

Better Testing Through Statistics

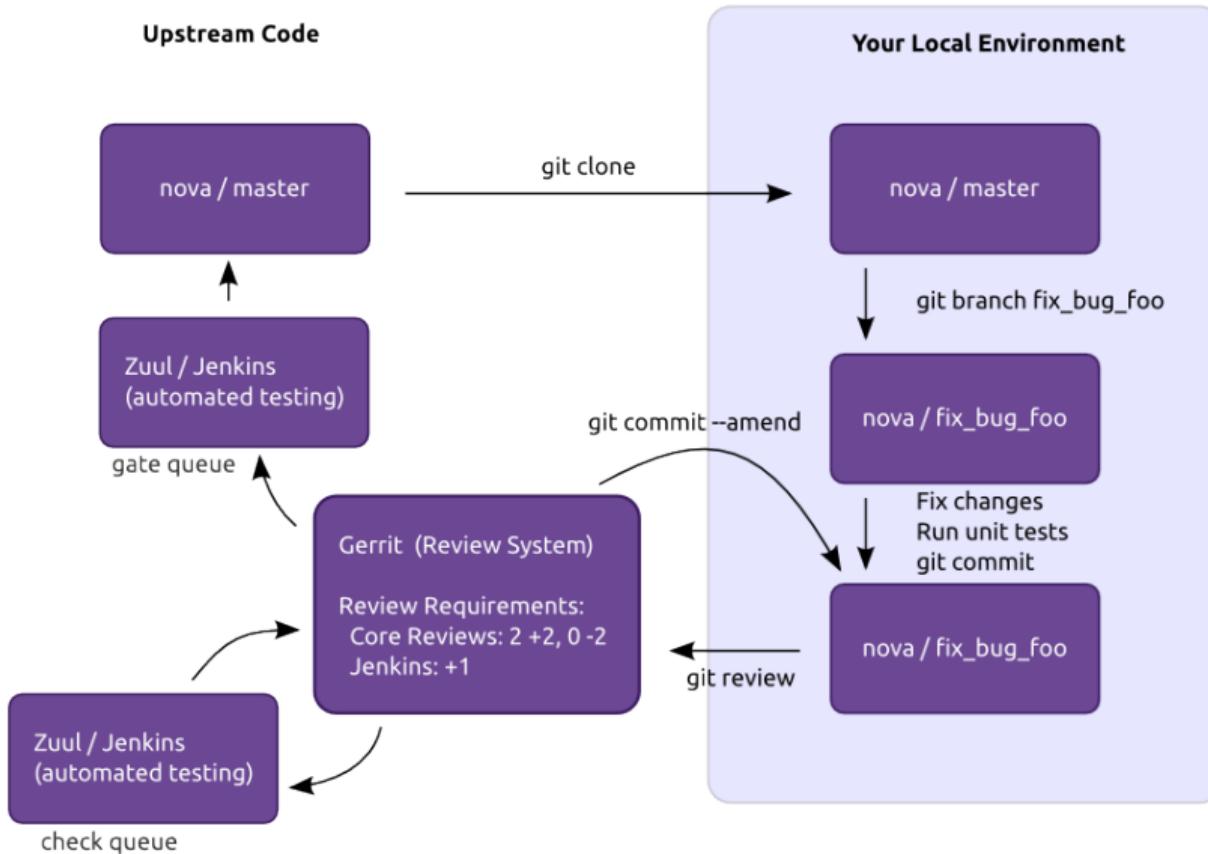
Matthew Treinish
mtreinisch@kortar.org
mtreinisch on Freenode

Masayuki Igawa
masayuki.igawa@gmail.com
masayukig on Freenode

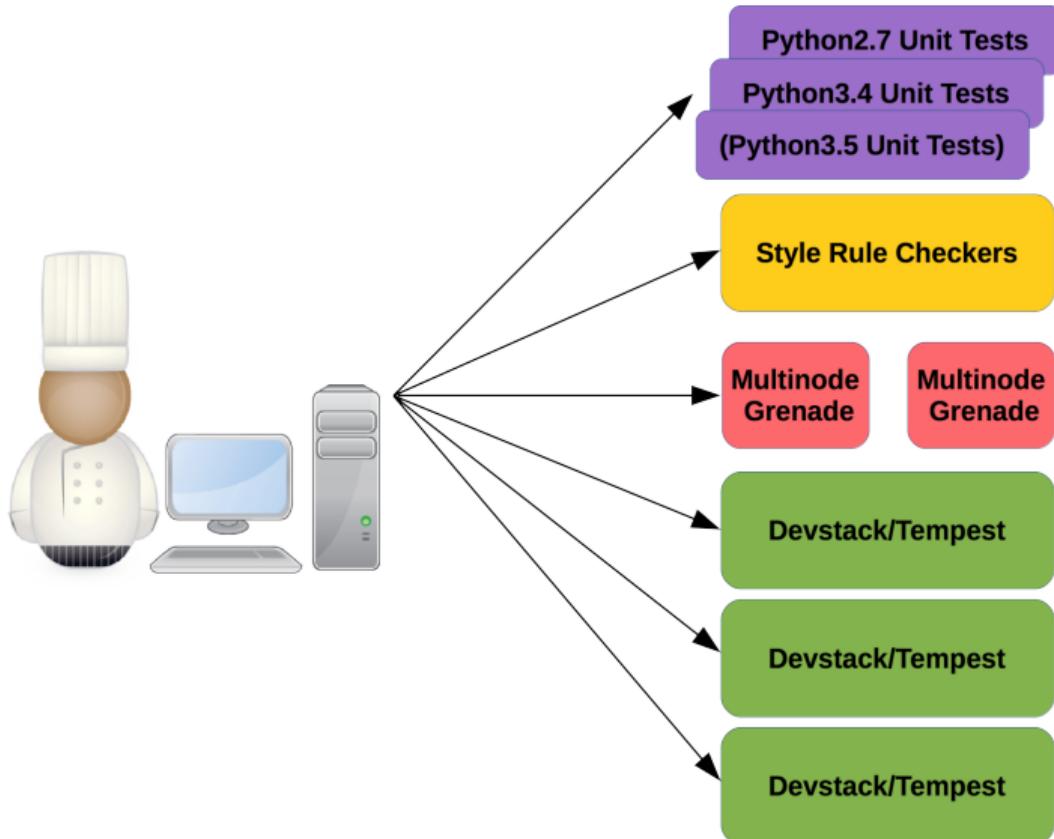
January 17, 2016

<https://github.com/masayukig/better-testing-through-statistics/tree/lca2017>

What is “the OpenStack Gate”?



What Happens when you push a change?



check

(268)

Newly uploaded patchsets enter this pipeline to receive an initial +1-1 Verified vote from Jenkins.

Change queue: openstack/neutron

	openstack/neutron 181574,23	unknown 3 hr 57 min
gate-neutron-docs:	SUCCESS	
gate-neutron-pep8:	SUCCESS	
gate-neutron-python27:	FAILURE	
gate-neutron-python34:	FAILURE	
gate-tempест-dsvm-neutron-full:	queued	
gate-grenade-dsvm-neutron:	SUCCESS	
gate-neutron-dsvm-api:	SUCCESS	
gate-neutron-dsvm-functional:	SUCCESS	
gate-neutron-dsvm-fullstack: (non-voting)	SUCCESS	
gate-rally-dsvm-neutron-neutron: (non-voting)	SUCCESS	
gate-tempест-dsvm-neutron-dvr:	SUCCESS	
gate-tempест-dsvm-neutron-identity-v3-only-full-rv: (non-voting)	SUCCESS	
gate-tempест-dsvm-neutron-linuxbridge:	SUCCESS	
gate-tempест-dsvm-neutron-pg-full: (non-voting)	SUCCESS	
gate-neutron-lbaasv2-dsvm-minimal:	SUCCESS	
gate-grenade-dsvm-neutron-multinode: (non-voting)	SUCCESS	
gate-grenade-dsvm-neutron-dvr-multinode: (non-voting)	SUCCESS	
gate-tempест-dsvm-neutron-multinode-full: (non-voting)	SUCCESS	
gate-tempест-dsvm-neutron-dvr-multinode-full: (non-voting)	SUCCESS	
gate-tempест-dsvm-ironic-pxe_ipa-rv: (non-voting)	SUCCESS	

Change queue: openstack/networking-generic-swift

	openstack/networking-generic-switch 308884,3	unknown 3 hr 52 min
gate-networking-generic-switch-docs:	queued	
gate-networking-generic-switch-pep8:	SUCCESS	
gate-networking-generic-switch-python27:	SUCCESS	
gate-networking-generic-switch-python34:	SUCCESS	
gate-networking-generic-switch-dsvm:	SUCCESS	

Change queue: openstack/neutron

	openstack/neutron 280595,12	unknown 3 hr 38 min
gate-neutron-docs:	SUCCESS	
gate-neutron-pep8:	SUCCESS	
gate-neutron-python27:	SUCCESS	
gate-neutron-python34:	SUCCESS	
gate-tempест-dsvm-neutron-full:	SUCCESS	
gate-grenade-dsvm-neutron:	SUCCESS	
gate-neutron-dsvm-api:	SUCCESS	
gate-neutron-dsvm-functional:	SUCCESS	
gate-neutron-dsvm-fullstack: (non-voting)	FAILURE	
gate-rally-dsvm-neutron-neutron: (non-voting)	queued	
gate-tempест-dsvm-neutron-dvr:	SUCCESS	
gate-tempест-dsvm-neutron-identity-v3-only-full-rv: (non-voting)	SUCCESS	
gate-tempест-dsvm-neutron-linuxbridge:	SUCCESS	
gate-tempест-dsvm-neutron-pg-full (non-voting)	SUCCESS	

gate

(26)

Changes that have been approved by core developers are enqueued in order in this pipeline, and if they pass tests in Jenkins, will be merged.

Change queue: integrated

	openstack/khova 307269,1	0 min 1 hr 10 min
gate-nova-docs:	SUCCESS	
gate-nova-pep8:	SUCCESS	
gate-nova-python27-db:	FAILURE	
gate-nova-python34-db:	FAILURE	
gate-nova-requirements:	queued	
gate-tempест-dsvm-full:	SUCCESS	
gate-tempест-dsvm-postgres-full:	SUCCESS	
gate-tempест-dsvm-neutron-full:	SUCCESS	
gate-grenade-dsvm:	SUCCESS	
gate-nova-releasenotes:	SUCCESS	
gate-nova-tox-db-functional:	SUCCESS	
gate-grenade-dsvm-multinode:	SUCCESS	
gate-tempест-dsvm-cells:	SUCCESS	
gate-tempест-dsvm-full-devstack-plugin-ceph:	SUCCESS	

openstack/khova
304730,1

	openstack/khova 304730,1	0 min 1 hr 10 min
gate-nova-docs:	SUCCESS	
gate-nova-pep8:	SUCCESS	
gate-nova-python27-db:	SUCCESS	
gate-nova-python34-db:	SUCCESS	
gate-tempест-dsvm-full:	SUCCESS	
gate-tempест-dsvm-postgres-full:	SUCCESS	
gate-tempест-dsvm-neutron-full:	SUCCESS	
gate-grenade-dsvm:	SUCCESS	
gate-nova-releasenotes:	SUCCESS	
gate-nova-tox-db-functional:	SUCCESS	
gate-grenade-dsvm-multinode:	SUCCESS	
gate-tempест-dsvm-cells:	SUCCESS	
gate-tempест-dsvm-full-devstack-plugin-ceph:	SUCCESS	

openstack/khova
303995,1

	openstack/khova 303995,1	0 min 1 hr 5 min
gate-nova-docs:	SUCCESS	
gate-nova-pep8:	SUCCESS	
gate-nova-python27-db:	SUCCESS	
gate-nova-python34-db:	SUCCESS	
gate-tempест-dsvm-full:	SUCCESS	
gate-tempест-dsvm-postgres-full:	SUCCESS	
gate-tempест-dsvm-neutron-full:	SUCCESS	
gate-grenade-dsvm:	SUCCESS	
gate-nova-releasenotes:	SUCCESS	
gate-nova-tox-db-functional:	SUCCESS	
gate-grenade-dsvm-multinode:	SUCCESS	
gate-tempест-dsvm-cells:	SUCCESS	
gate-tempест-dsvm-full-devstack-plugin-ceph:	SUCCESS	

openstack-dev/devstack
308791,1

	openstack-dev/devstack 308791,1	0 min 1 hr 5 min
gate-devstack-docs:	SUCCESS	

post

(79)

This pipeline runs jobs that operate after each change is merged.

Change queue: openstack/osl.concurrency

	openstack/osl.concurrency 342ef03	unknown 5 hr 2 min
oslo.concurrency-branch-tarball:	SUCCESS	
oslo.concurrency-docs:	queued	
oslo.concurrency-upstream-translation-update:	SUCCESS	
oslo.concurrency-coverage:	queued	

Change queue: openstack-infra/project-config

	openstack-infra/project-config 08001cc	unknown 5 hr 0 min
publish-infra-docs-index:	queued	
publish-specs-site:	queued	

Change queue: openstack-infra/project-config

	openstack-infra/project-config bd07b6c	unknown 4 hr 56 min
publish-infra-docs-index:	queued	
publish-specs-site:	queued	

Change queue: openstack/networking-vsphere

	openstack/networking-vsphere 1931ebe	unknown 4 hr 55 min
networking-vsphere-branch-tarball:	queued	

Change queue: openstack-infra/project-config

	openstack-infra/project-config d7f08ff	unknown 4 hr 54 min
publish-infra-docs-index:	queued	
publish-specs-site:	queued	

Change queue: openstack-infra/project-config

	openstack-infra/project-config 8cb6337	unknown 4 hr 52 min
publish-infra-docs-index:	queued	
publish-specs-site:	queued	

Change queue: openstack/stackalytics

	openstack/stackalytics 40f07b8	unknown 4 hr 7 min
hook-stackalytics-ics-rtfd:	SUCCESS	

Change queue: openstack/stackalytics

	openstack/stackalytics a5e5a37	unknown 4 hr 7 min
hook-stackalytics-ics-rtfd:	SUCCESS	

Change queue: openstack/inventance

	openstack/inventance	
--	----------------------	--

The Size of the Gate

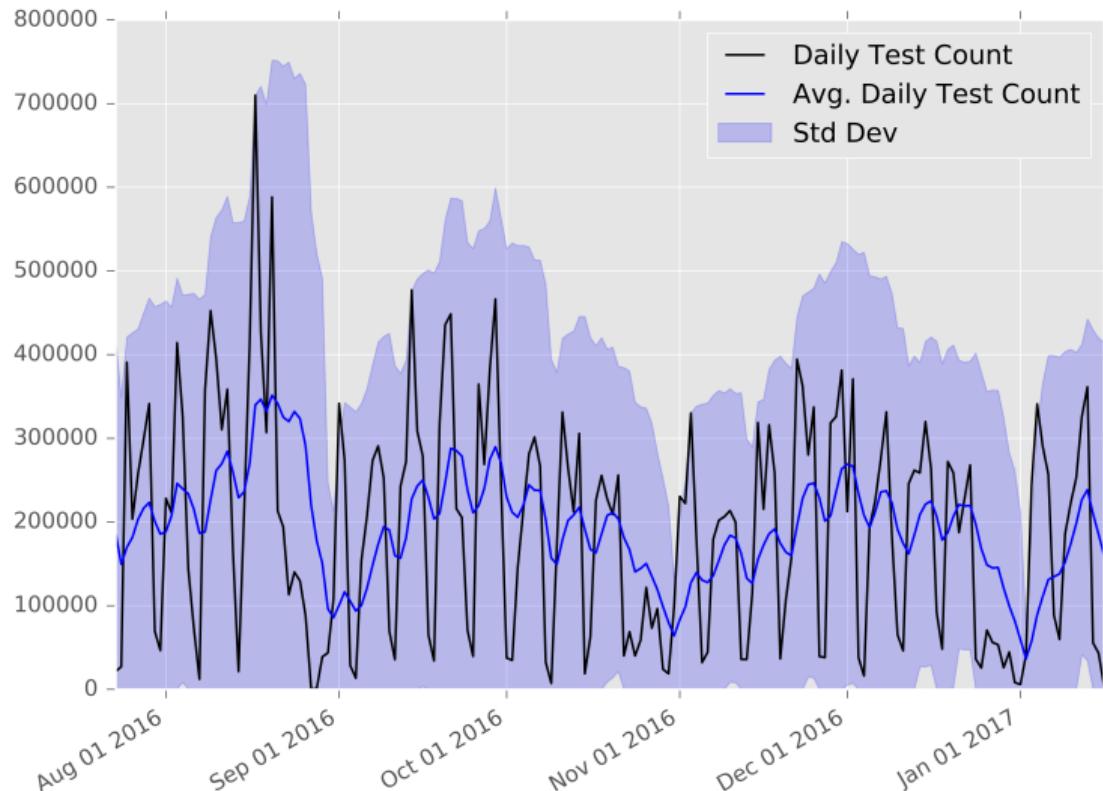
One Proposed Change Generates:

- ▶ 5–25 Devstacks
- ▶ ~10,000 integration tests (roughly 1.5k per devstack)
- ▶ ~151 2nd level guests created in each devstack cloud
- ▶ ~1 GB of logs uncompressed for each run

In aggregate:

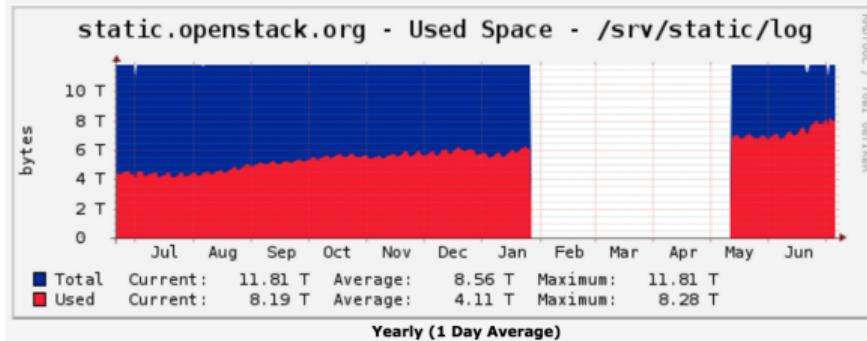
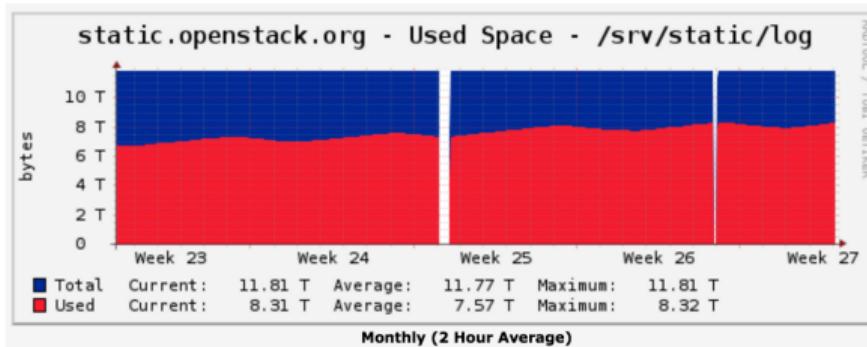
- ▶ ~12,500 jobs run in check and gate daily
- ▶ ~0.01% individual tempest test failure rate
- ▶ ~.77% tempest run failure rate

Number of Tempest Tests per Day in the Gate Queue:



Log Server

- Log Server: <http://logs.openstack.org/>
- Archive of all artifacts from all jobs for ~4 months
- ~8 TB of data compressed



Problems/Issues when Running in the Gate

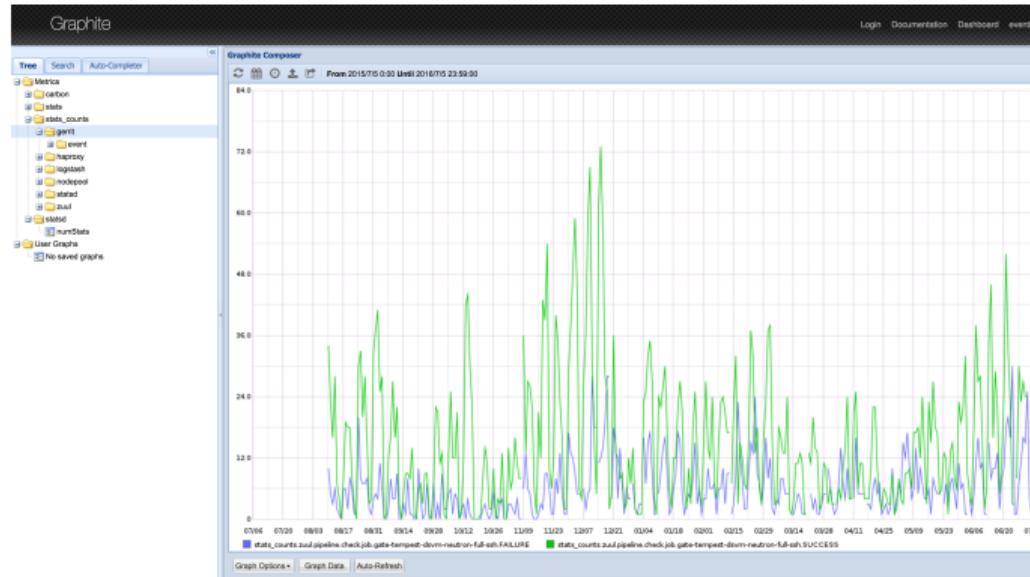
- ▶ Works fine for looking at results 1 at a time
- ▶ Difficult to find non-deterministic failures
- ▶ Difficult to find performance regressions
- ▶ Finding out how often something passes or fails is next to impossible

General Approach

- ▶ Look at things on the larger scale
- ▶ Use statistics and data mining to find previously unknown trends in OpenStack
- ▶ Make the data from test runs open and accessible to everyone
- ▶ Ensure there are APIs available for accessing everything

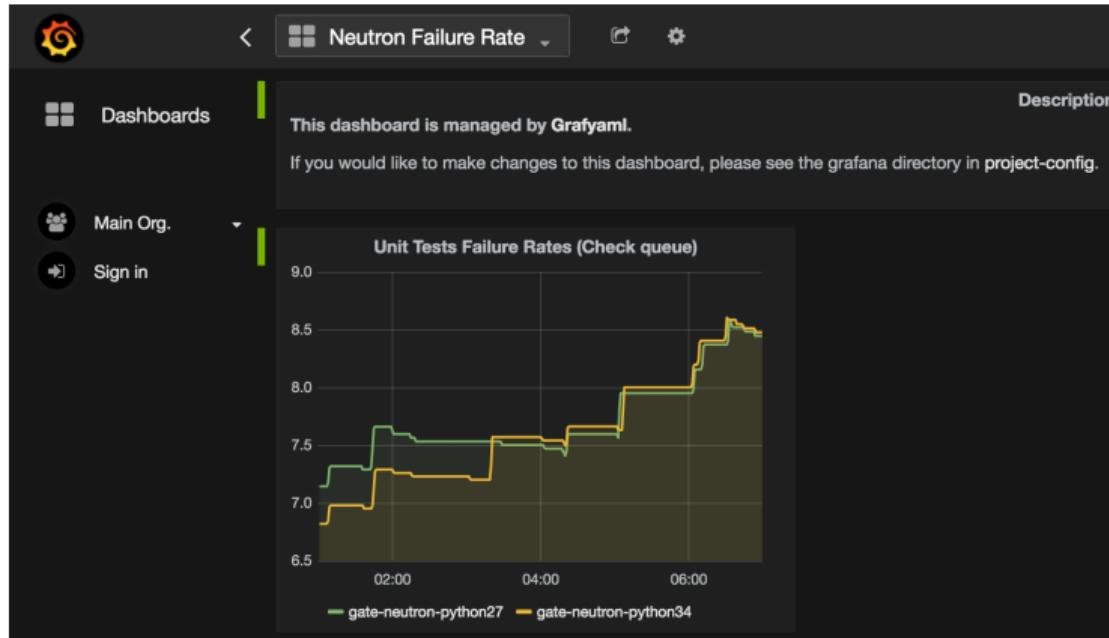
Graphite

- ▶ <http://graphite.openstack.org/>
- ▶ Infra services report to graphite
- ▶ Include job results
- ▶ Limited to job level data
- ▶ Time based, can't be linked to an individual job



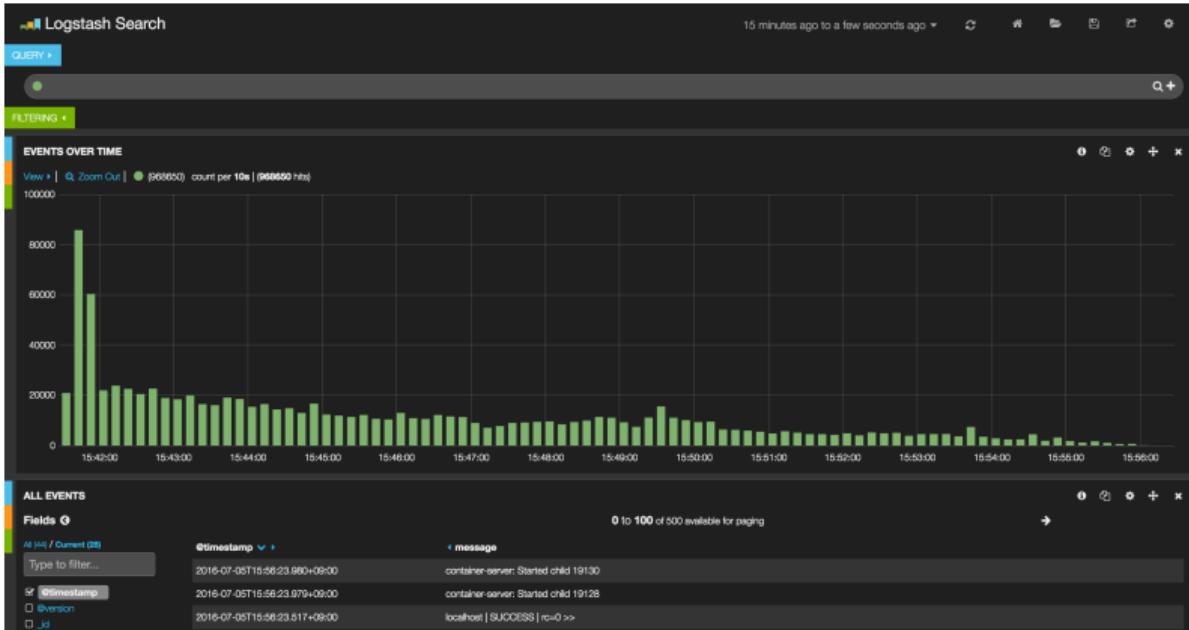
Grafana

- ▶ <http://grafana.openstack.org/>
- ▶ Provides a layer on top of graphite to easily make useful visualizations
- ▶ Adds a number of dashboards
- ▶ Some projects using this to track job failure rates



ELK

- ▶ Elasticsearch, Logstash, Kibana
- ▶ <http://logstash.openstack.org>
- ▶ Provides a search engine on top of are job artifacts
- ▶ Limited to 10 days of results

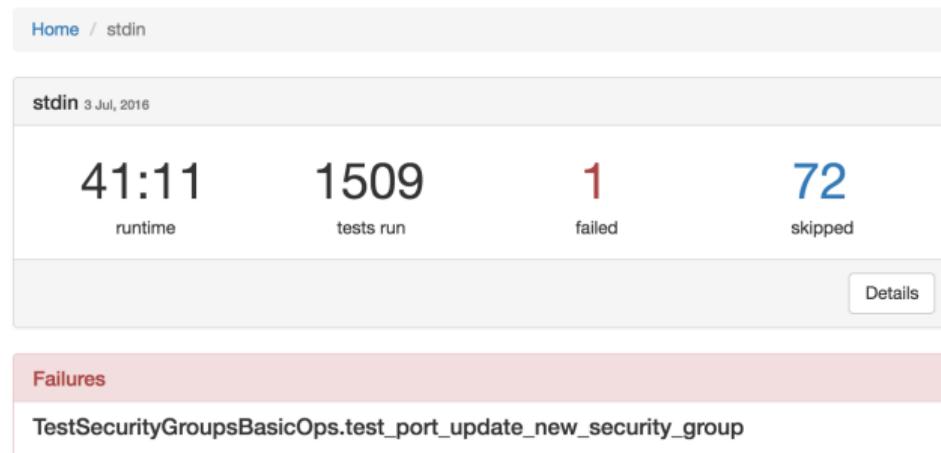


StackViz

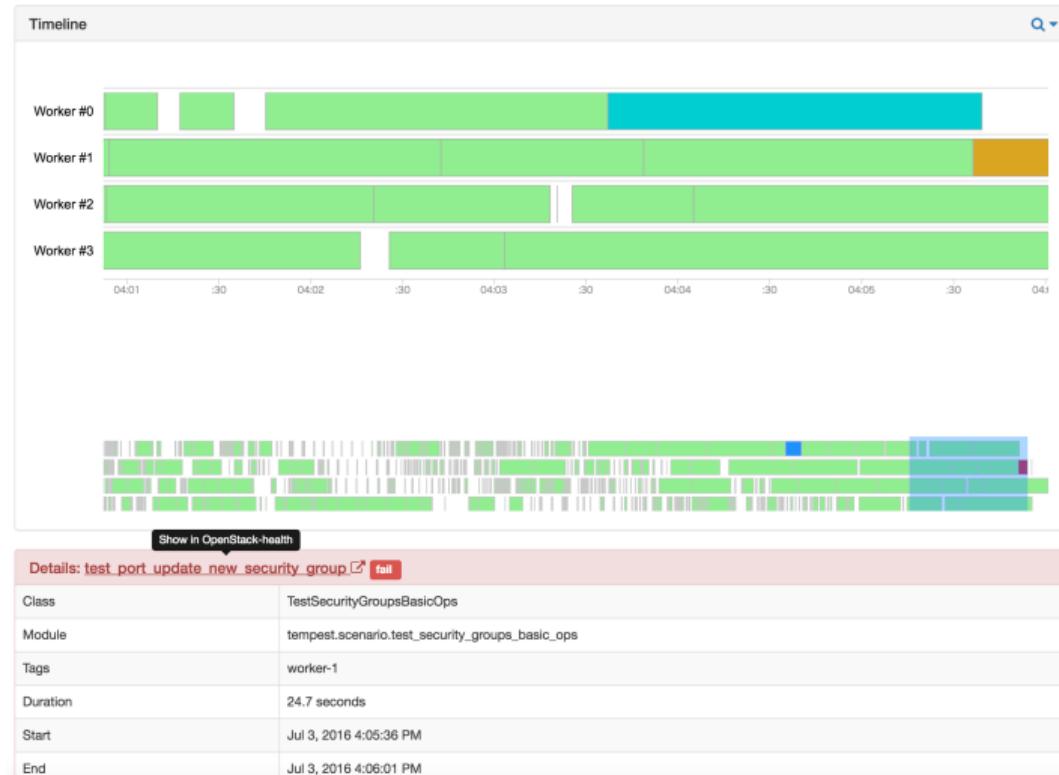
Visualization tool of individual CI build results

► git.openstack.org/cgit/openstack/stackviz

Datasets

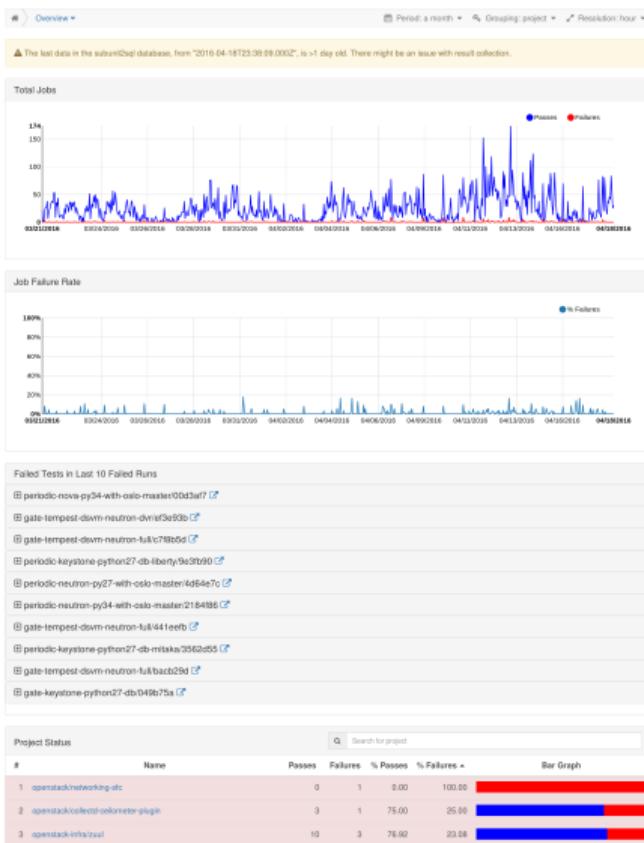


StackViz: Timeline



Using OpenStack Health

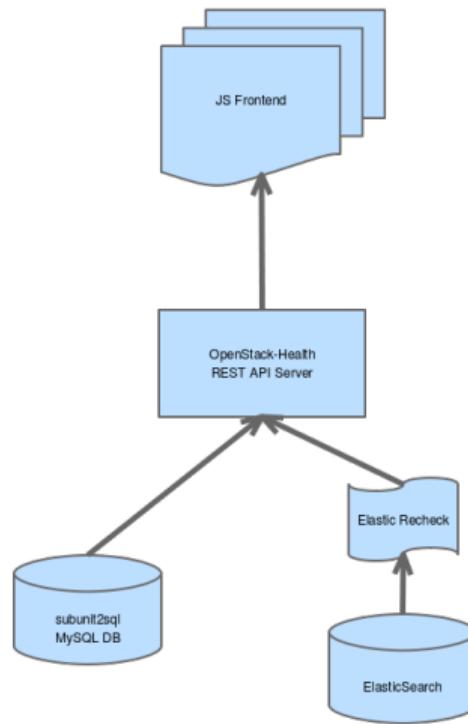
OpenStack Health is a dashboard for visualizing test results of OpenStack CI jobs.



openstack-health

- ▶ <http://status.openstack.org/openstack-health/#/>
- ▶ Designed to be a single point of access for all the data about the gate
- ▶ Currently can leverage subunit2sql and elastic-recheck

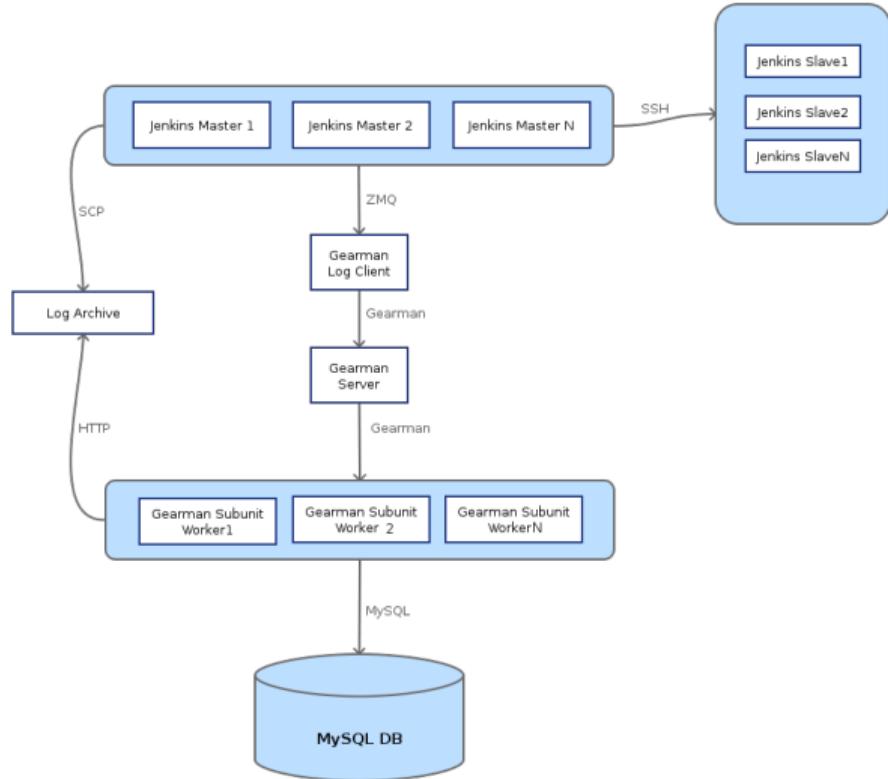
OpenStack-Health Architecture



subunit2sql

- ▶ Designed to store test results data in a sql database
- ▶ Provides a DB schema and a python API for interacting with the database
- ▶ CLI utilities for storing and retrieving results in the DB as subunit v2
- ▶ A public database of everything with subunit output run in OpenStack-Infra
- ▶ Used to store the results from test runs for 6 months

subunit2sql in OpenStack Infrastructure



Elastic Recheck

- ▶ Designed to answer the question “Have you seen this recently?”
- ▶ Leverages elastic search to identify failures with known fingerprints
- ▶ Contains a repository of elastic-search queries with known failures
- ▶ Has 2 parts:
 - ▶ A bot which watches changes and reports identified failures to gerrit and IRC
 - ▶ A dashboard which shows failure categorization

Elastic Recheck

<http://status.openstack.org/elastic-recheck/>

All Pipelines Gate Pipeline Uncategorized

The elastic-recheck project uses Elasticsearch to classify and track OpenStack gate failures. Documentation can be found [here](#). You can also learn more by reading this post on the Elasticsearch blog: [OpenStack elastic-recheck: powered by the elk stack](#)

Data Last Updated: Tue Jul 12 2016 13:00:01 GMT+0900 (JST)

Last Elastic Search Index Update: Tue Jul 12 2016 12:59:20 GMT+0900 (JST)

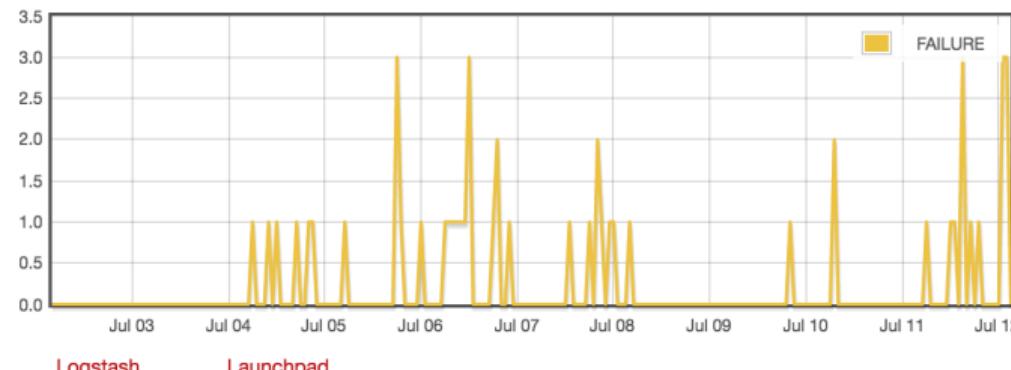
Delay in Elastic Search: Up to date

Cluster Health: green

Bug 1539271 - Libvirt live block migration migration stalls

14 fails in 24 hrs / 50 fails in 10 days

Projects: (nova - Confirmed)



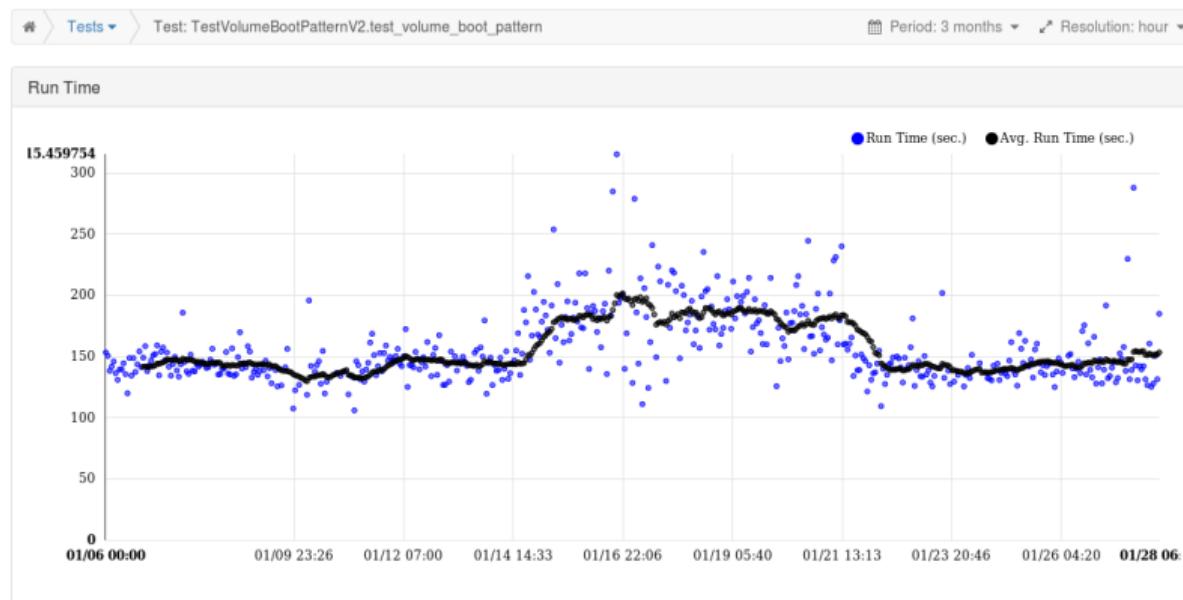
Data Driven Decision Making

- ▶ Determine when it's time to skip a test
- ▶ Identify tests that are actually catching bugs
- ▶ Determine if failures are isolated to region, config, etc.

Finding trends amongst the noise

- ▶ Catch performance regressions
- ▶ Identify relational trends in the data
- ▶ Find and identify non-deterministic bug/race conditions

TestVolumeBootPatternV2.test_volume_boot_pattern



Issues

- ▶ Too many varied data sources each with unique limitations:
 - ▶ Only Gate and Periodic Job data (subunit2sql)
 - ▶ No views for infra failure (subunit2sql)
 - ▶ Limited to 1 line for searching (Elastic Search)
- ▶ Limited contribution in this space

Future work

- ▶ Integrate all the things in openstack-health
- ▶ Use the data to optimize our test runner scheduler
- ▶ Enable automation around failure detection

Where to get more information

- ▶ openstack-dev ML openstack-dev@lists.openstack.org
- ▶ #openstack-qa on Freenode
- ▶ <http://git.openstack.org/cgit/openstack/openstack-health/>
- ▶ <http://git.openstack.org/cgit/openstack-infra/subunit2sql>
- ▶ <http://git.openstack.org/cgit/openstack-infra/elastic-recheck/>

Questions?