## Ruby master - Bug #12423

**Regexp: Heap Buffer Overflow in regparse.c : next_state_value()**

05/24/2016 06:02 PM - grajagandev (David Moore)

<table>
<thead>
<tr>
<th>Status:</th>
<th>Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
</tr>
<tr>
<td>Assignee:</td>
<td></td>
</tr>
<tr>
<td>Target version:</td>
<td>ruby 2.3.1p112 (2016-04-26 revision 54768) [i686-linux]</td>
</tr>
<tr>
<td>Backport:</td>
<td>2.1: WONTFIX, 2.2: DONE, 2.3: DONE</td>
</tr>
</tbody>
</table>

### Description

A crafted regular expression will cause a heap buffer overflow leading to invalid 4 byte reads/writes on 32-bit Ubuntu 14.04. The regular expression fails to close a character class and has an octal zero as the first character in the character class.

Despite the buffer overflow, ruby does not crash. This bug may have the same root cause as [12420](https://example.com).

```ruby
grajagandev# cat load-re.rb
File.open(ARGV[0]) do |f|  
  @re = Regexp.new("/\n" + File.read(f) + "/")  
end

grajagandev# xxd overflow-next_state_value
0000000: 5b5c 3630 3030 3030 3030 3030 3030 0a  
  \[600000000000.

grajagandev# ruby -v
ruby 2.3.1p112 (2016-04-26 revision 54768) [i686-linux]
grajagandev# uname -a
Linux x-Acer 3.19.0-25-generic #26~14.04.1-Ubuntu SMP Fri Jul 24 21:18:00 UTC 2015 i686 i686 i686 GNU/Linux
grajagandev# ruby load-re.rb overflow-next_state_value
=================================================================
==11873== ERROR: AddressSanitizer: heap-buffer-overflow on address 0xb470cd88 thread T0
READ of size 4 at 0xb470cd88 05/06/2020
```
0xb470cd88 is located 8 bytes to the left of 44-byte region [0xb470cd90, 0xb470cdb0)

allocated by thread T0 here:

Shadow bytes around the buggy address:

0x368e1960: fa fa 00 00 00 00 fa fa fd fd fd fd fa
0x368e1970: fa fa 00 00 00 00 fa fa 00 00 00 00 00 04
0x368e1980: fa fa 00 00 00 00 00 00 04 fa fa 00 00 00 00 00 00 04
0x368e1990: fa fa 00 00 00 00 00 00 04 fa fa 00 00 00 00 00 00 04
0x368e19a0: fa fa 00 00 00 00 00 04 fa fa 00 00 00 00 00 00 04

=>0x368e19b0: fa[fa]00 00 00 00 00 04 fa fa 00 00 00 00 00 04
0x368e19c0: fa fa 00 00 00 00 00 00 04 fa fa 00 00 00 00 00 00 04
0x368e19d0: fa fa 00 00 00 00 00 00 04 fa fa 00 00 00 00 00 00 04
0x368e19e0: fa fa fd fd fd fd fa fa fa fa fd fd fd fd fa

5/06/2020 2/6
Invalid read of size 4

Address 0x4cb6180 is 12 bytes after a block of size 44 alloc'd

Address 0x4cb6180 is 12 bytes after a block of size 44 alloc'd
Related issues:
Related to Ruby master - Bug #12420: Regexp: Segfault due to Invalid Read in ... Closed

Associated revisions
Revision 05c631ee - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@55163 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision 55163 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

Revision e45cf75f - 05/25/2016 11:51 AM - naruse (Yui NARUSE)

- re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@55165 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 55165 - 05/25/2016 11:51 AM - naruse (Yui NARUSE)

- re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

Revision 55165 - 05/25/2016 11:51 AM - naruse (Yui NARUSE)

- re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

Revision 55165 - 05/25/2016 11:51 AM - naruse (Yui NARUSE)

- re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

Revision 55165 - 05/25/2016 11:51 AM - naruse (Yui NARUSE)

- re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

Revision 1a7e2194 - 06/10/2016 08:45 AM - usa (Usaku NAKAMURA)
merge revision(s) 55163,55165: [Backport #12420] [Backport #12423]

  * regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

  * re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_2@55363 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 55363 - 06/10/2016 08:45 AM - usa (Usaku NAKAMURA)
merge revision(s) 55163,55165: [Backport #12420] [Backport #12423]

  * regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

  * re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]
Revision 68c4c9d2 - 06/19/2016 05:35 PM - nagachika (Tomoyuki Chikanaga)
merge revision(s) 55163,55165: [Backport #12420]

* regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

* re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_3@55458 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 55458 - 06/19/2016 05:35 PM - nagachika (Tomoyuki Chikanaga)
merge revision(s) 55163,55165: [Backport #12420]

* regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

* re.c (unescape_nonascii): scan hex up to only 3 characters. [Bug #12420] [Bug #12423]

History

#1 - 05/25/2016 09:45 AM - naruse (Yui NARUSE)
- Status changed from Open to Closed

Applied in changeset r55163.

- regparse.c (fetch_token_in_cc): raise error if given octal escaped character is too big. [Bug #12420] [Bug #12423]

#2 - 06/08/2016 05:41 AM - usa (Usaku NAKAMURA)
- Backport changed from 2.1: UNKNOWN, 2.2: UNKNOWN, 2.3: UNKNOWN to 2.1: WONTFIX, 2.2: REQUIRED, 2.3: REQUIRED

#3 - 06/10/2016 08:44 AM - usa (Usaku NAKAMURA)
- Related to Bug #12420: Regexp: Segfault due to Invalid Read in regparse.c : bbuf_free() added

#4 - 06/10/2016 08:45 AM - usa (Usaku NAKAMURA)
- Backport changed from 2.1: WONTFIX, 2.2: REQUIRED, 2.3: REQUIRED to 2.1: WONTFIX, 2.2: DONE, 2.3: REQUIRED

ruby_2_2 r55363 merged revision(s) 55163,55165.

#5 - 06/19/2016 05:37 PM - nagachika (Tomoyuki Chikanaga)
- Backport changed from 2.1: WONTFIX, 2.2: DONE, 2.3: REQUIRED to 2.1: WONTFIX, 2.2: DONE, 2.3: DONE

ruby_2_3 r55458 merged revision(s) 55163,55165.

Files

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>load-re.rb</td>
<td>79 Bytes</td>
<td>05/24/2016</td>
<td>grajagandev (David Moore)</td>
</tr>
<tr>
<td>overflow-next_state_value</td>
<td>15 Bytes</td>
<td>05/24/2016</td>
<td>grajagandev (David Moore)</td>
</tr>
</tbody>
</table>