Network Working Group
INTERNET-DRAFT
Updates: RFC 1864

Martin Hamilton
Loughborough University
February 1999

The Content-MD5-Origin: header

draft-hamilton-content-md5-origin-01.txt

Status of This Memo

Abstract

The Content-MD5: header specified in RFC 1864 has not been widely deployed, though this would be highly desirable for a number of reasons. The author conjectures that this lack of usage is due at least in part to the requirement that only originating user agents may add a Content-MD5: header. This proposal updates RFC 1864 to remove that requirement, and defines the header Content-MD5-Origin: for use by relaying hosts to indicate the point at which a Content-MD5: header was added.

1. Extending the scope of the Content-MD5: header

RFC 1864 specifies that :-

> The Content-MD5 field is generated by only an originating user agent. Message relays and gateways are expressly forbidden from generating a Content-MD5 field.

Whilst understandable, this restriction means that the technology must be pushed out to very large numbers of end users before it can be useful for the Internet community as a whole.

In order to maximize the deployment of the Content-MD5: header, it is essential that intermediate (relaying) systems be allowed to generate a Content-MD5: header, and that Content-MD5: headers may be generated for arbitrary message objects (rather than just leaf nodes of MIME objects).

The rationale for extending the RFC 1864 definition of Content-MD5: is that in addition to the basic message integrity check function, it provides a very effective means of protection for messaging systems against a number of common problems, such as

* loops - e.g. malfunctioning "vacation" programs or failure messages sent to mailing lists by broken server software

* multiple submissions - where the same message is injected over and over again, e.g. due to broken user agent or server software

* unsolicited bulk messaging - a special case of the above

It should be noted that Content-MD5: is not a complete solution in itself. For example, in some loop situations it is not uncommon for messages to include header information for diagnostic purposes. This would likely render the Content-MD5: digest value useless, since it would be different for each of the looping messages.

2. Introducing Content-MD5-Origin:

When a relaying host or system decides to create a Content-MD5: header, it should also add a Content-MD5-Origin: header, with its host name or Internet Protocol address as the right hand side.

For example :-

```
Content-MD5-Origin: plausible-deniability.lut.ac.uk
```

or

Content-MD5-Origin: [131.231.132.201]

Note that Internet Protocol addresses should be encapsulated within square brackets, as in the second example above.

Where a relaying host has multiple domain names and/or IP addresses (e.g. because it is multi-homed), any of these may be chosen arbitrarily. Most messaging systems provide a way for the administrator to indicate a host's "canonical" domain name or IP address - this is usually a good choice for the value of the Content-MD5-Orgin: header.

3. Security considerations

As noted in RFC 1864, Content-MD5: is no substitute for a strong cryptographic message integrity check. In this context, however, there is no authentication element to consider - the value of the Content-MD5: header is simply being used as a "key", typically into a hash database which may be used to eliminate duplicate/unwanted messages.

Implementations should take care not to assume that the value of the Content-MD5: header will always be 24 bytes or less - to avoid buffer overrun problems. It would also be unwise to assume that the characters in an arbitrary Content-MD5: header will be chosen from the base64 character set mandated by RFC 1864.

Relaying/receiving hosts should take care to check that the value supplied for the Content-MD5: header matches that calculated from a fresh iteration of the algorithm on the message. Supplying a bogus Content-MD5: header which was different every time would be an easy way to subvert simple-minded implementations. A future document may define a standard way for a relaying host to indicate that it has received an invalid Content-MD5: header.

4. Acknowledgements

Thanks to Nathaniel Borenstein for comments and feedback.

5. References

[1] Myers, J. and Rose, M. "The Content-MD5 Header Field." RFC 1864, October 1995.

6. Author's address

Martin Hamilton
Department of Computer Science
Loughborough University
Leics. LE11 3TU, UK

Email: martinh@gnu.org

This Internet Draft expires 31st August 1999.