Openstack features backporting

Backporting to Openstack Juno the RBD Object Map feature

Saverio Proto
saverio.proto@switch.ch

Zurich 9/03/2016 OpenStack Meetup

SWITCHengines

- Openstack public cloud with two regions in Switzerland
  - https://www.switch.ch/engines/

- SWITCHengines provides compute and storage services to Swiss academics. Users of participating institutions have on-demand access to create and manage virtual machines and storage. If you are interested in using the infrastructure, be sure to contact us at engines-support@switch.ch.

The Openstack Release cycle

- New stable release every 6 months
  - https://wiki.openstack.org/wiki/Releases
- Releases are EOF after 12 months

- Developers say you should upgrade your production Openstack deployment every 6 months. WOW !

- A lot of work to upgrade:
  - Test puppet modules
  - Test Openstack components (deprecation of config parameters)
  - Downtime during the upgrade
  - Solve problems at scale

The problem

- The system is slow to delete Volumes
  - Delete operations often crash leaving the system inconsistent
  - The larger the Volume the bigger the problem

- Ceph
  - Cinder is build with e Cinder RBD backend
  - We can reproduce in the backend that Ceph is slow to delete large rbd volumes, using the rbd command line client

Ceph version with solution

- Ceph Hammer has a release that fixes our problem.

- When the object map feature is enabled on an image, limiting the diff to the object extents will dramatically improve performance since the differences can be computed by examining the in-memory object map instead of querying RADOS for each object within the image.

- http://docs.ceph.com/docs/master/man/8/rbd/

Changes to ceph.conf

- Good news ! We are already running the latest Hammer release, just need to enable the feature we need in the config file !

- rbd default format = 2

- rbd default features = 13

- http://www.sebastien-han.fr/blog/2015/07/06/ceph-enable-the-object-map-feature/

RBD Clients do not honor ceph.conf

- Nova, cinder and glance have rbd specific code that ignores the new configuration in ceph.conf
  - As simple as when they parse ceph.conf they don't expect new features other than layering

- Of course we are running Juno and everything is fixed in Kilo.

- We know it takes 10 more weeks to upgrade to Kilo
  - We have to make our users happy sooner

Identify patches

- http://github.com/openstack/
- http://review.openstack.org
- We identified the Kilo patches to be backported to Juno

- We built new Ubuntu packages (Ubuntu Cloud Archive)

- Testing

- Finally fix the problem !
  - Our patches will never make it upstream, Juno is EOF

Read this story on our blog

- All details of this story on our blog
  - http://cloudblog.switch.ch/2016/02/08/backport-to-openstack-juno-the-ceph-rbd-object-map-feature/

- Lesson learned: **upgrade Openstack as fast as possible !**

Thank you – Questions ?

- Twitter:
  - @zioproto
  - @SWITCHengines

- http://cloudblog.switch.ch/
- https://www.switch.ch/engines/
- engines-support@switch.ch