Ruby trunk - Bug #14741

[].pack('M') quoted printable RFC2045 contains only LF, does not contain CRLF, Messages

05/08/2018 01:45 AM - uhrohraggy (Michael Moore)

Status:	Closed		
Priority:	Normal		
Assignee:			
Target version:			
ruby -v:	2.5.1p57	**Backport:**	2.3: UNKNOWN, 2.4: UNKNOWN, 2.5: UNKNOWN

Description

Important
You should contain at least the following items.

Reproduce process
your ruby version (ruby -v)
ruby 2.5.1p57 (2018-03-29 revision 63029) [x86_64-darwin17]

```
rails c
['aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabbbbbbbbcccc'].pack('M'
)
 => "aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa=\nbbbbbbbbcccc=\n"
```

Result of reproduce process:
Expected result and the reason why you expect:
Based on my (likely incorrect understanding of https://www.ietf.org/rfc/rfc2045.txt which [].pack('M') should be compliant with, all newlines should be formatted as CRLF and should only exist together, not separately. Relevant lines:

> 2.1. CRLF
>
> The term CRLF, in this set of documents, refers to the sequence of
> octets corresponding to the two US-ASCII characters CR (decimal value
> 13) and LF (decimal value 10) which, taken together, in this order,
> denote a line break in RFC 822 mail.
>
> 2.7. 7bit Data
>
> "7bit data" refers to data that is all represented as relatively
> short lines with 998 octets or less between CRLF line separation
> sequences [RFC-821]. No octets with decimal values greater than 127
> are allowed and neither are NULs (octets with decimal value 0). CR
> (decimal value 13) and LF (decimal value 10) octets only occur as
> part of CRLF line separation sequences.
>
> 2.8. 8bit Data
>
> "8bit data" refers to data that is all represented as relatively
> short lines with 998 octets or less between CRLF line separation
> sequences [RFC-821]), but octets with decimal values greater than 127
> may be used. As with "7bit data" CR and LF octets only occur as part
> of CRLF line separation sequences and no NULs are allowed.
>
> 2.10. Lines
>
> "Lines" are defined as sequences of octets separated by a CRLF
> sequences. This is consistent with both RFC 821 and RFC 822.
> "Lines" only refers to a unit of data in a message, which may or may
> not correspond to something that is actually displayed by a user
> agent.

Expected escaped sequence would be:

```
rails c
['aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabbbbbbbbcccc'].pack('M'
)
 => "aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa=\r\nbbbbbbbbcccc=
\r\n"
```

Related issues:

Related to Ruby trunk - Feature #14352: Array#pack("M") Quoted-Printable with... **Closed**

History

#1 - 05/08/2018 01:57 AM - shyouhei (Shyouhei Urabe)

- Related to Feature #14352: Array#pack("M") Quoted-Printable with binary mode added

#2 - 05/08/2018 02:05 AM - shyouhei (Shyouhei Urabe)

See also issue #14352. We have decided to reject this change. The behaviour is now documented.

To those who don't read Japanese let me summarize: Ruby's pack behabiour was from Perl's and Perl also don't emit CRLF. This is supposedly because in UNIX variants, emails are written in LF only and, to convert them to CRLF is done in sendmail(1). We don't want to break things here.

#3 - 05/08/2018 09:29 AM - matz (Yukihiro Matsumoto)

- Status changed from Open to Closed

#4 - 05/08/2018 12:48 PM - uhrohraggy (Michael Moore)

shyouhei (Shyouhei Urabe) wrote:

> See also issue #14352. We have decided to reject this change. The behaviour is now documented.
>
> To those who don't read Japanese let me summarize: Ruby's pack behabiour was from Perl's and Perl also don't emit CRLF. This is supposedly because in UNIX variants, emails are written in LF only and, to convert them to CRLF is done in sendmail(1). We don't want to break things here.

Understood, thanks shyouhei! Love how helpful the Ruby community is.

#5 - 05/09/2018 08:06 PM - shevegen (Robert A. Heiler)

> To those who don't read Japanese let me summarize:

Thanks for the summary in english, shyouhei!