Redmine - Patch #28940

Use Regexp#match? to reduce allocations of MatchData object

2018-06-04 01:24 - Pavel Rosický

Status:	Closed		**Start date:**	
Priority:	Normal		**Due date:**	
Assignee:	Go MAEDA		**% Done:**	0%
Category:	Performance		**Estimated time:**	0.00 hour
Target version:	4.1.0			

Description

since Rails 5.1+ we can use match? on a regex class safely even on older rubies
https://github.com/rails/rails/blob/5-1-stable/activesupport/lib/active_support/core_ext/regexp.rb

but the performance benefit is visible only on ruby 2.4+
https://bugs.ruby-lang.org/issues/8110

```
require 'benchmark/ips'
Benchmark.ips do |x|
  x.report('match') { /test\d/.match 'test5'.freeze }
  x.report('match?') { /test\d/.match? 'test5'.freeze }
  x.compare!
end

Comparison:
match?:  4493322.3 i/s
match:    926754.9 i/s - 4.85x  slower
```

rake test is about 5% faster

Related issues:

Related to Redmine - Patch #28939: replace regexp with casecmp	**Closed**

Associated revisions

Revision 18011 - 2019-03-27 03:15 - Go MAEDA

Use Regexp#match? to reduce allocations of MatchData object (#28940).

Patch by Pavel Rosický.

History

#1 - 2018-06-04 02:18 - Pavel Rosický

- File match_adapters.patch added

#2 - 2018-06-04 02:55 - Go MAEDA

It is a very interesting patch. But MailHandlerTest fails on my environment (Ruby 2.3).

```
$ ruby test/unit/mail_handler_test.rb
Run options: --seed 1974

# Running:

......................E

Error:
MailHandlerTest#test_truncate_emails_with_a_single_quoted_reply_should_truncate_the_email_at_the_delimiter_wit
h_the_quoted_reply_symbols_(>):
NoMethodError: undefined method `match?' for #<String:0x007f91d84ebe90>
Did you mean?  match
    test/unit/mail_handler_test.rb:1005:in `block (2 levels) in <class:MailHandlerTest>'
    test/test_helper.rb:93:in `with_settings'
    test/unit/mail_handler_test.rb:1001:in `block in <class:MailHandlerTest>'

bin/rails test test/unit/mail_handler_test.rb:1000
```

```
........E

Error:
MailHandlerTest#test_truncate_emails_with_multiple_quoted_replies_should_truncate_the_email_at_the_delimiter_w
ith_the_quoted_reply_symbols_(>):
NoMethodError: undefined method `match?' for #<String:0x007f91e3ad4d38>
Did you mean?  match
    test/unit/mail_handler_test.rb:1015:in `block (2 levels) in <class:MailHandlerTest>'
    test/test_helper.rb:93:in `with_settings'
    test/unit/mail_handler_test.rb:1011:in `block in <class:MailHandlerTest>'

bin/rails test test/unit/mail_handler_test.rb:1010

.........................................................

Finished in 21.213978s, 4.2896 runs/s, 20.2697 assertions/s.
91 runs, 430 assertions, 0 failures, 2 errors, 0 skips
```

The following sentence causes one of the error. Regexp.escape returns a String object. ActiveSupport defines Regexp#match? but does not String#match?.

```
    assert !Regexp.escape("--- Reply above. Do not remove this line. ---").match?(journal.notes)
```

#3 - 2018-06-04 03:14 - Pavel Rosický

You're right. I will revert lines with Regexp.escape tomorrow, but otherwise it should work.

#4 - 2018-06-04 03:22 - Mizuki ISHIKAWA

+1
Redmine uses a lot of match methods.
I think it's wonderful that the test will be 5% faster.

#5 - 2018-06-04 03:28 - Go MAEDA

- Related to Patch #28939: replace regexp with casecmp added

#6 - 2018-06-04 05:42 - Go MAEDA

- File match-adapters-v2.diff added

I have updated match-adapters.patch. The patch does not work due to syntax errors. "ActiveRecord::Base.connection.connection.adapter_name" must be replaced with "ActiveRecord::Base.connection.adapter_name" (one extra ".connection").

#7 - 2018-06-04 06:54 - Go MAEDA

- File match-v2.diff added

Pavel Rosický wrote:

> I will revert lines with Regexp.escape tomorrow,

Done. I am attaching an updated patch.

#8 - 2018-06-04 06:55 - Go MAEDA

- Target version set to Candidate for next major release

#9 - 2018-06-09 11:40 - Go MAEDA

- Target version changed from Candidate for next major release to 4.1.0

Setting target version to 4.1.0.

#10 - 2018-06-17 18:23 - Pavel Rosický

- File match_final.patch added

I combined all patches into a single file and I retested it on Ruby 2.3 and Ruby 2.4

#11 - 2019-03-21 23:01 - Pavel Rosický

- File match_final.patch added

#12 - 2019-03-25 15:49 - Go MAEDA

Pavel Rosický wrote:

> I combined all patches into a single file and I retested it on Ruby 2.3 and Ruby 2.4

Thank you for updating the patch. I have confirmed that the patch passes all test in r18003.

Regexp#match? is faster than Regexp#match or =~ in most cases because it does not allocate MatchData object. I think there is no reason not to merge this patch.

#13 - 2019-03-27 03:15 - Go MAEDA

- Subject changed from reduce allocations to Use Regexp#match? to reduce allocations of MatchData object

- Assignee set to Go MAEDA

Committed the patch. Thank you for posting many performance-tuning patches.

#14 - 2019-03-27 03:16 - Go MAEDA

- Status changed from New to Closed

Files

match.patch	34 KB	2018-06-03	Pavel Rosický
match_adapters.patch	3.23 KB	2018-06-04	Pavel Rosický
match-adapters-v2.diff	1.97 KB	2018-06-04	Go MAEDA
match-v2.diff	13.6 KB	2018-06-04	Go MAEDA
match_final.patch	27.7 KB	2018-06-17	Pavel Rosický
match_final.patch	28.3 KB	2019-03-21	Pavel Rosický