

HTTPAuth Working Group
Internet-Draft
Updates: [2617](#) (if approved)
Intended status: Experimental
Expires: November 17, 2013

J. Reschke
greenbytes
May 16, 2013

An Encoding Parameter for HTTP Basic Authentication
draft-ietf-httpauth-basicauth-enc-00

Abstract

The "Basic" authentication scheme defined in [RFC 2617](#) does not properly define how to treat non-ASCII characters. This has lead to a situation where user agent implementations disagree, and servers make different assumptions based on the locales they are running in. There is little interoperability for the non-ASCII characters in the ISO-8859-1 character set, and even less interoperability for any characters beyond that.

This document defines a backwards-compatible extension to "Basic", specifying the server's character encoding expectation, using a new authentication scheme parameter.

Editorial Note (To be removed by RFC Editor before publication)

Discussion of this draft takes place on the HTTPAuth working group mailing list (http-auth@ietf.org), which is archived at [<http://www.ietf.org/mail-archive/web/http-auth/current/maillist.html>](http://www.ietf.org/mail-archive/web/http-auth/current/maillist.html).

XML versions, latest edits and the issues list for this document are available from [<http://greenbytes.de/tech/webdav/#draft-ietf-httpauth-basicauth-enc>](http://greenbytes.de/tech/webdav/#draft-ietf-httpauth-basicauth-enc).

The changes in this draft are summarized in [Appendix C.9](#).

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 November 17, 2013.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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	4
2.	Notational Conventions	4
3.	The 'accept-charset' auth-param	4
4.	Examples	5
5.	Security Considerations	5
6.	IANA Considerations	5
7.	Acknowledgements	5
8.	References	6
8.1.	Normative References	6
8.2.	Informative References	7
Appendix A.	Deployment Considerations	7
A.1.	User Agents	7
A.1.1.	Alternative approach	7
A.2.	Origin Servers	8
Appendix B.	FAQ (to be removed by RFC Editor before publication)	8
B.1.	Why not simply switch the default encoding to UTF-8?	8
B.2.	What about Digest?	8
B.3.	Will existing UAs ignore the parameter?	8
Appendix C.	Change Log (to be removed by RFC Editor before publication)	8
C.1.	Since draft-reschke-basicauth-enc-00	8
C.2.	Since draft-reschke-basicauth-enc-01	8
C.3.	Since draft-reschke-basicauth-enc-02	9
C.4.	Since draft-reschke-basicauth-enc-03	9
C.5.	Since draft-reschke-basicauth-enc-04	9
C.6.	Since draft-reschke-basicauth-enc-05	9
C.7.	Since draft-reschke-basicauth-enc-06	9
C.8.	Since draft-reschke-basicauth-enc-07	9
C.9.	Since draft-reschke-basicauth-enc-08	9
Appendix D.	Open issues (to be removed by RFC Editor prior to publication)	9
D.1.	edit	9
D.2.	unorm	9
D.3.	terminology	10
D.4.	parmname2831	10

1. Introduction

The "Basic" authentication scheme defined in [Section 2 of \[RFC2617\]](#) does not properly define how to treat non-ASCII characters ([\[USASCII\]](#)): it uses the Base64 ([\[RFC4648\], Section 4](#)) encoding of the concatenation of username, separator character, and password without stating which character encoding to use.

This has lead to a situation where user agent implementations disagree, and servers make different assumptions based on the locales they are running in. There is little interoperability for the non-ASCII characters in the ISO-8859-1 character set ([\[USASCII\]](#), [\[ISO-8859-1\]](#)), and even less interoperability for any characters beyond that.

This document defines a backwards-compatible extension to "Basic", specifying the server's character encoding expectation, using a new auth-param for use in the Proxy-Authenticate and WWW-Authenticate header fields, as defined in [\[draft-ietf-httpbis-p7-auth\]](#).

2. Notational Conventions

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 [\[RFC2119\]](#).

3. The 'accept-charset' auth-param

In challenges, servers MAY use the "accept-charset" authentication parameter (case-insensitive) to express the character encoding they expect the user agent to use.

The only allowed value is "UTF-8", to be matched case-insensitively (see [\[RFC2978\], Section 2.3](#)), indicating that the server expects the UTF-8 character encoding to be used ([\[RFC3629\]](#)).

Other values are reserved for future use.

Note: The 'accept-charset' parameter cannot be included when sending credentials (e.g. in the Authorization or Proxy-Authorization header fields), as the "Basic" scheme uses a single token for credentials ('token68' syntax), not a parameter list ('#auth-param' syntax); see Section 2.1 of [\[draft-ietf-httpbis-p7-auth\]](#).

4. Examples

In the example below, the server prompts for authentication in the "foo" realm, using Basic authentication, with a preference for the UTF-8 character encoding:

```
WWW-Authenticate: Basic realm="foo", accept-charset="UTF-8"
```

Note that the parameter value can be either a token or a quoted string; in this case the server chose to use the quoted-string notation.

The user's name is "test", and his password is the string "123" followed by the Unicode character U+00A3 (POUND SIGN). Following [Section 1.2 of \[RFC2617\]](#), but using the character encoding UTF-8, the user-pass, converted to a sequence of octets, is:

```
't' 'e' 's' 't' ':' '1' '2' '3' pound
74 65 73 74 3A 31 32 33 C2 A3
```

Encoding this octet sequence in Base64 ([\[RFC4648\]](#), [Section 4](#)) yields:

```
dGVzdDoxMjPCow==
```

Thus the Authorization header field would be:

```
Authorization: Basic dGVzdDoxMjPCow==
```

Or, for proxy authentication:

```
Proxy-Authorization: Basic dGVzdDoxMjPCow==
```

5. Security Considerations

This document does not introduce any new security considerations beyond those defined for the "Basic" authentication scheme ([\[RFC2617\]](#), [Section 4](#)), and those applicable to the handling of UTF-8 ([\[RFC3629\]](#), [Section 10](#)).

6. IANA Considerations

There are no IANA Considerations related to this specification.

7. Acknowledgements

The internationalisation problem has been reported as a Mozilla bug back in the year 2000 (see https://bugzilla.mozilla.org/show_bug.cgi?id=41489) and also the

more recent <https://bugzilla.mozilla.org/show_bug.cgi?id=656213>). It was Andrew Clover's idea to address it using a new auth-param.

Thanks to Bjoern Hoehrmann, Amos Jeffries, James Manger, and Martin Thomson for providing feedback on this document.

8. References

8.1. Normative References

- [ISO-8859-1] International Organization for Standardization, "Information technology -- 8-bit single-byte coded graphic character sets -- Part 1: Latin alphabet No. 1", ISO/IEC 8859-1:1998, 1998.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2617] Franks, J., Hallam-Baker, P., Hostetler, J., Lawrence, S., Leach, P., Luotonen, A., and L. Stewart, "HTTP Authentication: Basic and Digest Access Authentication", [RFC 2617](#), June 1999.
- [RFC2978] Freed, N. and J. Postel, "IANA Charset Registration Procedures", [BCP 19](#), [RFC 2978](#), October 2000.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, [RFC 3629](#), November 2003.
- [USASCII] American National Standards Institute, "Coded Character Set -- 7-bit American Standard Code for Information Interchange", ANSI X3.4, 1986.
- [[draft-ietf-httpbis-p7-auth](#)] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Authentication", [draft-ietf-httpbis-p7-auth-22](#) (work in progress), February 2013.

8.2. Informative References

- [RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", [RFC 4648](#), October 2006.
- [XHR] Aubourg, J., Song, J., and H. Steen, "XMLHttpRequest", W3C Working Draft WD-XMLHttpRequest-20121206, December 2012, <<http://www.w3.org/TR/2012/WD-XMLHttpRequest-20121206/>>.
- Latest version available at
 <<http://www.w3.org/TR/XMLHttpRequest/>>.

Appendix A. Deployment Considerations

A.1. User Agents

User agents not implementing this specifications should continue to work as before, ignoring the new parameter.

User agents which already default to the UTF-8 encoding implement this specification by definition. Note that some user agents also have different defaults depending on whether the request originates from page navigation as opposed to a script-driven request using XMLHttpRequest [[XHR](#)].

Other user agents can keep their default behavior, and switch to UTF-8 when seeing the new parameter.

A.1.1. Alternative approach

On the other hand, the strategy below may already improve the user-visible behavior today:

- o In the first authentication request, choose the character encoding based on the user's credentials: if they do not need any characters outside the ISO-8859-1 character set, default to ISO-8859-1, otherwise use UTF-8.
- o If the first attempt failed and the encoding used was ISO-8859-1, retry once with UTF-8 encoding instead.

Note that there's a risk if the site blocks an account after multiple login failures (for instance, when it doesn't reset the counter after a successful login).

A.2. Origin Servers

Origin servers that do not support non-ASCII characters in credentials do not require any changes.

Origin servers that need to support non-ASCII characters, but can't use the UTF-8 encoding will not be affected; they will continue to function as well as before.

Finally, origin servers that need to support non-ASCII characters and can use the UTF-8 encoding can opt in as described above. In the worst case, they'll continue to see either broken credentials or no credentials at all (depending on how legacy clients handle characters they can not encode).

Appendix B. FAQ (to be removed by RFC Editor before publication)

B.1. Why not simply switch the default encoding to UTF-8?

There are sites in use today that default to a locale encoding, such as ISO-8859-1, and expect user agents to use that encoding. These sites will break if the user agent uses a different encoding, such as UTF-8.

B.2. What about Digest?

Although the solution proposed in this document may be applicable to "Digest" as well, any attempt to update this scheme may be an uphill battle hard to win.

B.3. Will existing UAs ignore the parameter?

It appears they will. See
<<http://greenbytes.de/tech/tc/httpauth/#simplebasicnewparam1>> and
<<http://greenbytes.de/tech/tc/httpauth/#simplebasicnewparam2>>.

Appendix C. Change Log (to be removed by RFC Editor before publication)

C.1. Since [draft-reschke-basicauth-enc-00](#)

Add and close issues "credparam" and "paramcase". Rewrite the deployment considerations.

C.2. Since [draft-reschke-basicauth-enc-01](#)

Note more recent Mozilla bugzilla entry; add behavior of existing UAs to FAQ (with pointer to test cases).

C.3. Since [draft-reschke-basicauth-enc-02](#)

Add and resolve issue "xhrtf8".

C.4. Since [draft-reschke-basicauth-enc-03](#)

Add and resolve issue "proxy".

C.5. Since [draft-reschke-basicauth-enc-04](#)

Add and resolve issues "paramname" and "sentparam". Add issues "terminology" and "unorm". Update HTTPbis reference.

C.6. Since [draft-reschke-basicauth-enc-05](#)

Update HTTPbis reference.

C.7. Since [draft-reschke-basicauth-enc-06](#)

Update HTTPbis and XHR references.

C.8. Since [draft-reschke-basicauth-enc-07](#)

"b64token" -> "token68" (ABNF term changed in HTTPbis P7). Change contact information from HTTPbis WG to HTTPAUTH WG. Add issue parmname2831. Changed intended status to "experimental".

C.9. Since [draft-reschke-basicauth-enc-08](#)

Made it a draft of the IETF HTTPauth Working Group.

[Appendix D.](#) Open issues (to be removed by RFC Editor prior to publication)**[D.1.](#) edit**

Type: edit

julian.reschke@greenbytes.de (2010-08-11): Umbrella issue for editorial fixes/enhancements.

[D.2.](#) unorm

Type: edit

julian.reschke@greenbytes.de (2012-02-02): We need a statement about unicode normalization forms.

D.3. terminology

Type: edit

julian.reschke@greenbytes.de (2012-02-02): Try to be consistent with the terminology defined in [RFC 6365](#).

D.4. parmname2831

In [Section 3](#):

Type: change

julian.reschke@greenbytes.de (2012-05-08): [RFC 2831](#) (Digest SASL Mechanism) defines a *very* similar parameter but calls it "charset". We may want to be consistent with that.

Author's Address

Julian F. Reschke
greenbytes GmbH
Hafenweg 16
Muenster, NW 48155
Germany

EMail: julian.reschke@greenbytes.de
URI: <http://greenbytes.de/tech/webdav/>

