Learn more from your logfiles using machine learning

[DEV1156]
Adam.Spiers@suse.com
Dirk.Mueller@suse.com

CC BY-NC 2.0 Thomas Hawk
We are SUSE OpenStack Cloud software engineers
We love green CI

### Cloud9 Gating

<table>
<thead>
<tr>
<th>S</th>
<th>W</th>
<th>Name</th>
<th>Last Success</th>
<th>Last Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-gating</td>
<td>6 hr 11 min - #325: Build1254</td>
<td>7 hr 31 min - #324: Build1253</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-gating-trigger</td>
<td>11 min - #403</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-sst min.x86_64</td>
<td>6 hr 11 min - #326: cloud-ardana-ci-sst23 (10.86.3.198)</td>
<td>7 hr 30 min - #325: cloud-ardana-ci-sst23 (10.86.2.30)</td>
</tr>
</tbody>
</table>

### Cloud9 Gerrit

<table>
<thead>
<tr>
<th>S</th>
<th>W</th>
<th>Name</th>
<th>Last Success</th>
<th>Last Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ardana-trackupstream</td>
<td>13 hr - #1258</td>
<td>23 days - #1211</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ardana-trigger-trackupstream</td>
<td>2 min 30 sec - #517</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-sst -3to-update-perf-x86_64</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-sst-min-centos-perf-x86_64</td>
<td>16 days - #2: cloud-ardana-ci-sst9 (10.86.2.194)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-sst-min-gerrit-x86_64</td>
<td>5 hr 49 min - #238: cloud-ardana-ci-sst11 (10.86.3.326)</td>
<td>9 hr 14 min - #279: cloud-ardana-ci-sst8 (10.86.4.54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>openstack-ardana-gerrit-cloud9</td>
<td>5 hr 51 min - #1842: 5096</td>
<td>8 hr 29 min - #1640: 5096 (std.min)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>openstack-ardana-gerrit-events</td>
<td>2 min 30 sec - #5716</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>openstack-ardana-heat</td>
<td>1 min - #9488: create-crowbar-poc</td>
<td>8 hr 27 min - #9432: create-cloud-ardana-ci-sst9</td>
</tr>
</tbody>
</table>

### Cloud9 Staging

<table>
<thead>
<tr>
<th>S</th>
<th>W</th>
<th>Name</th>
<th>Last Success</th>
<th>Last Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-dsc-3tp-x86_64</td>
<td>2 days 7 hr - #157: cloud-ardana-ci-sst3 (10.86.3.76)</td>
<td>3 hr 44 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloud-ardana9-job-demp-x86_64</td>
<td>1 day 14 hr - #276: cloud-ardana-ci-sst2 (10.86.2.30)</td>
<td>1 hr 33 min</td>
</tr>
</tbody>
</table>
We care about upstream OpenStack CI too
### OpenStack Health

<table>
<thead>
<tr>
<th>Job Name</th>
<th>Passes</th>
<th>Failures</th>
<th>% Failures</th>
</tr>
</thead>
<tbody>
<tr>
<td>openstack-ansible-odl-sfc-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>15</td>
<td>15</td>
<td>50.00</td>
</tr>
<tr>
<td>openstack-ansible-openvswitch-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>15</td>
<td>15</td>
<td>50.00</td>
</tr>
<tr>
<td>openstack-ansible-ovs-rsh-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>15</td>
<td>15</td>
<td>50.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-centos-7-openstack-ansible-os_neutron</td>
<td>35</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-distro_install-opensuse-423-openstack-ansible-os_neutron</td>
<td>34</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-distro_install-centos-7-openstack-ansible-os_neutron</td>
<td>31</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-ubuntu-bionic-openstack-ansible-os_neutron</td>
<td>26</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>20</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-ovn-centos-7-openstack-ansible-os_neutron</td>
<td>20</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-ovn-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>20</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-ovs-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>30</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-distro_install-opensuse-150-openstack-ansible-os_neutron</td>
<td>14</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-distro_install-ubuntu-xenial-openstack-ansible-os_neutron</td>
<td>4</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>openstack-ansible-functional-distro_install-centos-7-openstack-ansible-os_neutron</td>
<td>2</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Did you find it?
Idea: Reducing scrolling by pattern matching

warning (/(?i)warning/)
error /Traceback \(\text{most recent call last}\)/
error (/(?i)error/)
error (/(?i)bfail(ure|ed)?\b/)
error (/(?i)fatal/)
error /$h1!!/
Started by upstream project "cloud-mkcloud9-job-ha-linuxbridge-x86_64" build number 228
originally caused by:
  
  [EnvInject] - Loading node environment variables.

  Building remotely on cloud-swarm-mch1 (openstack-mkcloud-SLE12-x86_64 swarm openstack-mkcloud-SLE12-SP2-x86_64 openstack-mkcloud
  openstack-mkcloud-ha-x86_64) in workspace /root/workspace/openstack-mkcloud@2

  Set build name:

  New build name is '102889: cloud-mkcloud9-job-ha-linuxbridge-x86_64'

  [openstack-mkcloud@2] $ /bin/bash -x /tmp/jenkins13275971323501011787.sh
  
  '["!": "!/root/bin/update_automation.sh"]'

  /root/bin/update_automation

  Handling repo: github.com/SUSE-Cloud/automation with branch: master

  Your branch is up-to-date with 'origin/master'.

  HEAD is now at 564650010 Merge pull request #3287 from SUSE-Cloud/revert-3242-tempast_last

  Handling repo: github.com/openSUSEF/github-pr with branch: master

  Your branch is up-to-date with 'origin/master'.

  HEAD is now at 263e3fd let config define if white, black or both lists pass through

  [openstack-mkcloud@2] $ /bin/bash -x /tmp/jenkins205821586502912254.sh

  + export automationrepos=/root/github.com/SUSE-Cloud/automation

  + automationrepos=/root/github.com/SUSE-Cloud/automation


  + shopt -s extglob

  + cd /root/workspace/openstack-mkcloud@2

  + rm -rf ./log-parser-plugin-rules.txt ./mkcloud.config ./mkcloud.steps.log ./artifacts

  + '["!": "!/github_pr"]'

  + exit 0

  [openstack-mkcloud@2] $ /bin/bash -x /tmp/jenkins1745354864698202328.sh

  + export automationrepos=/root/github.com/SUSE-Cloud/automation

  + automationrepos=/root/github.com/SUSE-Cloud/automation

  + '["!": "!/github_pr"]'
Started by upstream project "cloud-mkcloud0-job-ha-linuxbridge-x86_64" build number 228
originally caused by:
Started by timer

[EnvInject] - Loading node environment variables.
Building remotely on cloud-swarmp-kush (openstack-mkcloud-SLE12-x86_64 swmp openstack-mkcloud-SLE12-SP2-x86_64 openstack-mkcloud openstack-mkcloud-ha-x86_64) in workspace /root/workspace/openstack-mkcloud
Set build name.
New build name is '4162889: cloud-mkcloud0-job-ha-linuxbridge-x86_64'
/openstack-mkcloud02.sh /bin/bash -x /tmp/jenkins142700712158010117787.sh
+ '[ 1 ]' -> /root/bin/update_automation '
+ /root/bin/update_automation
Handling repo: github.com/SUSE-Cloud/automation with branch: master
Already on 'master'
Your branch is up-to-date with 'origin/master'.
HEAD is now at 04669b19 Merge pull request 3237 from SUSE-Cloud/revert-3242-tempest_last
Handling repo: github.com/openSUSE/gh-pr with branch: master
Already on 'master'
Your branch is up-to-date with 'origin/master'.
HEAD is now at 205e3f6e let config define if white, black or both lists pass through
/openstack-mkcloud02.sh /bin/bash -x /tmp/jenkins12858235368642232256.sh
+ export automationrepos=root/github.com/SUSE-Cloud/automation
+ automationrepos=root/github.com/SUSE-Cloud/automation
+ shopt -s extglob
+ cd /root/workspace/openstack-mkcloud02
+ rm -r /log-parser-plugin-rules.txt ./mkcloud.config ./mkcloud.steps.log ./artifacts
+ '[ 1 ]' -> /github_pr '
+ exit 0
/openstack-mkcloud02.sh /bin/bash -x /tmp/jenkins17453548046982932258.sh
+ export automationrepos=root/github.com/SUSE-Cloud/automation
+ automationrepos=root/github.com/SUSE-Cloud/automation
+ '[ 1 ]' -> /root/workspace/openstack-mkcloud02/automation-git'
**Error (6)**

Starting proposal sat(default) at
Wed Mar 27 23:59:50 UTC 2019
3 Errors in this section
1 FESDB: Failed collecting logs for
K8s@honey.bunny.python2.
2 Failed to delete group with name or
ID testvm: No SecurityGroup found
testvm
3 1 of 1 groups failed to delete

**Warning (21)**

MCKCLOUD stage rebootstop
(1 Error in this section)
4 Running sync failed. Could not
find card matching card:nkccloud
vb2ha16mxrnxp.chi.64 in Cloud 9

**Info (56)**

### Parsed Console Output

```
++(mcloud-common.sh:510) getcloudver(): [ [ davelcloud9 =~ \(.*[[:digit:]]+\|[":\[\]{}\|]+\) ] ]
++(mcloud-common.sh:511) getcloudversion(): echo $9
+(mcloud-common.sh:533) lscloudver(): local ver=9
+(mcloud-common.sh:536) lscloudver(): [ [ 8 =~ \(8-9\)+\$ ] ]
+(mcloud-common.sh:548) lscloudver(): "\' 9 \ge 8 \'"
+(mcloud-common.sh:560) lscloudver(): return 0
+(qa_crowbarsetup.sh:41/1) oncontroller_testsetup():: openstack security group delete testvm
Failed to delete group with name or ID 'testvm': No SecurityGroup found for testvm
1 of 1 groups failed to delete.
+(qa_crowbarsetup.sh:41/1) oncontroller_testsetup()::
+(qa_crowbarsetup.sh:41/2) oncontroller_testsetup():: openstack security group create --description testvm testvm
```

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>created_at</td>
<td>2019-03-28T00:59:06Z</td>
</tr>
<tr>
<td>description</td>
<td>testvm</td>
</tr>
<tr>
<td>id</td>
<td>1d0b8701-a4e1-4d7a-0642-0f3c4f5c3a</td>
</tr>
<tr>
<td>name</td>
<td>testvm</td>
</tr>
<tr>
<td>project_id</td>
<td>82f74d4e69042fe0b0b2df5b932fe68</td>
</tr>
<tr>
<td>revision_number</td>
<td>1</td>
</tr>
<tr>
<td>rules</td>
<td>created_at='2019-03-28T00:59:06Z', direction='egress', ethertype='IPv6', id='10f8fa11-68ec-4439-a0f5-54a155d7759f', updated_at='2019-03-28T00:59:06Z'</td>
</tr>
</tbody>
</table>
```
```
Dealing with false positives

```bash
# Successful tempest run
ok /^ - (Expected Fail|Failed): 0$/
ok /Warning: Turning on 'gpg-auto-import-keys'/
ok /Warning: Permanently added .* to the list of known hosts/
ok /WARNING: Device for PV .* not found or rejected by a filter/
ok /WARNING: \w+ signature detected on .* offset \d+. Wipe it?/
ok /grep -v failed

# rpms containing "Error"
ok /perl-Error[-]|libsamba-errors|mariadb-errormessages/

# https://bugzilla.suse.com/show_bug.cgi?id=1030822
warning /Cleaning up (vip-admin-S+) on S+, removing fail-count-\1/

# https://bugzilla.suse.com/show_bug.cgi?id=971832
ok /Failed to try-restart vsftpd@.service: Unit name vsftpd@.service is not valid
```
Vision: Machine Learning

Job Results → Root Cause
Log-Classify

job-output.txt.gz (log link)

<table>
<thead>
<tr>
<th>Time</th>
<th>Status</th>
<th>Username</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>5903</td>
<td></td>
<td>ubuntu-xenial</td>
<td>skipping: [server2] =&gt; {item={u'execreloads': []}, u'enabled': True, u'config_overrides': {}, u'state': u'started', u'execstarts': u'/openstack/venvs/neutron-testing/bin/neutron-server --config-file /etc/neutron/neutron.conf --config-file /etc/neutron/plugins/ml2/ml2_conf.ini --log-file=/var/log/neutron/neutron-server.log ', u'service_name': u'neutron-server'})</td>
</tr>
<tr>
<td>5904</td>
<td></td>
<td>ubuntu-xenial</td>
<td></td>
</tr>
<tr>
<td>5905</td>
<td></td>
<td>ubuntu-xenial</td>
<td>TASK [os_neutron : Add RPC RabbitMQ vhost]</td>
</tr>
<tr>
<td>5906</td>
<td></td>
<td>ubuntu-xenial</td>
<td>fatal! [server1 -&gt; 10.1.0.2]: FAILED! =&gt; {&quot;changed&quot;: false, &quot;msg&quot;: &quot;Failed to find required executable rabbitmqctl in paths: /usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin&quot;}</td>
</tr>
<tr>
<td>5907</td>
<td></td>
<td>ubuntu-xenial</td>
<td></td>
</tr>
<tr>
<td>5908</td>
<td></td>
<td>ubuntu-xenial</td>
<td>NO MORE HOSTS LEFT</td>
</tr>
<tr>
<td>5909</td>
<td></td>
<td>ubuntu-xenial</td>
<td>to retry, use: --limit @/home/zuul/src/git.openstack.org/openstack/openstack-ansible-os_neutron/tests/test_retry</td>
</tr>
<tr>
<td>5916</td>
<td></td>
<td>ubuntu-xenial</td>
<td>server1 : ok=167 changed=73 unreachable=0 failed=1</td>
</tr>
<tr>
<td>5917</td>
<td></td>
<td>ubuntu-xenial</td>
<td>server2 : ok=167 changed=73 unreachable=0 failed=0</td>
</tr>
<tr>
<td>5918</td>
<td></td>
<td>ubuntu-xenial</td>
<td></td>
</tr>
</tbody>
</table>
Today's plan

- Intro to Machine Learning
- Log-Classify Implementation
- Demo
Why Machine Learning?
CI Logfiles: ML Challenges

• Each Instance of a CI Logfile execute the same steps
  – Install, Build, Test
  – Result is recorded (success, failures)

• The individual Logfiles are quickly evolving
  – Every check-in changes it 😊

• Each run has a lot of completely unique noise 😘
  – Timestamps, UUIDs, Passwords and
  – ordering due to parallel execution
## Learning model Variations

<table>
<thead>
<tr>
<th>Instance-based</th>
<th>Generalizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Directly store instances of training</td>
<td></td>
</tr>
<tr>
<td>• Derives hypotheses directly from training instances</td>
<td></td>
</tr>
<tr>
<td>• Model can be quickly react to new training input</td>
<td></td>
</tr>
<tr>
<td>• Model can be incrementally updated discarding old training input</td>
<td></td>
</tr>
<tr>
<td>• Abstracting a model from training data</td>
<td></td>
</tr>
<tr>
<td>• Requires much longer training phase</td>
<td></td>
</tr>
<tr>
<td>• Can not &quot;untrain&quot; previously learned data</td>
<td></td>
</tr>
</tbody>
</table>

**k-Nearest-Neighbor**

**Artificial Neural Networks (DL)**
Overfitting / Underfitting
# Machine Learning Variations

<table>
<thead>
<tr>
<th>Supervised</th>
<th>Unsupervised</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classification</strong></td>
<td><strong>Clustering</strong></td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>K-Means</td>
</tr>
<tr>
<td>NearestNeighbor</td>
<td>Hidden Markov Model</td>
</tr>
<tr>
<td>Support Vector Machines (SVM)</td>
<td>Neural Networks ...</td>
</tr>
<tr>
<td>Neural Networks ...</td>
<td></td>
</tr>
<tr>
<td><strong>Regression</strong></td>
<td></td>
</tr>
<tr>
<td>Decision Trees</td>
<td></td>
</tr>
<tr>
<td>Linear Regression</td>
<td></td>
</tr>
<tr>
<td>Neural Networks ...</td>
<td></td>
</tr>
</tbody>
</table>
Supervised Learning: Classification

Banana

Banana
Using machine learning for CI log files
Machine Learning Workflow

- **Build**: an individual CI log file
- **Baseline**: Collection of log files from good CI runs
- **Target**: The failed CI log run logfile to be analyzed
log-classify: Analogy using pictures
Generic Training Workflow

Baselines

Transform

Learning

Model
Generic Testing Workflow

Baselines

Target

Model

Transform

Learning

Testing
Log Input transformation example

Splitting by lines

Mar 11 02:43:28 localhost sudo[5195]: pam_unix(sudo:session): session opened for user root by (uid=5)

Tokenization

DATE localhost sudo pam_unix sudo session session opened for user root uid

Hashing

hash(DATE) hash(localhost) hash(sudo) hash(pam_unix) hash(sudo) hash(session) hash(opened) ...

Transformation

[0, ..., 0, 1, 0, ..., 0, 1, 0, ...]
**Input transformation: Replace irrelevant pieces with fixed strings**

<table>
<thead>
<tr>
<th>Token</th>
<th>Raw text</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATE</td>
<td>months/days/date</td>
</tr>
<tr>
<td>RNGU</td>
<td>UUIDs</td>
</tr>
<tr>
<td>RNGI</td>
<td>IPv4 or IPv6 addresses</td>
</tr>
<tr>
<td>RNGN</td>
<td>words that are exactly 32, 64 or 128 chars</td>
</tr>
<tr>
<td>RNGD</td>
<td>numbers of at least 3 digits</td>
</tr>
</tbody>
</table>
Example matrix of a CI logfile
k-Nearest Neighbors (k=1)

Baselines

Target
Example distance calculation in kNeighbors queries

Vector = controller  ephemeral_device  VARIABLE  IS  NOT  DEFINED

kneighbors(Vector) = 0.9

• VARIABLE IS NOT DEFINED is not part of the baseline
Limitations

- Nearest Neighbor performs linear search in model
- Complexity grows linearly with samples size
- Unfiltered Noise may distract from important information
- Logs containing too many features
Unique vectors over training set instances
Lookup time per sample size

Search time per sample size

 samples

0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
14000 15000 16000 17000 18000 19000 20000 21000

0.32
0.34
0.36
0.38
0.40
0.42
0.44
Introducing log-classify
Log-classify

http://scikit-learn.org/

not yet: https://www.tensorflow.org/

https://github.com/facebookresearch/pysparnn

- Python 3
- Multiple Text Extraction Models
- Assumes text, line based log-like input
log-classify: Installation

openSUSE Leap/Tumbleweed/SLE 15 SUSE Package Hub:

```bash
$ zypper install python3-logreduce
```

Others install from PyPI:

```bash
$ pip3 install --user logreduce
```

NOTE:

- **log-classify** is the new name
- Rename from **logreduce** hasn't been completed yet
### log-reduce

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>diff</code></td>
<td>Compare directories/files</td>
</tr>
<tr>
<td><code>dir</code></td>
<td>Train and run against local files/dirs</td>
</tr>
<tr>
<td><code>dir-train</code></td>
<td>Build a model for local files/dirs</td>
</tr>
<tr>
<td><code>dir-run</code></td>
<td>Run a model against local files/dirs</td>
</tr>
<tr>
<td><code>job</code></td>
<td>Train and run against CI logs</td>
</tr>
<tr>
<td><code>journal</code></td>
<td>Train and run against local journald</td>
</tr>
</tbody>
</table>
log-classify: Assumptions

- Baseline is built from **SUCCESS** input
- Model is run against **FAILURE** target
log-classify: DevStack

```bash
# logreduce diff logs/good.txt logs/bad.txt
...
0.527 | bad.txt:34245: 2018-10-09 05:56:51.021261 | controller |
    Details: {u'created': u'2018-10-09T05:11:20Z', u'code': 500,
    u'message': u'Exceeded maximum number of retries. Exhausted 
    all hosts available for retrying build failures for instance
d7046aa3-e885-4ed6-80e7-d7a7eff9f883.'}
...
97.98% reduction (from 35244 lines to 712)
```
Truncated singular value decomposition (SVD)
log-classify: Collecting baselines

$ logreduce dir-train model.clf baseline/*
Training on 8 logs took 12.090s at 1.426MB/s (20.831kl/s)

$ logreduce dir-run model.clf error.txt
Testing took 6.375s at 0.454MB/s (6.569kl/s)
99.72% reduction (from 41879 lines to 118)
log-classify: Influence of Baseline Size
log-classify: Journald

- Extract novelty in today's logs over yesterday:
  
  ```
  # logreduce journal --range day
  ```

- Build a model using previous month's logs and look for novelties:
  
  ```
  # logreduce journal-train --range month journal.clf
  ```
# logreduce journal-run --range day journal.clf

99.76% reduction (from 19677 lines to 48)

0.730 | cron - postdrop: warning: uid=16311: File too large

0.000 | smartd Device: /dev/sdb, 1 Offline uncorrectable sectors

# killall -SEGV automount

# logreduce journal-run --range day journal.clf

99.75% reduction (from 19679 lines to 50)

0.317 | systemd - DAEMON - autofs.service: Main process exited, code=dumped, s

0.314 | systemd - DAEMON - autofs.service: Failed with result 'core-dump'.
# logreduce dir-train nova.clf /var/log/nova/nova-compute.log-*xz
# logreduce dir-run nova.clf /var/log/nova/nova-compute.log

... 0.684 | INFO .. No calling threads waiting for msg_id : d3afd41a53bb4d14a5e42c
       0.619 | INFO .. Recovered from being unable to report status

... 93.15% reduction (from 6741 lines to 462)
# logreduce diff report-good/ report-bad/ --html report.html

Training took 51.364s at 1.543MB/s
Testing took 37.432s at 0.446MB/s
...
88.41% reduction (from 261091 lines to 30251)
neutron: set a failure-timeout on neutron-ha-tool #2063

dirkmueller wants to merge 1 commit into crowbar:master from dirkmueller:rabbit_l3_ha_failcount · Jump to bottom

We don't want the l3 ha service to be stopped after 3 weeks of weekly patching and rebooting the rabbitmq cluster. Set a timeout of a failure if it happened more than 10 minutes ago.

dirkmueller requested a review from aspiers as a code owner March 18, 2019 1:05 PM

chef/cookbooks/neutron/recipes/network_agents_ha.rb

```ruby
agent "systemd:neutron-l3-ha-service"
```

hound reviewed March 18, 2019 1:06 PM

Reviewers
- hound
- aspiers

Requested changes must be addressed to merge this pull request.

Assignees
No one—assign yourself

Labels
- needs backport to SOC7
- needs backport to SOC8
neutron: set a failure-timeout on neutron-ha-tool #2063

dirkmueller wants to merge 1 commit into crowbar:master from dirkmueller:rabbit_l3_ha_failcount

Aspiers added a commit to aspiers/presentation-namespace-regressions that referenced this pull request.

March 31, 2019 3:57 PM

Add more commits by pushing to the rabbit_l3_ha_failcount branch on dirkmueller/crowbar-openstack.

📊 1 review requesting changes by reviewers with write access.

**aspiers** requested changes

- Approve changes
- Dismiss review
- Re-request review

🔴 1 errored, 1 failing, and 3 successful checks

- Hound — 2 violations found.
- suse/mkcloud/ha — failed
- continuous-integration/travis-ci/pr — The Travis CI build passed
- suse/mkcloud — PR gating succeeded
- suse/mkcloud/testbuild — testbuild job succeeded

獒 Merging is blocked
log-classify: Other use cases

- Open Build Service build result files
- Your ideas?
Zuul Architecture
Post-Run Analysis

Builds

Test Instances

Zuul

Executor

Logserver
Post-Run Playbook

- job:
  name: base
  post-run:
    - classify-log
    - upload-log

- tasks:
  - name: Fetch or build the model
    command: log-classify job-train ...
  - name: Generate report
    command: log-classify job-run ...
  - name: Return report url
    zuul_return: report.html
<table>
<thead>
<tr>
<th>Job</th>
<th>Duration</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>gate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>post</td>
<td></td>
<td></td>
</tr>
<tr>
<td>promote</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-release</td>
<td></td>
<td></td>
</tr>
<tr>
<td>release</td>
<td></td>
<td></td>
</tr>
<tr>
<td>periodic</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Gate**

Changes that have been approved by core reviewers are queued in order in this pipeline, and if they pass tests, will be merged. For documentation on how to use with GitLab, please see https://docs.gitlab.com/ee/about/pipelines/

**Post**

This pipeline runs jobs that operate after each change is merged. Queue items are identified by the abbreviated hash (git log --format=onl) of the merge commit.

**Promote**

This pipeline runs jobs that operate after each change is merged in order to promote artifacts generated in the gate pipeline.

**Pre-release**

When a commit is tagged with a pre-release tag, this pipeline runs jobs that publish archives and documentation.

**Release**

When a commit is tagged as a release, this pipeline runs jobs that publish archives and documentation.

**Periodic**

Jobs in this queue are triggered on a timer.

Queue: openstack/kolla
Conclusions
Log-Classify is hosted on softwarefactory-project.io

Welcome

This platform hosts git repositories, a review system and CI components. Use the top menu to access hosted services. Below is the list of projects hosted on this platform. If you want to host a new project then propose a change on the config repository as explained in this documentation.

<table>
<thead>
<tr>
<th>Python3-SCL</th>
<th>ansible-network</th>
<th>Distributed-CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Python3 Software Collection</strong></td>
<td><strong>Supported roles developed by Ansible Network Engineering</strong></td>
<td><strong>A Distributed-CI Platform</strong></td>
</tr>
<tr>
<td>© Website: <a href="https://softwarefactory-project.io">https://softwarefactory-project.io</a></td>
<td>© Website: <a href="https://www.distributed-ci.io">https://www.distributed-ci.io</a></td>
<td></td>
</tr>
<tr>
<td>&lt; Documentation: <a href="https://softwarefactory-project.io/docs">https://softwarefactory-project.io/docs</a></td>
<td>&lt; Documentation: <a href="https://docs.distributed-ci.io">https://docs.distributed-ci.io</a></td>
<td></td>
</tr>
<tr>
<td>▪ Contact:</td>
<td>▪ Issue-tracker: <a href="https://github.com/redhat-cip/dci-control-server/issues">https://github.com/redhat-cip/dci-control-server/issues</a></td>
<td>▪ Contact:</td>
</tr>
<tr>
<td>▪ <a href="mailto:softwarefactory-dev@redhat.com">softwarefactory-dev@redhat.com</a></td>
<td>▪ Review dashboard: default</td>
<td>▪ <a href="mailto:softwarefactory-dev@redhat.com">softwarefactory-dev@redhat.com</a></td>
</tr>
<tr>
<td>▪ Mailing-list:</td>
<td>▪ Mailing-list:</td>
<td>▪ Mailing-list:</td>
</tr>
<tr>
<td>▪ <a href="mailto:softwarefactory-dev@redhat.com">softwarefactory-dev@redhat.com</a></td>
<td>▪ <a href="mailto:distributed-ci@redhat.com">distributed-ci@redhat.com</a></td>
<td>▪ <a href="mailto:distributed-ci@redhat.com">distributed-ci@redhat.com</a></td>
</tr>
</tbody>
</table>

**Associated git repositories (190)**

- scl/enable-py3-distgit
  Description: enable-py3 package
  Clone: git clone https://softwarefactory-project.io/scl/enable-py3-distgit

**Associated git repositories (19)**

- ansible-network/zuul-config
  Description:
  Clone: git clone https://softwarefactory-project.io/r/ansible-network/zuul-config
  Gisthub: last commits / browse files
  Code review: open reviews

- ansible-network/ansible-zuul-jobs
  Description:
  Clone: git clone https://softwarefactory-project.io/r/ansible-network/ansible-zuul-jobs
  Gisthub: last commits / browse files
  Code review: open reviews

**Associated git repositories (22)**
How to contribute

- Apache-2.0 Licensed
- #log-classify on Freenode
**Future plans**

- Handle Streaming logs
  - Adaptive model
  - Incremental model training
- Curate public domain datasets
- Fingerprint and detect archived anomalies
- More services: Jenkins build, Travis CI, ...
- More reporter: Logstash filter, ...
- Clustering (DBSCAN, k-means) for outlier detection
Q&A

Please fill out the survey!

https://pypi.org/project/logreduce/

Icons used in these diagrams are licensed under Creative Commons Attribution 3.0:

- https://fontawesome.com/license
- https://zuul-ci.org