

Redmine - Defect #24992

MailHandler: an unexpected error occurred when receiving email: invalid byte sequence in UTF-8

2017-02-03 19:10 - Deoren Moor

Status:	Needs feedback	Start date:	
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	Email receiving	Estimated time:	0.00 hour
Target version:		Affected version:	
Resolution:			
Description			
Problem description			
We have set RAILS_ENV to development for our email scraping, so we were fortunate enough to have this logged when we found that several emails had been rejected:			
MailHandler: an unexpected error occurred when receiving email: invalid byte sequence in UTF-8 Message 190 can not be processed			
The subject line of the email had this text:			
Boletín de entomología venezolana.			
Redmine installation info			
# bin/about			
ShowHide			
/var/lib/gems/2.3.0/gems/htmlentities-4.3.1/lib/htmlentities/mappings/expanded.rb:465: warning: key "inodot" is duplicated and overwritten on line 466			
sh: 1: darcs: not found sh: 1: hg: not found sh: 1: cvs: not found sh: 1: bzip: not found sh: 1: git: not found			
Environment:			
Redmine version	3.3.2.stable.16201		
Ruby version	2.3.1-p112 (2016-04-26) [x86_64-linux-gnu]		
Rails version	4.2.7.1		
Environment	production		
Database adapter	Mysql2		
SCM:			
Subversion	1.9.3		
Filesystem			
Redmine plugins:			
no plugin installed			
# bundle list			
ShowHide			
Gems included by the bundle:			
* actionmailer (4.2.7.1)			
* actionpack (4.2.7.1)			
* actionpack-action_caching (1.1.1)			
* actionpack-xml_parser (1.0.2)			
* actionview (4.2.7.1)			
* activejob (4.2.7.1)			
* activemodel (4.2.7.1)			

```
* activerecord (4.2.7.1)
* activesupport (4.2.7.1)
* addressable (2.4.0)
* arel (6.0.3)
* builder (3.2.2)
* bundler (1.11.2)
* coderay (1.1.1)
* concurrent-ruby (1.0.2)
* css_parser (1.4.5)
* erubis (2.7.0)
* globalid (0.3.7)
* htmlentities (4.3.1)
* i18n (0.7.0)
* jquery-rails (3.1.4)
* json (1.8.3)
* loofah (2.0.3)
* mail (2.6.4)
* mime-types (3.1)
* mime-types-data (3.2016.0521)
* mimemagic (0.3.2)
* mini_portile2 (2.1.0)
* minitest (5.9.1)
* mysql2 (0.3.21)
* net-ldap (0.12.1)
* nokogiri (1.6.8.1)
* protected_attributes (1.1.3)
* rack (1.6.4)
* rack-openid (1.4.2)
* rack-test (0.6.3)
* rails (4.2.7.1)
* rails-deprecated_sanitizer (1.0.3)
* rails-dom-testing (1.0.7)
* rails-html-sanitizer (1.0.3)
* railties (4.2.7.1)
* rake (11.3.0)
* rbpdf (1.19.0)
* rbpdf-font (1.19.0)
* rdoc (4.3.0)
* redcarpet (3.3.4)
* request_store (1.0.5)
* rmagick (2.16.0)
* roadie (3.2.0)
* roadie-rails (1.1.1)
* ruby-openid (2.3.0)
* sprockets (3.7.0)
* sprockets-rails (3.2.0)
* thor (0.19.1)
* thread_safe (0.3.5)
* tzinfo (1.2.2)
* yard (0.9.5)
```

I don't know what the fix actually is, but this [Stackoverflow.com](http://stackoverflow.com)¹ thread mentioned this:

The combination of using: `@file = IO.read(file).force_encoding("ISO-8859-1").encode("utf-8", replace: nil)` and `#encoding: UTF-8` solved the issue.

I don't have Ruby knowledge, but I'll start digging through the code on my end to see if I can find where to make those changes. Please let me know if you need additional information.

¹ <http://stackoverflow.com/questions/9607554/ruby-invalid-byte-sequence-in-utf-8>

Related issues:

Related to Redmine - Feature #31921: Changes to properly support 4 byte chara...

Closed

Has duplicate Redmine - Defect #27361: Failed when using Emoji

Closed

History

#1 - 2017-02-04 03:28 - Toshi MARUYAMA

- Status changed from New to Needs feedback

What is your MySQL encoding?

<http://stackoverflow.com/questions/tagged/utf8mb4>

#2 - 2017-02-04 03:46 - Deoren Moor

utf8

I used the directions on the [RedmineInstall](#) page:

```
CREATE DATABASE redmine CHARACTER SET utf8;
```

Should it be something else instead?

#3 - 2017-02-04 04:24 - Toshi MARUYAMA

<https://www.google.com/search?q=utf8+4byte+mysql>

#4 - 2017-02-04 06:10 - Deoren Moor

Toshi,

I understand that you wish to draw my attention to MySQL, utf8 and utf8mb4. Reading over the links you've provided, I understand that MySQL's UTF-8 implementation is only a subset (1-3 bytes) of the complete UTF-8 character set (1-4 bytes). utf8mb4 appears to be MySQL's complete implementation of UTF-8 for versions 5.5.3 and later. I also understand that current versions of PostgreSQL and SQLite3 support the full UTF8 character set.

Based on the [RedmineInstall](#) page, I understand that Redmine still supports MySQL 5.0 which doesn't support utf8mb4. Even if you can't update the installation docs to recommend that *all* MySQL users go with utf8mb4, perhaps by adding a note that all users of MySQL 5.5.3 and newer should use utf8mb4 for full UTF-8 character support would help prevent the need for later conversion.

I found this page:

<https://mathiasbynens.be/notes/mysql-utf8mb4>

which seems to indicate that converting from MySQL's utf8 to utf8mb4 is doable, but quite a chore. Are those steps complete? Is there support already in the upgrade/maintenance Redmine scripts for converting an existing utf8 character set db to utf8mb4?

I've never had to convert a database's character set before, so any tips/tricks you have would be welcome.

Thanks.

EDIT

I found this post which seems a bit easier to follow:

<https://pento.net/2014/04/07/wordpress-and-utf-8/>

#5 - 2017-03-22 18:31 - Rol Sch

Deoren Moor wrote:

We have set RAILS_ENV to **development** for our email scraping, so we were fortunate enough to have this logged when we found that several emails had been rejected:

Hello Deoren, we seem to be running into a similar issue, but I don't know where to look for the relevant log file. There is no error in the redmine log and none in the rake output. In which log file did you find the error?

Thanks!

#6 - 2017-03-22 18:43 - Deoren Moor

Rol Sch wrote:

Deoren Moor wrote:

We have set RAILS_ENV to **development** for our email scraping, so we were fortunate enough to have this logged when we found that several emails had been rejected:

Hello Deoren, we seem to be running into a similar issue, but I don't know where to look for the relevant log file. There is no error in the redmine log and none in the rake output. In which log file did you find the error?
Thanks!

In our case we have Redmine installed via a SVN working copy at /opt/redmine/. There the default directory layout applies. Since we use the development RAILS environment (RAILS_ENV) for our scraping script, the /opt/redmine/log/development.log file is created and holds details related to the scraping process.

#7 - 2017-03-23 11:24 - Rol Sch

Thanks! I believe now that we do have the same issue:

MailHandler: an unexpected error occurred when receiving email: Mysql2::Error: Incorrect string value: '\xC5\x82owy ...' for column 'description' at row 1: INSERT INTO `issues` (`project_id`, `author_id`, `priority_id`, `tracker_id`, `status_id`, `subject`, `description`, `start_date`, `updated_on`, `created_on`, `lock_version`) VALUES (2, 2, 4, 1, 1, 'Test', 'Dodaj szczegółowy opis błędu\r\n\r\n*****\r\n\r\nVersion: 3.9.2.126\r\n\r\n*****', '2017-03-23', '2017-03-23 10:59:41', '2017-03-23 10:59:41', 0)
Message 31 can not be processed

#8 - 2017-03-23 13:00 - Rol Sch

The issue is indeed UTF8/UTF8mb4.

I found a good guide how to convert the mysql redmine database including a conversion script here:

<http://www.pflaeging.net/blog/archives/938>

#9 - 2017-04-13 16:40 - Library Systems Dept

Rol Sch wrote:

The issue is indeed UTF8/UTF8mb4.

I found a good guide how to convert the mysql redmine database including a conversion script here:

<http://www.pflaeging.net/blog/archives/938>

Do you know if this is proven to be a safe option? For example, if/when Redmine updates are released in the future, do those database schema migrations expect the character set to be configured specifically as utf8 or are those changes separate from the character set?

#10 - 2017-11-03 16:05 - Toshi MARUYAMA

- Has duplicate Defect #27361: Failed when using Emoji added

#11 - 2018-11-02 06:47 - Juozapis Juozapauskiksi

Rol Sch wrote:

The issue is indeed UTF8/UTF8mb4.

I found a good guide how to convert the mysql redmine database including a conversion script here:

<http://www.pflaeging.net/blog/archives/938>

I believe, that URL address of article changed to: <https://www.pflaeging.net/blog/redmine-8bit-sicher-f%C3%BCr-stickie~flow-installieren/> or <https://www.pflaeging.net/blog/redmine-und-emojis-probleme-mit-utf8-und-sonderzeichen/>

#12 - 2019-07-31 09:55 - Tobias Fischer

Related with more discussion: [#18866](#) and [#21398](#)

#13 - 2019-08-17 18:13 - Marius BĂLTEANU

- Related to Feature #31921: Changes to properly support 4 byte characters (emoji) when database is MySQL added