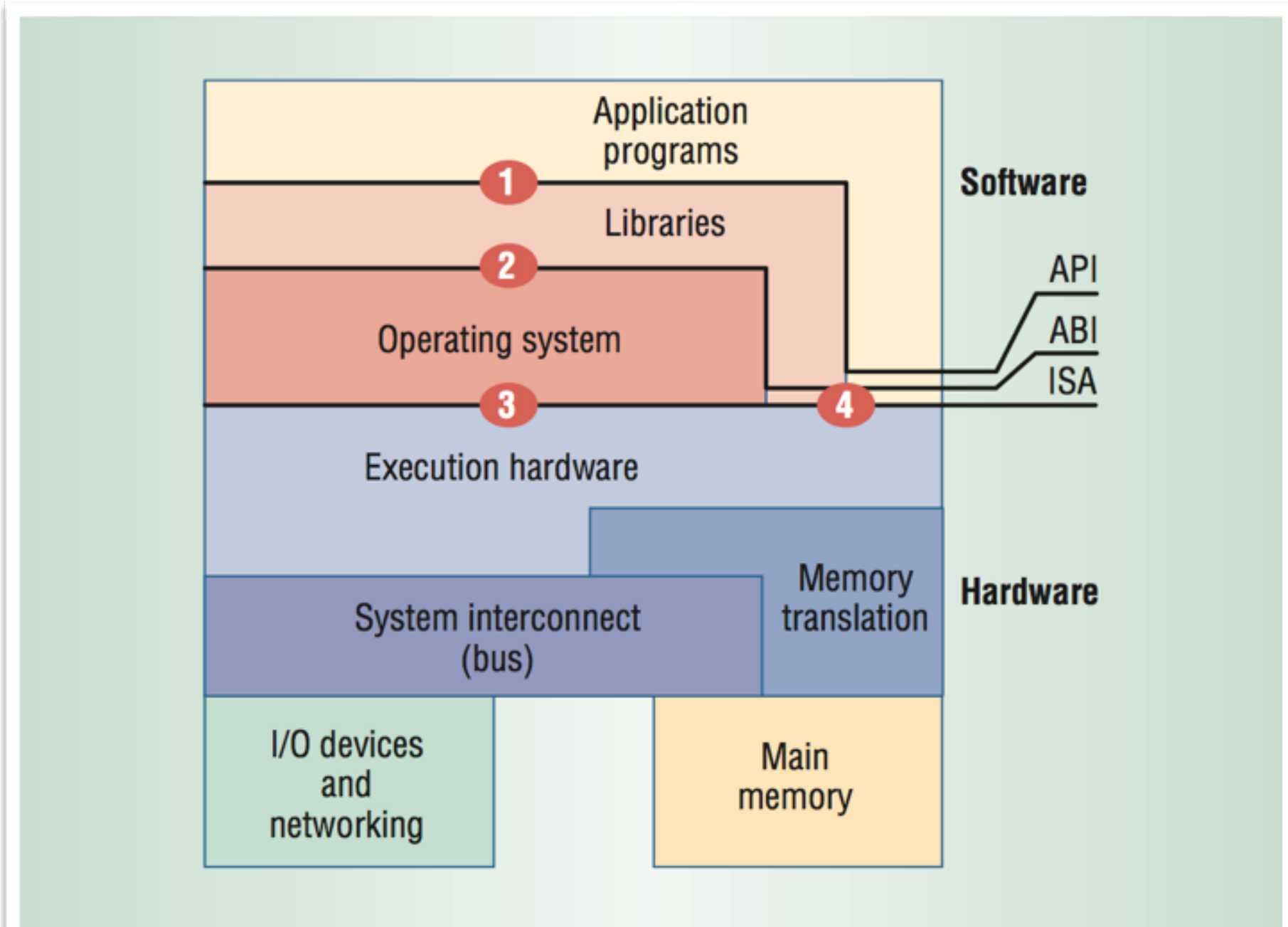
Cloud Computing #8 - Datacenter OS

Johan Eker

Outline

- What is a Datacenter OS?
- OpenStack
- Kubernetes
- Resource Management

What is an OS?


Application programs
1
Libraries
2
Operating system
3
4
Execution hardware
System interconnect (bus)
Memory translation
I/O devices and networking
Main memory
Software
API
ABI
ISA
Hardware

What is an OS?

Operating system

From Wikipedia, the free encyclopedia

An **operating system** (**OS**) is software that manages computer hardware and software resources and provides common services for computer programs. The operating system is an essential component of the system software in a computer system. Application programs usually require an operating system to function.

- Manage hardware resources such as
  - CPU, RAM, disk, I/O, etc.
- Provide services
  - Communication, authentication, synchronization, etc.
- An application program
  - A set of threads that communicate using e.g. shared memory
  - Uses OS services for disk, network, etc.

A Datacenter OS

- Manage hardware resources such as
  - Compute servers (CPU, RAM), storage servers, (disks), I/O, network, etc.
- Some services
  - Communication, authentication, synchronization, etc.
- Application program
  - A set of VM that communicate using message passing
  - Use OS services for disk, communication, etc.
- Multiple tenants

OS for the Cloud

- OpenStack
- Kubernetes
- Mesosphere
- VMware vCloud Air

OpenStack

- OpenStack is a set of open source software projects to setup a cloud
- Launched in 2010 by Rackspace and NASA with initial contributions
- OpenStack consortium has 125+ members
  - (Few tier 1 players)
- Open source under Apache license
- A lot of Python code
- A number of distributions (SUSE, RH, Mirantis, etc.)

openstack™
CLOUD SOFTWARE

- Command-line interfaces (nova, neutron, swift, and so on)
- Cloud Management Tools (Rightscale, Enstratius, and so on..)
- GUI tools (Dashboard, Cyberduck, iPhone client, and so on..)

Internet

Dashboard

Object Store

Image Store

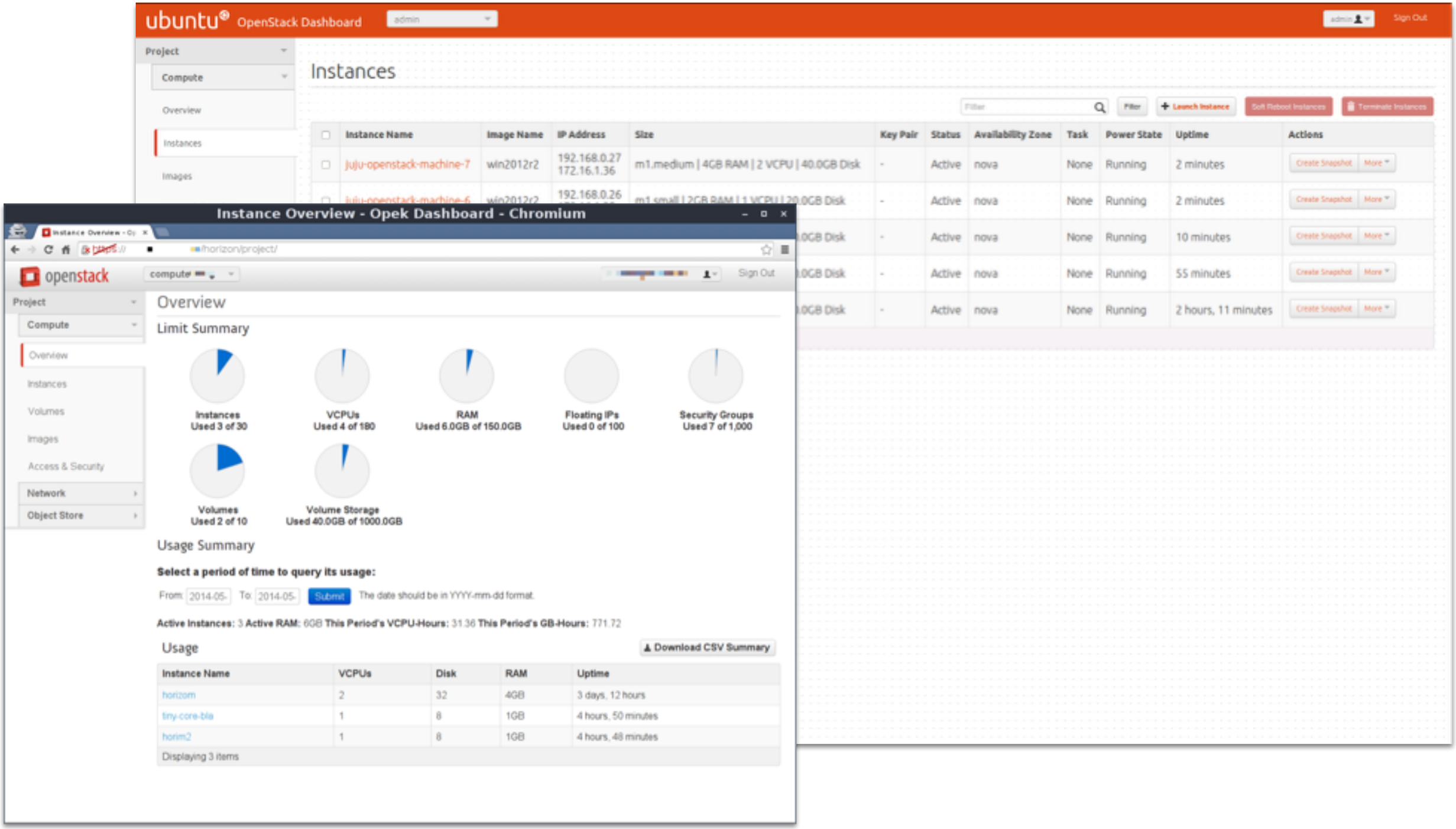
Compute

Block Storage

Networking

Authentication

Dashboard: “Horizon”



Compute: “Nova”

- Manage and automate pools of computer resources
  - Life cycle of VM instances
  - Keeps track of resources (virtual & real)
- Nova does not provide any virtualization capabilities, by itself; instead, it uses libvirt API to interact with supported hypervisors.
- Hypervisor agnostic
  - Xen, XenServer/XCP, KVM, UML, VMware vSphere and Hyper-V, LXC containers, Dockers
- REST-based API
  - Asynchronous eventually consistent communication
- Decides where to allocate instances (Nova-Schedule)
- API compatible with the EC2 API of AWS

Many ways to use the OpenStack REST API

```
#!/usr/bin/env python
 import time
 from credentials import get_nova_credentials_v2
 from novaclient.client import Client

 try:
     credentials = get_nova_credentials_v2()
     nova_client = Client(**credentials)

     image = nova_client.images.find(name="cirros")
     flavor = nova_client.flavors.find(name="m1.tiny")
     net = nova_client.networks.find(label="private")
     nics = [{'net-id': net.id}]
     instance = nova_client.servers.create(name="vm2", image=image,
                                 flavor=flavor, key_name="keypair-1", nics=nics)
     print("Sleeping for 5s after create command")
     time.sleep(5)
     print("List of VMs")
     print(nova_client.servers.list())
finally:
     print("Execution Completed")
```

```
$ nova boot --flavor FLAVOR_ID --image IMAGE_ID --key-name KEY_NAME \
--user-data USER_DATA_FILE --security-groups SEC_GROUP_NAME --meta KEY=VALUE \
INSTANCE_NAME
```

Object Store: “Swift”

- Swift is a highly available, distributed, eventually consistent object/blob store
  - Unstructured data store. Swift simply stores bits. Swift is not a database. Swift is not a block-level storage system. Swift stores blobs of data.
- Scales to thousands of servers with tens of thousands of hard drives. Horizontally scalable w/o no single point of failure.
- Redundancy by multiple copies in different availability zones.

Object Store: “Swift”

- Swift provides a REST API over HTTP
- A swift storage URL looks like
  - swift.example.com/v1/account/container/object

List of all containers: GET http://swift.example.com/v1/account/

Create new container: PUT http://swift.example.com/v1/account/new_container

List all object in a container: GET http://swift.example.com/v1/account/container/

Create new object: PUT http://swift.example.com/v1/account/container/new_object.

- API similar to AWS S3

Block Storage: Cinder

- Persistent block storage for VMs
- Three services:
  - *Volumes* (virtual raw block devices)
  - *Snapshots* (quick)
  - *Backups* (full copy stored in Swift).
- Implemented in top of: Ceph, GlusterFS, XFS, NFS, NetApp, SMB, etc (long list of drivers)
- API similar to the AWS Elastic Block Storage (EBS)

Orchestration: “Heat”

- Automated configuration of cloud resources in an application:
  - Servers, Load Balancers, Databases, Block Storage, DNS, Auto Scaling, Init scripts,
- A Heat template describes the infrastructure for a cloud application
- Autoscaling service
- Heat manages the whole lifecycle of the application - launch, update, terminate
- Compatible with AWS CloudFormation
- Version control of distributed applications

Metering: “Ceilometer”

- Provide counters for utilisation of the physical and virtual resources comprising deployed clouds
- Keep database of metering data
- Setup conditions for triggering actions
- For billing, scaling, etc
- Think: AWS CloudWatch, AWS CloudMetrics

Networking: “Neutron”

- Networking as a service
- Manages IP addresses, (static/dynamic/floating)
- Users can create their own networks, control traffic, and connect servers and devices
- Based on software-defined networking (SDN) to provide high levels of multi-tenancy and scale

openstack™
CLOUD SOFTWARE

- Command-line interfaces (nova, neutron, swift, and so on)
- Cloud Management Tools (Rightscale, Enstratius, and so on..)
- GUI tools (Dashboard, Cyberduck, iPhone client, and so on..)

Internet

Dashboard
"Horizon"

Object
Store
"Swift"

Image
Store
"Glance"

Compute
"Nova"

Block
Storage
"Cinder"

Networking
"Neutron"

Authentication
"Keystone"

openstack
CLOUD SOFTWARE

- Command-line interfaces (nova, neutron, swift, and so on)
- Cloud Management Tools (Rightscale, Enstratius, and so on..)
- GUI tools (Dashboard, Cyberduck, iPhone client, and so on..)

Internet

HTTP(S)
OpenStack Object API
OpenStack Identity API
OpenStack Image API
HTTP(S)
OpenStack Dashboard
Horizon
OpenStack Compute API
Amazon Web Services EC2 API
VNC/VMRC
OpenStack Block Storage API
OpenStack Networking API
OpenStack Networking API
OpenStack Object API
OpenStack Image API
OpenStack Identity API
OpenStack Compute API / Admin API
OpenStack Block Storage API

swift-proxy
OpenStack Object API
memcached
account
container
object
account DB
container DB
object store
OpenStack Object Store

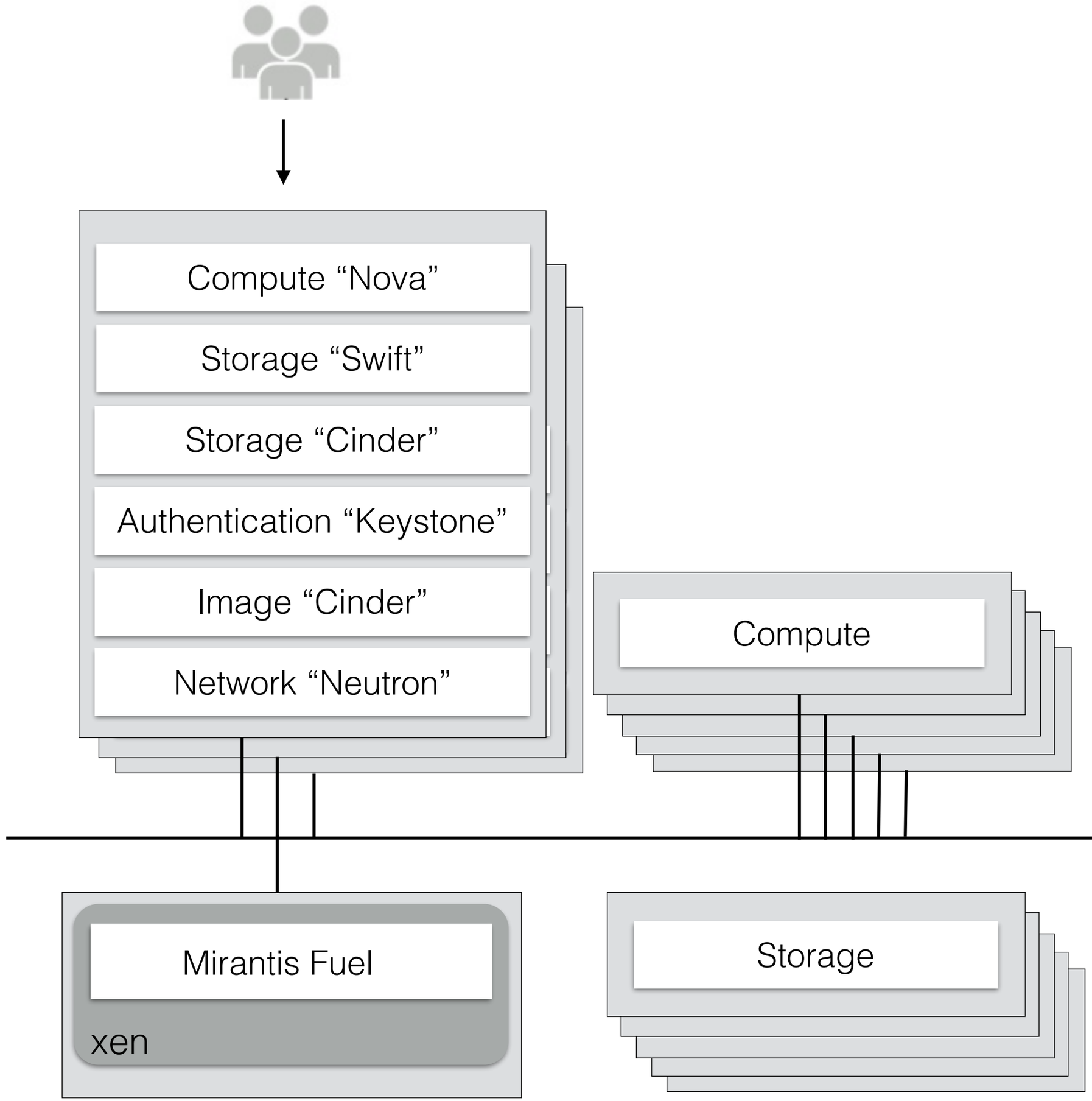
glance-api
OpenStack Image API
glance-registry
glance database
OpenStack Image Service
OpenStack Identity API

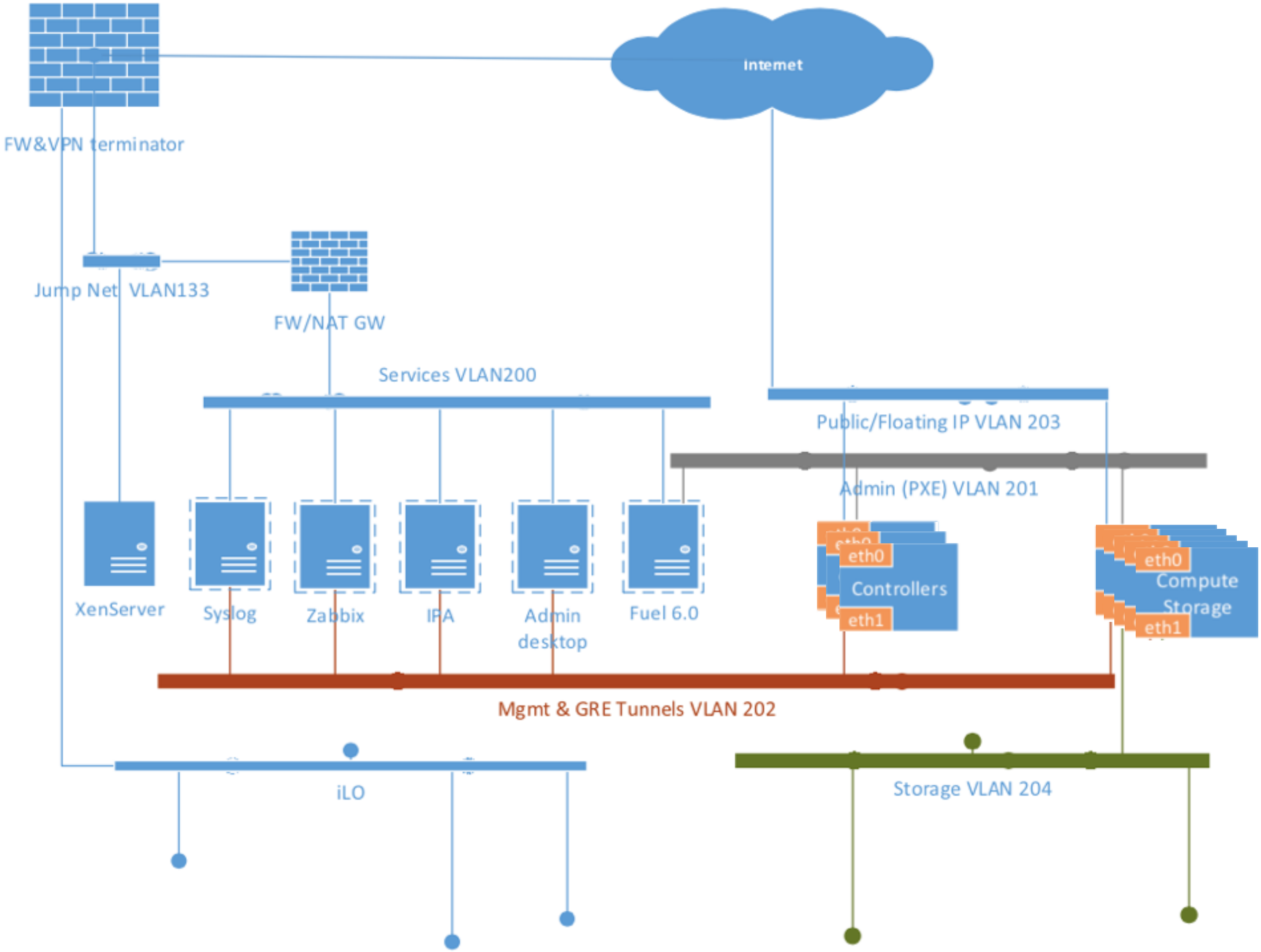
OpenStack Image API
nova-api (OS, EC2, Admin)
nova-console
nova-cert/ objectstore
nova-compute
libvirt, XenAPI, etc.
hypervisor
nova database
Queue
nova-conductor
nova-consoleauth
nova-scheduler
OpenStack Identity API
OpenStack Compute

cinder-api
cinder-volume
AMQP
volume provider
cinder database
AMQP
cinder-scheduler
OpenStack Identity API
OpenStack Block Storage

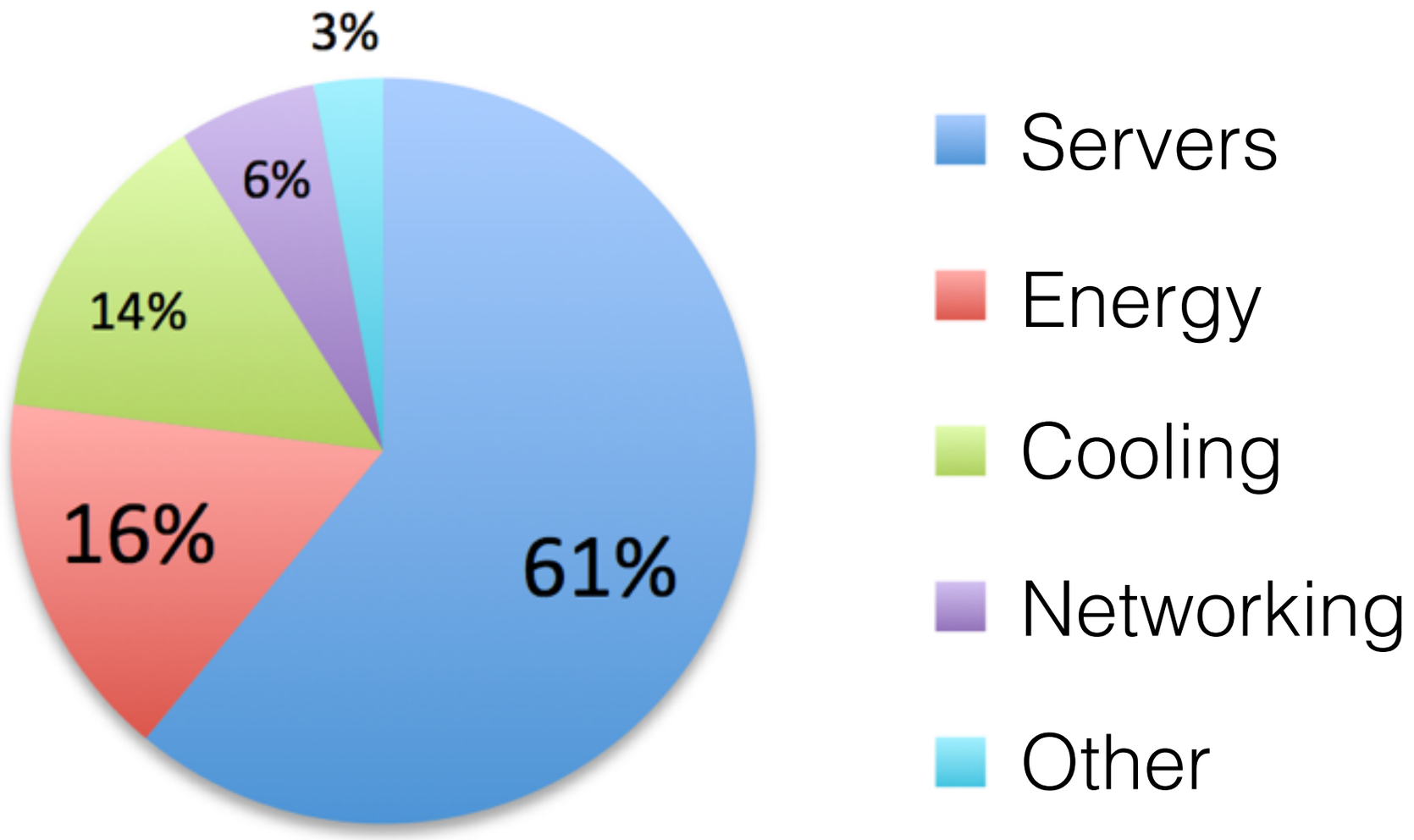
neutron-server
neutron agent(s)
Queue
neutron plugin(s)
network provider
neutron database
OpenStack Networking

OpenStack Identity API
keystone (service & admin APIs)
OpenStack Identity Service
token backend
catalog backend
policy backend
identity backend
OpenStack Identity API
OpenStack Identity API
OpenStack Identity API


Compute “Nova”
Storage “Swift”
Storage “Cinder”
Authentication “Keystone”
Image “Cinder”
Network “Neutron”
Compute
Mirantis Fuel
xen
Storage

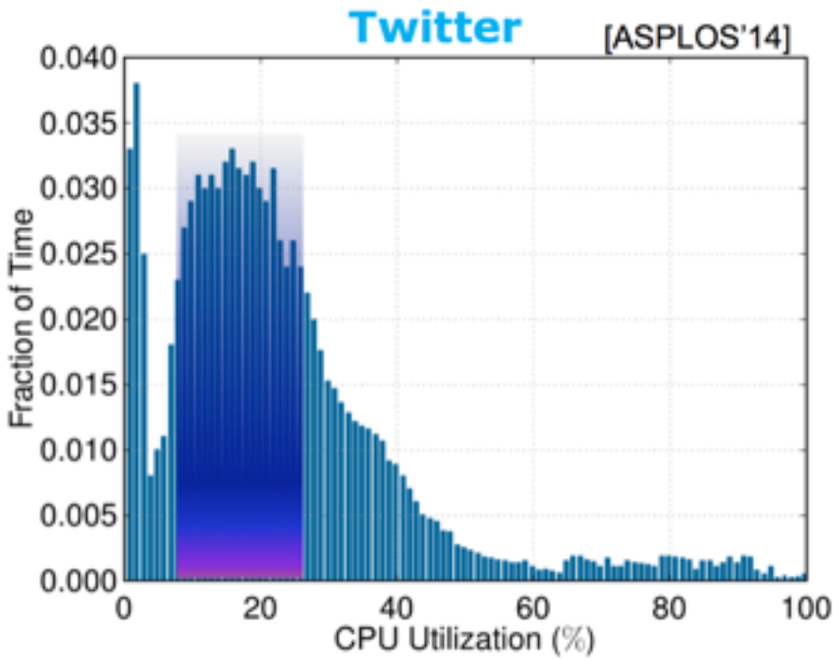

Internet
FW&VPN terminator
Jump Net VLAN133
FW/NAT GW
Services VLAN200
Public/Floating IP VLAN 203
Admin (PXE) VLAN 201
XenServer
Syslog
Zabbix
IPA
Admin desktop
Fuel 6.0
eth0
Controllers
eth1
eth0
Compute
Storage
eth1
Mgmt & GRE Tunnels VLAN 202
iLO
Storage VLAN 204

Total Cost of Cloud Ownership

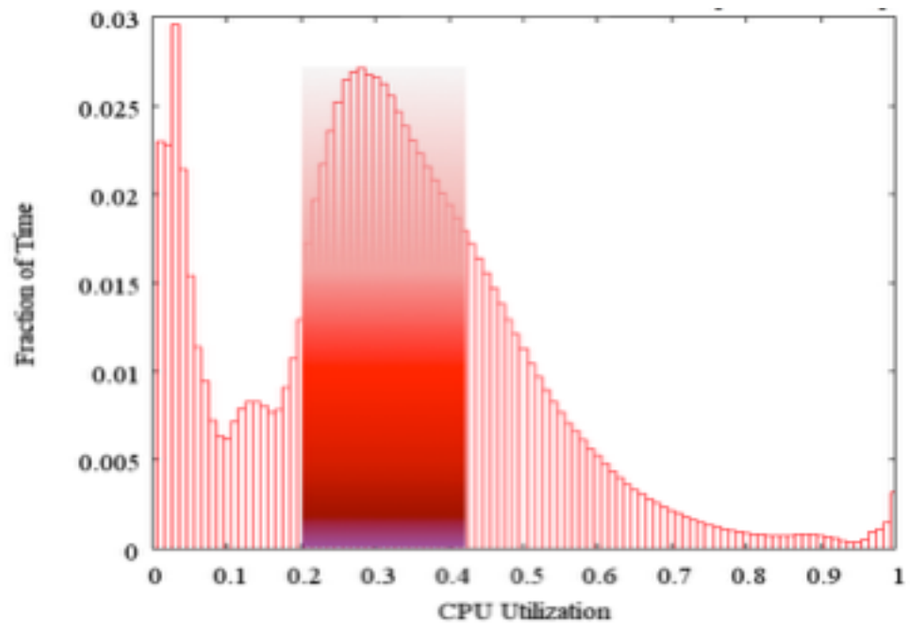


[J. Hamilton, http://mvdirona.com]

Resource Scheduling

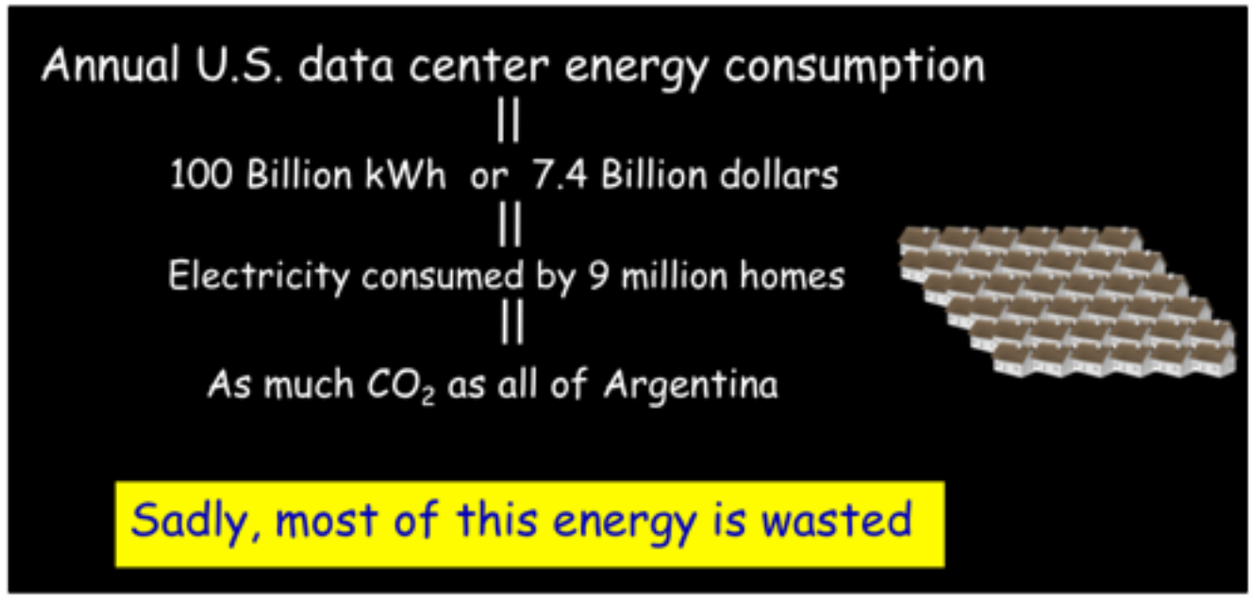


Dynamic cluster management using Mesos



Dynamic cluster management using Borg

FB in Luleå will consume 1 TWh/year



Next

- RESTful APIs (Ola Angelsmark)

- Resource Management (Jonas Dürango)