

Redmine - Patch #29359

Switch to mini_mime from mime-types

2018-08-10 21:44 - Pavel Rosický

Status:	Closed	Start date:	
Priority:	Normal	Due date:	
Assignee:	Go MAEDA	% Done:	0%
Category:	Performance	Estimated time:	0.00 hour
Target version:	4.0.0		

Description

Redmine uses unbound cache for mime types which is eliminated now. Also mime-types is too heavy and Redmine uses only a tiny piece of it, we just need a simple ext -> mime lookup.

<https://github.com/mikel/mail/pull/1059> (mail 2.6.x still uses mime-types)

<https://github.com/teamcapybara/capybara/pull/1884>

```
require 'benchmark/ips'
```

```
Benchmark.ips do |x|
  x.report('known') { Redmine::MimeType.of("file.patch") }
  x.report('not known') { Redmine::MimeType.of("file.zip") }
end
```

trunk

```
known 284.431k (± 1.9%) i/s - 1.422M in 5.002378s
not known 285.699k (± 3.0%) i/s - 1.446M in 5.065113s
```

patch

```
known 657.114k (± 1.8%) i/s - 3.327M in 5.064494s
not known 392.574k (± 2.2%) i/s - 1.965M in 5.006827s
```

boot allocations boot retained

```
mime-types      109796    31165
mini_mime       398        62
```

what do you think?

Related issues:

Related to Redmine - Defect # 29365: MailHandlerTest#test_add_issue_with_loca...	Closed
Related to Redmine - Patch # 29383: Update mini_mime gem (~> 1.0.1)	Closed
Related to Redmine - Feature # 29443: Update mail gem (~> 2.7.1)	Closed

Associated revisions

Revision 17468 - 2018-08-13 01:37 - Go MAEDA

Replace mime-types gem with more efficient mini_mime gem (#29359).

Contributed by Pavel Rosický.

Revision 17470 - 2018-08-13 13:38 - Go MAEDA

Don't require mime-types (#29359).

Revision 17471 - 2018-08-15 03:23 - Go MAEDA

Test that Redmine::MimeType.of is case-insensitive (#29359).

Revision 17472 - 2018-08-15 15:20 - Go MAEDA

Test that Redmine::MimeType.css_class_of, main_mimetype_of and is_type? are case-insensitive (#29359).

Revision 17473 - 2018-08-15 16:01 - Go MAEDA

Refactor Redmine::MimeType.of. Uses MiniMime.lookup_by_extension instead of lookup_by_filename (#29359).

History

#1 - 2018-08-11 08:03 - Go MAEDA

- File *replace-mime-types-with-mini_mime.diff* added
- Category changed from *Gems support* to *Performance*

Thank you for posting the interesting improvement.

Your patch is very fast, but I think an application should not have codes which rely on the internal structure of third-party libraries. Your patch accesses `mini_mime`'s instance variable '@db'.

I updated your patch not to manipulate an internal variable of `mini_mime`. It is 15% slower than your patch but 70% faster than the current `Redmine::MimeType.of`.

without patches:

known	226.301k (± 5.0%) i/s -	1.131M in	5.010006s
unknown	225.491k (± 4.6%) i/s -	1.131M in	5.024550s

replace-mime-types-with-mini_mime.diff: (faster):

known	382.020k (± 6.2%) i/s -	1.923M in	5.052268s
unknown	382.602k (± 6.6%) i/s -	1.908M in	5.008955s

mime_type.rb.patch by Pavel Rosický (fastest):

known	431.349k (± 5.4%) i/s -	2.167M in	5.038918s
unknown	432.307k (± 5.5%) i/s -	2.172M in	5.040046s

#2 - 2018-08-11 08:27 - Go MAEDA

Sorry, there was a problem with my benchmark script. After fixing the problem, I found that my patch is very slow for unknown extensions.

without patches:

```
known 170.009k (± 7.2%) i/s - 859.560k in 5.083952s
unknown 169.572k (± 6.2%) i/s - 844.816k in 5.000900s
```

replace-mime-types-with-mini_mime.diff:

```
known 244.827k (± 5.3%) i/s - 1.228M in 5.031343s
unknown 79.055k (± 5.5%) i/s - 394.128k in 5.000338s
```

mime_type.rb.patch by Pavel Rosický:

```
known 256.660k (± 5.7%) i/s - 1.291M in 5.047902s
unknown 150.151k (± 5.5%) i/s - 753.483k in 5.033494s
```

```
require 'benchmark/ips'
require 'mini_mime'
require './lib/redmine/mime_type'
```

```
Benchmark.ips do |x|
  x.report('known') { Redmine::MimeType.of('file.txt') }
  x.report('unknown') { Redmine::MimeType.of('file.zip') }
end
```

#3 - 2018-08-11 09:37 - Go MAEDA

- File *replace-mime-types-with-mini_mime-v2.diff* added

I updated my patch to preserve existing cache mechanism. Now it is fast also for unknown extensions.

```
known 232.028k (± 4.9%) i/s - 1.170M in 5.052806s
unknown 236.938k (± 4.6%) i/s - 1.199M in 5.071712s
```

#4 - 2018-08-11 09:42 - Pavel Rosický

Lets ask sam saffron if he could add it to the public api.

#5 - 2018-08-11 10:02 - Pavel Rosický

@known_types is a class variable. There're two problems:

It isn't threadsafe.

It's unlimited and because it's related to user's actions it could grow indefinitely. This memory won't be released. That's why I was trying to avoid it or limit it. Minimime has the upper bound for caches. Please correct me if I'm wrong.

#6 - 2018-08-12 09:10 - Go MAEDA

Pavel Rosický wrote:

@known_types is a class variable. There're two problems:

It isn't threadsafe.

It's unlimited and because it's related to user's actions it could grow indefinitely. This memory won't be released.

Although the patch still has the same problem with the current implementation, simply replacing mime-types with mini_mime will reduce memory usage and improve performance.

I suggest that we discuss the problem as a new issue, after committing the patch. Is it OK with you?

#7 - 2018-08-12 09:24 - Pavel Rosický

It's ok, thanks

#8 - 2018-08-12 09:58 - Go MAEDA

- Subject changed from Switch to mini_mime to Switch to mini_mime from mime-types

- Target version set to 4.1.0

#9 - 2018-08-12 11:43 - Go MAEDA

- File replace-mime-types-with-mini_mime-v3.diff added

#10 - 2018-08-13 01:37 - Go MAEDA

- Status changed from New to Closed

- Assignee set to Go MAEDA

Committed. Thank you for your contribution.

#11 - 2018-08-13 01:45 - Go MAEDA

- Target version changed from 4.1.0 to 4.0.0

#12 - 2018-08-13 01:56 - Go MAEDA

- Related to Defect #29365: MailHandlerTest#test_add_issue_with_localized_attributes fails with mail gem 2.7.0 added

#13 - 2018-08-13 11:18 - Pavel Rosický

please also remove

require 'mime/types'

#14 - 2018-08-13 11:28 - Marius BALTEANU

- Status changed from Closed to Reopened

#15 - 2018-08-13 13:39 - Go MAEDA

- Status changed from Reopened to Closed

Pavel Rosický wrote:

```
| please also remove  
| [...]
```

Committed. Thanks.

#16 - 2018-08-14 12:09 - Go MAEDA

- File 29359-remove-known-types-hash.patch added

- Status changed from Closed to Reopened

mini_mime 1.0.1 supports lookup_by_extension. We can remove the cache mechanism using known_types hash by making use of the new method. Attaching a patch to remove the hash.

Even without the cache mechanism, Redmine::MimeTypes.of is faster than the original code.

r17467 (before applying replace-mime-types-with-mini_mime-v3.diff):

```
Warming up -----  
  known  20.028k i/100ms  
 unknown 12.677k i/100ms  
Calculating -----  
  known  246.855k (± 6.0%) i/s - 1.242M in 5.049235s  
 unknown 138.044k (± 5.9%) i/s - 697.235k in 5.068005s
```

with the patch 29359-remove-known-types-hash.patch:

```
Warming up -----  
  known  20.249k i/100ms  
 unknown 11.787k i/100ms  
Calculating -----  
  known  264.152k (± 8.0%) i/s - 1.316M in 5.015411s  
 unknown 142.650k (± 8.6%) i/s - 719.007k in 5.076025s
```

#17 - 2018-08-14 13:18 - Pavel Rosický

29359-remove-known-types-hash.patch - .downcase shouldn't be removed, both EXTENSIONS[] and lookup_by_extension expect lower-case and the benchmark looks odd, in r17467 both known/unknown scenarios were cached in a class var, so I would expect similar performance

#18 - 2018-08-14 17:19 - Go MAEDA

- File 29359-remove-known-types-hash-v2.patch added

Pavel Rosický wrote:

| 29359-remove-known-types-hash.patch - .downcase shouldn't be removed, both EXTENSIONS[] and lookup_by_extension expect lower-case

Fixed.

| and the benchmark looks odd, in r17467 both known/unknown scenarios were cached in a class var, so I would expect similar performance

You are right. Maybe I pasted a wrong result. The following results are correct ones. Redmine::MimeType.of will be faster than r17467 in many cases.

r17467 (before switching to mini_mime):

```
found in array (file.patch)
    164.476k (± 6.5%) i/s - 832.300k in 5.081748s
found in gem (file.zip)
    167.251k (± 5.7%) i/s - 841.580k in 5.047872s
not found in gem (file.go)
    168.700k (± 5.7%) i/s - 851.440k in 5.063132s
```

with the patch:

```
found in array (file.patch)
    418.181k (± 6.0%) i/s - 2.099M in 5.036624s
found in gem (file.zip)
    183.820k (± 6.9%) i/s - 923.232k in 5.047995s
not found in gem (file.go)
    147.241k (± 6.7%) i/s - 734.046k in 5.007225s
```

benchmark script:

```
require 'benchmark/ips'
require 'mini_mime'
require './lib/redmine/mime_type'
```

```
Benchmark.ips do |x|
  x.report('found in array (file.patch)') { Redmine::MimeType.of('file.patch') }
  x.report('found in gem (file.zip)') { Redmine::MimeType.of('file.zip') }
```

```
x.report('not found in gem (file.go)') { Redmine::MimeType.of('file.go') }  
end
```

#19 - 2018-08-14 17:53 - Pavel Rosický

my results (ruby 2.5.0):

with the patch (mini_mime)

```
found in array (file.patch)  
    1.079M (± 1.9%) i/s - 5.407M in 5.011560s  
found in gem (file.zip)  
    486.991k (± 2.6%) i/s - 2.456M in 5.047320s  
not found in gem (file.go)  
    428.810k (± 2.3%) i/s - 2.174M in 5.072125s
```

r17467 (mime_types)

```
found in array (file.patch)  
    293.348k (± 1.6%) i/s - 1.477M in 5.035304s  
found in gem (file.zip)  
    295.426k (± 1.8%) i/s - 1.500M in 5.080399s  
not found in gem (file.go)  
    293.534k (± 5.3%) i/s - 1.477M in 5.048996s
```

r17467 (mime_types without @known_types cache)

```
found in array (file.patch)  
    298.984k (± 5.2%) i/s - 1.499M in 5.027456s  
found in gem (file.zip)  
    60.804k (± 1.6%) i/s - 304.803k in 5.014262s  
not found in gem (file.go)  
    65.405k (± 2.9%) i/s - 332.575k in 5.089481s
```

#20 - 2018-08-15 02:09 - Go MAEDA

- Related to Patch #29383: Update mini_mime gem (~> 1.0.1) added

#21 - 2018-08-15 16:02 - Go MAEDA

- Status changed from Reopened to Closed

Committed 29359-remove-known-types-hash-v2.patch.

#22 - 2018-08-27 13:54 - Go MAEDA

- Related to Feature #29443: Update mail gem (~> 2.7.1) added

Files

Gemfile.patch	381 Bytes	2018-08-10	Pavel Rosický
mime_type.rb.patch	1.43 KB	2018-08-10	Pavel Rosický
replace-mime-types-with-mini_mime.diff	1.38 KB	2018-08-11	Go MAEDA
replace-mime-types-with-mini_mime-v2.diff	1.17 KB	2018-08-11	Go MAEDA
replace-mime-types-with-mini_mime-v3.diff	1.14 KB	2018-08-12	Go MAEDA
29359-remove-known-types-hash.patch	1.13 KB	2018-08-14	Go MAEDA
29359-remove-known-types-hash-v2.patch	1.14 KB	2018-08-14	Go MAEDA