GlusterFS as Storage for OpenShift Container Platform Applications

K8s Meetup / Munich / Nov 2018

Elvir Kuric, Red Hat

kubernetes

Who Am I?

- Elvir Kuric [Dipl.Ing. El.]
- Working for Red Hat - Perf / Scale Engineering Team
- (home)based in Munich
- fan of hiking

Jochberg (-> https://www.bergtour-online.de)

Agenda

- What is OpenShift Container Platform
  - Why to use OCP and what it offers more if compared with other K8S solutions
- What is RHOCS (Red Hat OpenShift Container Storage)
  - We will spend most of time at this point particularly covering
    - RHOCS Install
    - Underlying technology
    - What to do / not to do with RHOCS
    - Demo
- Questions and Open discussion
  - Questions regarding RHOCS
  - What I can answer will answer, what not will get back to you
  - I am not professional presented - so your feedback is valuable for me

What is OpenShift Container Platform

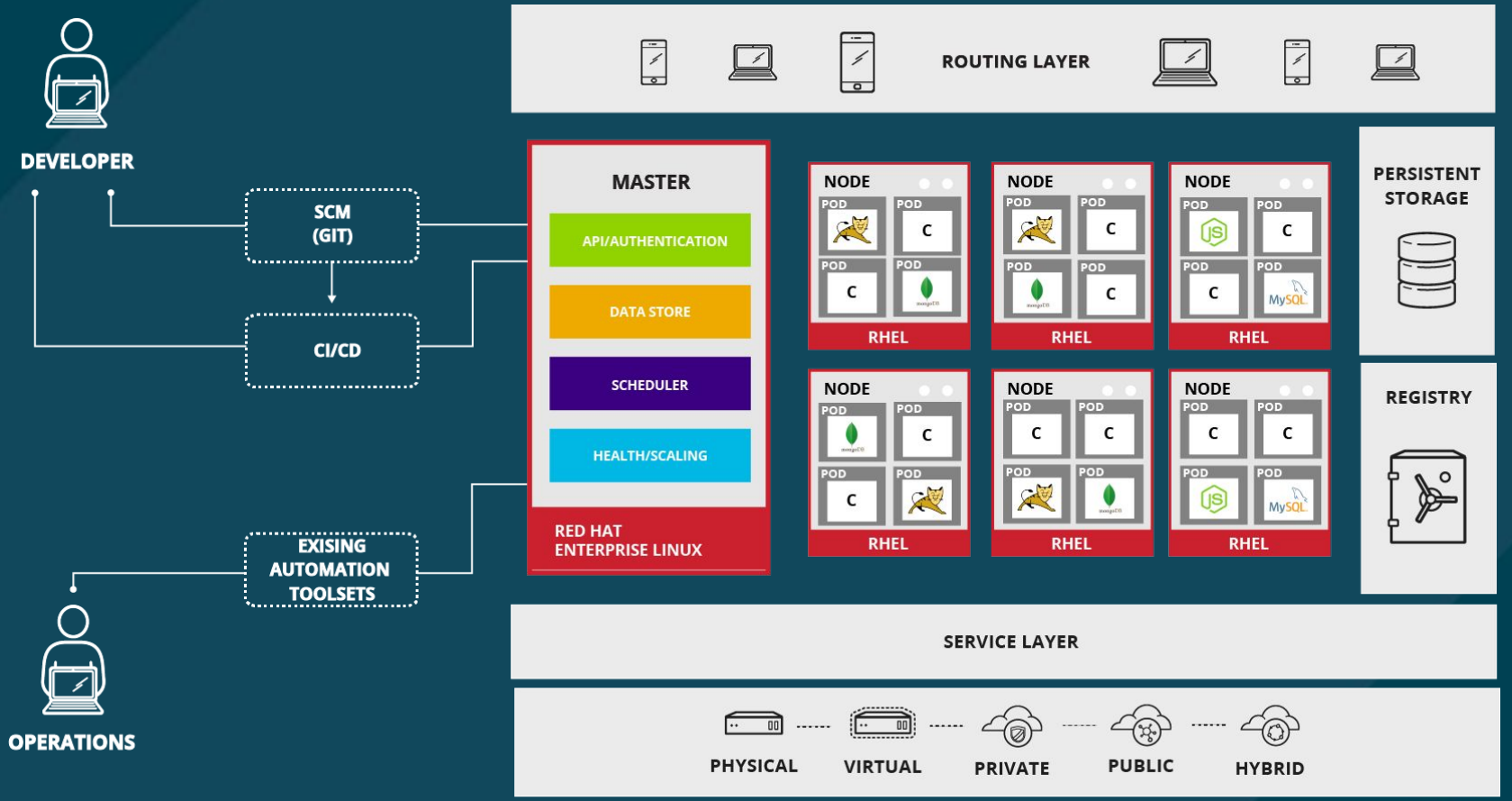
- Red Hat *version* of kubernetes
- Built on top of RHEL (RHEL Atomic / CoreOS)
- Benefits of RHEL OS/ RHEL kernel / support teams
- Possible to install OCP
  - Public clouds (GCP , AWS, Azure, ….)
  - Openstack / Baremetal (on-premise)
  - Virtual platforms (VMware / RHEV)

What is OpenShift Container Platform

- Openshift-installer (openshift-ansible)
- Installs logging / metrics by default
- Implements RBAC - Role Based Access Controls / Selinux support

OpenShift (OCP) Cluster in one picture

10,000 foot overview



redhat

Kubernetes Storage

- List is possible storage solutions for kubernetes is long
  - Public Cloud solutions (storage on aws, gpc, azure) - all of them will offer some kind of storage and they can be directly used by kubernetes cluster
    - sometimes this is best approach
  - Gluster / ceph / cinder / nfs … (public / private cloud)
  - Many others storage vendors (Quobyte , Portwrox, Netapp…)
  - CSI - Container Storage Interface will make process of adding new storage type easier

Glusterfs Storage for Openshift

- Glusterfs
  - How it is implemented in Openshift
  - How to install it (we will run this live)
    - Node selections - label / daemon set / recommended node size
- Gluster CSI driver is released
  https://github.com/gluster/gluster-csi-driver/releases

Glusterfs Storage

- Install on OCP
- Run uninstall.yaml - part of openshift-ansible
    - This will clean up disks for storage (wipe FS signatures, partitions on devices planned for gluster)
  - Install with pointing to config.yaml
    - glusterfs_devices='["/dev/sdb" , “/dev/sdc”]'
    - Important: Ensure that devices planned for storage are really ones you want!!!

Glusterfs Storage

```
# oc get pods -n glusterfs
```

NAME	READY	STATUS	RESTARTS	AGE
glusterblock-storage-provisioner-dc-1-65hjr	1/1	Running	2	14d
glusterfs-storage-g8rkk	1/1	Running	1	7d
glusterfs-storage-gtsh7	1/1	Running	2	7d
glusterfs-storage-xpttm	1/1	Running	2	7d
heketi-storage-1-rglps	1/1	Running	0	1d

Glusterfs storage

```
# oc get sc

NAME                                PROVISIONER                AGE

glusterfs-storage                   kubernetes.io/glusterfs    14d

glusterfs-storage-block               gluster.org/glusterblock   20d
```

Daemonset

- https://gist.github.com/ekuric/6550e6f59ed3f653a4d8d5e3a7510ae7
- It gives info what from OCP nodes is used by gluster pods - helpful with debugging

Glusterblock provisioner

- glusterblock-storage-provisioner pod
- It serves as “glusterblock” provisioner

Heketi

- heketi-storage - new pod which provides RESTful API on top of gluster
- JWT (Json Web Tokens) for auth
- # oc logs -f heketi-storage-<pod>

  is first place to check when debugging
- heketi-client is client package (man heketi-cli)
- Possible also to run it in so called “standalone” mode

Heketi

- heketi-storage - new pod which provides RESTful API on top of gluster
- JWT (Json Web Tokens) for auth
- # oc logs -f heketi-storage-<pod> is first place to check when debugging
- heketi-cli is client - it is part of heketi-client package

  (man heketi-cli)
- Possible also to run it in so called “standalone” mode

Heketi

- Get secret

```
oc get secret -n glusterfs heketi-storage-admin-secret -o yaml
| grep key | awk '{print $2}' | base64 --decode
```

- `heketi-cli -s <server_ip> --user admin --secret <secret> volume list`
- `heketi-cli -s <server_ip> --user admin --secret <secret> topology info`
- `heketi-cli is also interface for adding new nodes / devices`

Heketi

- Uses BoltDB as data store (bbolt)
- Database itself is located on glustervolume (oc exec -n glusterfs glusterfs-strorage-pod -- gluster v list)
- during installation, glustervolume : heketidbstorage is created and used as storage for bolt database.
- heketidbstorage volume is delete protected

GlusterFS scaling

- One TSP (Trusted Storage Pool) is 4 nodes
  - 3 nodes are enough for 3-way replication, but 4th node is used to preserve volume creation in case on node is down (3 way replica)
  - 1000 PVC per one TSP
  - It is recommended to add new TSP if there is need for more PVCs (use different node labels for new TSP)

Upgrades

- We have “daemonset”
- Rolling updates are option
- Prior upgrade
    - Ensure new are images are accessible (either pre-pull images in advance, or ensure registry where images are is accessible)

“Image” and “updateStrategy: type: RollingUpdate

Apps

- MongoDB (with db-workload profile)
- Postgresql DB (with db-workload profile) https://docs.openshift.com/container-platform/3.11/scaling_performance/optimizing_on_glusterfs_storage.html
- Container Registry

Profiles

```
# oc exec glusterpod  -  ls -l /var/lib/glusterd/groups

-rw-r--r--. 1 root root 227 Nov 14 09:45 db-workload

-rw-r--r--. 1 root root 516 Nov 14 09:45 gluster-block

-rw-r--r--. 1 root root 197 Nov 14 09:45 metadata-cache

-rw-r--r--. 1 root root 158 Nov 14 09:45 nl-cache

-rw-r--r--. 1 root root 406 Nov 14 09:45 virt
```

Demo

- Demo of Installation
- Demo of Applications using Glusterfs on OCP

Questions?

Thank you!

@elvirkuric

ekuric@redhat.com