OpenStack, Big Data Processing and ML at RaaS-IS (UAIC)

RaaS-IS POC – SMIS: 124759

Cluci Marius-Iulian
(FEAA)

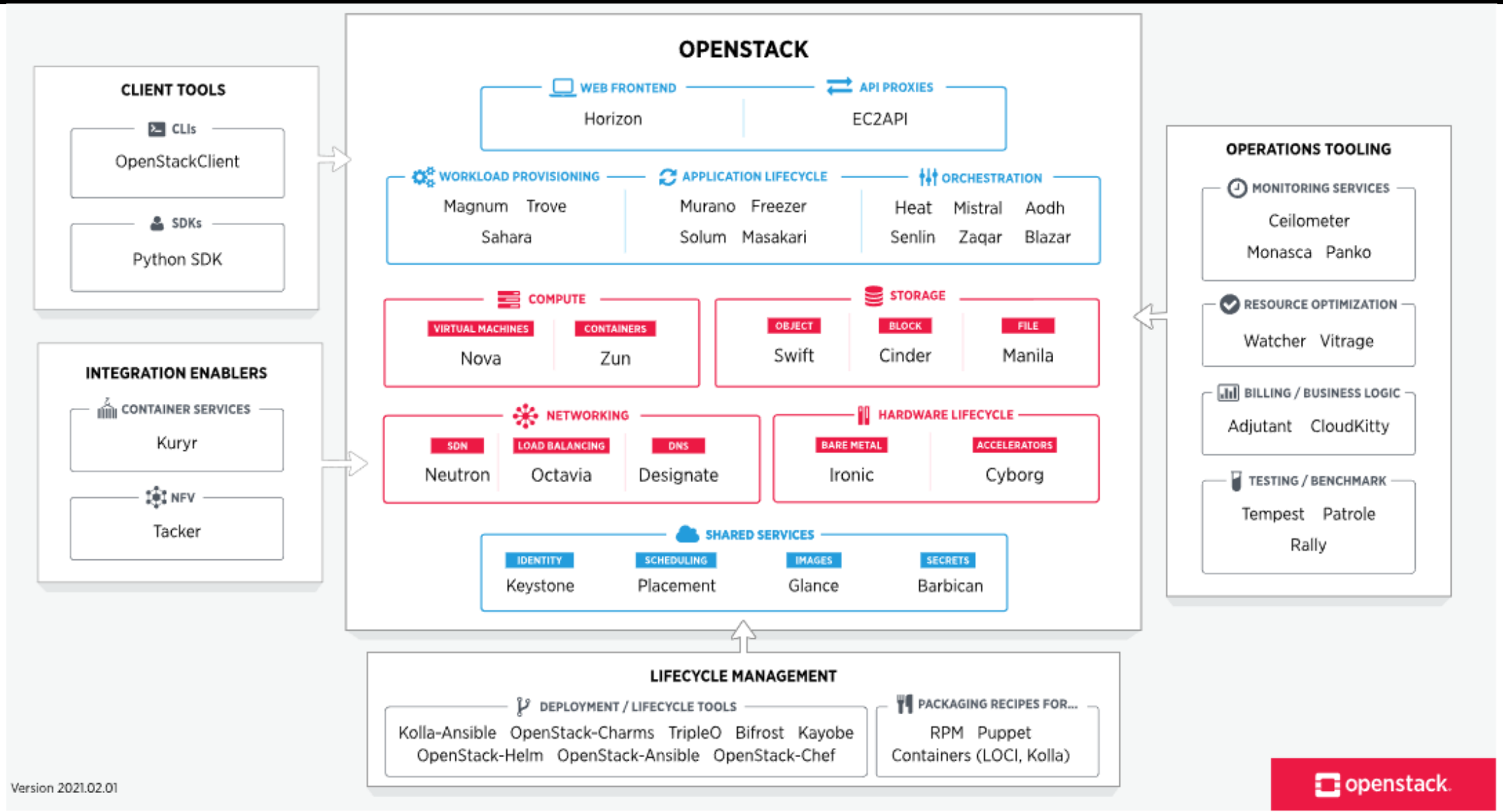
What we do

- Mainly: Provide IaaS and PaaS to researches
- Optionally: Help students train on better hardware and distributed computing

Summary

- Introduction
- Openstack services
- Install and configure MAAS
- Install and configure Juju
- Deploying OpenStack with Juju
- Openstack Operations
  - Interacting with Openstack cloud
  - Shared services
  - Openstack services
- Orchestation for Big Data | DEMO
  - OpenStack Heat
  - Kubernetes
  - Sahara

What is OpenStack



Hardware

Compute Nodes

- **16 x compute node HPE Synergy 480 Gen10**:
- 2 x Intel Xeon-Gold 6240 (2.6GHz/18-core)
- 128 GB RAM @ 2933 MHz | 2 x 300GB SAS 12G Enterprise 10K in RAID 1
- 2 x 25 Gbps Ethernet and 2 x 32 Gbps FC

Management and Controllers

- **4 x rack server HPE ProLiant DL360**
- 2 x Intel Xeon-Gold 6240 (2.6GHz/18-core)
- 128 GB RAM @ 2933 MHz | 2 x 300GB SAS 12G Enterprise 10K in RAID 1
- 2 x 25 Gbps Ethernet and 4 x 1 Gbps Ethernet

Hardware

SAN (storage area network):

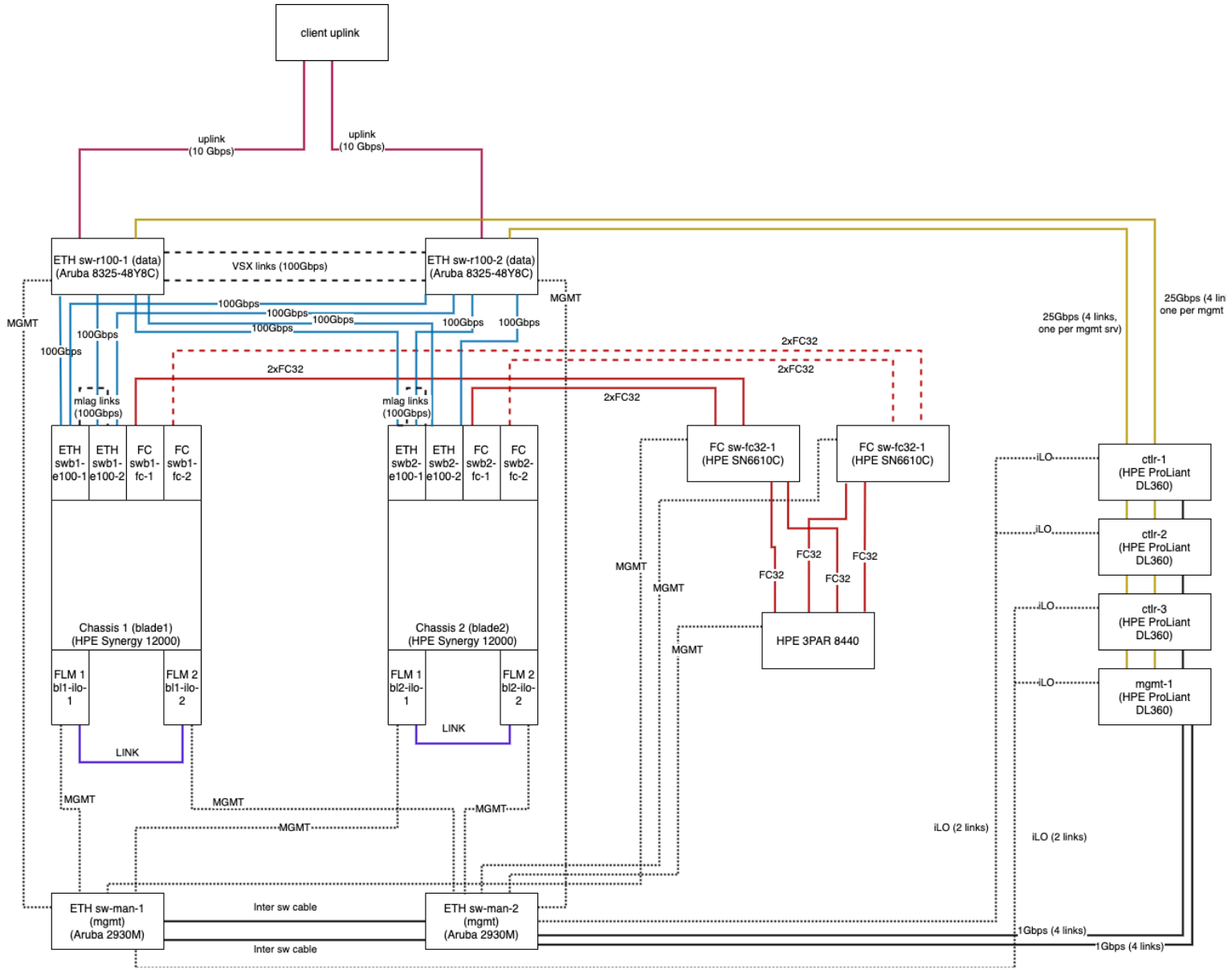
HPE 3PAR 8440

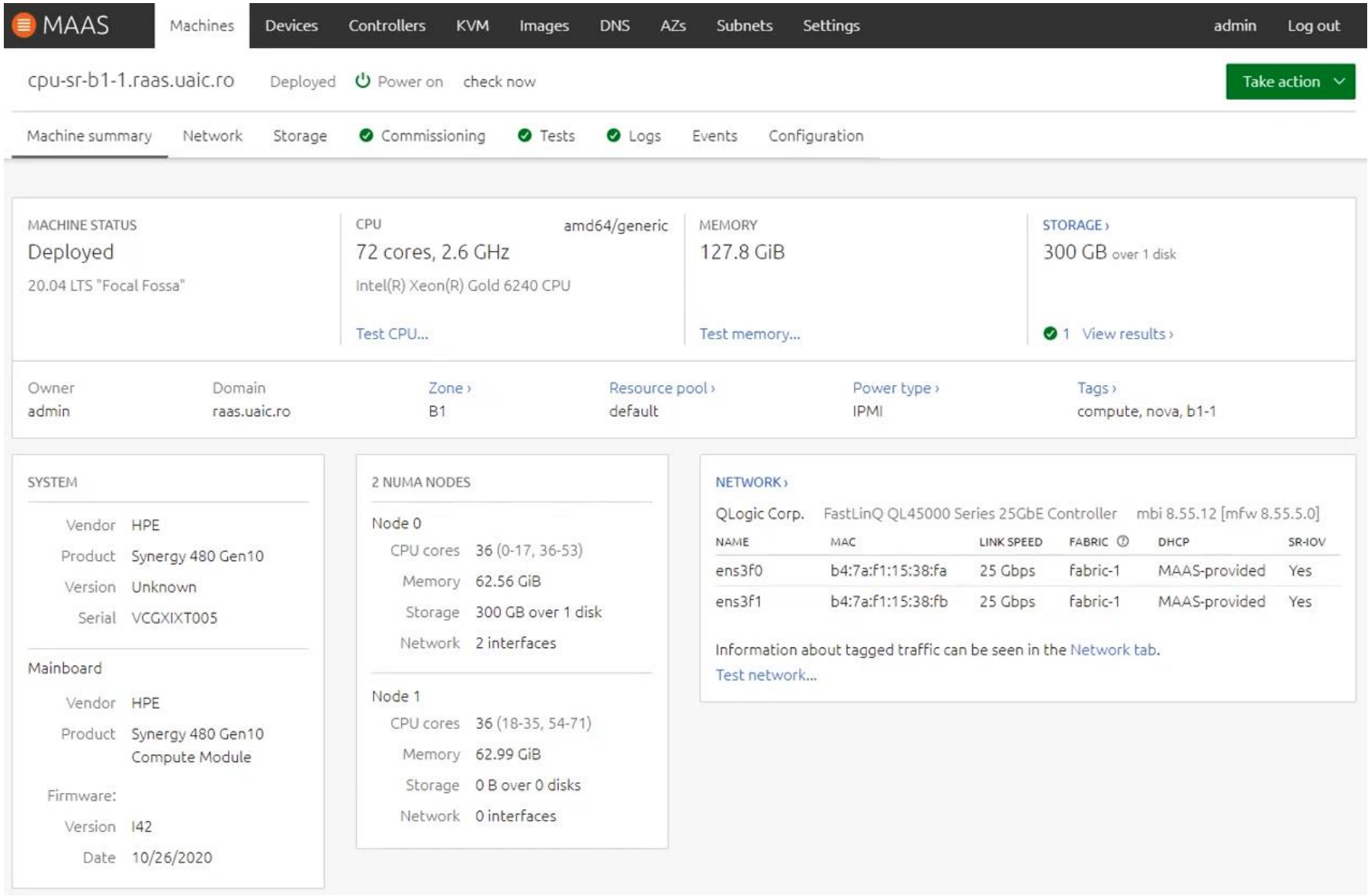
- 162K IOPS Mixed (read/write), 110K write
- 3.5 GB/s read, 2.4 GB/s write
- Capacity 760 TB RAW, 550 TB USABLE
- IO 4 x 32 Gbps FC and 4 x 16 Gbps

Network Diagram

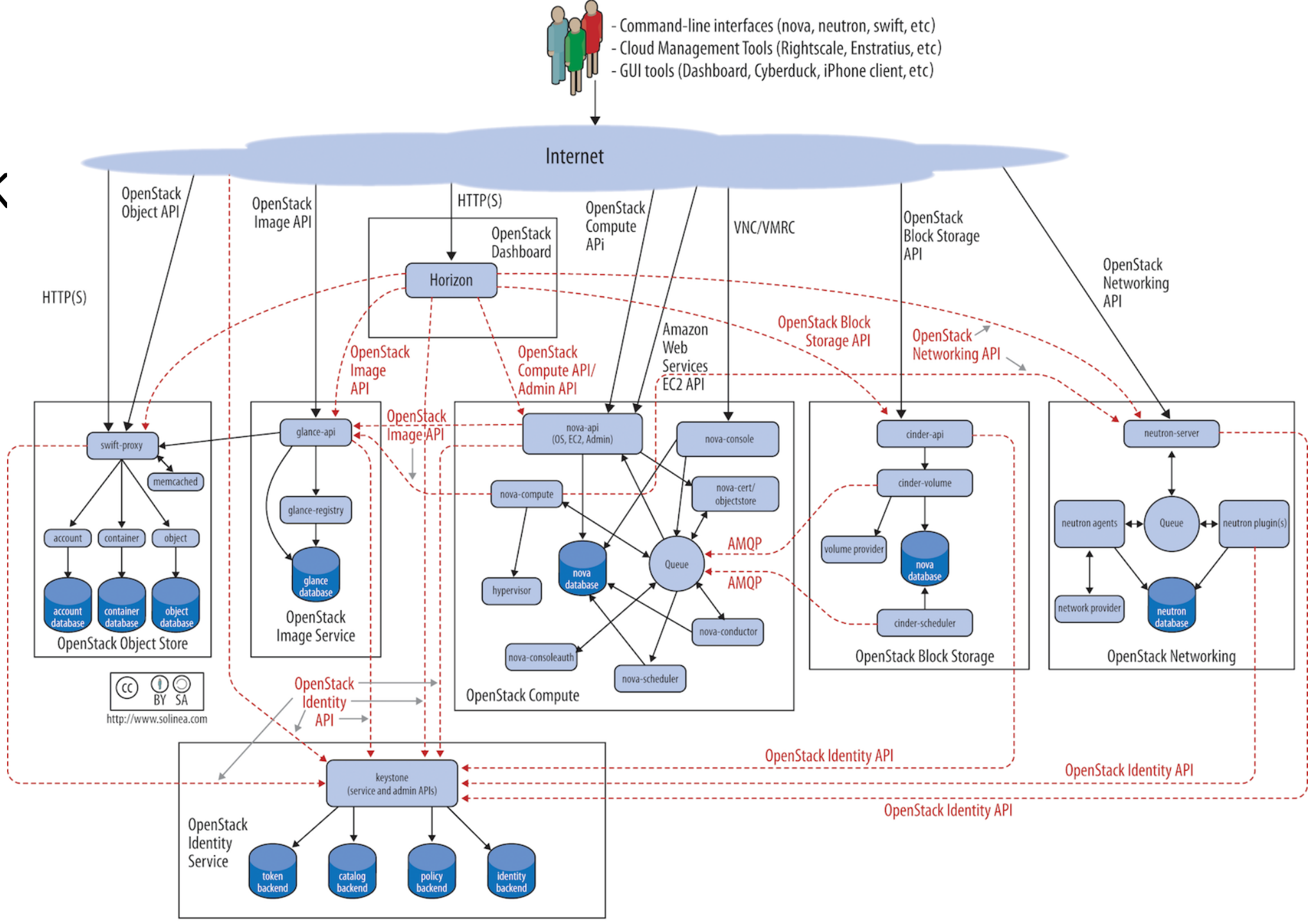
Architecture wise, OpenStack can have 4 different networks:

- Management: services talk to each other here
- Data/internal: VM to VM communication
- Exterior: exterior networking
- Storage: storage networking

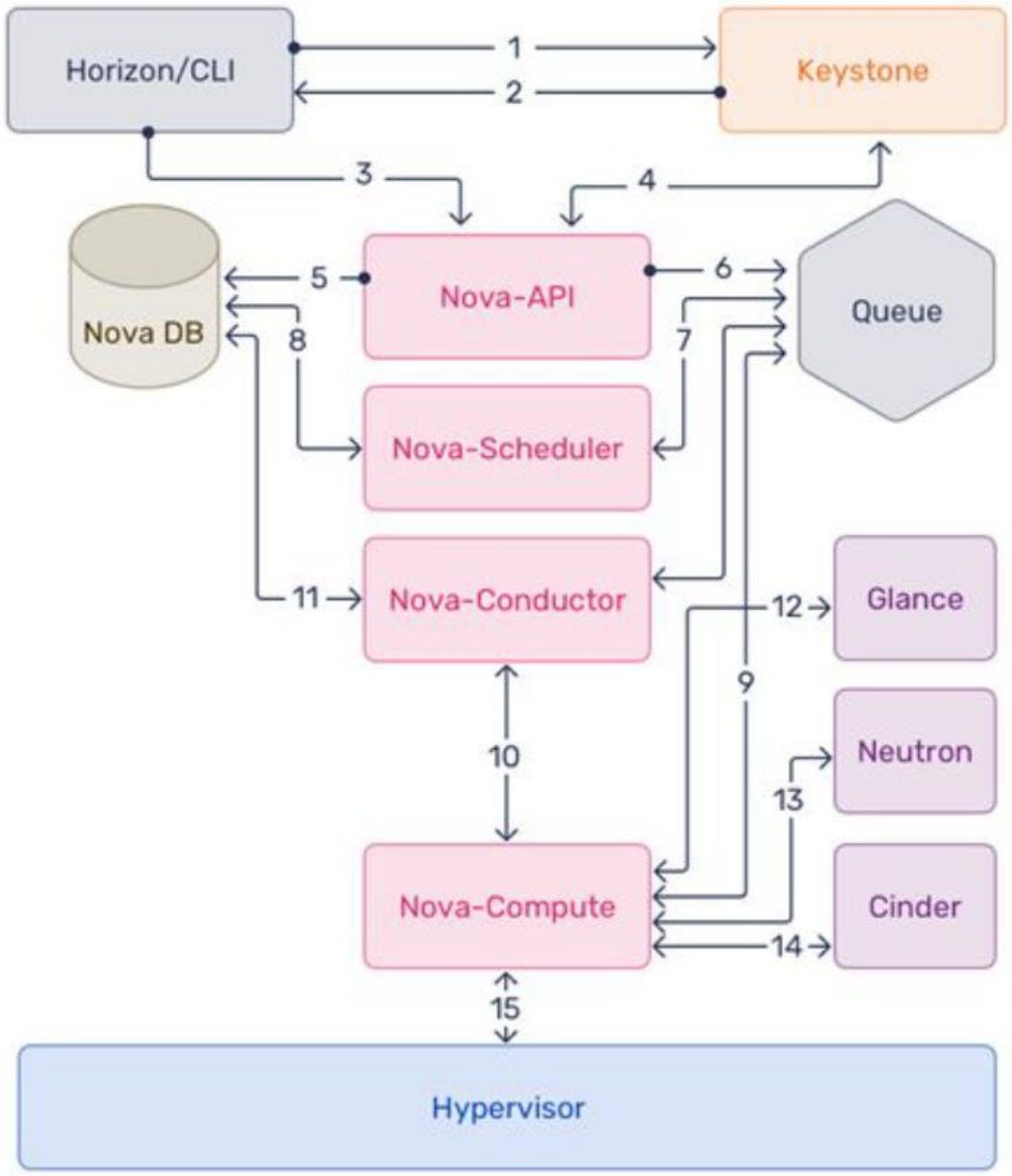



MAAS
Machines
Devices
Controllers
KVM
Images
DNS
AZs
Subnets
Settings
admin
Log out
cpu-sr-b1-1.raas.uaic.ro
Deployed
Power on
check now
Take action
Machine summary
Network
Storage
Commissioning
Tests
Logs
Events
Configuration
MACHINE STATUS
Deployed
20.04 LTS "Focal Fossa"
CPU
amd64/generic
72 cores, 2.6 GHz
Intel(R) Xeon(R) Gold 6240 CPU
Test CPU...
MEMORY
127.8 GiB
Test memory...
STORAGE ›
300 GB over 1 disk
1 View results ›
Owner
admin
Domain
raas.uaic.ro
Zone ›
B1
Resource pool ›
default
Power type ›
IPMI
Tags ›
compute, nova, b1-1
SYSTEM
Vendor HPE
Product Synergy 480 Gen10
Version Unknown
Serial VCGXIXT005
Mainboard
Vendor HPE
Product Synergy 480 Gen10 Compute Module
Firmware:
Version I42
Date 10/26/2020
2 NUMA NODES
Node 0
CPU cores 36 (0-17, 36-53)
Memory 62.56 GiB
Storage 300 GB over 1 disk
Network 2 interfaces
Node 1
CPU cores 36 (18-35, 54-71)
Memory 62.99 GiB
Storage 0 B over 0 disks
Network 0 interfaces
NETWORK ›
QLogic Corp. FastLinQ QL45000 Series 25GbE Controller mbi 8.55.12 [mfw 8.55.5.0]
NAME MAC LINK SPEED FABRIC DHCP SR-IOV
ens3f0 b4:7a:f1:15:38:fa 25 Gbps fabric-1 MAAS-provided Yes
ens3f1 b4:7a:f1:15:38:fb 25 Gbps fabric-1 MAAS-provided Yes
Information about tagged traffic can be seen in the Network tab.
Test network...

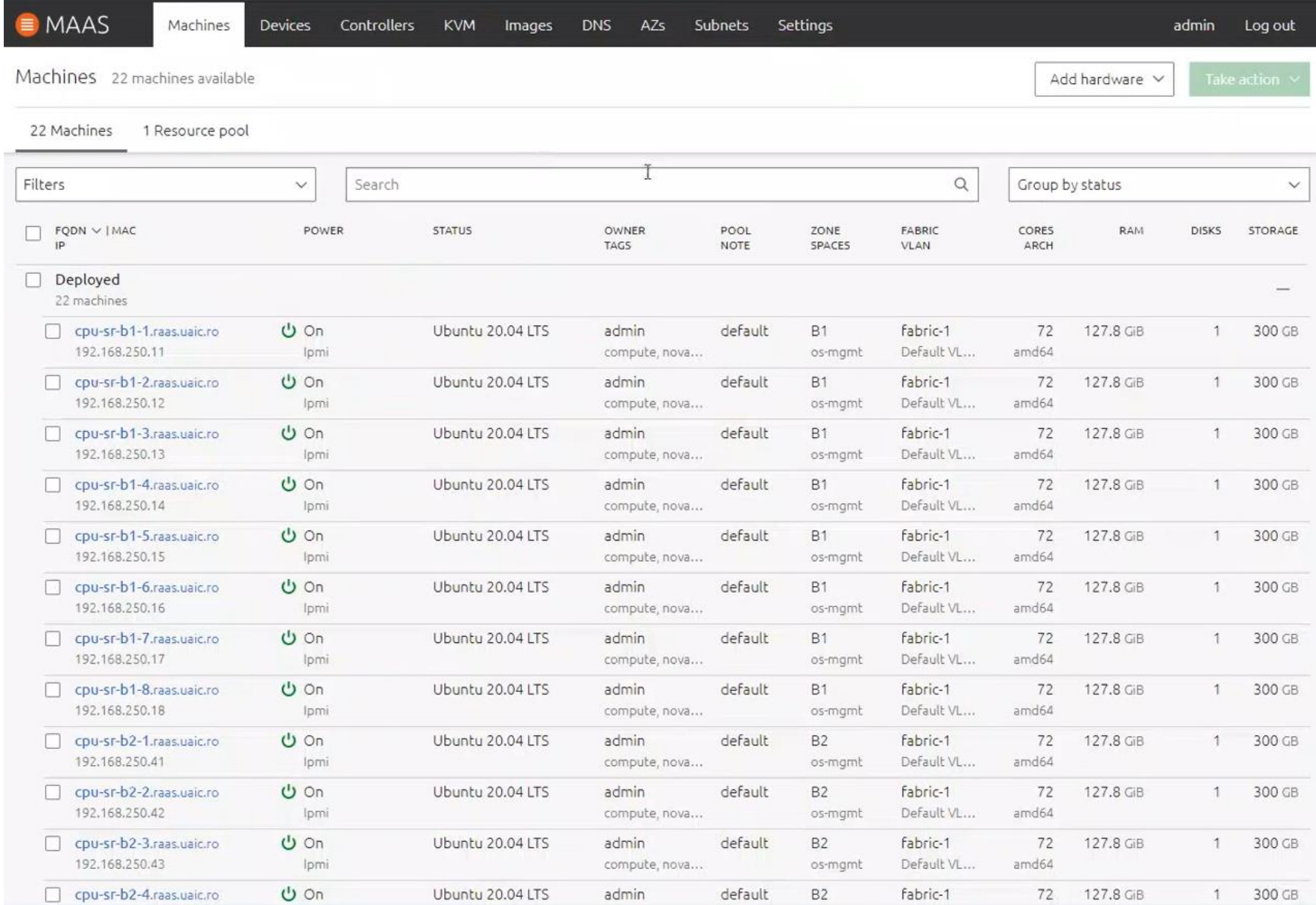
OpenStack


- Command-line interfaces (nova, neutron, swift, etc)
- Cloud Management Tools (Rightscale, Enstratius, etc)
- GUI tools (Dashboard, Cyberduck, iPhone client, etc)
Internet
OpenStack Object API
OpenStack Image API
HTTP(S)
OpenStack Dashboard
OpenStack Compute APi
VNC/VMRC
OpenStack Block Storage API
OpenStack Networking API
HTTP(S)
Horizon
OpenStack Image API
OpenStack Compute API/ Admin API
Amazon Web Services EC2 API
OpenStack Block Storage API
OpenStack Networking API
swift-proxy
memcached
account
container
object
account database
container database
object database
OpenStack Object Store
glance-api
glance-registry
glance database
OpenStack Image Service
OpenStack Image API
nova-api (OS, EC2, Admin)
nova-console
nova-compute
nova-cert/ objectstore
hypervisor
nova database
Queue
AMQP
AMQP
nova-conductor
nova-consoleauth
nova-scheduler
OpenStack Compute
cinder-api
cinder-volume
volume provider
nova database
cinder-scheduler
OpenStack Block Storage
neutron-server
neutron agents
Queue
neutron plugin(s)
network provider
neutron database
OpenStack Networking
BY SA
http://www.solinea.com
OpenStack Identity API
keystone (service and admin APIs)
OpenStack Identity Service
token backend
catalog backend
policy backend
identity backend
OpenStack Identity API
OpenStack Identity API
OpenStack Identity API

Instance creation workflow



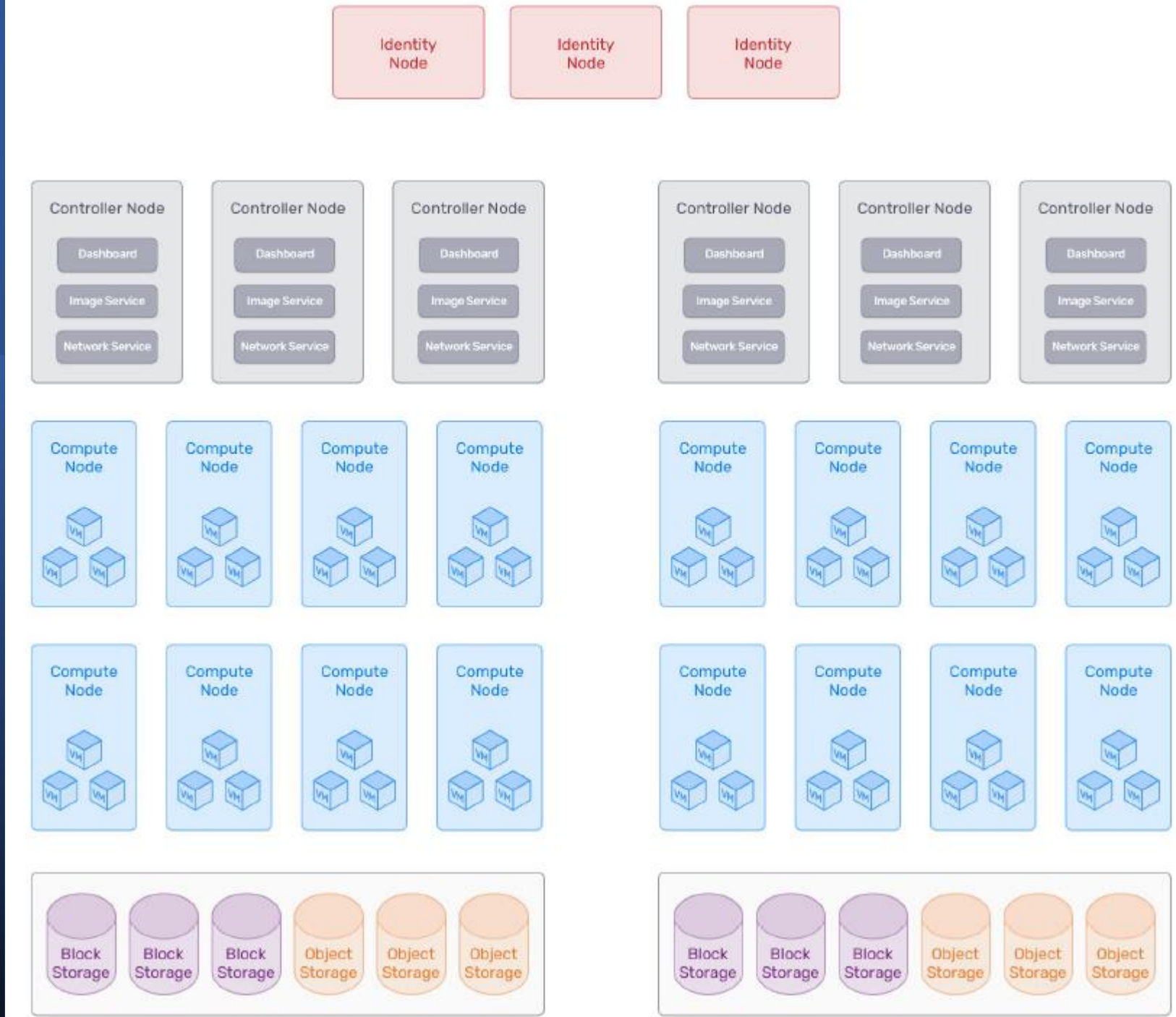
MAAS config


MAAS
Machines
Devices
Controllers
KVM
Images
DNS
AZs
Subnets
Settings
admin
Log out
Machines 22 machines available
Add hardware
Take action
22 Machines
1 Resource pool
Filters
Search
Group by status
FQDN | MAC IP
POWER
STATUS
OWNER TAGS
POOL NOTE
ZONE SPACES
FABRIC VLAN
CORES ARCH
RAM
DISKS
STORAGE
Deployed
22 machines
cpu-sr-b1-1.raas.uaic.ro 192.168.250.11 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-2.raas.uaic.ro 192.168.250.12 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-3.raas.uaic.ro 192.168.250.13 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-4.raas.uaic.ro 192.168.250.14 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-5.raas.uaic.ro 192.168.250.15 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-6.raas.uaic.ro 192.168.250.16 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-7.raas.uaic.ro 192.168.250.17 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b1-8.raas.uaic.ro 192.168.250.18 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B1 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b2-1.raas.uaic.ro 192.168.250.41 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B2 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b2-2.raas.uaic.ro 192.168.250.42 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B2 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b2-3.raas.uaic.ro 192.168.250.43 On Ipmi Ubuntu 20.04 LTS admin compute, nova... default B2 os-mgmt fabric-1 Default VL... 72 amd64 127.8 GiB 1 300 GB
cpu-sr-b2-4.raas.uaic.ro On Ubuntu 20.04 LTS admin default B2 fabric-1 72 127.8 GiB 1 300 GB

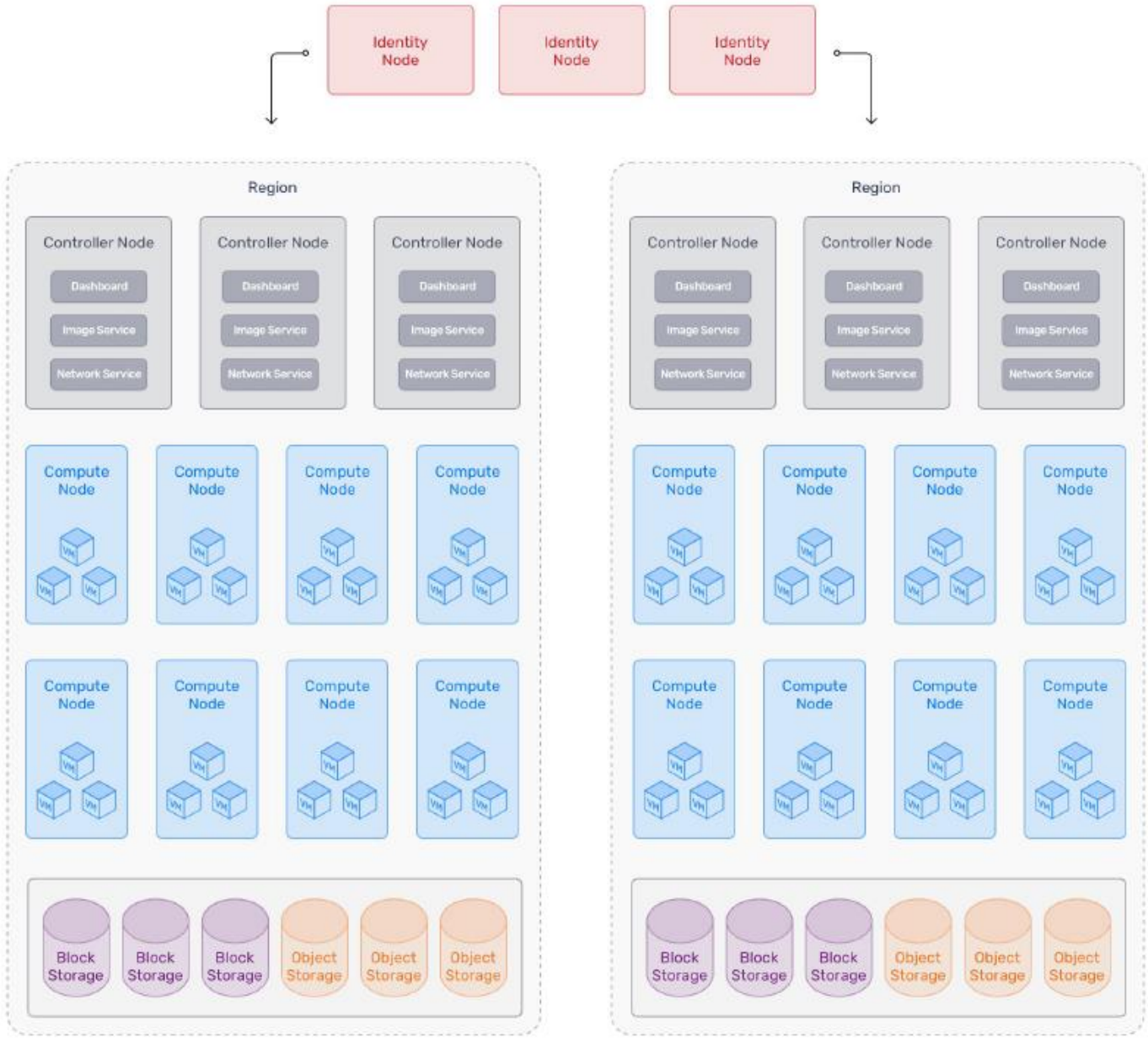
MAAS config

ctrl-sr-u1-1.raas.uaic.ro 192.168.250.61 (+1)	On Ipmi	Ubuntu 20.04 LTS	admin controller, ope...	default	CTRL 2 spaces	fabric-1 Default VL...	48 amd64	127.8 GiB	2	2.7 TB
ctrl-sr-u1-2.raas.uaic.ro 192.168.250.62 (+1)	On Ipmi	Ubuntu 20.04 LTS	admin controller, ope...	default	CTRL 2 spaces	fabric-1 Default VL...	48 amd64	127.8 GiB	2	2.7 TB
ctrl-sr-u1-3.raas.uaic.ro 192.168.250.63 (+1)	On Ipmi	Ubuntu 20.04 LTS	admin controller, ope...	default	CTRL 2 spaces	fabric-1 Default VL...	48 amd64	127.8 GiB	2	2.7 TB
juju-ctrl.maas 192.168.250.249 (PXE)	On Virsh	Ubuntu 20.04 LTS	admin virtual, juju, juj...	default	MGMT os-mgmt	fabric-1 Default VL...	4 amd64	8 GiB	1	128.8 GB
monitoring-os.raas.uaic.ro 192.168.250.152 (PXE)	On Virsh	Ubuntu 20.04 LTS	admin virtual, monito...	default	MGMT os-mgmt	fabric-1 Default VL...	8 amd64	16 GiB	1	515.4 GB
openldap.raas.uaic.ro 192.168.250.248 (PXE)	On Virsh	Ubuntu 20.04 LTS	admin virtual, ldap	default	MGMT os-mgmt	fabric-1 Default VL...	2 amd64	4 GiB	1	21.5 GB

Cloud Diagram

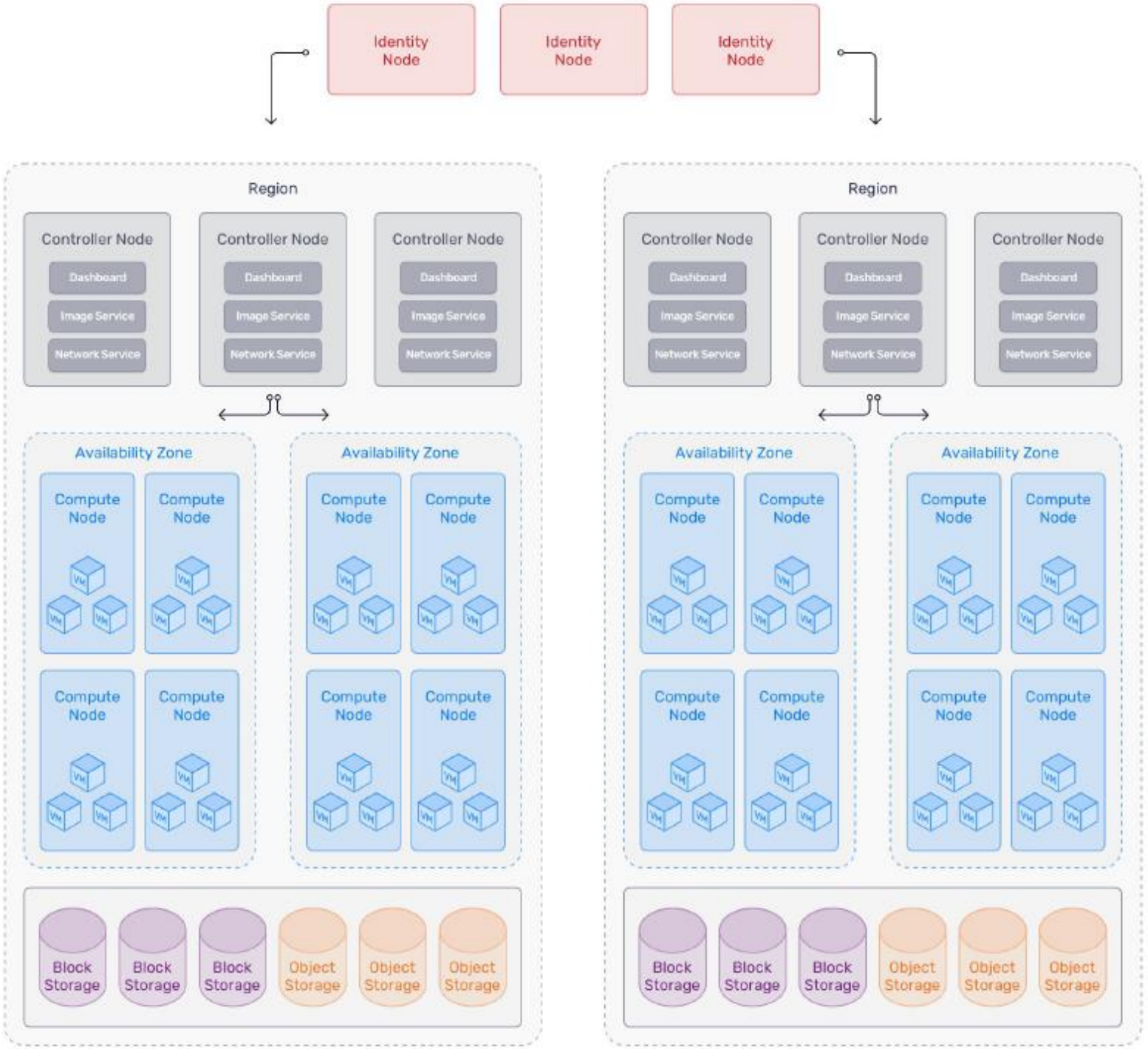


Regions

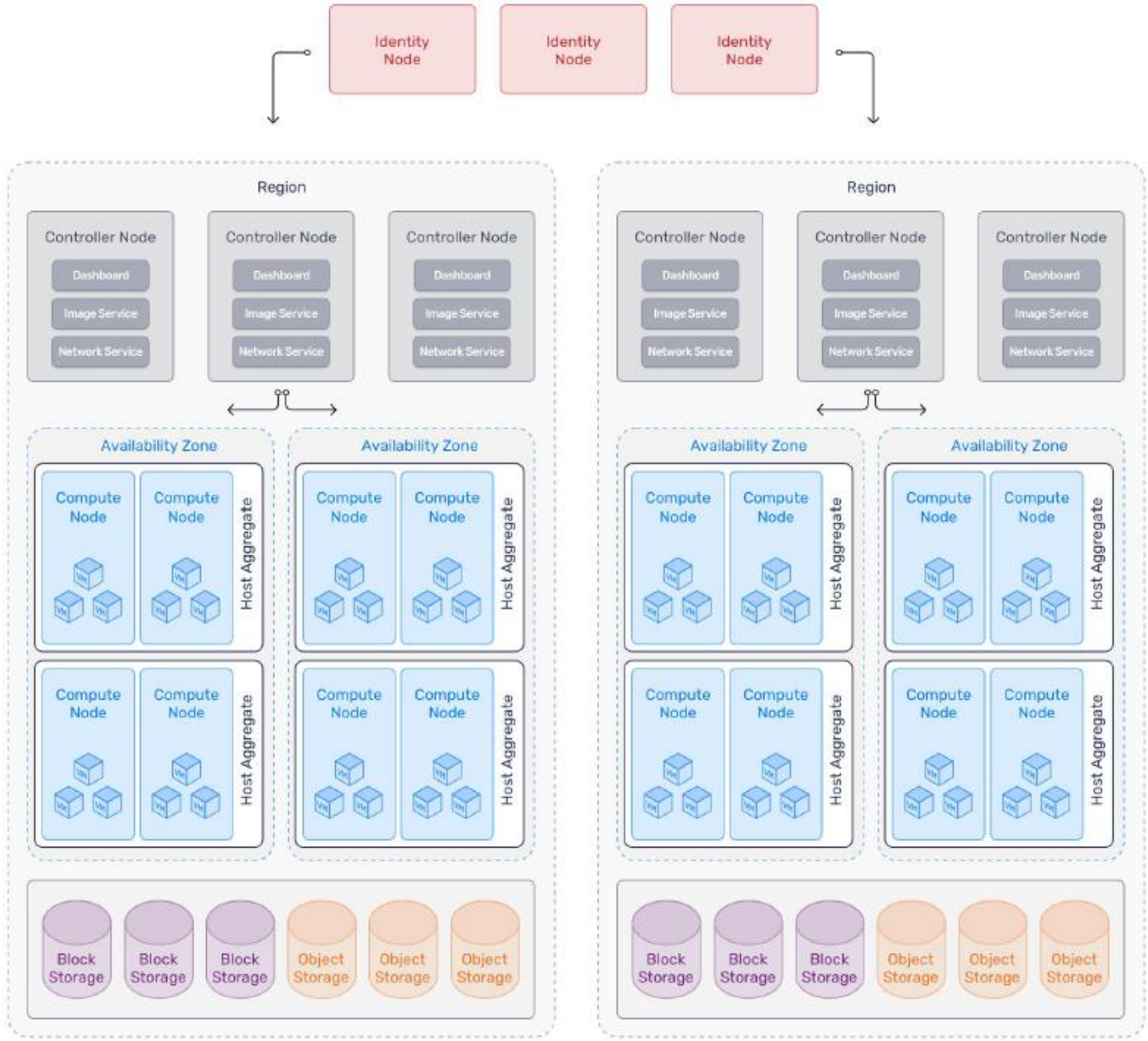

Identity Node
Identity Node
Identity Node
Region
Controller Node
Dashboard
Image Service
Network Service
Controller Node
Dashboard
Image Service
Network Service
Controller Node
Dashboard
Image Service
Network Service
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Block Storage
Block Storage
Block Storage
Object Storage
Object Storage
Object Storage
Region
Controller Node
Dashboard
Image Service
Network Service
Controller Node
Dashboard
Image Service
Network Service
Controller Node
Dashboard
Image Service
Network Service
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Compute Node
Block Storage
Block Storage
Block Storage
Object Storage
Object Storage
Object Storage

Availability Zones

Availability zones are a logical subdivision of cloud block storage, compute and network services. They provide a way for cloud operators to logically segment their compute based on arbitrary factors like location (country, datacenter, rack), network layout and/or power source.



Host Aggregates

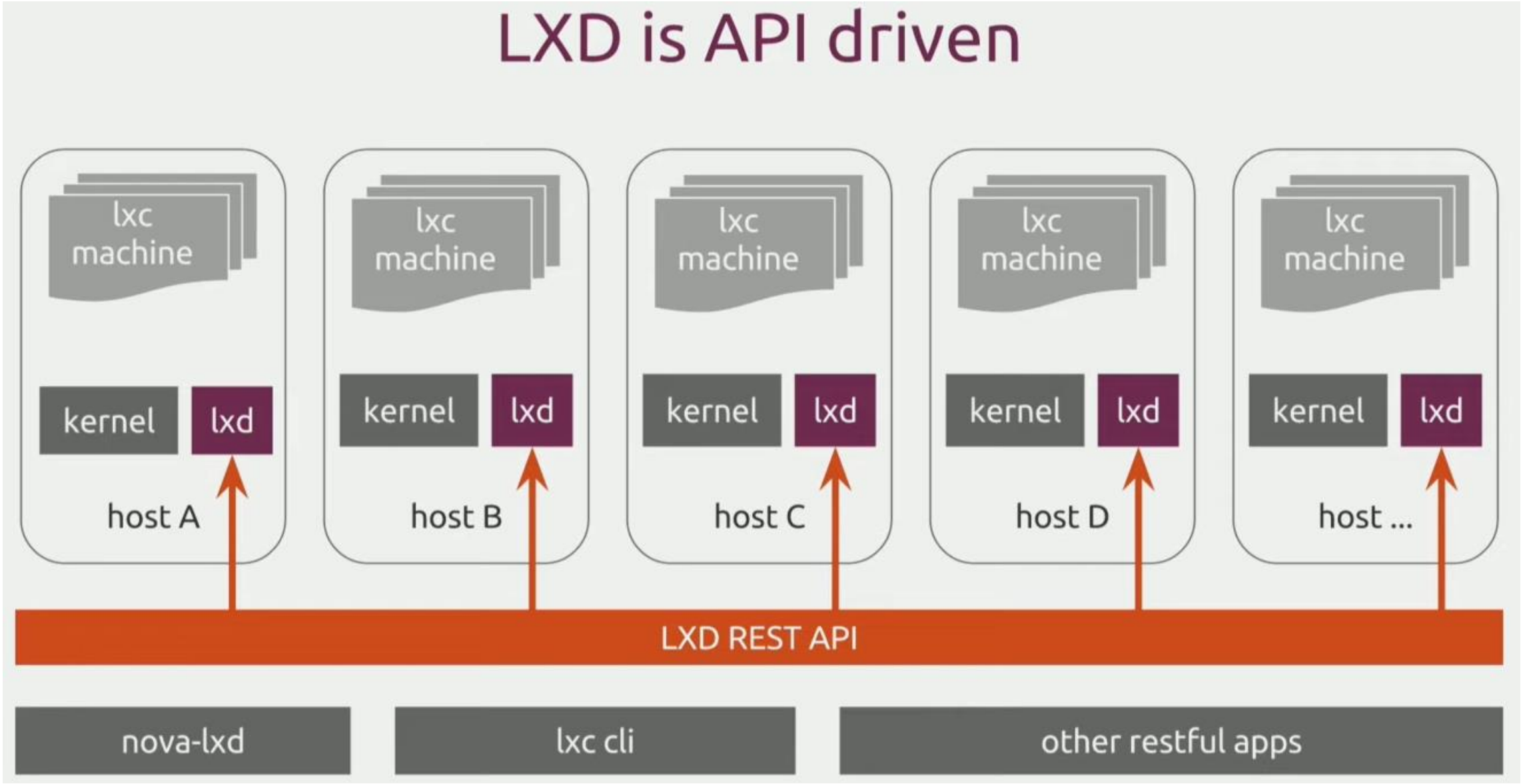


LXD vs LXC | Advantages

Application containers like Docker
host a single process on a filesystem

Machine containers from LXD
boot a full OS on their filesystem

LXD vs LXC | Advantages



RaaS-IS JUJU configuration

- Deploys a set of applications in a single step
- The bundle YAML file contains:

➔ All applications required for the bundle

➔ Any application specific configuration options

➔ All machine specifications

➔ All relations required

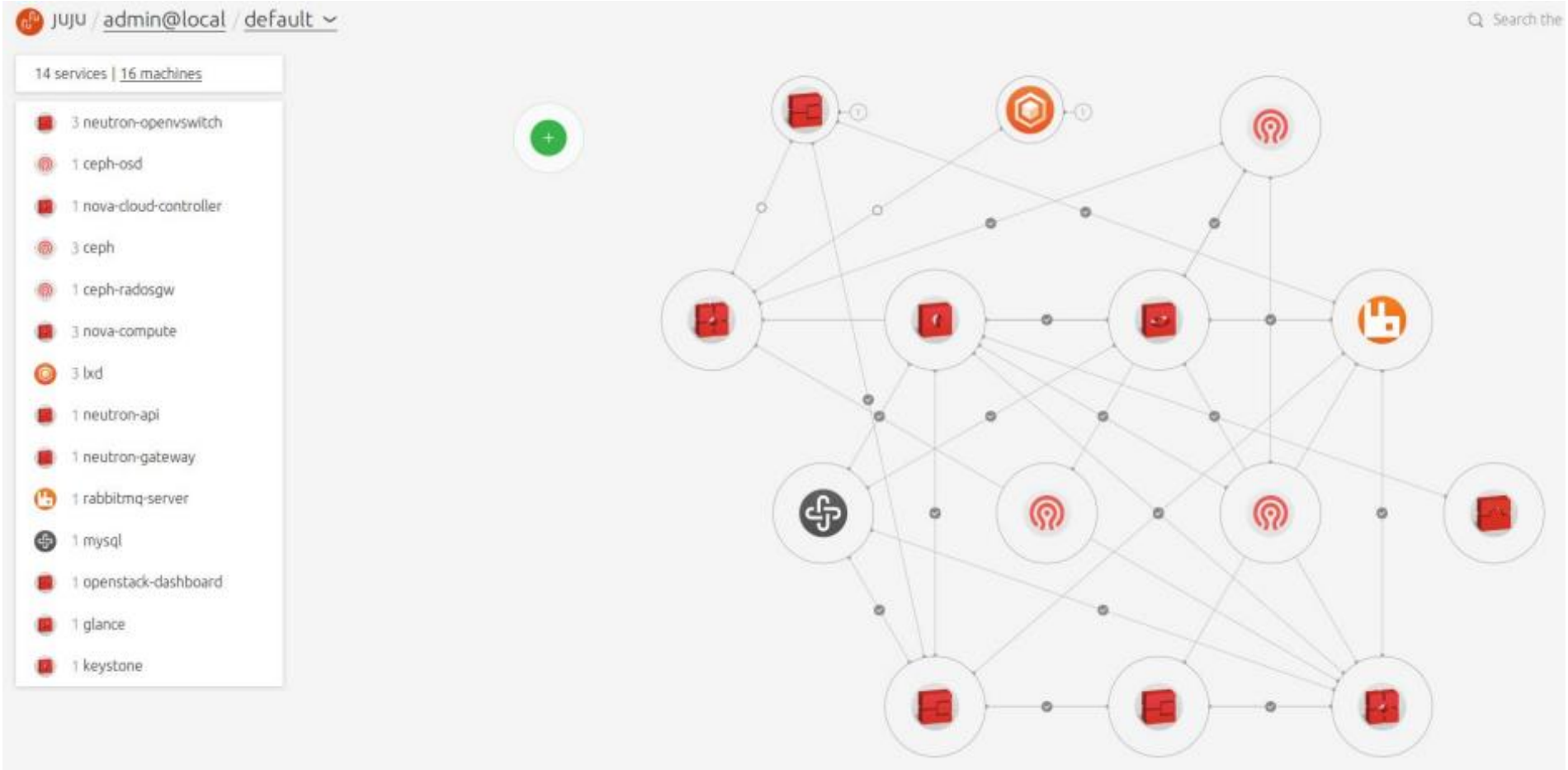
- Bundles can be:

➔ Local

➔ In the charm store

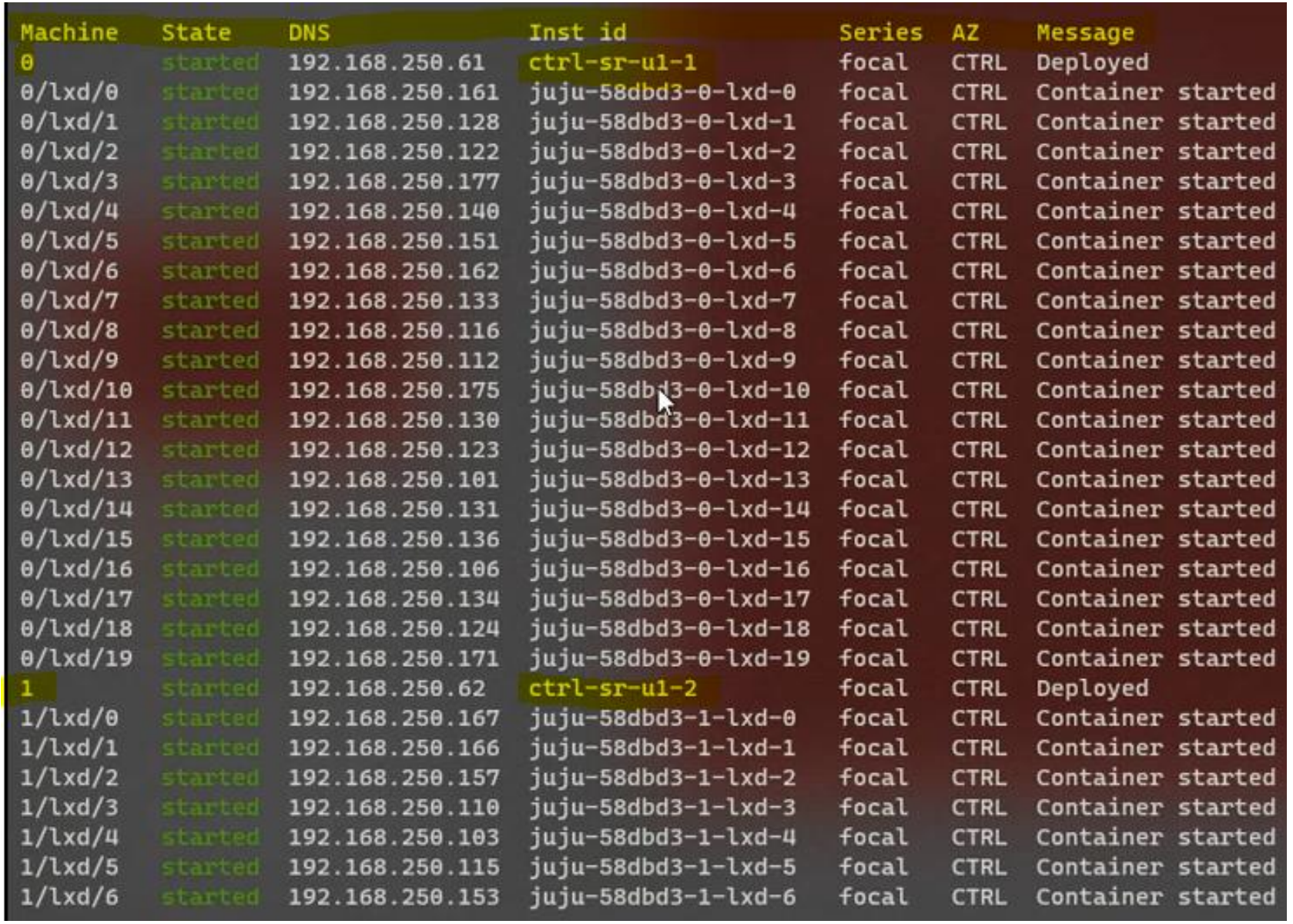
- Bundle files can contain one or more bundle definitions

Juju example



Unit	Workload	Agent	Machine	Public address	Ports	Message
barbican/0*	active	idle	0/lxd/0	192.168.250.161	9311/tcp,9312/tcp	Unit is ready
barbican-hacluster/2	active	idle		192.168.250.161		Unit is ready and clustered
barbican-mysql-router/2	active	idle		192.168.250.161		Unit is ready
barbican-vault/2	active	idle		192.168.250.161		Unit is ready
telegraf/47	active	idle		192.168.250.161	9103/tcp	Monitoring barbican/0 (source version/commit dec0633)
barbican/1	active	idle	1/lxd/0	192.168.250.167	9311/tcp,9312/tcp	Unit is ready
barbican-hacluster/0*	active	idle		192.168.250.167		Unit is ready and clustered
barbican-mysql-router/0*	active	idle		192.168.250.167		Unit is ready
barbican-vault/0*	active	idle		192.168.250.167		Unit is ready
telegraf/31	active	idle		192.168.250.167	9103/tcp	Monitoring barbican/1 (source version/commit dec0633)
barbican/2	active	idle	2/lxd/0	192.168.250.107	9311/tcp,9312/tcp	Unit is ready
barbican-hacluster/1	active	idle		192.168.250.107		Unit is ready and clustered
barbican-mysql-router/1	active	idle		192.168.250.107		Unit is ready
barbican-vault/1	active	idle		192.168.250.107		Unit is ready
telegraf/38	active	idle		192.168.250.107	9103/tcp	Monitoring barbican/2 (source version/commit dec0633)
ceph-mon/0	active	idle	0/lxd/1	192.168.250.128		Unit is ready and clustered
ceph-mon/1*	active	idle	1/lxd/1	192.168.250.166		Unit is ready and clustered
ceph-mon/2	active	idle	2/lxd/1	192.168.250.139		Unit is ready and clustered
ceph-osd/0	active	idle	3	192.168.250.11		Unit is ready (1 OSD)
ceph-osd/1	active	idle	4	192.168.250.12		Unit is ready (1 OSD)
ceph-osd/2	active	idle	5	192.168.250.13		Unit is ready (1 OSD)
ceph-osd/3	active	idle	11	192.168.250.41		Unit is ready (1 OSD)
ceph-osd/4	active	idle	12	192.168.250.42		Unit is ready (1 OSD)
ceph-osd/5*	active	idle	13	192.168.250.43		Unit is ready (1 OSD)
ceph-radosgw/0*	active	idle	0/lxd/2	192.168.250.122	80/tcp,443/tcp	Unit is ready
ceph-rgw-hacluster/1	active	idle		192.168.250.122		Unit is ready and clustered
ceph-radosgw/1	active	idle	1/lxd/2	192.168.250.157	80/tcp,443/tcp	Unit is ready

See all LXC containers with their services



Machine	State	DNS	Inst id	Series	AZ	Message
0	started	192.168.250.61	ctrl-sr-u1-1	focal	CTRL	Deployed
0/lxd/0	started	192.168.250.161	juju-58dbd3-0-lxd-0	focal	CTRL	Container started
0/lxd/1	started	192.168.250.128	juju-58dbd3-0-lxd-1	focal	CTRL	Container started
0/lxd/2	started	192.168.250.122	juju-58dbd3-0-lxd-2	focal	CTRL	Container started
0/lxd/3	started	192.168.250.177	juju-58dbd3-0-lxd-3	focal	CTRL	Container started
0/lxd/4	started	192.168.250.140	juju-58dbd3-0-lxd-4	focal	CTRL	Container started
0/lxd/5	started	192.168.250.151	juju-58dbd3-0-lxd-5	focal	CTRL	Container started
0/lxd/6	started	192.168.250.162	juju-58dbd3-0-lxd-6	focal	CTRL	Container started
0/lxd/7	started	192.168.250.133	juju-58dbd3-0-lxd-7	focal	CTRL	Container started
0/lxd/8	started	192.168.250.116	juju-58dbd3-0-lxd-8	focal	CTRL	Container started
0/lxd/9	started	192.168.250.112	juju-58dbd3-0-lxd-9	focal	CTRL	Container started
0/lxd/10	started	192.168.250.175	juju-58dbd3-0-lxd-10	focal	CTRL	Container started
0/lxd/11	started	192.168.250.130	juju-58dbd3-0-lxd-11	focal	CTRL	Container started
0/lxd/12	started	192.168.250.123	juju-58dbd3-0-lxd-12	focal	CTRL	Container started
0/lxd/13	started	192.168.250.101	juju-58dbd3-0-lxd-13	focal	CTRL	Container started
0/lxd/14	started	192.168.250.131	juju-58dbd3-0-lxd-14	focal	CTRL	Container started
0/lxd/15	started	192.168.250.136	juju-58dbd3-0-lxd-15	focal	CTRL	Container started
0/lxd/16	started	192.168.250.106	juju-58dbd3-0-lxd-16	focal	CTRL	Container started
0/lxd/17	started	192.168.250.134	juju-58dbd3-0-lxd-17	focal	CTRL	Container started
0/lxd/18	started	192.168.250.124	juju-58dbd3-0-lxd-18	focal	CTRL	Container started
0/lxd/19	started	192.168.250.171	juju-58dbd3-0-lxd-19	focal	CTRL	Container started
1	started	192.168.250.62	ctrl-sr-u1-2	focal	CTRL	Deployed
1/lxd/0	started	192.168.250.167	juju-58dbd3-1-lxd-0	focal	CTRL	Container started
1/lxd/1	started	192.168.250.166	juju-58dbd3-1-lxd-1	focal	CTRL	Container started
1/lxd/2	started	192.168.250.157	juju-58dbd3-1-lxd-2	focal	CTRL	Container started
1/lxd/3	started	192.168.250.110	juju-58dbd3-1-lxd-3	focal	CTRL	Container started
1/lxd/4	started	192.168.250.103	juju-58dbd3-1-lxd-4	focal	CTRL	Container started
1/lxd/5	started	192.168.250.115	juju-58dbd3-1-lxd-5	focal	CTRL	Container started
1/lxd/6	started	192.168.250.153	juju-58dbd3-1-lxd-6	focal	CTRL	Container started

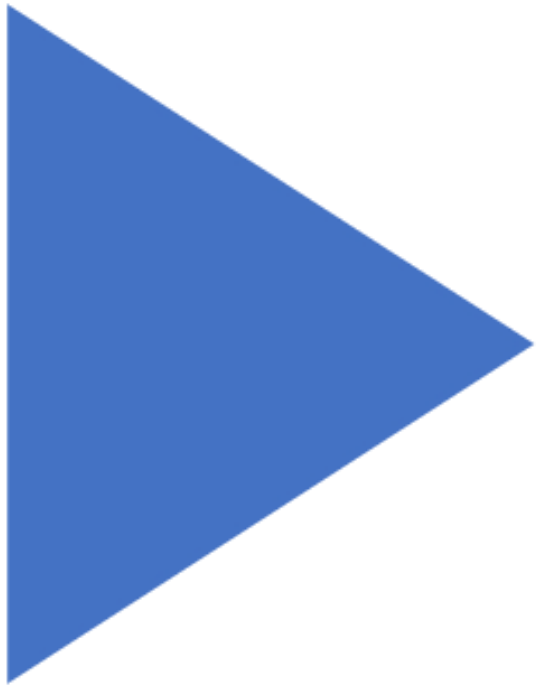
All modules communicate

Relation provider	Requirer	Interface	Type
barbican:juju-info	telegraf:juju-info	juju-info	subordinate
cinder-api:juju-info	telegraf:juju-info	juju-info	subordinate
glance:juju-info	telegraf:juju-info	juju-info	subordinate
grafana:juju-info	telegraf:juju-info	juju-info	subordinate
heat:juju-info	telegraf:juju-info	juju-info	subordinate
keystone-hacluster:ha	keystone:ha	hacluster	subordinate
keystone-hacluster:hanode	keystone-hacluster:hanode	hacluster	peer
keystone-ldap:domain-backend	keystone:domain-backend	keystone-domain-backend	subordinate
keystone-mysql-router:shared-db	keystone:shared-db	mysql-shared	subordinate
keystone:cluster	keystone:cluster	keystone-ha	peer
keystone:identity-service	barbican:identity-service	keystone	regular
keystone:identity-service	ceph-radosgw:identity-service	keystone	regular
keystone:identity-service	cinder-api:identity-service	keystone	regular
keystone:identity-service	cinder-volume:identity-service	keystone	regular
keystone:identity-service	glance:identity-service	keystone	regular
keystone:identity-service	heat:identity-service	keystone	regular
keystone:identity-service	magnum:identity-service	keystone	regular
keystone:identity-service	neutron-api:identity-service	keystone	regular
keystone:identity-service	nova-cloud-controller:identity-service	keystone	regular
keystone:identity-service	openstack-dashboard:identity-service	keystone	regular
keystone:identity-service	placement:identity-service	keystone	regular
keystone:identity-service	sahara:identity-service	keystone	regular
keystone:juju-info	telegraf:juju-info	juju-info	subordinate
magnum:juju-info	telegraf:juju-info	juju-info	subordinate
mysql-innodb-cluster:db-router	keystone-mysql-router:db-router	mysql-router	regular
mysql-innodb-cluster:juju-info	telegraf:juju-info	juju-info	subordinate
neutron-api:juju-info	telegraf:juju-info	juju-info	subordinate
nova-cloud-controller:juju-info	telegraf:juju-info	juju-info	subordinate
nova-compute:juju-info	telegraf:juju-info	juju-info	subordinate
openstack-dashboard:juju-info	telegraf:juju-info	juju-info	subordinate
placement:juju-info	telegraf:juju-info	juju-info	subordinate
rabbitmq-server:amqp	telegraf:amqp	rabbitmq	subordinate
rabbitmq-server:juju-info	telegraf:juju-info	juju-info	subordinate
sahara:juju-info	telegraf:juju-info	juju-info	subordinate
telegraf:dashboards	grafana:dashboards	grafana-dashboard	regular
telegraf:prometheus-client	prometheus:target	http	regular
vault:certificates	keystone:certificates	tls-certificates	regular

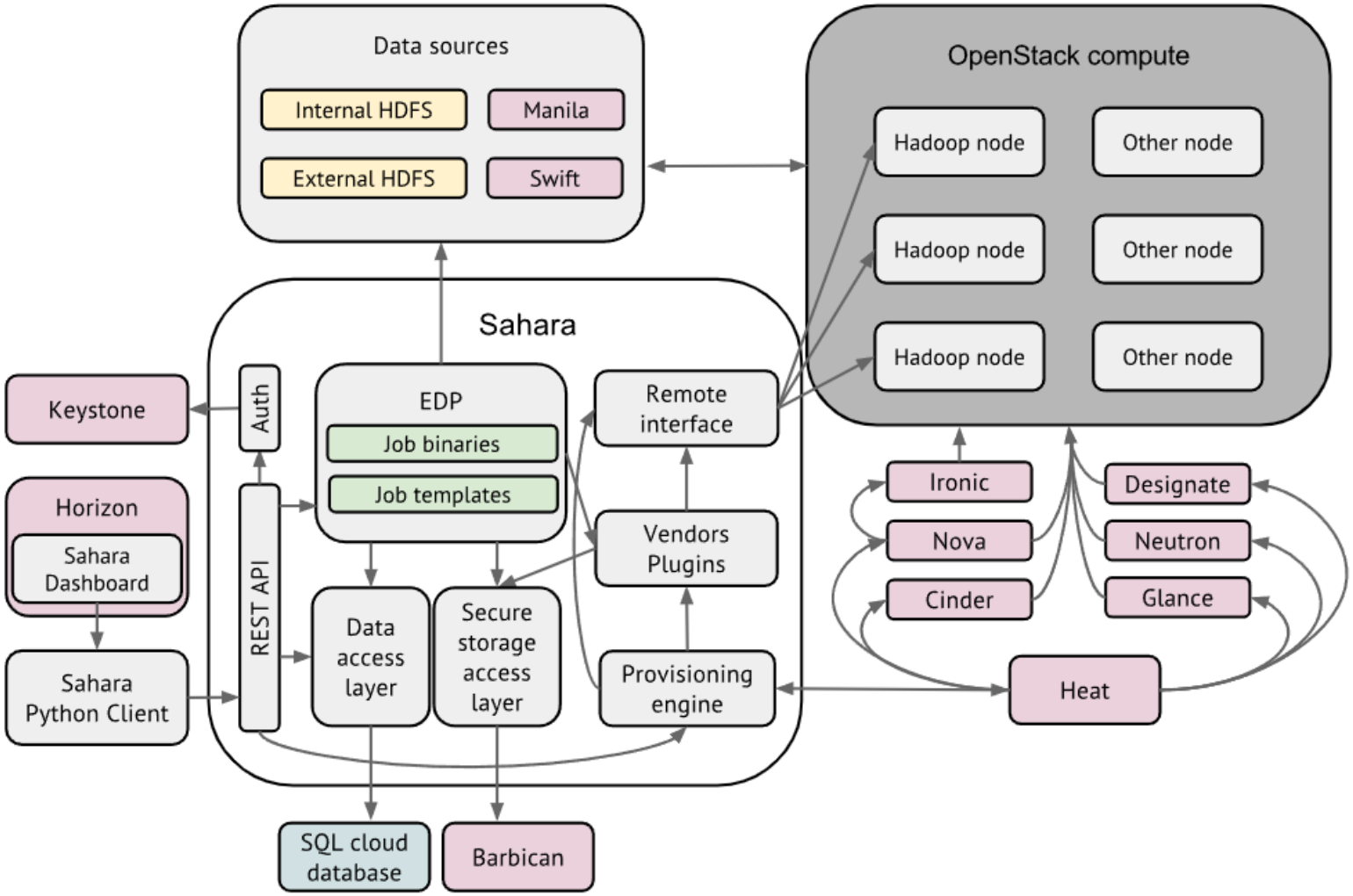
OpenStack Heat

- Orchestration tool
- Deploys resources based on templates
- Provides both an OpenStack-native REST API and a CloudFormation-compatible Query API.
- Can be used by Horizon and openstack-clients

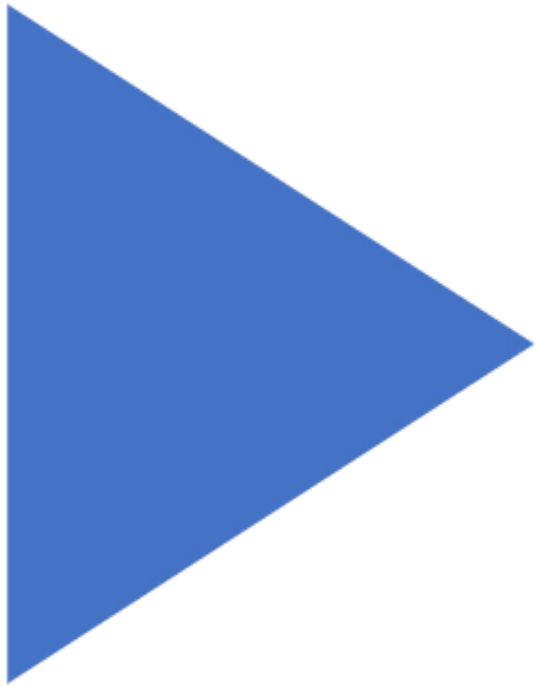
Demo



OpenStack Sahara - arhitecture



Demo



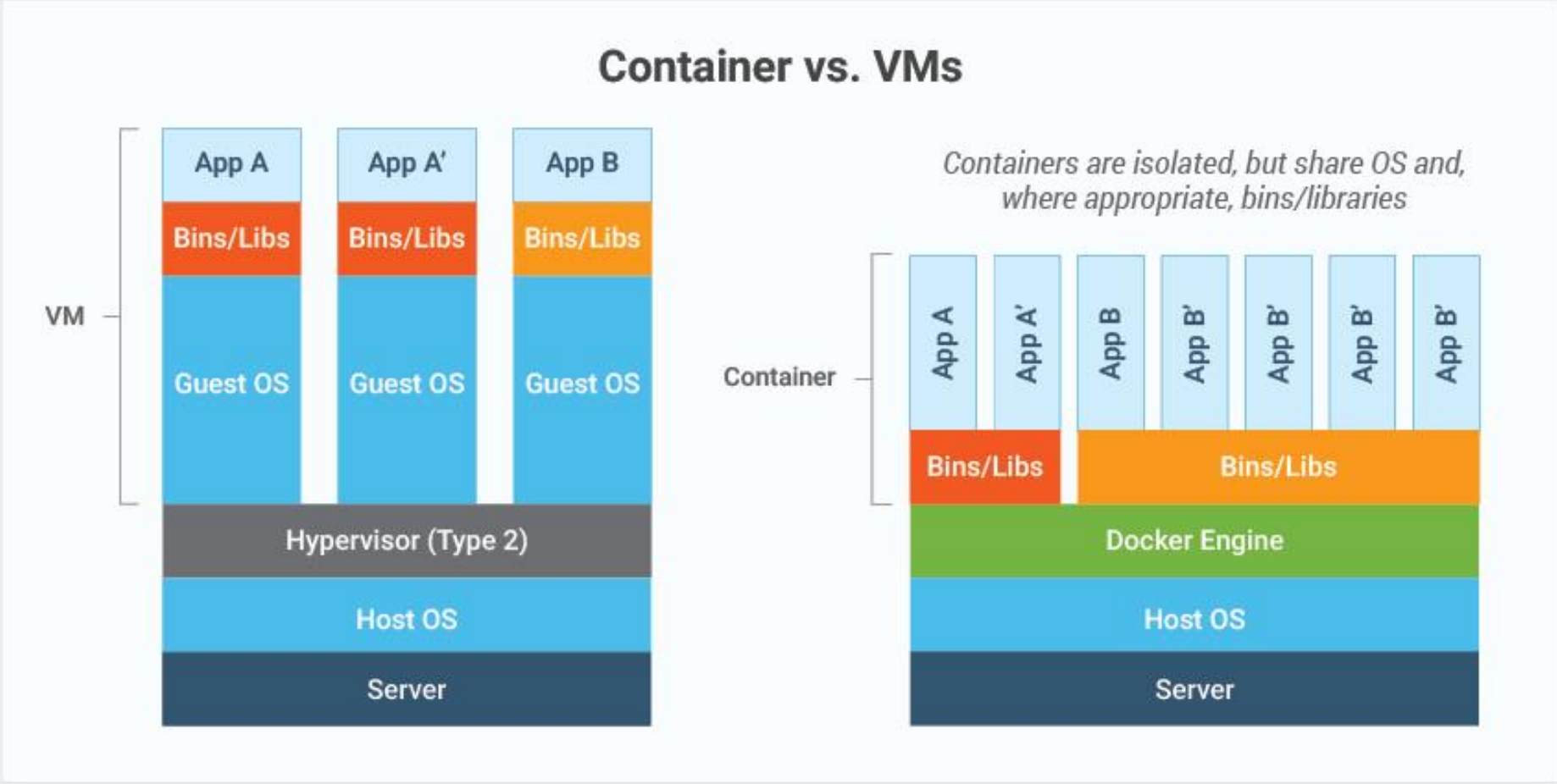
Kubernetes (using Magnum*)

- OpenStack can use *Zun* and *Magnum* for containers

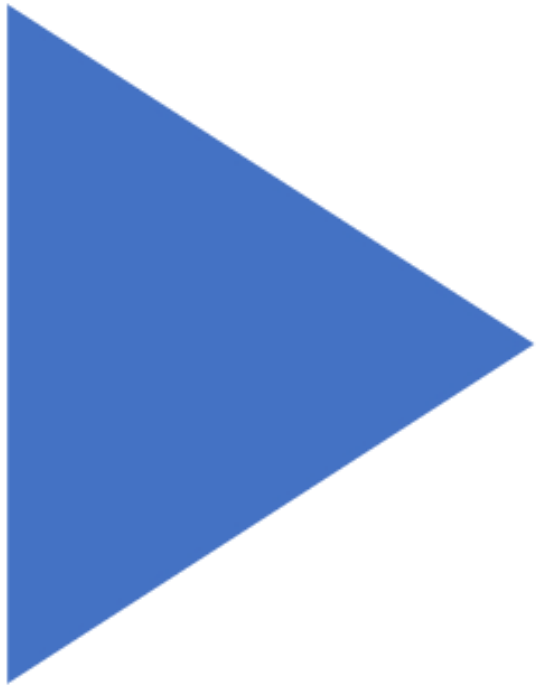
- Magnum makes container orchestration engines such as Docker Swarm, Kubernetes, and Apache Mesos available as first-class resources in OpenStack. Magnum uses Heat to orchestrate an OS image which contains Docker and Kubernetes and runs that image in either virtual machines or bare metal in a cluster configuration.
- We are going to use **Container Orchestration Engine (COE)**

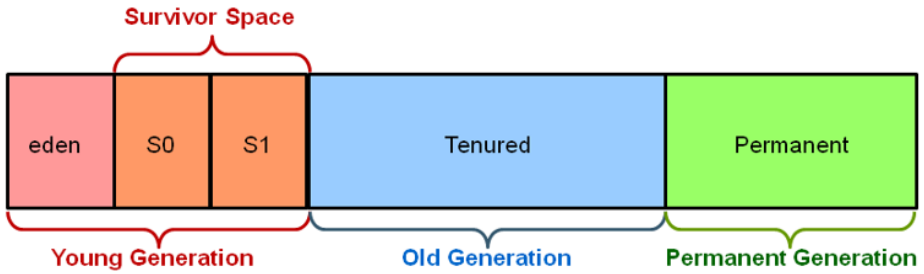
Put in practice I

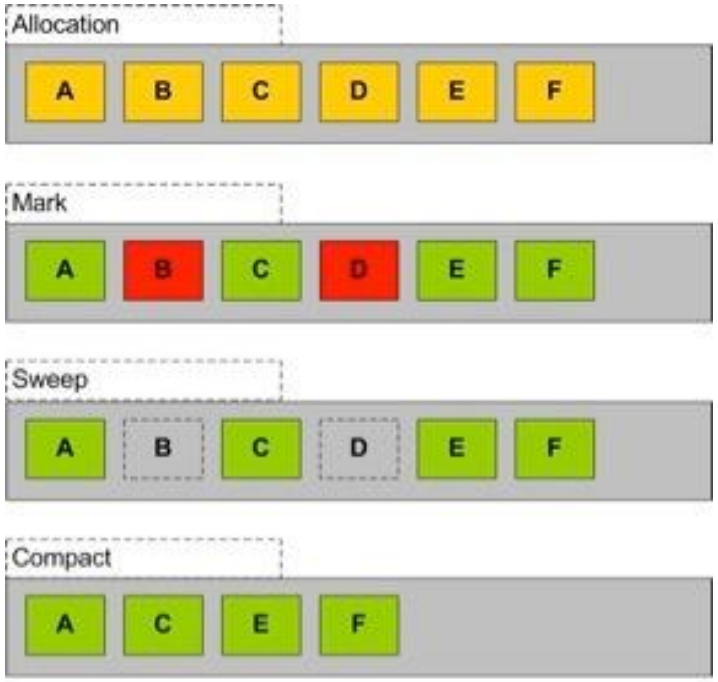
Why Containers and not VMs?

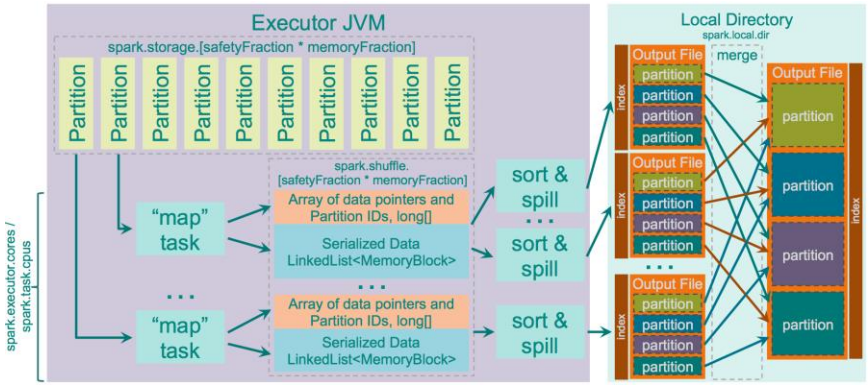


Demo



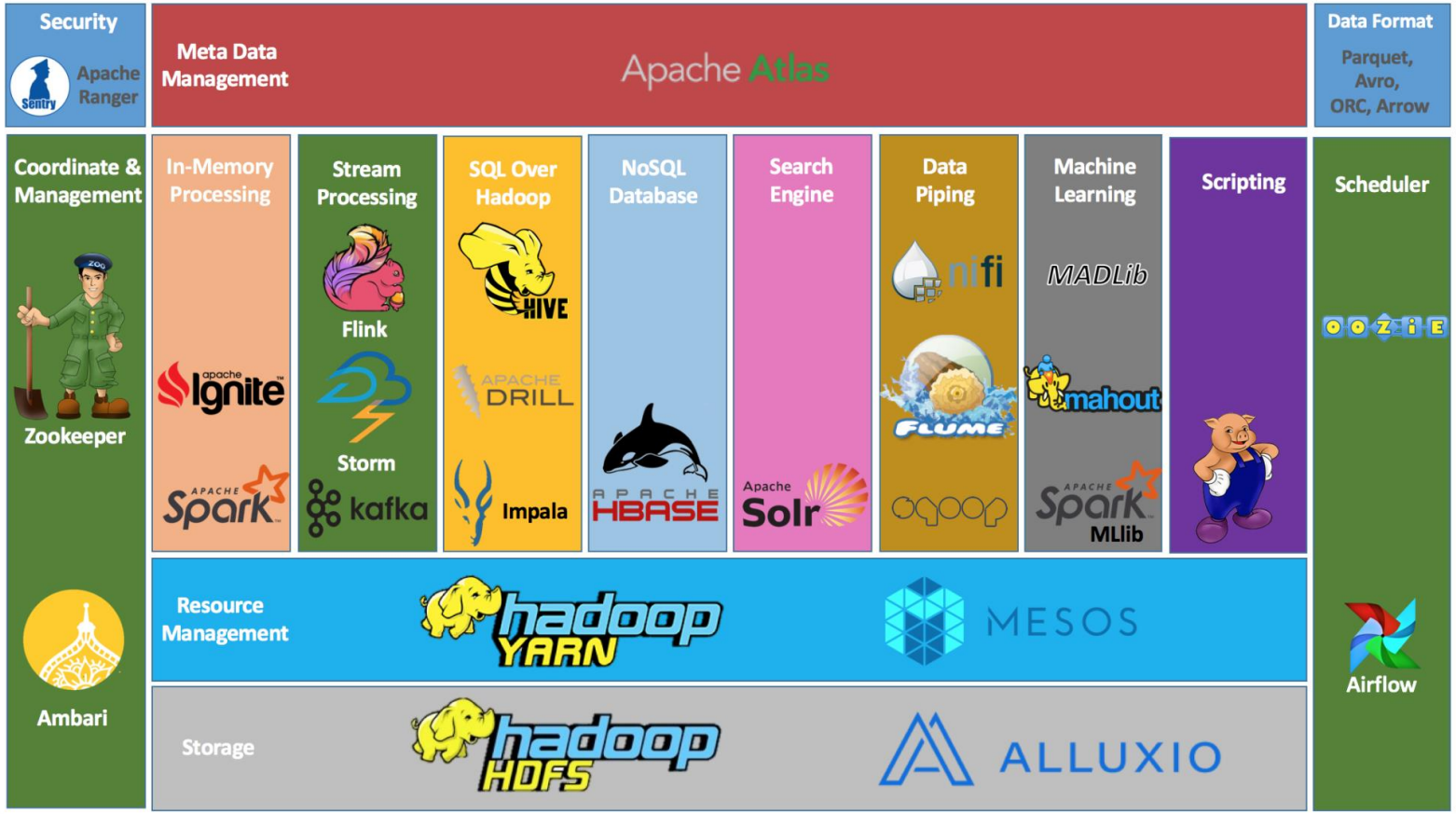






Apache Spark & Hadoop challenges

- JVM
- Shuffles
- Sort and Spills
- Data skewness



Apache Spark & Hadoop challenges

- **Optimize Hadoop I/O**
- **HDFS Short-Circuit Local Reads**
- **Centralized Cache Management & memory storage**
- **Multihomed Networks**
- **HDFS Erasure Coding**
- **Disk balancer**
- **Rack awareness**

TO DOs

- JupyterHub or Jupyter Lab in containers for data science and ML (maybe as a SaaS)
- OpenStack SDK Application for better operations to non experienced users
- More practical example for students (Microservices & DevOps)
- New GPU servers
- Adopt more Apache Foundation FOSS
- Focus on K8s
- Maybe Ansible too…

References

- LXD vs LXC images borrowed from *Canonical - Using containers to create the World's fastest OpenStack* presentation
- Text from OpenStack Docs