Redmine - Defect #23055
Error with Fetch commits with Mercurial repository when log has invalid char
2016-06-13 13:31 - mole mole

<table>
<thead>
<tr>
<th>Status:</th>
<th>Closed</th>
<th>Start date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
<td>Due date:</td>
</tr>
<tr>
<td>Assignee:</td>
<td>Toshi MARUYAMA</td>
<td>% Done: 0%</td>
</tr>
<tr>
<td>Category:</td>
<td>SCM</td>
<td>Estimated time: 0.00 hour</td>
</tr>
<tr>
<td>Target version:</td>
<td>4.2.0</td>
<td>Affected version: 3.1.3</td>
</tr>
<tr>
<td>Resolution:</td>
<td>Fixed</td>
<td></td>
</tr>
</tbody>
</table>

Description
Error with Fetch commits automatically with mercurial repository

When I add a mercurial repository, redmine start to scan and add to Mysql. This take several minutes. When redmine finish to index .Redmine show a error page. In production.log show this message. It happen with a huge repository about 70335 commits

production.log

Started GET "/projects/iset/repository" for 192.168.1.164 at 2016-06-13 10:56:04 +0000
Processing by RepositoriesController#show as HTML
  Parameters: {"id"=>"iset"}
  Current user: admin (id=1)
Completed 500 Internal Server Error in 2039ms (ActiveRecord: 4.8ms)

NoMethodError (undefined method `[]' for nil:NilClass):
  lib/redmine/scm/adapters/mercurial_adapter.rb:194:in `each_revision'
  app/models/repository/mercurial.rb:187:in `block in fetch_changesets'
  app/models/repository/mercurial.rb:186:in `step'
  app/models/repository/mercurial.rb:186:in `fetch_changesets'
  app/controllers/repositories_controller.rb:114:in `show'
  lib/redmine/sudo_mode.rb:63:in `sudo_mode'

If you exec a manual auto_fetch to refresh data with rake -f Rakefile RAIL_ENV=production redmine:fetch_changesets --trace
Console shows:

    /usr/local/bundle/gems/htmlentities-4.3.1/lib/htmlentities/mappings/expanded.rb:465: warning: duplicated key at line 466
    ignored: "inodot"
    ** Invoke redmine:fetch_changesets (first_time)
    ** Invoke environment (first_time)
    ** Execute environment
    ** Execute redmine:fetch_changesets
    rake aborted!
NoMethodError: undefined method `[]' for nil:NilClass
  /usr/src/redmine/lib/redmine/scm/adapters/mercurial_adapter.rb:194:in `each_revision'
  /usr/src/redmine/app/models/repository/mercurial.rb:187:in `block in fetch_changesets'
  /usr/src/redmine/app/models/repository/mercurial.rb:186:in `step'
  /usr/src/redmine/app/models/repository/mercurial.rb:186:in `fetch_changesets'
  /usr/src/redmine/app/models/repository.rb:344:in `block (2 levels) in fetch_changesets'
  /usr/local/bundle/gems/activerecord-4.2.5/lib/active_record/relation/delegation.rb:46:in `each'
  /usr/local/bundle/gems/activerecord-4.2.5/lib/active_record/relation/delegation.rb:46:in `each'
/usr/src/redmine/app/models/repository.rb:342:in `block in fetch_changesets'
/usr/local/bundle/gems/activerecord-4.2.5/lib/active_record/relation/delegation.rb:46:in `each'
/usr/local/bundle/gems/activerecord-4.2.5/lib/active_record/relation/delegation.rb:46:in `each'
/usr/src/redmine/app/models/repository.rb:341:in `fetch_changesets'
/usr/src/redmine/lib/tasks/redmine.rake:47:in `block (2 levels) in <top (required)>
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:240:in `call'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:240:in `block in execute'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:235:in `each'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:235:in `execute'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:179:in `block in invoke_with_call_chain'
/usr/lib/ruby/2.2.0/monitor.rb:211:in `mon_synchronize'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:172:in `invoke_with_call_chain'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/task.rb:165:in `invoke'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:150:in `block in invoke_task'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:106:in `block (2 levels) in top_level'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:106:in `each'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:106:in `block in top_level'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:115:in `run_with_threads'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:100:in `top_level'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:78:in `block in run'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:176:in `standard_exception_handling'
/usr/local/bundle/gems/rake-10.5.0/lib/rake/application.rb:75:in `run'
/usr/local/bundle/gems/rake-10.5.0/bin/rake:33:in `<top (required)>'
/usr/local/bundle/bin/rake:16:in `load'
/usr/local/bundle/bin/rake:16:in `<main>'
Tasks: TOP => redmine:fetch_changesets

Redmine version 3.1.3.stable
Ruby version 2.2.4-p230 (2015-12-16) [x86_64-linux-gnu]
Rails version 4.2.5
Environment production
Database adapter Mysql2
SCM:
  Subversion 1.8.13
  Mercurial 3.1.2
  Git 2.1.4
  Bazaar 2.7.0
  Filesystem
Redmine plugins:
  no plugin installed

Associated revisions

Revision 19606 - 2020-03-22 17:15 - Toshi MARUYAMA

scm: mercurial: change template "escape" to "urlescape" for "author" and "desc" (#23055)

Revision 19607 - 2020-03-22 17:34 - Toshi MARUYAMA

2020-04-16
What output of scm_stderr_log_file?

source:tags/3.2.3/config/configuration.yml.example#L138

I haven't any exit on scm_stderr_log_file.

But in production.log when do the first scan logged this:

Changeset Load (0.8ms)  SELECT `changesets`.* FROM `changesets` WHERE `changesets`.`repository_id` = 11 AND `changesets`.`scmid` = '6e892cac598ab919853ca54c7745e783b70dcb8' ORDER BY changesets.id DESC LIMIT 1

Changeset Exists (0.6ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10555' AND `changesets`.`id` != 37510 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.5ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '6e892cac598ab919853ca54c7745e783b70dcb8' AND `changesets`.`id` != 37510 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10598' AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '6289f36f5adf23615f6fc2f49ab5828beff29b' AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Load (0.4ms)  SELECT `changesets`.* FROM `changesets` WHERE `changesets`.`repository_id` = 11 AND `changesets`.`scmid` = '6289f36f5adf23615f6fc2f49ab5828beff29b' ORDER BY changesets.id DESC LIMIT 1

Changeset Exists (0.6ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10598' AND `changesets`.`id` != 37553 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.5ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '6289f36f5adf23615f6fc2f49ab5828beff29b' AND `changesets`.`id` != 37553 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10599' AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '2babe5c449c6274f052a8bda9637c247ea210e23' AND `changesets`.`repository_id` = 11) LIMIT 1

SQL (0.5ms)  INSERT INTO `changesets` (`repository_id`, `revision`, `scmid`, `committer`, `commit_date`, `committed_on`, `comments`, `user_id`) VALUES (11, '10598', '6289f36f5adf23615f6fc2f49ab5828beff29b', 'pepe<eduardo@pepe.es>', '2011-07-15', '2011-07-15 07:42:42', 'TextoStats ahora senyala el estilo del caracter donde esta el cursor.', 6)

SQL (0.3ms)  INSERT INTO `changesets` (`repository_id`, `revision`, `scmid`, `committer`, `commit_date`, `committed_on`, `comments`, `user_id`) VALUES (11, '10598', '6289f36f5adf23615f6fc2f49ab5828beff29b', 'pepe<eduardo@pepe.es>', '2011-07-15', '2011-07-15 07:42:42', 'TextoStats ahora senyala el estilo del caracter donde esta el cursor.', 6)

(34.2ms)  COMMIT

(0.2ms) BEGIN

Changeset Load (0.7ms)  SELECT `changesets`.* FROM `changesets` WHERE `changesets`.`repository_id` = 11 AND `changesets`.`scmid` = '6289f36f5adf23615f6fc2f49ab5828beff29b' ORDER BY changesets.id DESC LIMIT 1

Changeset Exists (0.6ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10598' AND `changesets`.`id` != 37553 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.5ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '6289f36f5adf23615f6fc2f49ab5828beff29b' AND `changesets`.`id` != 37553 AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`revision` = BINARY '10599' AND `changesets`.`repository_id` = 11) LIMIT 1

Changeset Exists (0.4ms)  SELECT 1 AS one FROM `changesets` WHERE (`changesets`.`scmid` = BINARY '2babe5c449c6274f052a8bda9637c247ea210e23' AND `changesets`.`repository_id` = 11) LIMIT 1

SQL (0.5ms)  INSERT INTO `changesets` (`repository_id`, `revision`, `scmid`, `committer`, `commit_date`, `committed_on`, `comments`, `user_id`) VALUES (11, '10599', '2babe5c449c6274f052a8bda9637c247ea210e23', 'pepe<eduardo@pepe.es>', '2011-07-15', '2011-07-15 07:42:42', 'TextoStats ahora senyala el estilo del caracter donde esta el cursor.', 6)
I haven't any exit on scm_stderr_log_file.

How about ".scm_stderr.log"?

What output of above "hg" on command line?

I haven't any exit on scm_stderr_log_file. -> I try to say: I haven't any error related with this issue. Only have this message:

not trusting file /home/share/mirror/hg/all/.hg/hgrc from untrusted user root, group www-data

It can be solve with adding in hgrc.

[trusted]
users =
hg in command line:

Mercurial Distributed SCM

basic commands:

add           add the specified files on the next commit
annotate      show changeset information by line for each file
clone         make a copy of an existing repository
commit        commit the specified files or all outstanding changes
diff          diff repository (or selected files)
export        dump the header and diffs for one or more changesets
forget        forget the specified files on the next commit
init          create a new repository in the given directory
log           show revision history of entire repository or files
merge         merge working directory with another revision
pull          pull changes from the specified source
push          push changes to the specified destination
remove        remove the specified files on the next commit
serve         start stand-alone webserver
status        show changed files in the working directory
summary       summarize working directory state
update        update working directory (or switch revisions)

use "hg help" for the full list of commands or "hg -v" for details

#5 - 2016-06-14 16:00 - Toshi MARUYAMA

What is output of following command?

'hg' '-R' '/home/share/mirror/hg/all' '--encoding' 'utf-8' '--config'
'extensions.redminehelper=/usr/src/redmine/lib/redmine/scm/adapters/mercurial/redminehelper.py' '--config' 'diff.git=false' 'log' '--debug' '-C' '--style'
'/usr/src/redmine/lib/redmine/scm/adapters/mercurial/hg-template-1.0.tmpl' '-r' '10600:10699'

#6 - 2016-06-14 16:18 - mole mole

- File hgcommand.log added

The command 'hg' '-R' '/home/share/mirror/hg/all' '--encoding' 'utf-8' '--config'
'extensions.redminehelper=/usr/src/redmine/lib/redmine/scm/adapters/mercurial/redminehelper.py' '--config' 'diff.git=false' 'log' '--debug' '-C' '--style'
'/usr/src/redmine/lib/redmine/scm/adapters/mercurial/hg-template-1.0.tmpl' '-r' '10600:10699'

Can be executed without errors. Attach output.

#7 - 2016-06-14 16:48 - Toshi MARUYAMA
What is output?

```sql
select revision from changesets where id =
(select max(id) from changesets where repository_id = 11)
```

#8 - 2016-06-14 17:09 - mole mole

revision
10599

#9 - 2016-06-14 17:22 - Toshi MARUYAMA

Try reduce FETCH_AT_ONCE.

source:tags/3.1.3/app/models/repository/mercurial.rb#L30

I think a revision from 10600 to 10699 has problem.

#10 - 2016-06-14 17:31 - Toshi MARUYAMA

I think your revision has problem.
- no date
- no commit log

If you find this revision, try this patch.

```diff
diff --git a/lib/redmine/scm/adapters/mercurial_adapter.rb b/lib/redmine/scm/adapters/mercurial_adapter.rb
--- a/lib/redmine/scm/adapters/mercurial_adapter.rb
+++ b/lib/redmine/scm/adapters/mercurial_adapter.rb
@@ -212,8 +212,8 @@
模块Redmine

    yield Revision.new({
        :revision => le['revision'],
-        :scmid => le['node'],
-        :time => Time.parse(le['date']['__content__']),
-        :message => le['msg']['__content__'],
+        :time => (Time.parse(le['date']['__content__']) rescue Time.at(0)),
+        :message => (le['msg']['__content__'] rescue ''),
        :paths => paths,
        :parents => parents_ary)
    }
```

#11 - 2016-06-14 18:02 - mole mole

- File issue.jpg added

It didn't works. Same result.

Maybe a wrong code of EOL?. I see in sublime editor a DC3 code.
I thing that double !! can break it.

The message in mercurial is:

```
-dimensiones de los geosets de m3Font funcionando bien con shader.!!
```

and the output of script have a DC3 code when in original have a !!:

```
</msg>
```

Try upgrading Mercurial. Latest version is 3.8.3.

It didn't works. Same result.

Please report to Mercurial BTS. Broken XML is Mercurial bug.

https://bz.mercurial-scm.org/

The character is ^S I try to reproduce in small repo

You can reproduce this issue with this repo.

It is obviously Mercurial bug.

```
$ LANG=C /WEB-DOWN/hg-repo/hg-crew/hg log --version
Mercurial Distributed SCM (version 3.8.3)
(see https://mercurial-scm.org for more information)
```
Copyright (C) 2005-2016 Matt Mackall and others
This is free software; see the source for copying conditions. There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.

$ /WEB-DOWN/hg-repo/hg-crew/hg log --debug --style
/REDMINE/MY-REDMINE/test-trunk/lib/redmine/scm/adapters/mercurial/hg-template-1.0.tmpl > hg.xml

$ xmllint hg.xml
hg.xml:10: parser error : PCDATA invalid Char value 19
<msg>-Test</msg>
^

hg.xml:17: parser error : Premature end of data in tag log line 2

#19 - 2016-06-15 13:16 - Toshi MARUYAMA

Sorry, we need add "</log>".

$ /WEB-DOWN/hg-repo/hg-crew/hg log --debug --style
/REDMINE/MY-REDMINE/test-trunk/lib/redmine/scm/adapters/mercurial/hg-template-1.0.tmpl > hg.xml

$ echo "</log>" >> hg.xml

$ xmllint hg.xml
hg.xml:23: parser error : PCDATA invalid Char value 19
<msg>-Test</msg>
^

#20 - 2016-06-15 13:28 - mole mole

Using Mercurial xml template generate a correct xml

$ hg log --debug --style=xml > hg.xml

$ xmllint hg.xml

It runs correctly

PS: --style is a deprecated parameter in mercurial

#21 - 2016-06-15 13:44 - mole mole

I changed in hg-template-1.0.tmpl {desc|escape} by {desc|revscape} and xmllint runs correctly...
Sorry, it's probably we must use `xml_escape`.

I see this in this file sample of mercurial

```
/usr/share/mercurial/templates/map-cmdline.xml
```

- File `hg-template-1.0.tmpl` added

Attach `hg-template-1.0.tmpl` modified

- File `hg-template-1.0.tmpl` added

Sorry, is this file

Subject changed from Error with Fetch commits automatically with mercurial repository to Error with Fetch commits with Mercurial repository when log has invalid char

- Status changed from Needs feedback to Confirmed
- Assignee set to Toshi MARUYAMA

Thanks for your help.

I can also confirm this. Any workaround?

Environment:
- Redmine version: 3.2.0.stable
- Ruby version: 2.3.1-p112 (2016-04-26) [x86_64-linux-gnu]
- Rails version: 4.2.5
- Environment: production
- Database adapter: PostgreSQL

SCM:
- Mercurial: 3.7.3

Status changed from Confirmed to Closed
I choose `urlescape` because `paths` uses `urlescape` and decodes it.

source:trunk/lib/redmine/scm/adapters/mercurial_adapter.rb@19606#L204