Redmine - Patch #4409

Remove email body via a delimiter

2009-12-16 00:20 - Eric Davis

Status:	Closed	**Start date:**	2009-12-16
Priority:	Normal	**Due date:**	
Assignee:	Eric Davis	**% Done:**	0%
Category:	Email receiving	**Estimated time:**	0.00 hour
Target version:	0.9.0		

Description

My clients and I use the email receiving features of Redmine all the time but it's a pain if someone forgets to remove the replies or has a long signature line. This patch adds a setting to allow Administrators to configure delimiters for the incoming email body. When the email is read, the body content that is after these delimiters are removed.

For example, with the delimiter of "cut" the following email would only have "Here is the email body for issue 4." as the issue/journal body.

```
Here is the email body for issue 4.

CUT

Sent from my iPhone
```

Multiple delimiters can be added easily by separating them with a line in the settings.

I can commit this patch for 0.9 if there are no problems. I just wanted to get some feedback first.

Related issues:

Has duplicate Redmine - Feature #3199: Ability to remove mail signature when ...	**Closed**	**2009-04-17**

Associated revisions

Revision 3226 - 2009-12-23 19:35 - Jean-Philippe Lang

Adds a setting to remove incoming emails body after a delimiter (#4409).

History

#1 - 2009-12-16 00:43 - Shane Pearlman

Its amazing how this little feature makes such a difference when trying to read an issue history. The ability to manage the filter via the admin is nice as many system parse emails different (like mac mail killing ---). +1.

#2 - 2009-12-16 01:17 - ciaran jessup

A definite +1 from me also, my devs had requested this exact same feature ;)

#3 - 2009-12-16 19:25 - Mr Embedded

This looks interesting for my purposes as well.   I would like to test it but how should this patch be applied?

#4 - 2009-12-18 21:05 - Jean-Philippe Lang

A few comments:

- this should be applied to forum replies as well
- the setting name is a bit vague, I think that mail_handler_body_delimiters would be better

```
regex = Regexp.new(token + '.*', Regexp::IGNORECASE | Regexp::MULTILINE)
```

- entering an invalid regexp in this setting will raise an error, this should be rescued
- this will remove content after a line that **starts** with the delimiter. IMHO, it's not a desirable behaviour. Since we accept a regexp, why not let the user append .* if he really wants this?
- we should add an additional setting to turn on/off regexp matching (something like mail_handler_body_delimiters_regexp set to off by default).

#5 - 2009-12-20 18:46 - Eric Davis

Jean-Philippe Lang wrote:

A few comments:

- this should be applied to forum replies as well

Agreed.

- the setting name is a bit vague, I think that mail_handler_body_delimiters would be better

Thanks, I had trouble coming up with a good description for it so mail_handler_body_delimiters would be good.

- entering an invalid regexp in this setting will raise an error, this should be rescued

Good catch, I can write a test for that case.

- this will remove content after a line that **starts** with the delimiter. IMHO, it's not a desirable behaviour. Since we accept a regexp, why not let the user append .* if he really wants this?

I think this is the standard behavior. Most other systems with a similar feature have a section saying "all content after HERE will be removed". I can remove the .* and add a note on the user interface to let the admin know they will need to add .* if they want to remove all content after the regex.

- we should add an additional setting to turn on/off regexp matching (something like mail_handler_body_delimiters_regexp set to off by default).

So would this setting enable/disable the body removal? So if it's disabled, the body is added as is. If it's enabled, the body is stripped according to the regular expressions in mail_handler_body_delimiters.

#6 - 2009-12-20 21:43 - Felix Schäfer

Eric Davis wrote:

> Jean-Philippe Lang wrote:
>
> > - this will remove content after a line that **starts** with the delimiter. IMHO, it's not a desirable behaviour. Since we accept a regexp, why not let the user append .* if he really wants this?
>
> I think this is the standard behavior. Most other systems with a similar feature have a section saying "all content after HERE will be removed". I can remove the .* and add a note on the user interface to let the admin know they will need to add .* if they want to remove all content after the regex.

According to the signature block article on Wikipedia, a signature begins with "-- \n", so I think the appended .* should not be the default.

#7 - 2009-12-20 23:34 - Jean-Philippe Lang

> I think this is the standard behavior. Most other systems with a similar feature have a section saying "all content after HERE will be removed". I can remove the .* and add a note on the user interface to let the admin know they will need to add .* if they want to remove all content after the regex.

I didn't try your patch but it looks like if you set the delimiter to '--' for example, it will strip any text after '--' anywhere if the email body.

The following email:

```
This is my body ------ and this should not be stripped.

--
JP
```

would become:

```
This is my body
```

And I think the regexp is definitly not needed. The setting should be regexp escaped.

#8 - 2009-12-21 00:36 - Eric Davis

Felix Schäfer wrote:

> According to the [signature block](#) article on Wikipedia, a signature begins with "@--
>
>
> ", so I think the appended .*@ should not be the default.

Without .* it will convert the following email like so, which isn't useful at all:

```
This is my body content

--
JPL
Maintainer of Redmine.org
Ruby on Rails developer
user@example.com
```

to:

```
This is my body content

JPL
Maintainer of Redmine.org
Ruby on Rails developer
user@example.com
```

Jean-Philippe Lang wrote:

> I didn't try your patch but it looks like if you set the delimiter to '--' for example, it will strip any text after '--' anywhere if the email body.

Correct, the administrator would have to be careful with what delimiters are used. I have two sites in production using --- already and has only stripped to much content once out of a few thousand emails.

> And I think the regexp is definitly not needed. The setting should be regexp escaped.

I'm not following, how would you match the content then?

#9 - 2009-12-21 10:16 - Felix Schäfer

Eric Davis wrote:

> Felix Schäfer wrote:
>
> > According to the [signature block](#) article on Wikipedia, a signature begins with "@--
> >
> >
> > ", so I think the appended .*@ should not be the default.
>
> Without .* it will convert the following email like so, which isn't useful at all:

Yeah, sorry, I was more thinking in terms of lines… What I meant was that it should only cut content starting with a line containing *only* given string or regex.

#10 - 2009-12-21 15:15 - Eric Voisard

How about the token having to be on a line by istelf. I don't know in Ruby but in Perl it becomes:

```
$token="--";
$_ = "This is my body\n--\nthis si crap";

s/^$token\n.*//m;

print;
```

I'd not make the regexp ignoring case. I.e. "END\n" must be matched, but not "End\n",
.* will remove the rest anyway.

#11 - 2009-12-21 15:54 - Eric Voisard

Little change because '.' doesn't match newlines by default in Perl, and bodies like the following one would not be processed correctly

```
$token="--";
```

```
$_ = "This is my body\n--\nthis si crap\nthis line too";

s/^$token.*//ms;

#this would work too:
#s/^$token(.|\s)*//m;

print;
```

Sorry for talking Perl in a Ruby project, but there should be an equivalence...
Eric

#12 - 2009-12-21 16:31 - Eric Voisard

Sorry to pollute this issue :-/ I omitted to put '\n' back into my last regexps.
It should indeed be s/^$token\n.*//ms; or s/^$token\n(.|\s)*//m;
well you get the idea...

#13 - 2009-12-23 19:31 - Jean-Philippe Lang

- Status changed from 7 to Closed

- Target version set to 0.9.0

Feature added in [r3226](#).
Delimiters are strings (not regexp) and must match the entire line.

#14 - 2009-12-28 15:29 - Felix Schäfer

Works a charm, thanks a lot :-)

#15 - 2010-09-02 13:33 - Dinis Quelhas

Guys, I'm sure I'm being thick here, but could you please give me an example on how to use this great feature?
I've tried updating an issue from Gmail. The reply is something like:

On Thu, Sep 2, 2010 at 12:22 PM, <redmine@romulus.thepieguide.com> wrote:

```
Issue #3195 has been updated by Dinis Quelhas.

trying to truncate with gmail
     Feature #3195: T1 Test Issue
```

.....

If I go to Settings->Incoming Emails->Truncate emails after one of these lines and add "On Thu, Sep 2, 2010 at 12:22 PM, < redmine@romulus.thepieguide.com> wrote:" it works just fine but if I add something like "Issue [#3195](#) has been updated by Dinis Quelhas" it doesn't work.

I would like to have something a bit more generic for all of them like:

"Feature #*:" to apply to all incoming emails. Is this possible?

Thanks in advance for your help :-)

#16 - 2010-09-24 16:29 - sebastian ovide

most of mail clients would append somethink like "john wrote" or "on xx/xx/xx ..." etc... is there any way to add some kind of reg expression ? in case of that regular expression is not valid, just ignore the expression !

in my case I would cut the body as soon as I find "redmine@myserver.com"... but it is never a whole line !

#17 - 2010-10-14 14:19 - Raghu Sunderam

Regexp based delimiters should be added.
Please revert the change soon.

Files

0001-Remove-the-body-of-incoming-emails-after-specific-de.patch	10.7 KB	2009-12-15	Eric Davis