openstack.

November 2018

Provide TurnKey container clusters on OpenStack

Spyros Trigazis @strigazi, Feilong Wang @feilongwang

Who we are

- Spyros Trigazis, @strigazi on Freenode & Twitter
  Magnum PTL for Queens, Rocky and Stein
  Computing Engineer at CERN

- Feilong Wang, @feilongwang on Twitter
  Core contributor of Magnum
  Head of R&D at Catalyst Cloud

openstack.

OpenStack Magnum

What is Magnum?

- OpenStack API service for creation of container clusters
- Single-tenant clusters
- Credential management
- OpenStack integration, cloud provider
- Lifecycle operations
- Kubernetes, Docker Swarm, Mesos, DC/OS

MAGNUM

an OpenStack Community Project

Magnum Terminology - *Cluster Template*

- Set of parameters describing a cluster (base for cluster creation)

```
+-----------------------+------------------------------------------------+
| Field                 | Value                                          |
+-----------------------+------------------------------------------------+
| insecure_registry     | -                                              |
| labels                | {u'kube_dashboard_enabled': u'false',          |
|                       | u'prometheus_monitoring': u'true',             |
|                       | u'kube_tag': u'v1.11.2-1',                     |
|                       | u'flannel_backend': u'vxlan'}                  |
| updated_at            | -                                              |
| floating_ip_enabled   | False                                          |
| fixed_subnet          | -                                              |
| master_flavor_id      | m2.medium                                      |
| uuid                  | afee31b7-6f35-42d3-8a21-9328edd5acf3           |
| no_proxy              | -                                              |
| https_proxy           | -                                              |
| tls_disabled          | False                                          |
| keypair_id            | -                                              |
| public                | True                                           |
| http_proxy            | -                                              |
| docker_volume_size    | -                                              |
| server_type           | vm                                             |
| external_network_id   | -                                              |
| cluster_distro        | fedora-atomic                                  |
| image_id              | 55e22657-74e5-46d9-ba28-47980986b42c           |
| volume_driver         | -                                              |
| registry_enabled      | False                                          |
| docker_storage_driver | overlay                                        |
| apiserver_port        | -                                              |
| name                  | kubernetes-alpha                               |
| created_at            | 2018-11-91T10:47:17+00:00                      |
| network_driver        | flannel                                        |
| fixed_network         | -                                              |
| coe                   | kubernetes                                     |
| flavor_id             | m2.medium                                      |
| master_lb_enabled     | False                                          |
| dns_nameserver        | 8.8.8.8                                        |
+-----------------------+------------------------------------------------+
```

Magnum Terminology - *Cluster*

- Configurable number of *master nodes*
- Configurable number of *worker nodes*
- Deployed as Heat Stacks
- A trustee user and a trust
- A Certificate Authority
  - Stored in Barbican or Magnum DB
- 3 cluster orchestrator engines
  - Kubernetes, Swarm, Mesos / DC/OS
- Multiple OS options
  - Fedora Atomic, CoreOS, Ubuntu, Centos
- VM or Baremetal
- Cluster scaling up/down

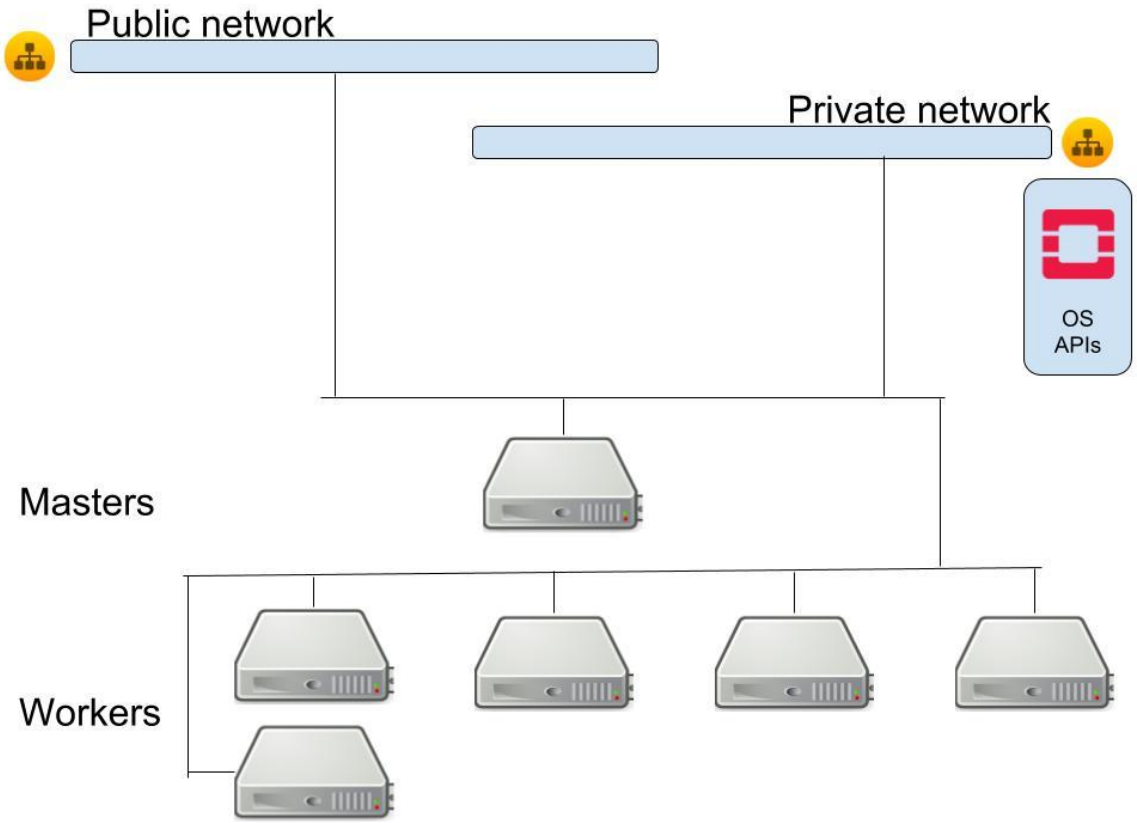
Field	Value
status	CREATE_COMPLETE
cluster_template_id	27d0fef7-3a03-4a83-ae27-6c219a84e589
node_addresses	[u'yyy.yyy.yyy.yyy']
uuid	89f79322-b574-4ea5-8169-606888d38b6f
stack_id	7cbca34c-afe3-43f6-9443-d2cfc1232996
status_reason	Stack CREATE completed successfully
created_at	2018-04-30T14:08:26+00:00
updated_at	2018-04-30T14:19:46+00:00
coe_version	v1.9.3
labels	{u'kube_tag': u'v1.10.1’}
faults	
keypair	strigazi-lxplus
api_address	https://xxx.xxx.xxx.xxx:6443
master_addresses	[u'xxx.xxx.xxx.xxx']
create_timeout	60
node_count	1
discovery_url	https://discovery.etcd.io/bc41b65fe11669d
master_count	1
container_version	1.12.6
name	strigazi-kube
master_flavor_id	m2.medium
flavor_id	m2.medium

Magnum existing features

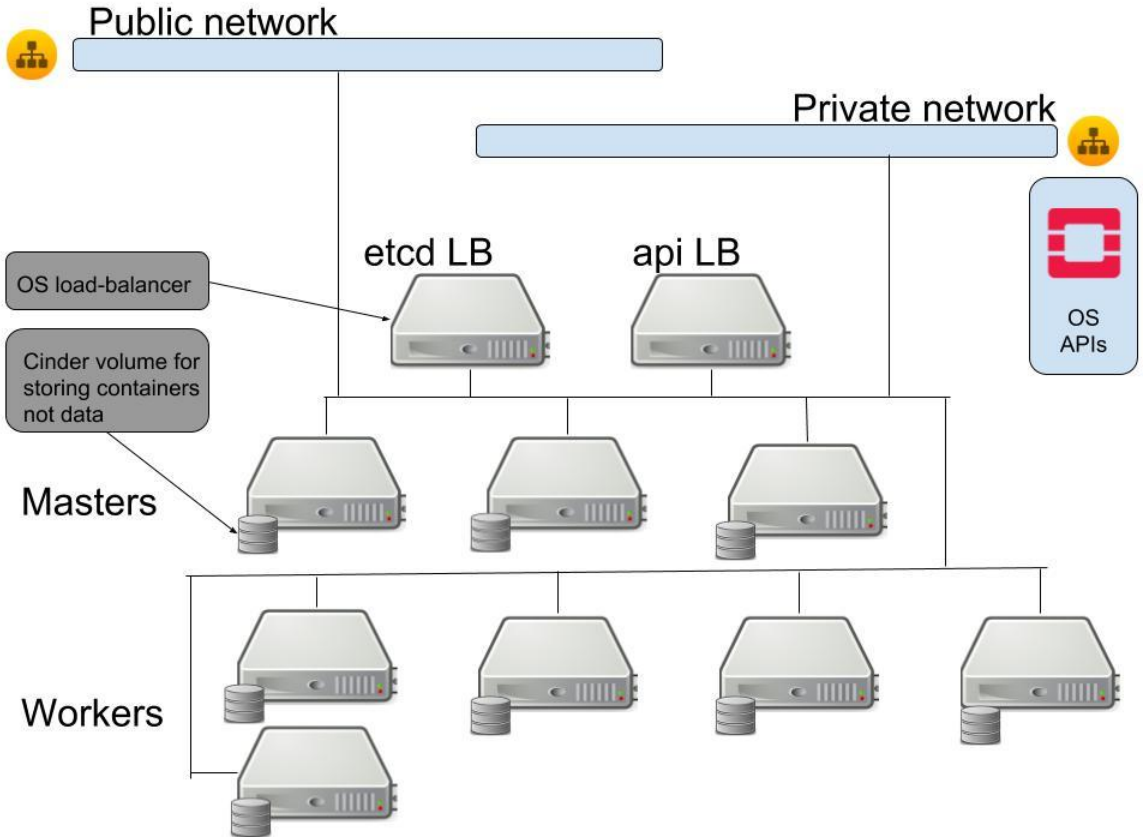
- Per cluster certificate authority
  - Each COE API is TLS-protected
    - Docker daemon
    - Kubernetes apiserver
- Scale up or down
- Load balancer (Octavia) on front of multi-master COE APIs for HA
- Simplified cluster creation:
  - Master and node flavor
  - Docker volume size
  - Labels
- Cluster availability zone selection

```
$ openstack coe cluster create --cluster-template swarm-mode-ha \
                               --flavor m2.medium \
                               --master-flavor m2.large \
                               --master-count 3 \
                               --node-count 32 \
                               --labels availability-zone=cern-geneva-a \
                               my-swarm-cluster
Request to create cluster ad418271-5232-466b-a4db-768a7ecae526 accepted
```

Default 5-node cluster

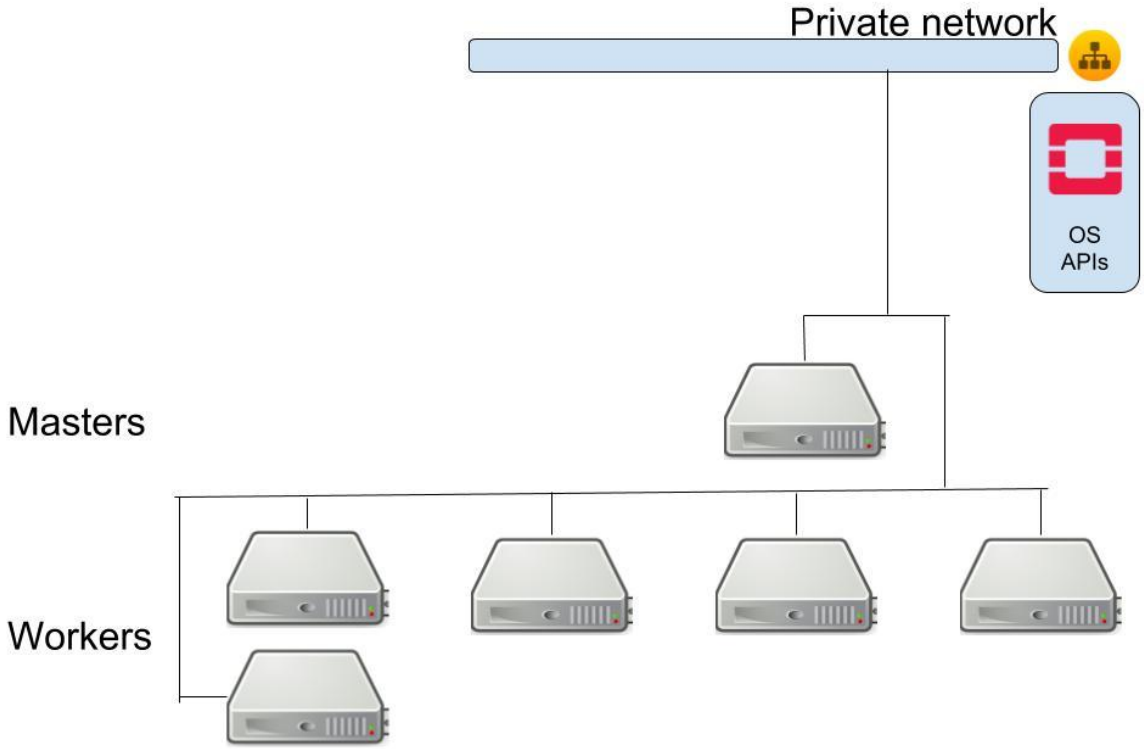


Full feature cluster

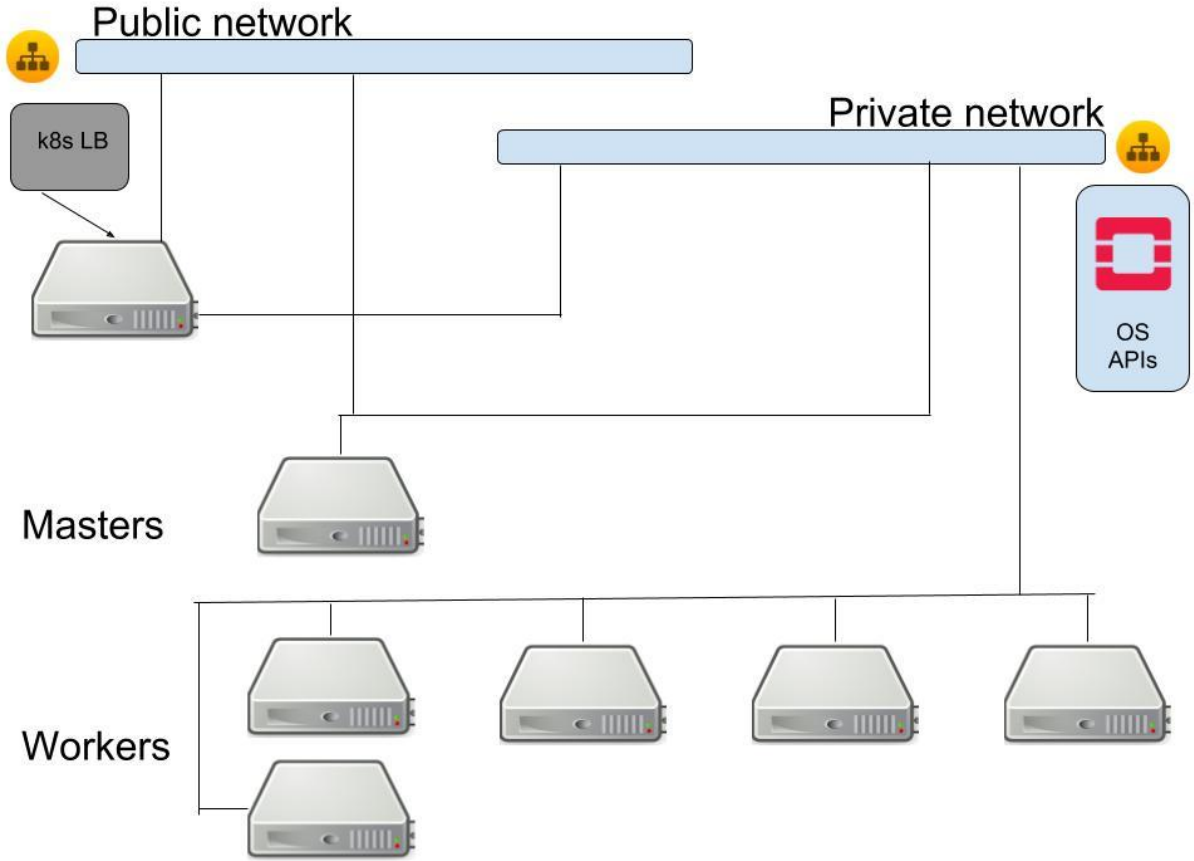


openstack.

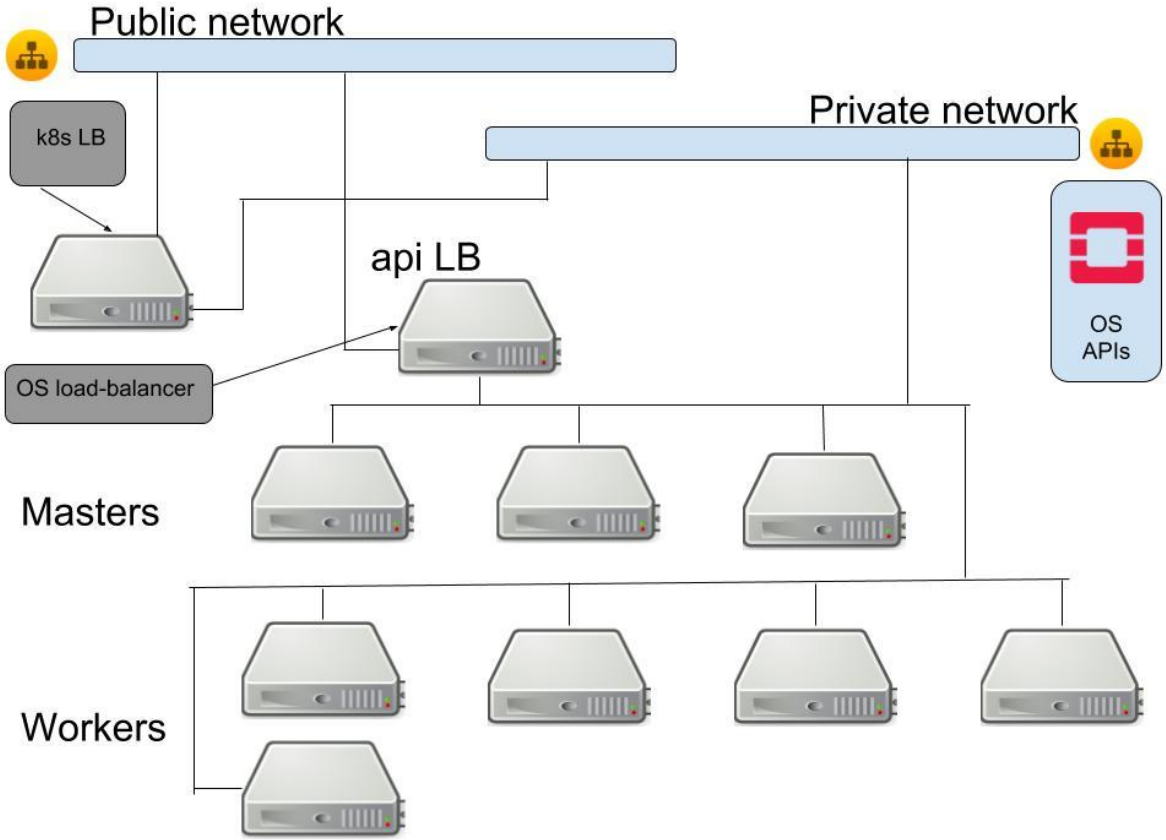
Minimal isolated cluster



Optimal single master cluster



Optimal multi master cluster



Magnum Kubernetes Features

- Calico as a network driver
- CoreDNS pod autoscaler
- Role Based Access Control - RBAC
- Kubernetes dashboard
- Monitoring stack, heapster, influxDB and grafana
- Traefik ingress controller
- Support for versions v1.9.x (queens), 1.11.x (rocky) 1.12.x (not default)

openstack.

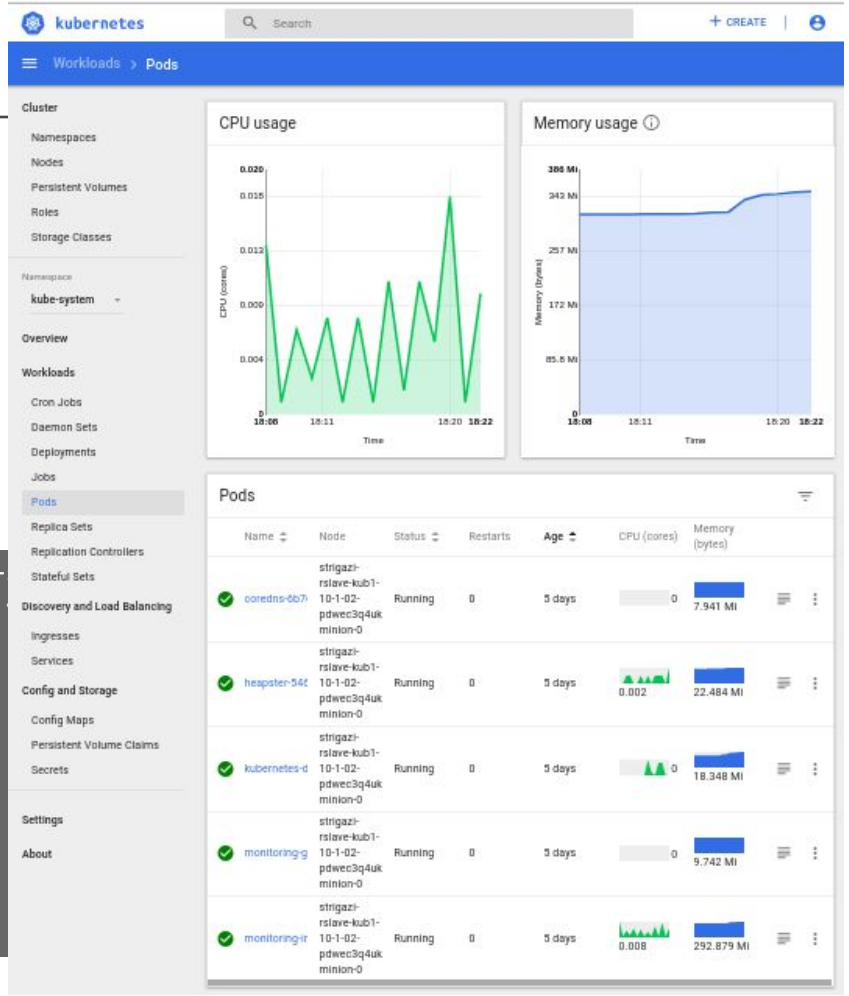
Usage

- https://docs.openstack.org/magnum/latest/user/
- Operators: manage cluster templates
- End user: create clusters, custom templates

```
$ openstack coe cluster create --cluster-template kubernetes --
Request to create cluster ad418271-5232-466b-a4db-768a7ecae526
$ ...
$ $(openstack coe cluster config kubernetes)

$ kubectl get componentstatuses
NAME                 STATUS    MESSAGE              ERROR
etcd-0               Healthy   {"health": "true"}
scheduler            Healthy   ok
controller-manager   Healthy   ok

$ kubectl proxy
Starting to serve on 127.0.0.1:8001
```



Goal/Work for Stein

- Rolling Upgrades
- Auto healing
- Node groups
- K8s-keystone auth integration
- Prometheus Operator
- FEK (Fluentd, Elasticsearch and Kibana) support
- Heat-container-agent on worker nodes
- More strict security rules for worker nodes
- Self-hosted flannel
- Deploy Tiller
- Release k8s docker images in CI

Catalyst Cloud experiences

- Don’t use overlay + docker_volume_size at least from v1.11.x
- Heat-container-agent’s multi regions bug
- v1.11.x missing IPs bug
- Build your own k8s images?

CERN Cloud experiences

- spectre/meltdown and L1TF reboots campaigns
  - Revealed network configuration issues
- Cloud Provider high-load on Nova/Neutron impact
  - Followed here: https://github.com/kubernetes/kubernetes/issues/61144
- Central Health monitoring
- Scale/Configure of the heat API
  - Configure the number of db connections properly
- Control the version of kubernetes explicitly
- Use stock operating system

openstack.

Demo!

openstack.

THANKS.

Questions?