Ruby master - Bug #13926
Non UTF response headers raise an Argument error since 2.4.2p198
09/21/2017 06:01 PM - petehamilton (Pete Hamilton)

| Status:  | Closed  |
| Priority: | Normal  |
| Assignee: |         |
| Target version: |         |
| ruby -v: | ruby 2.4.2p198 (2017-09-14 revision 59899) [x86_64-darwin16] |
| Backport: | 2.2: DONTNEED, 2.3: DONE, 2.4: DONE |

Description
When setting headers using Net::HTTPHeader#add_field or Net::HTTPHeader[] in v2.4.2, an ArgumentError (invalid byte sequence in UTF-8) is raised.

In 2.4.1, this behaviour didn't exist and it looks like it was introduced in one of the revisions associated with https://bugs.ruby-lang.org/issues/13852, where the header value is matched against a regular expression to prevent newlines.

Previously, Net::HTTP would accept non-UTF8 header values and just return them as invalid UTF8 strings. It was then on the user of Net::HTTP to handle this. With this change, there's now no way for the user to handle the case where they receive non-UTF8 header values as Net::HTTP raises an error.

RFC2616 allowed an HTTP header field content to be made up of any non-whitespace octets. Because of this RFC7230 makes an allowance for all characters in the ISO-8859-1 charset (both lower and extended ASCII characters).

Specifically, this section of RFC7230 suggests that although ideally response header values would be compatible with UTF-8, we can't assume this to be the case.

Historically, HTTP has allowed field content with text in the ISO-8859-1 charset [ISO-8859-1], supporting other charsets only through use of [RFC2047] encoding. In practice, most HTTP header field values use only a subset of the US-ASCII charset [USASCII]. Newly defined header fields SHOULD limit their field values to US-ASCII octets. A recipient SHOULD treat other octets in field content (obs-text) as opaque data.

Not entirely sure where to go from here or what the fix is but given this is a behaviour change, it'd be great to hear your thoughts.

Associated revisions
Revision 56f91c6e - 09/25/2017 02:28 PM - naruse (Yui NARUSE)
HTTPHeader#add_field should allow binary [Bug #13926]
git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@60021 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 60021 - 09/25/2017 02:28 PM - naruse (Yui NARUSE)
HTTPHeader#add_field should allow binary [Bug #13926]

Revision 60021 - 09/25/2017 02:28 PM - naruse (Yui NARUSE)
HTTPHeader#add_field should allow binary [Bug #13926]

Revision 60021 - 09/25/2017 02:28 PM - naruse (Yui NARUSE)
HTTPHeader#add_field should allow binary [Bug #13926]

Revision d2b5c16e - 12/20/2017 03:59 PM - nagachika (Tomoyuki Chikanaga)
merge revision(s) 60021: [Backport #13926]
HTTPHeader#add_field should allow binary [Bug #13926]
git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_4@61373 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

08/28/2022 1/3
HTTPHeader#add_field should allow binary [Bug #13926]

Revision ecb7182f - 01/31/2018 11:47 AM - usa (Usaku NAKAMURA)
merge revision(s) 60021: [Backport #13926]

HTTPHeader#add_field should allow binary [Bug #13926]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_3@62133 b2dd03c8-39d4-4d8f-98f7-823fe69b080e

Revision 62133 - 01/31/2018 11:47 AM - usa (Usaku NAKAMURA)
merge revision(s) 60021: [Backport #13926]

HTTPHeader#add_field should allow binary [Bug #13926]

History

#1 - 09/22/2017 04:03 AM - shevegen (Robert A. Heiler)
Can you add a link to Net::HTTPHeader#add_header?
I was trying to find it but I can not find it at https://ruby-doc.org/stdlib/libdoc/net/http/rdoc/Net/HTTPHeader.html - I do find &= though
https://ruby-doc.org/stdlib-2.4.2/libdoc/net/http/rdoc/Net/HTTPHeader.html#method-i-5B-5D-3D and the documentation does not mention anything about Encoding.
I think either way, the behaviour in regards to Encoding should be noted down in the documentation as well. I assume that the change introduced a regression but I have absolutely no idea; the documentation should mention this somewhere though, either at the method, or at the main page or Net::HTTPHeader imo.

#2 - 09/22/2017 08:29 AM - petehamilton (Pete Hamilton)
shevegen (Robert A. Heiler) wrote:

Can you add a link to Net::HTTPHeader#add_header?
I was trying to find it but I can not find it at https://ruby-doc.org/stdlib/libdoc/net/http/rdoc/Net/HTTPHeader.html

Apologies, I meant Net::HTTPHeader#add_field (https://ruby-doc.org/stdlib-2.4.2/libdoc/net/http/rdoc/Net/HTTPHeader.html#method-i-add_field). I will update the ticket description.

#3 - 09/22/2017 08:29 AM - petehamilton (Pete Hamilton)
- Description updated

#4 - 09/25/2017 02:29 PM - naruse (Yui NARUSE)
- Status changed from Open to Closed

Applied in changeset trunk|r60021.

HTTPHeader#add_field should allow binary [Bug #13926]

#5 - 09/25/2017 02:29 PM - naruse (Yui NARUSE)
- Backport changed from 2.2: UNKNOWN, 2.3: UNKNOWN, 2.4: UNKNOWN to 2.2: DONTNEED, 2.3: REQUIRED, 2.4: REQUIRED

#6 - 12/20/2017 03:59 PM - nagachika (Tomoyuki Chikanaga)
- Backport changed from 2.2: DONTNEED, 2.3: REQUIRED, 2.4: REQUIRED to 2.2: DONTNEED, 2.3: REQUIRED, 2.4: DONE

ruby_2_4 r61373 merged revision(s) 60021.

#7 - 01/31/2018 11:47 AM - usa (Usaku NAKAMURA)
- Backport changed from 2.2: DONTNEED, 2.3: REQUIRED, 2.4: DONE to 2.2: DONTNEED, 2.3: DONE, 2.4: DONE

ruby_2_3 r62133 merged revision(s) 60021.