

RPM Support - Issue #3535

Error syncing Oracle EPEL repository

03/28/2018 08:03 PM - steveh1138

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:	bmbouter	Estimated time:	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	2. Medium	Groomed:	No
Version:		Sprint Candidate:	No
Platform Release:	2.16.2	Tags:	Pulp 2
OS:	CentOS 7	Sprint:	Sprint 38
Triaged:	Yes	Quarter:	

Description

When attempting to sync the official public Oracle EPEL repository I receive the following error in Foreman/Katello. As this is a pulp error I was directed over to your project. Foreman / Katello 1.16.0 / 3.5.1

PLP0000: Invalid block tag: '}'gmail{*}com>'

Here is the error output I receive

```
{ "contents_changed"=>true,
  "pulp_tasks"=>
  [ {"exception"=>nil,
    "task_type"=>"pulp.server.managers.repo.sync.sync",
    "_href"=>"/pulp/api/v2/tasks/57edd0fc-dcf8-453f-a5ac-2ede318d2311/",
    "task_id"=>"57edd0fc-dcf8-453f-a5ac-2ede318d2311",
    "tags"=>
    [ "pulp:repository:4d73ea9e-7776-486e-811b-d914dc823105",
      "pulp:action:sync" ],
    "finish_time"=>"2018-03-28T04:15:52Z",
    "_ns"=>"task_status",
    "start_time"=>"2018-03-28T04:02:38Z",
    "traceback"=>nil,
    "spawned_tasks"=>
    [ { "_href"=>"/pulp/api/v2/tasks/18a119c8-9f71-4d05-972b-d9932e635ec8/",
      "task_id"=>"18a119c8-9f71-4d05-972b-d9932e635ec8" } ],
    "progress_report"=>
    { "yum_importer"=>
      { "content"=>
        { "items_total"=>10,
          "state"=>"FINISHED",
          "error_details"=>[],
          "details"=>
          { "rpm_total"=>10, "rpm_done"=>10, "drpm_total"=>0, "drpm_done"=>0,
            "size_total"=>15409527,
            "size_left"=>0,
            "items_left"=>0 },
          "comps"=>{ "state"=>"FINISHED" },
          "purge_duplicates"=>{ "state"=>"FINISHED" },
          "distribution"=>
          { "items_total"=>0,
            "state"=>"FINISHED",
            "error_details"=>[],
            "items_left"=>0 },
          "errata"=>{ "state"=>"FINISHED" },
          "metadata"=>{ "state"=>"FINISHED" } } },
      "queue"=>"reserved_resource_worker-2@ussl-foreman.usanainc.com.dq",
```

```

"state"=>"finished",
"worker_name"=>"reserved_resource_worker-2@ussl-foreman.usanainc.com",
"result"=>
  {"result"=>"success",
   "importer_id"=>"yum_importer",
   "exception"=>nil,
   "repo_id"=>"4d73ea9e-7776-486e-811b-d914dc823105",
   "traceback"=>nil,
   "started"=>"2018-03-28T04:02:40Z",
   "_ns"=>"repo_sync_results",
   "completed"=>"2018-03-28T04:15:50Z",
   "importer_type_id"=>"yum_importer",
   "error_message"=>nil,
   "summary"=>
     {"content"=>{"state"=>"FINISHED"},
      "comps"=>{"state"=>"FINISHED"},
      "purge_duplicates"=>{"state"=>"FINISHED"},
      "distribution"=>{"state"=>"FINISHED"},
      "errata"=>{"state"=>"FINISHED"},
      "metadata"=>{"state"=>"FINISHED"}},
    "added_count"=>10,
    "removed_count"=>0,
    "updated_count"=>2,
    "id"=>"5abb16f7a55b78667ce16f7a",
    "details"=>
      {"content"=>
        {"size_total"=>15409527,
         "items_left"=>0,
         "items_total"=>10,
         "state"=>"FINISHED",
         "size_left"=>0,
         "details"=>
           {"rpm_total"=>10, "rpm_done"=>10, "drpm_total"=>0, "drpm_done"=>0},
          "error_details"=>[]},
         "comps"=>{"state"=>"FINISHED"},
         "purge_duplicates"=>{"state"=>"FINISHED"},
         "distribution"=>
           {"items_total"=>0,
            "state"=>"FINISHED",
            "error_details"=>[],
            "items_left"=>0},
          "errata"=>{"state"=>"FINISHED"},
          "metadata"=>{"state"=>"FINISHED"}},
        "error"=>nil,
        "_id"=>{"$oid"=>"5abb134bde37317496f9af38"},
        "id"=>"5abb134bde37317496f9af38"},
      {"exception"=>nil,
       "task_type"=>"pulp.server.managers.repo.publish.publish",
       "_href"=>"/pulp/api/v2/tasks/18a119c8-9f71-4d05-972b-d9932e635ec8/",
       "task_id"=>"18a119c8-9f71-4d05-972b-d9932e635ec8",
       "tags"=>
         ["pulp:repository:4d73ea9e-7776-486e-811b-d914dc823105",
          "pulp:action:publish"],
       "finish_time"=>"2018-03-28T04:20:36Z",
       "_ns"=>"task_status",
       "start_time"=>"2018-03-28T04:20:02Z",
       "traceback"=>
         "Traceback (most recent call last):\n File \"/usr/lib/python2.7/site-packages/celery/app/trace.py", line 240, in trace_task\n   R = retval = fun(*args, **kwargs)\n File \"/usr/lib/python2.7/site-packages/pulp/server/async/tasks.py", line 527, in __call__\n   return super(Task, self).__call__(*args, **kwargs)\n File \"/usr/lib/python2.7/site-packages/pulp/server/async/tasks.py", line 107, in __call__\n   return super(PulpTask, self).__call__(*args, **kwargs)\n File \"/usr/lib/python2.7/site-packages/celery/app/trace.py", line 438, in __protected_call__\n   return self.run(*args, **kwargs)\n File \"/usr/lib/python2.7/site-packages/pulp/server/controllers/repository.py", line 1106, in publish\n   result = check_publish(repo_obj, dist_id, dist_inst, transfer_repo, conduit, call_config)\n File \"/usr/lib/python2.7/site-packages/pulp/server/controllers/repository.py", line 1200, in check_publish\n   result = _do_publish(repo_obj, dist_id, dist_inst,

```

```

transfer_repo, conduit, call_config)\n File \"/usr/lib/python2.7/site-packages/pulp/server/contr
ollers/repository.py\", line 1252, in _do_publish\n     publish_report = publish_repo(transfer_repo
, conduit, call_config)\n File \"/usr/lib/python2.7/site-packages/pulp/server/async/tasks.py\", l
ine 730, in wrap_f\n     return f(*args, **kwargs)\n File \"/usr/lib/python2.7/site-packages/pulp_
rpm/plugins/distributors/yum/distributor.py\", line 174, in publish_repo\n     return self._publish
er.process_lifecycle()\n File \"/usr/lib/python2.7/site-packages/pulp/plugins/util/publish_step.p
y\", line 572, in process_lifecycle\n     super(PluginStep, self).process_lifecycle()\n File \"/us
r/lib/python2.7/site-packages/pulp/plugins/util/publish_step.py\", line 163, in process_lifecycle\
n     step.process()\n File \"/usr/lib/python2.7/site-packages/pulp/plugins/util/publish_step.py\"
, line 239, in process\n     self._process_block(item=item)\n File \"/usr/lib/python2.7/site-packa
ges/pulp/plugins/util/publish_step.py\", line 301, in _process_block\n     self.process_main(item=i
tem)\n File \"/usr/lib/python2.7/site-packages/pulp_rpm/plugins/distributors/yum/publish.py\", li
ne 485, in process_main\n     context.add_unit_metadata(unit)\n File \"/usr/lib/python2.7/site-pac
kages/pulp_rpm/plugins/distributors/yum/metadata/other.py\", line 42, in add_unit_metadata\n     se
lf.metadata_file_handle.write(unit.render_other(self.checksum_type))\n File \"/usr/lib/python2.7/
site-packages/pulp_rpm/plugins/db/models.py\", line 850, in render_other\n     return self._render(
metadata, context)\n File \"/usr/lib/python2.7/site-packages/pulp_rpm/plugins/db/models.py\", lin
e 879, in _render\n     t = Template(template)\n File \"/usr/lib/python2.7/site-packages/django/te
mplate/base.py\", line 125, in __init__\n     self.nodelist = compile_string(template_string, origi
n)\n File \"/usr/lib/python2.7/site-packages/django/template/base.py\", line 153, in compile_stri
ng\n     return parser.parse()\n File \"/usr/lib/python2.7/site-packages/django/template/base.py\"
, line 276, in parse\n     self.invalid_block_tag(token, command, parse_until)\n File \"/usr/lib/p
ython2.7/site-packages/django/template/base.py\", line 332, in invalid_block_tag\n     raise self.e
rror(token, \"Invalid block tag: '%s'\" % command)\nTemplateSyntaxError: Invalid block tag: '}'mai
l{*}com>>'\n\",
    \"spawned_tasks\"=>[],
    \"progress_report\"=>
    {\"4d73ea9e-7776-486e-811b-d914dc823105\"=>
    [{\"num_success\"=>1,
      \"description\"=>\"Copying files\",
      \"step_type\"=>\"save_tar\",
      \"items_total\"=>1,
      \"state\"=>\"FINISHED\",
      \"error_details\"=>[],
      \"details\"=>\"\",
      \"num_failures\"=>0,
      \"step_id\"=>\"9c354dd3-4837-4ea8-a3e1-7a03861942ff\",
      \"num_processed\"=>1},
    {\"num_success\"=>1,
      \"description\"=>\"Initializing repo metadata\",
      \"step_type\"=>\"initialize_repo_metadata\",
      \"items_total\"=>1,
      \"state\"=>\"FINISHED\",
      \"error_details\"=>[],
      \"details\"=>\"\",
      \"num_failures\"=>0,
      \"step_id\"=>\"23f96ff2-525e-451b-bfda-b558d113f17f\",
      \"num_processed\"=>1},
    {\"num_success\"=>0,
      \"description\"=>\"Publishing Distribution files\",
      \"step_type\"=>\"distribution\",
      \"items_total\"=>0,
      \"state\"=>\"FINISHED\",
      \"error_details\"=>[],
      \"details\"=>\"\",
      \"num_failures\"=>0,
      \"step_id\"=>\"848c3bd5-0fed-4a5e-bc76-05b886952c19\",
      \"num_processed\"=>0},
    {\"num_success\"=>299,
      \"description\"=>\"Publishing RPMs\",
      \"step_type\"=>\"rpms\",
      \"items_total\"=>17849,
      \"state\"=>\"FAILED\",
      \"error_details\"=>
      [\"traceback\"=>
        \" File \"/usr/lib/python2.7/site-packages/pulp/plugins/util/publish_step.py\", line
239, in process\n     self._process_block(item=item)\n\n File \"/usr/lib/python2.7/site-packages/p

```

```

ulp/plugins/util/publish_step.py", line 301, in _process_block\n    self.process_main(item=item)\n\n    File \"/usr/lib/python2.7/site-packages/pulp_rpm/plugins/distributors/yum/publish.py", line\n485, in process_main\n    context.add_unit_metadata(unit)\n\n    File \"/usr/lib/python2.7/site-pack\nages/pulp_rpm/plugins/distributors/yum/metadata/other.py", line 42, in add_unit_metadata\n    sel\nf.metadata_file_handle.write(unit.render_other(self.checksum_type))\n\n    File \"/usr/lib/python2.7\n/site-packages/pulp_rpm/plugins/db/models.py", line 850, in render_other\n    return self._render\n(metadata, context)\n\n    File \"/usr/lib/python2.7/site-packages/pulp_rpm/plugins/db/models.py",\nline 879, in _render\n    t = Template(template)\n\n    File \"/usr/lib/python2.7/site-packages/djan\ngo/template/base.py", line 125, in __init__\n    self.nodelist = compile_string(template_string,\norigin)\n\n    File \"/usr/lib/python2.7/site-packages/django/template/base.py", line 153, in compi\nle_string\n    return parser.parse()\n\n    File \"/usr/lib/python2.7/site-packages/django/template/\nbase.py", line 276, in parse\n    self.invalid_block_tag(token, command, parse_until)\n\n    File \n"/usr/lib/python2.7/site-packages/django/template/base.py", line 332, in invalid_block_tag\n    r\naise self.error(token, "Invalid block tag: '%s'" % command)\n",\n    "error"=>"Invalid block tag: 'gmail{*}com&gt;'",\n    "details"=>"",\n    "num_failures"=>1,\n    "step_id"=>"c772a9bf-e9f4-43b7-bde8-302909f9dd9b",\n    "num_processed"=>300},\n{"num_success"=>0,\n "description"=>"Publishing Delta RPMs",\n "step_type"=>"drpms",\n "items_total"=>1,\n "state"=>"NOT_STARTED",\n "error_details"=>[],\n "details"=>"",\n "num_failures"=>0,\n "step_id"=>"d1431e03-4eb4-474f-aec2-bd850013fd91",\n "num_processed"=>0},\n{"num_success"=>0,\n "description"=>"Publishing Errata",\n "step_type"=>"errata",\n "items_total"=>1,\n "state"=>"NOT_STARTED",\n "error_details"=>[],\n "details"=>"",\n "num_failures"=>0,\n "step_id"=>"d148e811-003c-4402-bbd7-db155bd0d486",\n "num_processed"=>0},\n{"num_success"=>0,\n "description"=>"Publishing Comps file",\n "step_type"=>"comps",\n "items_total"=>1,\n "state"=>"NOT_STARTED",\n "error_details"=>[],\n "details"=>"",\n "num_failures"=>0,\n "step_id"=>"6231f6da-1020-4fd4-92fd-131e57dce3e7",\n "num_processed"=>0},\n{"num_success"=>0,\n "description"=>"Publishing Metadata.",\n "step_type"=>"metadata",\n "items_total"=>1,\n "state"=>"NOT_STARTED",\n "error_details"=>[],\n "details"=>"",\n "num_failures"=>0,\n "step_id"=>"1eb2afec-4d4e-4273-bb61-d7280cb86e2a",\n "num_processed"=>0},\n{"num_success"=>0,\n "description"=>"Closing repo metadata",\n "step_type"=>"close_repo_metadata",\n "items_total"=>1,\n "state"=>"NOT_STARTED",\n "error_details"=>[],\n "details"=>"",\n "num_failures"=>0,

```

```

    "step_id"=>"ec3585fb-9eb5-4731-b14a-ab7ab3944713",
    "num_processed"=>0},
  {"num_success"=>0,
   "description"=>"Generating sqlite files",
   "step_type"=>"generate sqlite",
   "items_total"=>1,
   "state"=>"NOT_STARTED",
   "error_details"=>[],
   "details"=>"",
   "num_failures"=>0,
   "step_id"=>"f2f4bedc-2662-4107-b08e-91593d7ad2cb",
   "num_processed"=>0},
  {"num_success"=>0,
   "description"=>"Generating HTML files",
   "step_type"=>"repoview",
   "items_total"=>1,
   "state"=>"NOT_STARTED",
   "error_details"=>[],
   "details"=>"",
   "num_failures"=>0,
   "step_id"=>"2b9388b9-622a-4e91-a85b-6f9cc743e389",
   "num_processed"=>0},
  {"num_success"=>0,
   "description"=>"Publishing files to web",
   "step_type"=>"publish_directory",
   "items_total"=>1,
   "state"=>"NOT_STARTED",
   "error_details"=>[],
   "details"=>"",
   "num_failures"=>0,
   "step_id"=>"777f72ca-23b9-4a38-9257-1025123a0e2d",
   "num_processed"=>0},
  {"num_success"=>0,
   "description"=>"Writing Listings File",
   "step_type"=>"initialize_repo_metadata",
   "items_total"=>1,
   "state"=>"NOT_STARTED",
   "error_details"=>[],
   "details"=>"",
   "num_failures"=>0,
   "step_id"=>"1b2738a1-31d0-47b3-b658-f866245b97a8",
   "num_processed"=>0},
  {"num_success"=>0,
   "description"=>"Writing Listings File",
   "step_type"=>"initialize_repo_metadata",
   "items_total"=>1,
   "state"=>"NOT_STARTED",
   "error_details"=>[],
   "details"=>"",
   "num_failures"=>0,
   "step_id"=>"3488c45f-94cb-43ab-976c-bbcf65275c5f",
   "num_processed"=>0}}},
  "queue"=>"reserved_resource_worker-0@ussl-foreman.usanainc.com.dq",
  "state"=>"error",
  "worker_name"=>"reserved_resource_worker-0@ussl-foreman.usanainc.com",
  "result"=>nil,
  "error"=>
    {"code"=>"PLP0000",
     "data"=>{},
     "description"=>"Invalid block tag: 'gmail{*}com&gt;'",
     "sub_errors"=>[]},
  "_id"=>{"$oid"=>"5abb16f8de37317496f9af98"},
  "id"=>"5abb16f8de37317496f9af98"},
  "poll_attempts"=>{"total"=>37, "failed"=>1}}

```

Here are the versions of every pulp package currently installed

```
python-kombu-3.0.33-8.pulp.el7.noarch
rubygem-smart_proxy_pulp-1.3.0-1.el7.noarch
python-pulp-streamer-2.13.4-1.el7.noarch
pulp-puppet-tools-2.13.4-1.el7.noarch
python-pulp-ostree-common-1.2.1-3.el7.noarch
python-pulp-puppet-common-2.13.4-1.el7.noarch
python-pulp-docker-common-2.4.1-2.el7.noarch
python-isodate-0.5.0-4.pulp.el7.noarch
python-pulp-client-lib-2.13.4-1.el7.noarch
python-pulp-common-2.13.4-1.el7.noarch
python-pulp-rpm-common-2.13.4-1.el7.noarch
pulp-selinux-2.13.4-1.el7.noarch
pulp-puppet-plugins-2.13.4-1.el7.noarch
pulp-rpm-admin-extensions-2.13.4-1.el7.noarch
python-pulp-bindings-2.13.4-1.el7.noarch
pulp-katello-1.0.2-1.el7.noarch
pulp-ostree-plugins-1.2.1-3.el7.noarch
pulp-ostree-admin-extensions-1.2.1-3.el7.noarch
python-pulp-repoauth-2.13.4-1.el7.noarch
pulp-docker-plugins-2.4.1-2.el7.noarch
pulp-client-1.0-1.noarch
pulp-rpm-plugins-2.13.4-1.el7.noarch
pulp-server-2.13.4-1.el7.noarch
python-pulp-oid_validation-2.13.4-1.el7.noarch
pulp-admin-client-2.13.4-1.el7.noarch
```

Associated revisions

Revision 4a9965fc - 06/12/2018 08:26 PM - bmbouter

Sanitize Invalid Changelogs Too

The diff below causes the selection of input to be sanitized to be greedy instead of non-greedy. This causes it work sanitize more data in the cases of invalid entries such as Changelogs with too many sections in one rpm.

This fix is very safe for several reasons:

1. Escaping all data for the template layer should be safe. This is essentially user-provided data given that we got it from the internet, so it probably all should be escaped. Since this fixes this issue I'm not going to rewrite the sanitizer now to keep it low-risk.
2. The extra-greediness added is still limited by the end tag matched

<https://pulp.plan.io/issues/3535> closes #3535

Revision 62edf22a - 06/18/2018 08:09 PM - bmbouter

Sanitize Invalid Changelogs Too

The diff below causes the selection of input to be sanitized to be greedy instead of non-greedy. This causes it work sanitize more data in the cases of invalid entries such as Changelogs with too many sections in one rpm.

This fix is very safe for several reasons:

1. Escaping all data for the template layer should be safe. This is essentially user-provided data given that we got it from the internet, so it probably all should be escaped. Since this fixes this issue I'm not going to rewrite the sanitizer now to keep it low-risk.
2. The extra-greediness added is still limited by the end tag matched

<https://pulp.plan.io/issues/3535> closes #3535

(cherry picked from commit 4a9965fc100f1414d3be0845138fd66981b0ba31)

History

#1 - 03/28/2018 08:09 PM - daviddavis

- Project changed from File Support to RPM Support

#2 - 03/29/2018 11:22 AM - ttereshc

- Description updated

Could you provide a feed URL to help with reproducing your issue?

If the remote repository is not publicly available, it will be helpful if you can find the metadata of a package which causes the error. I would look for a substring }gmail{*}com in the primary.xml file of the remote repository.

Thanks

#3 - 03/29/2018 11:32 AM - ttereshc

FWIW, my guess would be that a bug is somewhere around a logic for a substitution of special characters.

https://github.com/pulp/pulp_rpm/blob/f0ffdac0390ecf78ac2c02a2f5a2a840dda7fdf1/plugins/pulp_rpm/plugins/db/models.py#L929

#4 - 03/29/2018 05:45 PM - steveh1138

This is the URL for the repository I'm attempting to sync

http://yum.oracle.com/repo/OracleLinux/OL7/developer_EPEL/x86_64

#5 - 04/03/2018 04:39 PM - dalley

- Triaged changed from No to Yes

- Sprint set to Sprint 35

#6 - 04/09/2018 03:21 AM - mansari

- Status changed from NEW to ASSIGNED

- Assignee set to mansari

#7 - 04/23/2018 04:26 PM - rchan

- Sprint changed from Sprint 35 to Sprint 36

#8 - 05/11/2018 11:02 PM - rchan

- Sprint changed from Sprint 36 to Sprint 37

#9 - 05/14/2018 06:10 PM - rchan

- Status changed from ASSIGNED to NEW

- Assignee deleted (mansari)

#10 - 05/23/2018 02:38 PM - terjetrane

I'm seeing the same thing using Red Hat Satellite 6.3 which has pulp-server-2.13.4.9-1.el7sat.noarch

The error occurs twice when trying to sync EPEL6 (from a mirror at http://fedora.uib.no/epel/6/x86_64/)

PLP0000: Invalid block tag: '}gmail{*}com>'

Grepping through the xml files in repodata I find these to lines in

9bae1663c2fef450ad66c4113bc1d70c1acc0852bfc6bbe07d7c711b883c7da-other.xml:

```
<changelog author="Neal Gompa <ngompa13{%}gmail{*}com> - 0.11.2-1" date="1437048000">- Update to 0.11.2</changelog>
```

```
<changelog author="Neal Gompa <ngompa13{%}gmail{*}com> - 0.3.3-1" date="1274702400">- Update to OggConvert 0.3.3</changelog>
```

Seems he is replacing @ with {%} to avoid spammers harvesting his email address, and the {%} sequence is then interpreted as "openblock" for a Django template or something, that expects a tag (like 'load', 'comment', 'for', 'if', etc.) and gets confused when it sees '}gmail{*}com>' instead.

I believe this worked for our previous version of Satellite, 6.2 that used pulp 2.8.7

#11 - 05/25/2018 03:31 PM - dkliban@redhat.com

If the above is true, then the problem is most likely in this method[0].

The problem is described here[1]. Need to escape the django syntax character somehow.

[0] https://github.com/pulp/pulp_rpm/blob/master/plugins/pulp_rpm/plugins/db/models.py#L887

[1] <https://docs.djangoproject.com/en/1.10/ref/templates/api/#imitations-with-string-literals>

#12 - 06/01/2018 11:04 PM - bmbouter

- Status changed from NEW to ASSIGNED

- Assignee set to bmbouter

#13 - 06/01/2018 11:12 PM - rchan

- Sprint changed from Sprint 37 to Sprint 38

#14 - 06/08/2018 10:51 PM - bmbouter

I was not able to reproduce the emitted exception. I verified the exception was not present in the logs and also in the raw json returned from the server. I tested the following:

Pulp 2.16.1 source checkout on F28 w/ Django 1.11.13
Satellite 6.3, on EL7 w/ Django 1.6.11

I tried syncing from http://archive.linux.duke.edu/pub/epel/6Server/x86_64/ which should be effectively the same as http://fedora.uib.no/epel/6/x86_64/

I verified that the `<changelog author="Neal Gomba <ngompa13{%}gmail{*}com> - 0.11.2-1" date="1437048000">- Update to 0.11.2</changelog>` was present in the other.xml published data so we have confirmation that in the test cases Pulp's template processing handled this entry.

I'm going to try http://fedora.uib.no/epel/6/x86_64/ just to be sure it also doesn't produce the issue.

#15 - 06/08/2018 11:05 PM - bmbouter

I was also not able to reproduce when syncing from http://fedora.uib.no/epel/6/x86_64/ either.

Specifically to test w/ pulp-admin I create the repo and sync it with:

```
pulp-admin rpm repo create --repo-id epel6 --download-policy=on_demand --feed="http://fedora.uib.no/epel/6/x86_64/"
pulp-admin rpm repo sync run --repo-id epel6
```

I'm going to try http://yum.oracle.com/repo/OracleLinux/OL7/developer_EPEL/x86_64/ next with:

```
pulp-admin rpm repo create --repo-id OL7 --download-policy=on_demand --feed="http://yum.oracle.com/repo/OracleLinux/OL7/developer_EPEL/x86_64/"
pulp-admin rpm repo sync run --repo-id OL7
```

#16 - 06/08/2018 11:14 PM - bmbouter

Actually this does reproduce w/ the oracle linux repo.

Task Failed

Invalid block tag on line 7: '}%gmail{*}com>'. Did you forget to register or load this tag?

#17 - 06/11/2018 11:08 PM - bmbouter

Actually the escaping code is not even performing that transform. The exception occurs when Template(foo) tries to instantiate sanitized data and it contains `{%}%gmail{*}com>not {% templatetag openblock %}%gmail{*}com>`. I'm looking into why this data isn't sanitized like it should be.

#18 - 06/11/2018 11:47 PM - bmbouter

This issue occurs because the regex that is used to select the `<changelog>...</changelog>` entry for sanitization matches non greedily. The package python2-pika-doc (details below) has multiple changelogs. Here is the snippet stored in Pulp's db that is being published:

- merge changelog- Rebuild for Python 3.6- Honor %_smp_ncpus_max setting on testing- enable parallel testing with nose - enable python-twisted-core and python-tornado on epel - drop obsolete Group tag- https://fedoraproject.org/wiki/Changes/Automatic_Provides_for_Python_RPM_Packages- drop some duplications - validate b0rken tests on epel7- add %check with execution of both unit and acceptance tests - enable adapters for both tornado and twisted - generate additional documentation, split into subpackage- Fix builds by defining python3_pkgversion if it doesn't exist - Add missing BRs for py3-other variant (for EPEL 7)- Actually make the python 3 bcond work- Upgrade to version 0.10.0 - Refactor to meet current Fedora guidelines - Add Python 3 subpackage (with EPEL 7 compatibility)- Rebuilt for https://fedoraproject.org/wiki/Fedora_24_Mass_Rebuild- Rebuilt for https://fedoraproject.org/wiki/Fedora_23_Mass_Rebuild- Rebuilt for https://fedoraproject.org/wiki/Fedora_21_Mass_Rebuild- Rebuilt for https://fedoraproject.org/wiki/Fedora_20_Mass_Rebuild- Rebuilt for https://fedoraproject.org/wiki/Fedora_19_Mass_Rebuild- Rebuilt for https://fedoraproject.org/wiki/Fedora_18_Mass_Rebuild- Bump pika release version to fix upgrade path for f17 -> f18- Patch pika/adapters/blocking_connection.py- Rebuilt for https://fedoraproject.org/wiki/Fedora_17_Mass_Rebuild- Upgrade to version 0.9.5- Upgrade to version 0.9.4- Upgrade to version 0.9.3- Initial Package

package details:

```
<package arch="noarch" name="python2-pika-doc" pkgid="{{ pkgid }}"><version epoch="0" rel="9.el7" ver="0.10.0" /
```

#19 - 06/12/2018 08:25 PM - bmbouter

The diff below causes the section selected for sanitization to be greedy instead of non-greedy and resolves the issue so sync+publish on that repo work. This fix is very safe for several reasons:

1. Escaping all data for the template layer should be safe. This is essentially user-provided data given that we got it from the internet so it probably all should be escaped. Since this fixes this issue I'm not going to rewrite the sanitizer now to keep it low-risk.

2. The extra-greediness added is still limited by the end tag matched

I'm making a PR from this now.

```
diff --git a/plugins/pulp_rpm/plugins/db/models.py b/plugins/pulp_rpm/plugins/db/models.py
index 4d49ff2c..77f455a2 100644
--- a/plugins/pulp_rpm/plugins/db/models.py
+++ b/plugins/pulp_rpm/plugins/db/models.py
@@ -903,7 +903,7 @@ class RpmBase(NonMetadataPackage):
     """
     start_tag_pattern = r'<%s.*?(?<!/)>' % tag_name
     end_tag_pattern = r'</%s>' % tag_name
-    complete_tag_pattern = r'(%s)(.*?)(%s)' % (start_tag_pattern, end_tag_pattern)
+    complete_tag_pattern = r'(%s)(.*)(%s)' % (start_tag_pattern, end_tag_pattern)
     tag_re = re.compile(complete_tag_pattern, flags=re.DOTALL)
     template = tag_re.sub(RpmBase._generate_tag_replacement_str, template)
     return template
```

#21 - 06/13/2018 03:16 PM - bmbouter

- Status changed from ASSIGNED to MODIFIED

Applied in changeset [4a9965fc100f1414d3be0845138fd66981b0ba31](#).

#22 - 06/15/2018 02:03 PM - dkliban@redhat.com

- Platform Release set to 2.16.2

#24 - 06/18/2018 08:09 PM - bmbouter

Applied in changeset [62edf22a6632be8f2632799836c6234466fb2f95](#).

#25 - 07/09/2018 04:25 PM - ipanova@redhat.com

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE

#26 - 04/15/2019 10:12 PM - bmbouter

- Tags Pulp 2 added