

Use Regexp#match? to reduce allocations of MatchData object

2022/05/09 18:33 - Admin Redmine

ステータス:	Closed	開始日:	2022/05/09
優先度:	通常	期日:	
担当者:		進捗率:	0%
カテゴリ:	Performance_53	予定工数:	0.00時間
対象バージョン:	4.1.0_127	作業時間:	0.00時間
Redmineorg_URL:	https://www.redmine.org/issues/28940	status_id:	5
category_id:	53	tracker_id:	3
version_id:	127	plus1:	1
issue_org_id:	28940	affected_version:	
author_id:	123153	closed_on:	
assigned_to_id:	332	affected_version_id:	
comments:	14		

説明

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/activerecord/lib/active_record/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

journals

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_with_the_quoted_reply_symbols(>):
NoMethodError: undefined method `match?' for #
Did you mean? match
test/unit/mail_handler_test.rb:1005:in `block (2 levels) in '
test/test_helper.rb:93:in `with_settings'
test/unit/mail_handler_test.rb:1001:in `block in '
```

```
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_with_the_quote_d_reply_symbols(>):
NoMethodError: undefined method `match?' for #
Did you mean? match
test/unit/mail_handler_test.rb:1015:in `block (2 levels) in '
test/test_helper.rb:93:in `with_settings'
test/unit/mail_handler_test.rb:1011:in `block in '

```
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)
```

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

+1
Redmine uses a lot of match methods.

I think it's wonderful that the test will be 5% faster.

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@").

Pavel Rosický wrote:

I will revert lines with Regexp.escape tomorrow,

Done. I am attaching an updated patch.

Setting target version to 4.1.0.

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

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.

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

related_issues

relates,Closed,28939,replace regexp with casecmp

履歴

#1 - 2022/05/10 17:04 - Admin Redmine

- カテゴリ を Performance_53 にセット

- 対象バージョン を 4.1.0_127 にセット