Ruby - Bug #5217

lineno is broken when source code has about 7000 lines

08/23/2011 11:42 PM - mame (Yusuke Endoh)

Status:	Closed		
Priority:	Normal		
Assignee:	ko1 (Koichi Sasada)		
Target version:			
ruby -v:	-	**Backport:**	

Description

遠藤です。

asakusa.rb で出た話だそうですが (なひさんからの伝聞) ´soap4r と
simplecov を組み合わせると [BUG] bug が出るそうです。

いろいろ話を聞かせてもらった結果、以下のようにすると再現できました。

```
$ ruby -e 'puts "p\n" * 7000; puts "p([1"; puts "])"' > t.rb

$ cat t.rb
p
p
p
```

... 7000 行ほど p ...

```
p
p

$ ./ruby -rcoverage -e 'Coverage.start; load "t.rb"'
```

応急措置ですが、以下のパッチで直ります。
(rb_bug("bug") というアホメッセージは私の所業のようです。スミマセン)
害はないと思うので、これはこれでコミットしておこうと思います。

```
diff --git a/thread.c b/thread.c
index 6970d8f..57a6962 100644
--- a/thread.c
+++ b/thread.c
@@ -4764,7 +4764,7 @@ update_coverage(rb_event_flag_t event, VALUE proc, VALUE self, ID id, VALUE klas
long line = rb_sourceline() - 1;
long count;
if (RARRAY_PTR(coverage)[line] == Qnil) {
```

- [redacted]

- [redacted]

```
    }
    count = FIX2LONG(RARRAY_PTR(coverage)[line]) + 1;
    if (POSFIXABLE(count)) {
```

しかしそもそもの原因は、iseq のバイトコードの大きさが unsigned short
で管理されているために、バイトコードの大きさが 65536 を超えると行番号を
見失ってしまうという問題のような気がします。
私も最初勘違いしたのですが、行番号が 16 ビットという問題とは別であること
に注意。65536 行ほど大きくなくても 7000 行くらいで発症します。

これにより、rb_vm_get_sourceline の挙動が怪しくなっています。
これは coverage に限らず、

set_trace_func(proc {|type, file, line,| p line if type == "line" })
p
p
p

... 7000 行ほど p ...

p
p
p

というコードの出力が、

2
3
4
5
...
6550
6551
6552
6553
7001
7001
7001
7001
7001
...

となります。行番号が飛んでいることがわかります。

直そうと思えば、以下のようにすれば直ります。しかしあえて short を選んで
いるのは省メモリ化のためだと思いますので、これは既知の制限ということで
しょうか > ささださん

```
diff --git a/iseq.h b/iseq.h
index beeacbb..9c19501 100644
--- a/iseq.h
+++ b/iseq.h
@@ -44,9 +44,9 @@ struct rb_compile_option_struct {
};

struct iseq_insn_info_entry {

-    unsigned short position;
-    unsigned short line_no;
-    unsigned short sp;

+    unsigned long position;
+    unsigned long line_no;
+    unsigned long sp;
     };

struct iseq_catch_table_entry {
```

--
Yuskue Endoh mame@tsg.ne.jp

Associated revisions

Revision 7049d9c8 - 08/24/2011 06:31 AM - ko1 (Koichi Sasada)

- iseq.h, iseq.c, compile.c: Change the line number data structure
  to solve an issue reported at [ruby-dev:44413] [Ruby 1.9 - Bug #5217].
  Before this fix, each instruction has an information including
  line number (iseq::iseq_insn_info_table). Instead of this data

structure, recording only line number changing places
(iseq::iseq_line_info_table).
The order of entries in iseq_line_info_table is ascending order of
iseq_line_info_table_entry::position. You can get a line number
by an iseq and a program counter with this data structure.
This fix reduces memory consumption of iseq (bytecode).
On my measurement, a rails application consumes 21.8MB for
iseq with this fix on the 32bit CPU. Without this fix, it
consumes 24.7MB for iseq [ruby-dev:44415].

- proc.c: ditto.
- vm_insnhelper.c: ditto.
- vm_method.c: ditto.
- vm.c (rb_vm_get_sourceline): change to use rb_iseq_line_no().

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@33046 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

History

#1 - 08/23/2011 11:43 PM - mame (Yusuke Endoh)

- Assignee set to ko1 (Koichi Sasada)

#2 - 08/24/2011 07:23 AM - ko1 (Koichi Sasada)

- ruby -v changed from ruby 1.9.4dev (2011-08-23 trunk 33019) [i686-linux] to -

□ささだです.

(2011/08/23 7:43), Yusuke Endoh wrote:

> 直そうと思えば、以下のようにすれば直ります。しかしあえて short を選んで
> いるのは省メモリ化のためだと思いますので、これは既知の制限ということで
> しょうか > ささださん

□どうもすみません.

□色々考えて，遠藤さんに IRC でアドバイスしてもらった，「改行位置が変
わったところだけ情報を付ける」という方法でやり直しました.

□ある大きな Rails アプリで，ロード後にバイトコード全部のサイズを測った
ところ (*1)，
改善前: 24,699,850 B (約 24.7MB)
改善後: 21,835,244 B (約 21.8MB)
となりました. 32bit 環境です: ruby 1.9.4dev (2011-08-22 trunk 33022)
[i686-linux].

*1: objspace を使うと，こんな風に測れます.

```
require 'objspace'
p ObjectSpace.memsize_of_all(RubyVM::InstructionSequence)
exit!
```

これを，`rack-*/lib/rack/handler/webrick.rb の self.run`
に無理矢理書き加えるというかっこ悪さ.
もうちょっとなんとかならないかな.

□バイトコードのサイズ自体は，fork した後で共有できるとかあるので，あん
まり削減しても嬉しくないような気もしますが，まぁちょっとは効果があったと
いうことで.

□パッチ自体は結構大きくなりました.
http://www.atdot.net/sp/view/wpheql/readonly?lang=diff

□コミットしてもいいでしょうか. って，trunk ならさくっと入れちゃえばいい
かなぁ.

--
// SASADA Koichi at atdot dot net

#3 - 08/24/2011 10:08 AM - nahi (Hiroshi Nakamura)

Koichi Sasada wrote:

> □色々考えて，遠藤さんに IRC でアドバイスしてもらった，「改行位置が変

わったところだけ情報を付ける」という方法でやり直しました.

ありがとうございました。

[]パッチ自体は結構大きくなりました.
http://www.atdot.net/sp/view/wpheql/readonly?lang=diff

[]コミットしてもいいでしょうか. って, trunk ならさくっと入れちゃえばいいかなぁ.

今のままでは評価を依頼しにくいので、入れてくださいな。

#4 - 08/24/2011 12:53 PM - mame (Yusuke Endoh)

遠藤です。

2011年8月24日10:08 Hiroshi Nakamura nakahiro@gmail.com:

> Koichi Sasada wrote:
>
> > [] 色々考えて, 遠藤さんに IRC でアドバイスしてもらった, 「改行位置が変わったところだけ情報を付ける」という方法でやり直しました.
>
> ありがとうございました。

おお。ささださんが一晩でやってくれました。

> [] パッチ自体は結構大きくなりました.
> http://www.atdot.net/sp/view/wpheql/readonly?lang=diff
>
> [] コミットしてもいいでしょうか. って, trunk ならさくっと入れちゃえばいいかなぁ.

trunk ならいいんじゃないですかね。
1.9.3 はいじらなくていいと思います。

> 今のままでは評価を依頼しにくいので、入れてくださいな。

元問題 ([BUG] bug) の話でしたら、trunk には前述の「応急措置」パッチをコミット済み (r33030) なので、ささださんのパッチに関わらず、もう [BUG] にはならなくなってると思います。

Yugui さん、このパッチ (r33030) は 1.9.3 にバックポート希望です。

--
Yusuke Endoh mame@tsg.ne.jp

#5 - 08/24/2011 01:53 PM - ko1 (Koichi Sasada)

(2011/08/23 20:43), Yusuke ENDOH wrote:

> おお。ささださんが一晩でやってくれました。

[]こっちは朝だったけどね. コミットします.

--
// SASADA Koichi at atdot dot net

#6 - 08/24/2011 03:31 PM - ko1 (Koichi Sasada)

- Status changed from Open to Closed

- % Done changed from 0 to 100

This issue was solved with changeset r33046.
Yusuke, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.

- iseq.h, iseq.c, compile.c: Change the line number data structure
  to solve an issue reported at [ruby-dev:44413] [Ruby 1.9 - Bug #5217].
  Before this fix, each instruction has an information including
  line number (iseq::iseq_insn_info_table). Instead of this data
  structure, recording only line number changing places
  (iseq::iseq_line_info_table).
  The order of entries in iseq_line_info_table is ascending order of
  iseq_line_info_table_entry::position. You can get a line number
  by an iseq and a program counter with this data structure.
  This fix reduces memory consumption of iseq (bytecode).
  On my measurement, a rails application consumes 21.8MB for
  iseq with this fix on the 32bit CPU. Without this fix, it
  consumes 24.7MB for iseq [ruby-dev:44415].
- proc.c: ditto.
- vm_insnhelper.c: ditto.
- vm_method.c: ditto.
- vm.c (rb_vm_get_sourceline): change to use rb_iseq_line_no().

#7 - 08/24/2011 06:23 PM - kosaki (Motohiro KOSAKI)

> 元問題 ([BUG] bug) の話でしたら、trunk には前述の「応急措置」
> パッチをコミット済み (r33030) なので、ささださんのパッチに関わらず、
> もう [BUG] にはならなくなってると思います。
>
> Yugui さん、このパッチ (r33030) は 1.9.3 にバックポート希望です。

脱線ですが、そろそろbackport projectに1.9.3を追加してチケットで管理しませんか?
メールでバックポート依頼はあぶなっかしいと思います。

#8 - 08/24/2011 07:53 PM - ursm (Keita Urashima)

浦嶌と申します。
Asakusa.rb でなひさんに相談に乗っていただきました。

[BUG] が出ていたコードを r33046 で試してみたところ、正常に動いているように見えました。
念のため以下のパッチを当てた状態でも試してみましたが、落ちるようなことはありませんでした。

ご対応ありがとうございました。

```
diff --git a/thread.c b/thread.c
index 57a6962..880e5f8 100644
--- a/thread.c
+++ b/thread.c
@@ -4764,7 +4764,7 @@ update_coverage(rb_event_flag_t event, VALUE
proc, VALUE self, ID id, VALUE klas
long line = rb_sourceline() - 1;
long count;
if (RARRAY_PTR(coverage)[line] == Qnil) {
-    [redacted]
-    [redacted]
+    }
     count = FIX2LONG(RARRAY_PTR(coverage)[line]) + 1;
     if (POSFIXABLE(count)) {
```

#9 - 08/24/2011 09:53 PM - mame (Yusuke Endoh)

2011年8月24日18:09 KOSAKI Motohiro kosaki.motohiro@gmail.com:

> 元問題 ([BUG] bug) の話でしたら、trunk には前述の「応急措置」
> パッチをコミット済み (r33030) なので、ささださんのパッチに関わらず、
> もう [BUG] にはならなくなってると思います。
>
> Yugui さん、このパッチ (r33030) は 1.9.3 にバックポート希望です。

backport project 1.9.3

--
Yusuke Endoh mame@tsg.ne.jp

#10 - 08/24/2011 09:59 PM - mame (Yusuke Endoh)

2011 8 24 19:30 Keita Urashima ursm@ursm.jp:

Asakusa.rb

[BUG] r33046

--
Yusuke Endoh mame@tsg.ne.jp

#11 - 08/24/2011 11:36 PM - mame (Yusuke Endoh)

2011 8 24 10:08 Hiroshi Nakamura nakahiro@gmail.com:

Koichi Sasada wrote:

IRC

http://www.atdot.net/sp/view/wpheql/readonly?lang=diff

trunk

trunk
1.9.3

([BUG] bug) trunk
(r33030)
[BUG]

Yugui (r33030) 1.9.3

--
Yusuke Endoh mame@tsg.ne.jp

#12 - 08/24/2011 11:53 PM - yugui (Yuki Sonoda)

2011/8/24 KOSAKI Motohiro kosaki.motohiro@gmail.com:

> 脱線ですが、そろそろbackport projectに1.9.3を追加してチケットで管理しませんか?
> メールでバックポート依頼はあぶなっかしいと思います。

done

--
Yuki Sonoda (Yugui)
yugui@yugui.jp
http://yugui.jp