defined Terms

The following terms are used in this document:

URI

TOC

2. Requirements

Provides a strong reference to a static data object.

Does not provide a means of resolution.

Allows an authenticated data source to provide an authenticated reference to a static data object.

Intended applications include creating references from

Web pages delivered over HTTP/TLS

DNS resource records signed using DNSSEC

Data values embedded in certificates, CRLs, OCSP tokens and other signed data objects.

TOC

2.1. Examples of Use

TOC

2.1.1. Simple Digest

For example, the following digest URI specifies a reference to the text "Hello World !" using the SHA-2 algorithm with 256 bit output:

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc

TOC

2.1.2. Truncated Digest

Message Digest algorithms are designed to provide protection against a collision attack. Due to the birthday paradox, this requires that the digest length be twice the length of a encryption or authentication key to achieve the same work factor.

The digest portion MAY be truncated at any 32 bit boundary. If a truncated digest is used the query separator '?' MUST be specified.

di:sha-256;B_K97zTtFuOhug27fke4_Q?

TOC

2.1.3. Digest with Meta-Data

The semantics of a digest being used to establish a secure reference from an authenticated source to an external source may be a function of associated meta data such as the content type. This data MAY be specified by means of a parameter:

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc?ct=text/plain

TOC

2.1.4. Digest with Locator

A digest identifier MAY provide a location from which the referenced content MAY be available. Note however that since it is statistically unlikely that a given identifier will correspond to more than one content sequence, the actual location from which the data is retrieved is immaterial.

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc?http=di.example.com

The corresponding content MAY be retrieved from the URL:

http://di.example.com/.well-known/di/sha-256/B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc

A digest identifier MAY specify multiple locations from which the content MAY be obtained:

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc?http=one.example.com&http=two.example.com

Asserts that the content may be retrieved from either of the following URIs:

http://one.example.com/.well-known/di/sha-256/B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc http://two.example.com/.well-known/di/sha-256/B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc

TOC

2.1.5. Digest with Decryption Key

A digest identifier MAY provide a key for decrypting the referenced data.

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc? enc=aes-cbc:Fw3x20nEKfq6FDGzq7ttlQ

TOC

3. The di (DIGEST) URI TYPE

TOC

3.1. Syntax

The DIGEST URI Type has the following format:

"di:" algorithm ";" digest ["?" tag "=" value ["&" tag "=" value] *]

TOC

3.1.1. Encoding Considerations

TOC

3.1.1.1. Use of base64url Encoding

Section 2.2 of **[RFC4395]** states [URI schemes that are not intended for use with relative URIs SHOULD avoid use of the forward slash "/" character, which is used for hierarchical delimiters, and the complete path segments "." and ".." (dot-segments).]

Consequently the encoding of the digest value is effected using the base64url encoding specified in Section 5 of **[RFC4648]** which avoids the use of the forward slash '/' character.

TOC

3.1.1.2. Query Parameter Encoding

The query segment of a URI is NOT hierarchical. Thus escape encoding of slash '/' characters is NOT required.

Section 3.4 of **[RFC3986]** states [The characters slash ("/") and question mark ("?") may represent data within the query component.]

Consequently no special escaping mechanism is required for the query parameter portion of the URI. URI escaping is however frequently imposed automatically by scripting environments. Thus to ensure interoperability, implementations SHOULD NOT generate URIs that employ URI character escaping and implementations MUST accept and URIs that employ URI character escaping.

TOC

3.2. Semantics

TOC

3.2.1. Digest Algorithms

Implementations MUST support the sha-256 algorithm as specified in **[RFC4055]**.

Implementations MAY support other algorithms specified in the Data Structure for the Security Suitability of Cryptographic Algorithms registry 'Cryptographic Algorithms' **[RFC5698]**.

TOC

3.2.2. Parameters

TOC

3.2.2.1. Content Type (ct)

The Content Type "ct" parameter specifies the MIME Content Type of the associated data as defined in **[RFC4288]**

TOC

3.2.2.2. HTTP Locator (http & https)

The http and https parameters are used to specify a possible means of resolving the referenced content. Mulltiple locator parameters MAY be used to specify alternative sources for accessing the content.

The http and https parameters take a single argument, the domain name to be used for resolution. To permit the use of digest URIs in ASCII-only environments, the ASCII encoding (aka punycode) of the domain name MUST be used.

To resolve such a location reference, a client first transforms the digest URI to obtain a http or https url as follows:

URL = prefix + domain + "/.well-known/di/" + algorithm + digest

Where:

prefix
: Is the string "http://" for the http parameter type and is the string "https://" for the https parameter type.

domain
: Is the value associated with the parameter.

algorithm
: Is the algorithm portion of the digest URI.

digest
: Is the digest portion of the digest URI.

Implementations MUST NOT disclose any other data. In particular implementations MUST NOT disclose the query parameter portion of the URI.

TOC

3.2.2.3. Encryption and MIME Encryption Specifiers (enc & menc)

The encryption and MIME encryption specifiers are used to provide a means of obtaining the plaintext of a reference to encrypted content.

The enc specifier is used when the encrypted object consists of the encrypted content alone. The menc spcifier is used when the encrypted object consists of a MIME header containing metadata followed by the binary object encoding.

The encryption specifiers both take an agrument of the form:

algorithm ":" base64url (key) [":" base64url (iv)]

Where

algorithm
: Is the algorithm used to encrypt the associated content

key
: Is the value of the cryptographic key

iv (optional)
: Is the value of the cryptographic Initialization Vector.
: If the IV is not spcified for a block cipher mode that requires one, the IV MUST be

prepended to the encrypted content.
[Actually the IV does not provide any additional security for this application but explaining the reason would be more effort than it is worth and what I really care about is saving bytes in the identifier, not the resulting data package.]

TOC

4. Security Considerations

TOC

4.1. Integrity

No secret information is required to generate a DIGEST URI. Therefore a DIGEST URI only provides a proof of integrity for the referenced object and the proof of integrity provided is only as good as the proof of integrity for the DIGEST URI value.

TOC

4.2. Confidentiality

Disclosure of a DIGEST URI value does not necessarily entail disclosure of the referenced object but may enable an attacker to determine the contents of the referenced object by reference to a search engine or other data repository.

TOC

4.3. Weak Digest Algorithm

[The digest algorithm MUST be strong]

[For most use cases collision resistance is a requirement]

TOC

5. IANA Considerations

TOC

5.1. Assignment of URI Scheme di

The procedures for registration of a URI scheme are specified in **RFC 4395** [RFC4395]. The following is the proposed assignment template.

URI scheme name: di

Status: Permanent

URI scheme syntax. See **Section 3.1**

URI scheme semantics. See **Section 3.2**

Encoding considerations. See **Section 3.1.1**

Applications/protocols that use this URI scheme name: General applicability with initial use cases provided by WEBSEC and DECADE

Interoperability considerations: TBS

Security considerations: See **Section 4**

Contact: IETF / Phillip Hallam-Baker

Author/Change controller: IETF / Phillip Hallam-Baker

References: As specified in this document

TOC

5.2. Assignment of Well Known URI prefix di

The procedures for registration of a Well Known URI entry are specified in **RFC 5785** [RFC5785]. The following is the proposed assignment template.

URI suffix: di

Change controller: IETF

Specification document(s): This document

Related information: None

TOC

5.3. Specification of Additional Cryptographic Algorithms

[Added in case it is decided to specify truncated forms of the cryptographic digests]

The procedures for registration of a Cryptographic Algorithm identifier are specified in **RFC 5698** [RFC5698]. The following is the proposed assignment template.

Textual name of the algorithm: SHA-128

OID of the algorithm: [TBS]

Reference: This document.

TOC

5.4. Creation of di parameter registry

This specification creates a new IANA registry entitled "DI URI Parameter Definitions".

The policy for future assignments to the registry is "RFC Required".

TOC

6. References

TOC

6.1. Normative References

[RFC2119] Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, “Uniform Resource Identifier (URI): Generic Syntax,” STD 66, RFC 3986, January 2005 (TXT, HTML, XML).

[RFC4055] Schaad, J., Kaliski, B., and R. Housley, “Additional Algorithms and Identifiers for RSA Cryptography for use in the Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile,” RFC 4055, June 2005 (TXT).

[RFC4288] Freed, N. and J. Klensin, “Media Type Specifications and Registration Procedures,” BCP 13, RFC 4288, December 2005 (TXT).

[RFC4395] Hansen, T., Hardie, T., and L. Masinter, “Guidelines and Registration Procedures for New URI Schemes,” BCP 35, RFC 4395, February 2006 (TXT).

[RFC4648] Josefsson, S., “The Base16, Base32, and Base64 Data Encodings,” RFC 4648, October 2006 (TXT).

[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 (TXT).

[RFC5698] Kunz, T., Okunick, S., and U. Pordesch, “Data Structure for the Security Suitability of Cryptographic Algorithms (DSSC),” RFC 5698, November 2009 (TXT).

[RFC5785] Nottingham, M. and E. Hammer-Lahav, “Defining Well-Known Uniform Resource Identifiers (URIs),” RFC 5785, April 2010 (TXT).

TOC

6.2. Non Normative References

[NIST-ALGS] National Institute of Standards and Technology, “**Cryptographic Algorithm Registration**,” March 2009.

[RFC3642] Legg, S., “**Common Elements of Generic String Encoding Rules (GSER) Encodings**,” RFC 3642, October 2003 (**TXT**).

TOC

Appendix A. Test Vectors

TOC

A.1. Example: Hello World !

The Digest URI of the text file "Hello World !" is computed as follows:

Scheme `di`

Algorithm Identifier: `sha-256`

sha-256 ("Hello World !") `07 f2 bd ef 34 ed 16 e3 a1 ba 0d bb 7e 47 b8 fd 98 1c e0 cc b3 e1 bf e5 64 d8 2c 42 3c ba 7e 47`

BASE64URI (sha-256 ("Hello World !")): `B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc`

Content Type Parameter 'text/plain' `ct=text/plain`

Depending on the context, the digest URI MAY be specified using the digest value alone or the digest value plus content type parameter:

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc

di:sha-256;B_K97zTtFuOhug27fke4_Zgc4Myz4b_lZNgsQjy6fkc?ct=text/plain

TOC

Authors' Addresses

Phillip Hallam-Baker
Comodo Group Inc.
Email: phillip@comodo.com

Rob Stradling
Comodo CA Ltd.
Email: rob.stradling@comodo.com