encoding of String#unpack for 'H', 'h', 'B' and 'b'

09/23/2012 05:03 AM - Eregon (Benoit Daloze)

Status: Closed  
Priority: Normal  
Assignee:  
Target version: 2.0.0  
ruby -v: ruby 2.0.0dev (2012-07-15 trunk 36395)  
[x86_64-darwin10.8.0]  

Description

Originally mentioned in [#6799](https://bugs.ruby-lang.org/issues/6799), I think the encoding of the resulting Strings for the 'H', 'h', 'B' and 'b' directives of String#unpack should be encoded in US-ASCII, because they will always contain 0-9a-f (0-1 for 'B','b') chars. This is already done for the 'm', 'M' and 'u' directives in Array#pack.

It would also fix SecureRandom hexdigests as mentioned in [#6799](https://bugs.ruby-lang.org/issues/6799).

P.S.: Does anyone know the rationale behind which method to use for 'u','m','M' and 'H','h','B','b'?
To have a base64 representation you need [s].pack('m') (and so #pack is used for 'u','m','M' to obtain the specified format), but for 'H','h','B','b' you need s.unpack(directive). I think the latter makes more sense when compared to other types such as 'l', in which #unpack is used to obtain the desired format.

Associated revisions

Revision b0e40509 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.  
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]  
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

P.S.: Does anyone know the rationale behind which method to use for 'u','m','M' and 'H','h','B','b'?
To have a base64 representation you need [s].pack('m') (and so #pack is used for 'u','m','M' to obtain the specified format), but for 'H','h','B','b' you need s.unpack(directive). I think the latter makes more sense when compared to other types such as 'l', in which #unpack is used to obtain the desired format.

Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.  
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]  
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.  
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]  
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.  
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]  
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.  
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]  
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]
Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

Revision 37269 - 10/19/2012 01:13 PM - Eregon (Benoit Daloze)

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

History

#1 - 09/23/2012 05:09 AM - Eregon (Benoit Daloze)

I would of course also add Aaron's test for SecureRandom hexdigests if this is fine to be merged.

#2 - 09/25/2012 10:58 AM - tenderlovemaking (Aaron Patterson)

I would really like this to be added! (Thank you for writing the patch)

#3 - 10/18/2012 07:33 AM - Eregon (Benoit Daloze)

Ping!

I would like to merge this.
If there are no objections, may I merge this in a week?

#4 - 10/19/2012 02:47 PM - naruse (Yui NARUSE)

Eregon (Benoit Daloze) wrote:

Ping!

I would like to merge this.
If there are no objections, may I merge this in a week?

OK, here you go

#5 - 10/19/2012 10:13 PM - Eregon (Benoit Daloze)

- Status changed from Open to Closed
- % Done changed from 0 to 100

This issue was solved with changeset r37269.
Benoit, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

set encoding to ASCII for appropriate String#unpack modifiers

- pack.c (pack_unpack): set encoding of the 'H','h','B' and 'B' modifiers to US-ASCII.
- test/ruby/test_pack.rb: tests for the above. [ruby-core:47653][Bug #7050]
- test/test_securerandom.rb: tests for SecureRandom.hex from tenderlove. [ruby-core:46792][Bug #6799]

#6 - 10/19/2012 10:53 PM - Eregon (Benoit Daloze)

On 19 October 2012 07:47, naruse (Yui NARUSE) naruse@airemix.jp wrote:

Eregon (Benoit Daloze) wrote:

Ping!

I would like to merge this.
If there are no objections, may I merge this in a week?

OK, here you go
Thank you!

Files

<table>
<thead>
<tr>
<th>File Name</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001-set-encoding-to-ASCII-for-appropriate-String-unpack-.patch</td>
<td>2.74 KB</td>
<td>09/23/2012</td>
<td>Eregon (Benoit Daloze)</td>
</tr>
</tbody>
</table>