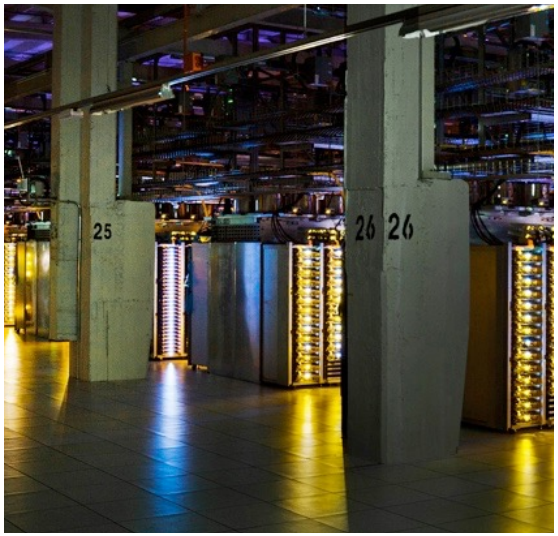
Building a private cloud with OpenStack



OpenStack
Introduction

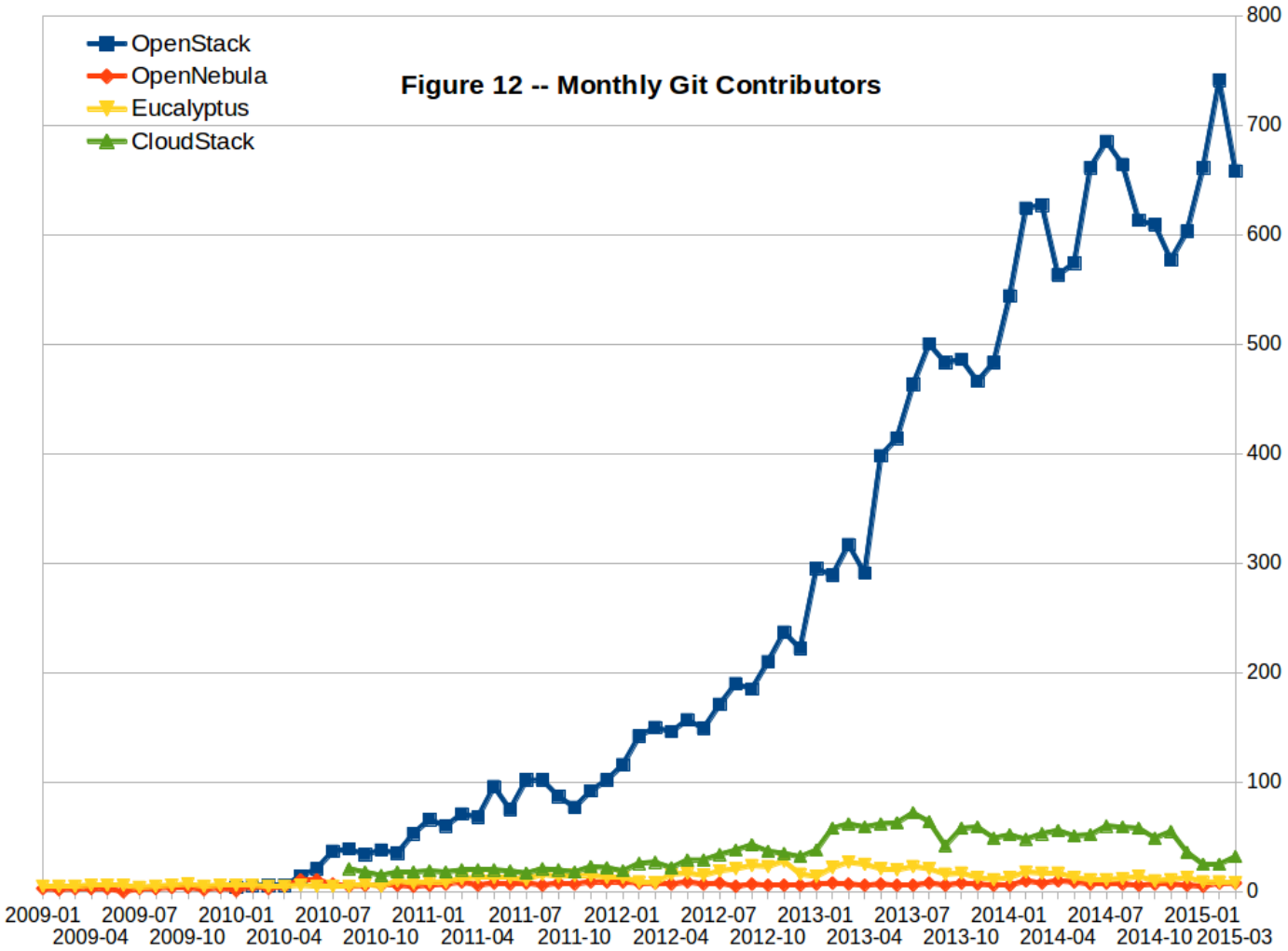
- Software for creating private and public clouds (Infrastructure-as-a-Service), released as Open Source (Apache license)
  - Initiated by Rackspace and NASA in 2010
    - NASA contributes compute platform Nebula
    - Rackspace contributes storage platform Cloud Files
  - Today OpenStack Foundation is supported by >500 companies
    - Current Platinum members:
      - AT&T
      - Canonical (Ubuntu Linux)
      - HP Enterprise
      - IBM
      - Intel
      - Rackspace
      - Red Hat
      - SUSE
  - "OpenStack is the Linux of cloud computing."

OpenStack
History

- 2010-07 Rackspace and NASA launch OpenStack Open Source cloud software initiative
- 2010-11 first release, *Austin*
- 2011-04 Canonical starts to distribute OpenStack in Ubuntu 11.04
- 2012-04 Rackspace starts offering public cloud service based on OpenStack
- 2012-09 OpenStack Foundation is launched as an independent body
- 2012-12 HP Enterprise starts offering public cloud service based on OpenStack
- 2013-03 IBM/Softlayer start offering public cloud service based on OpenStack
- 2013-07 Red Hat starts to offer commercial support for OpenStack
- 2014-09 Oracle starts to distribute OpenStack in Oracle Linux

OpenStack and the competition

OpenStack vs. OpenNebula vs. Eucalyptus vs. CloudStack



Top 10 contributing orgs for **OpenStack** in 2015Q1 by email domain

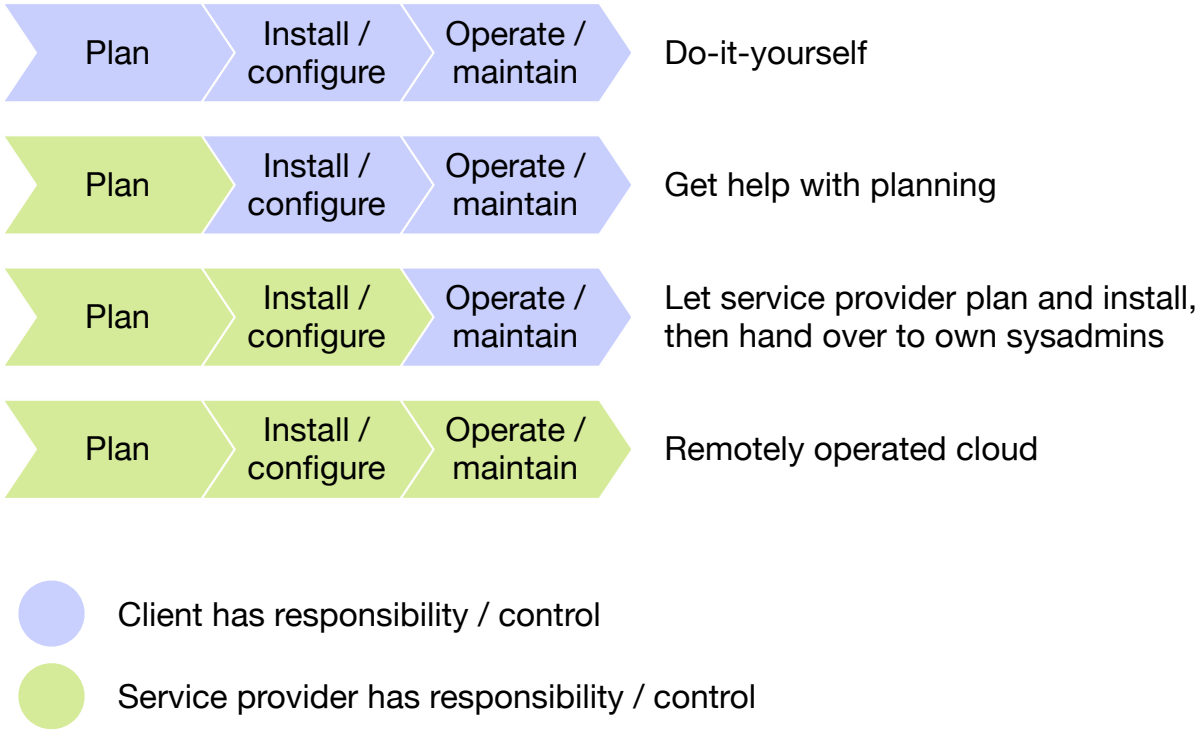
9.6% gmail.com
7.3% redhat.com
5.0% ibm.com
4.7% mirantis.com
4.6% hp.com
1.6% rackspace.com
1.4% intel.com
1.2% yahoo-inc.com
1.1% doughellmann.com
0.8% cisco.com

Source: Qingye Jiang, Open Source IaaS Community Analysis, http://www.qyjohn.net/?p=3801

OpenStack
Deployment options

- Deploy your own **private** cloud
  - Many options available from do-it-yourself to remotely operated
- Use a **public** cloud
  - Rackspace Public Cloud (since 2012-04)
  - HP Enterprise (since 2012-12)
  - IBM/Softlayer (since 2013-03)
  - and many others
- Open Source nature of OpenStack avoids vendor lock-in!

Private cloud deployment options

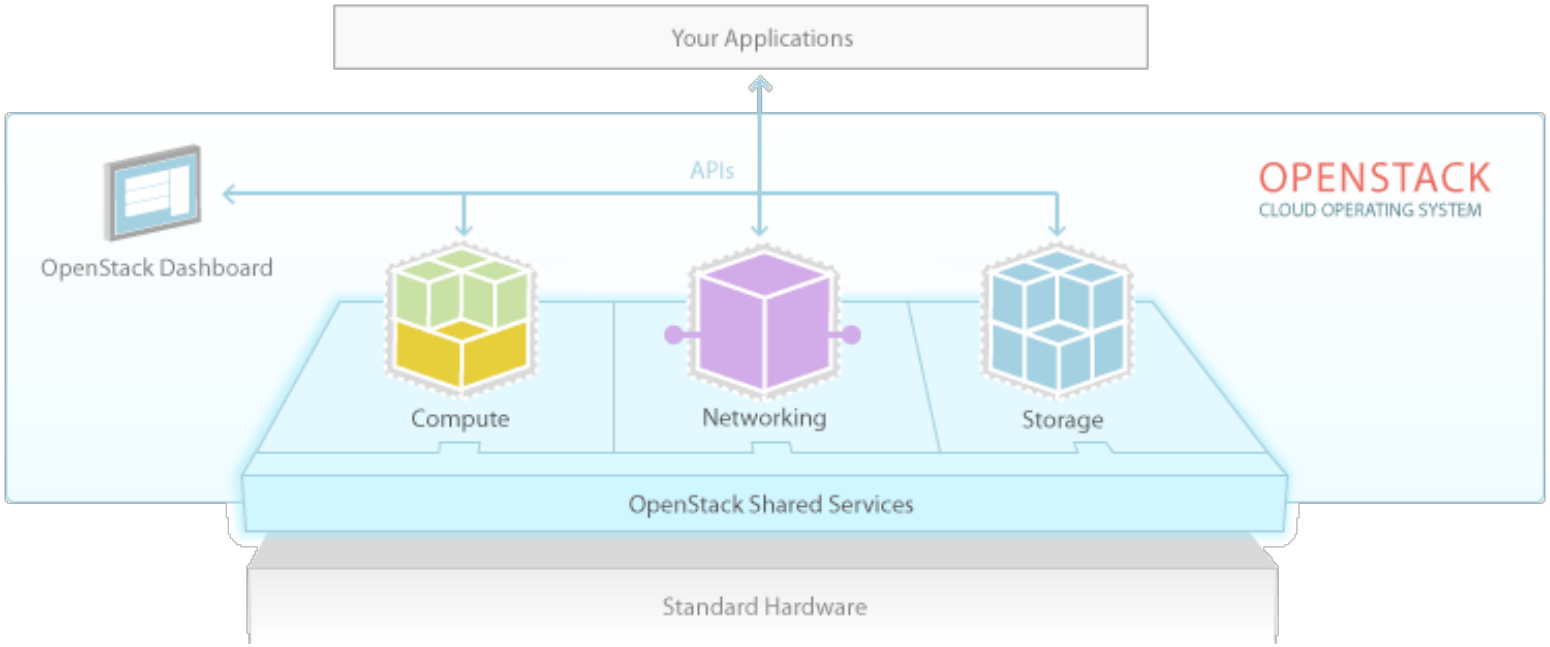


OpenStack
Users

- Among the most notable users of OpenStack
  - AT&T
  - Bloomberg
  - CERN
  - Cisco Webex
  - Disney
  - Fidelity
  - Swisscom
  - Walmart

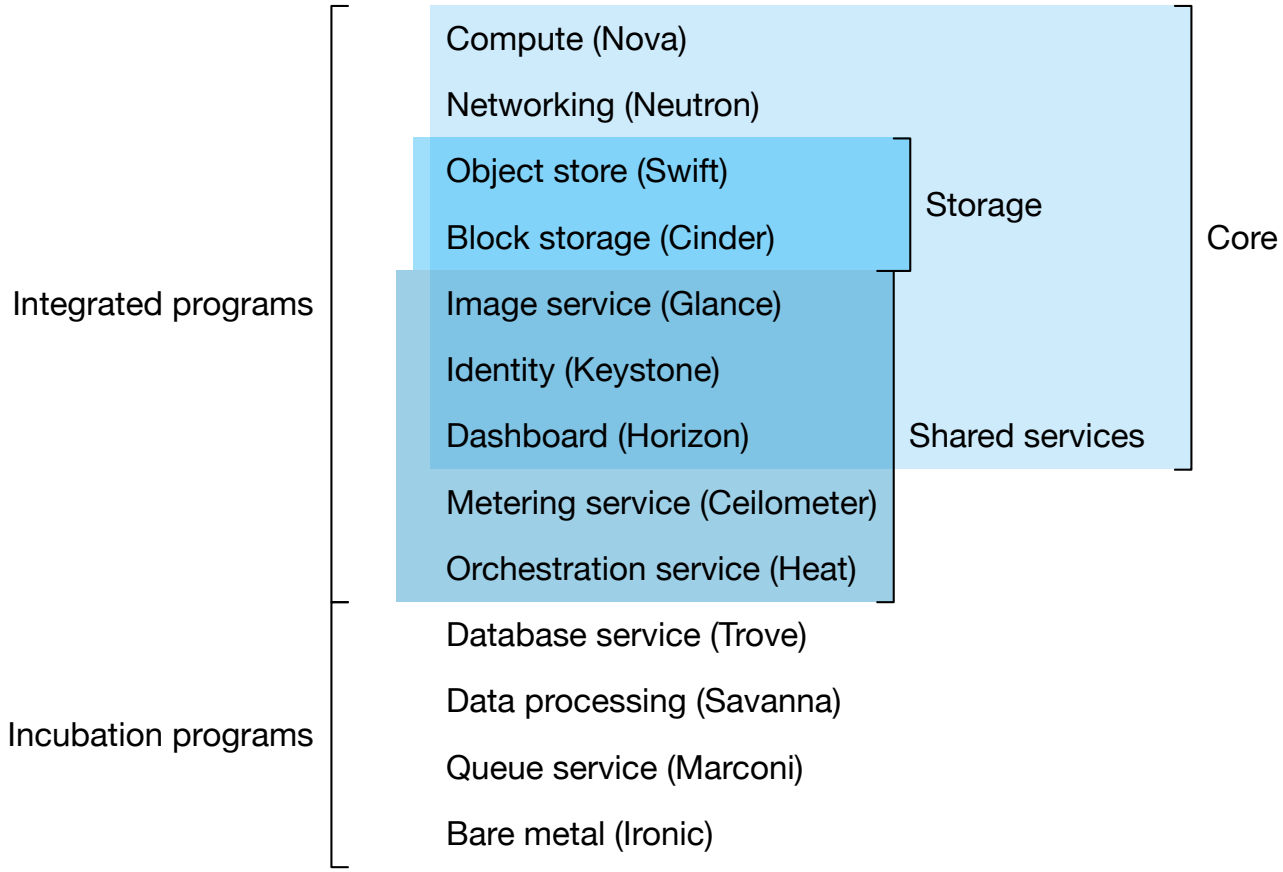
OpenStack
Main components

- OpenStack components are developed in their own sub-projects called *programs*
  - **Compute** — Nova
  - **Storage**
    - Object storage — Swift
    - Block storage — Cinder
- **Networking** — Neutron
- **Dashboard** — Horizon
- **Identity** (authentication and authorization) — Keystone



OpenStack

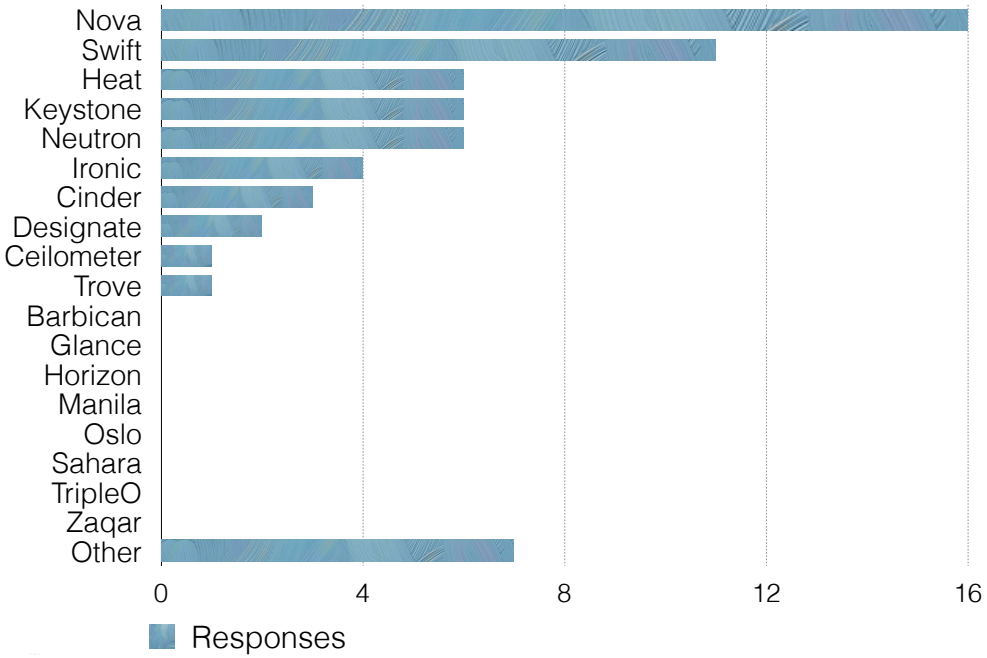
Integrated programs and incubation programs



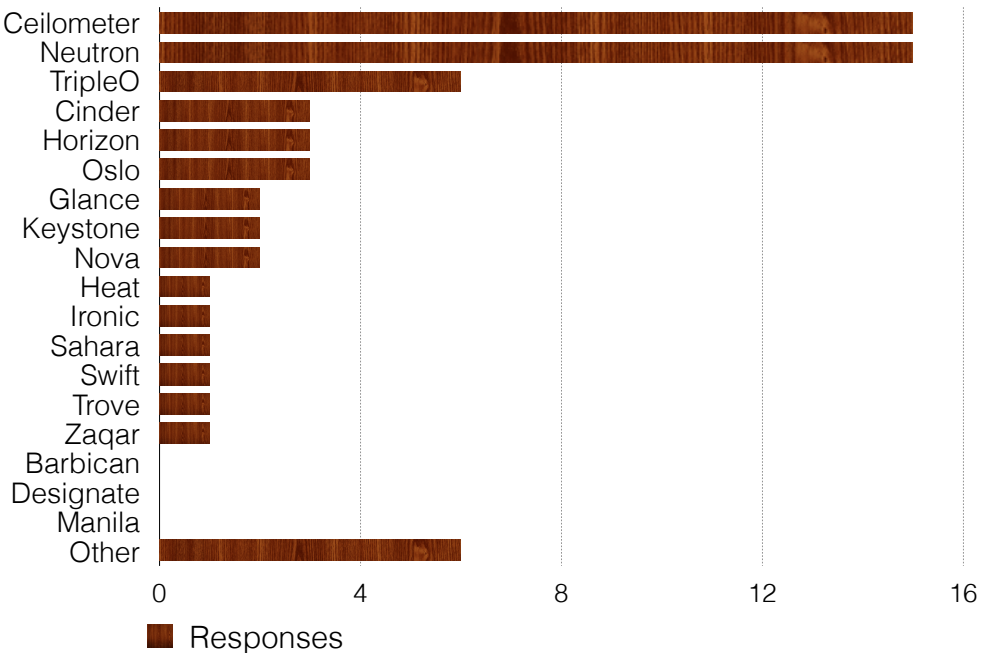
OpenStack

Not all projects are created equal

What is your **most** favorite OpenStack project?



What is your **least** favorite OpenStack project?

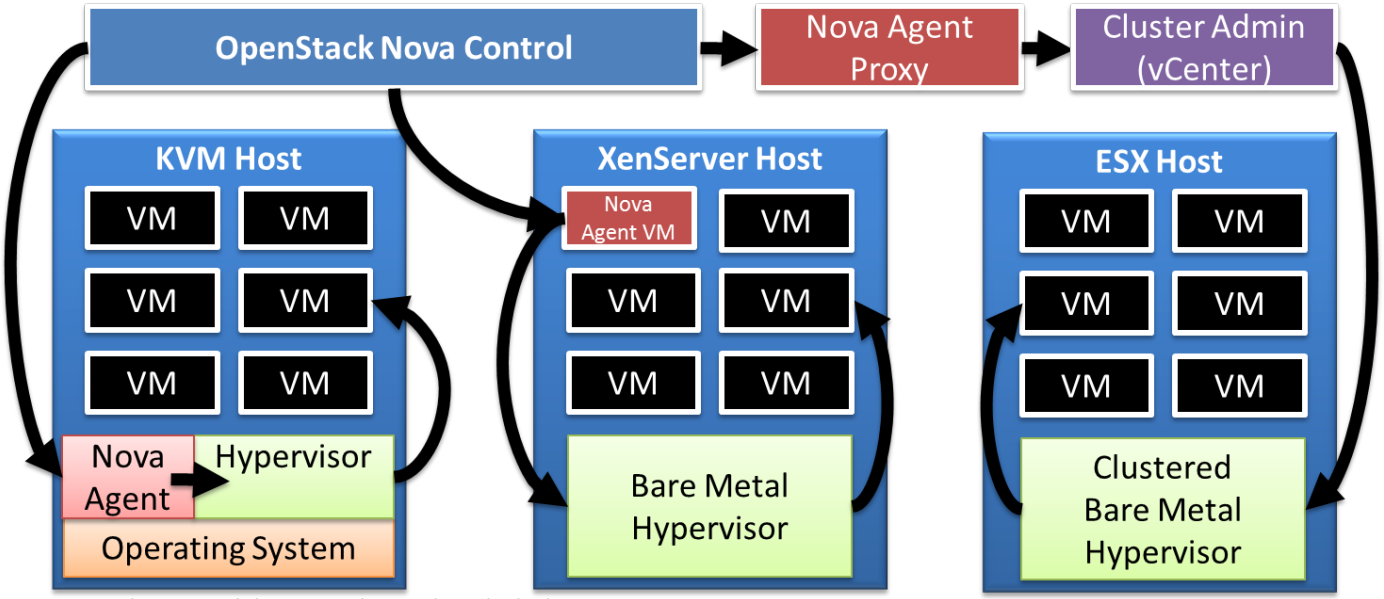


OpenStack / Amazon Web Services translation chart

Generic term	OpenStack term	AWS term
Compute service	Nova	EC2 (includes compute service)
Object storage	Swift	S3
Block storage	Cinder	Elastic Block Store (EBS)
Image service	Glance	Amazon Machine Image (AMI)
Virtual machine	Instance	Instance
Virtual disk	Volume	Volume
Firewall configuration	Security group	Security group
Fixed IP address	Floating IP address	Elastic IP address

OpenStack
Compute — Nova

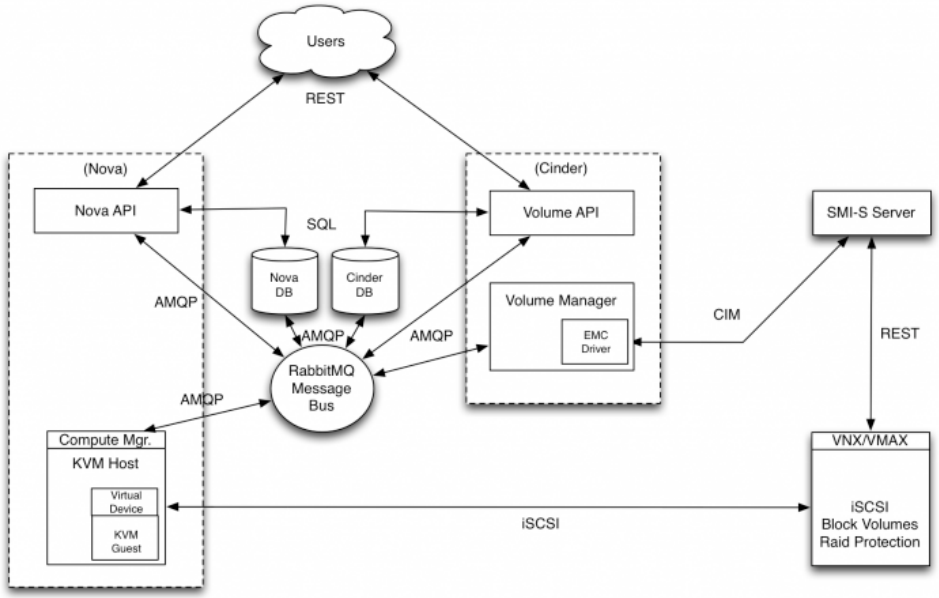
- Manages virtual machines in compute servers
  - Receives requests from users
  - Creates / starts / stops / releases virtual machines
  - Determines on which server to create a VM requested by a user (scheduling)
- Relies on hypervisors to implement the virtual machines
  - Can make use of KVM, Xen, ESX, Hyper-V, …



Source: Dell, http://en.community.dell.com/techcenter/b/techcenter/archive/2011/06/01/virtualizing-approaches-for-openstack-nova-looking-at-the-many-ways-to-skin-the-cactus-kvm-v-xenserver-v-esx.aspx

OpenStack
Block Storage — Cinder

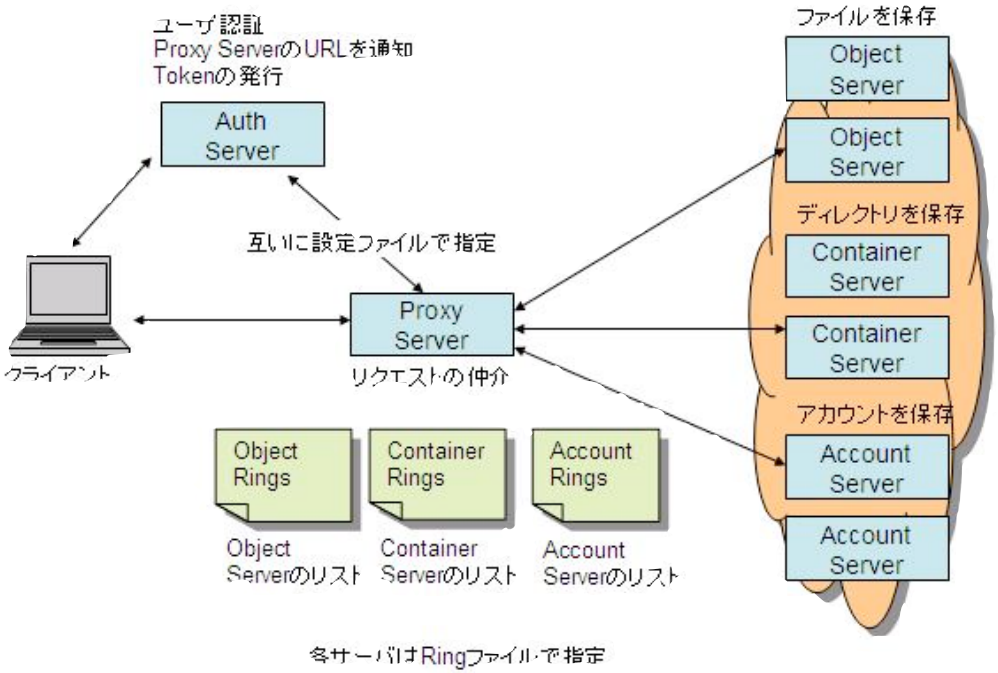
- Service for offering virtual disks (volumes) that can be attached to virtual machines.
  - The operating system in the VM sees a block device and puts a file system on it.
- Functions:
  - Create volume
  - Create volume from snapshot
  - Create volume from VM image
  - Save volume in VM image
  - Attach / detach volume to / from VM
  - Create / delete snapshot
- Can use different storage technologies underneath
  - Local disks
  - SAN (Fiber Channel, iSCSI)
  - NAS (NFS, CIFS)



Source: Rackspace, http://www.rackspace.com/knowledge_center/article/implementing-openstack-cinder-with-emc-storage-on-the-rackspace-private-cloud-software

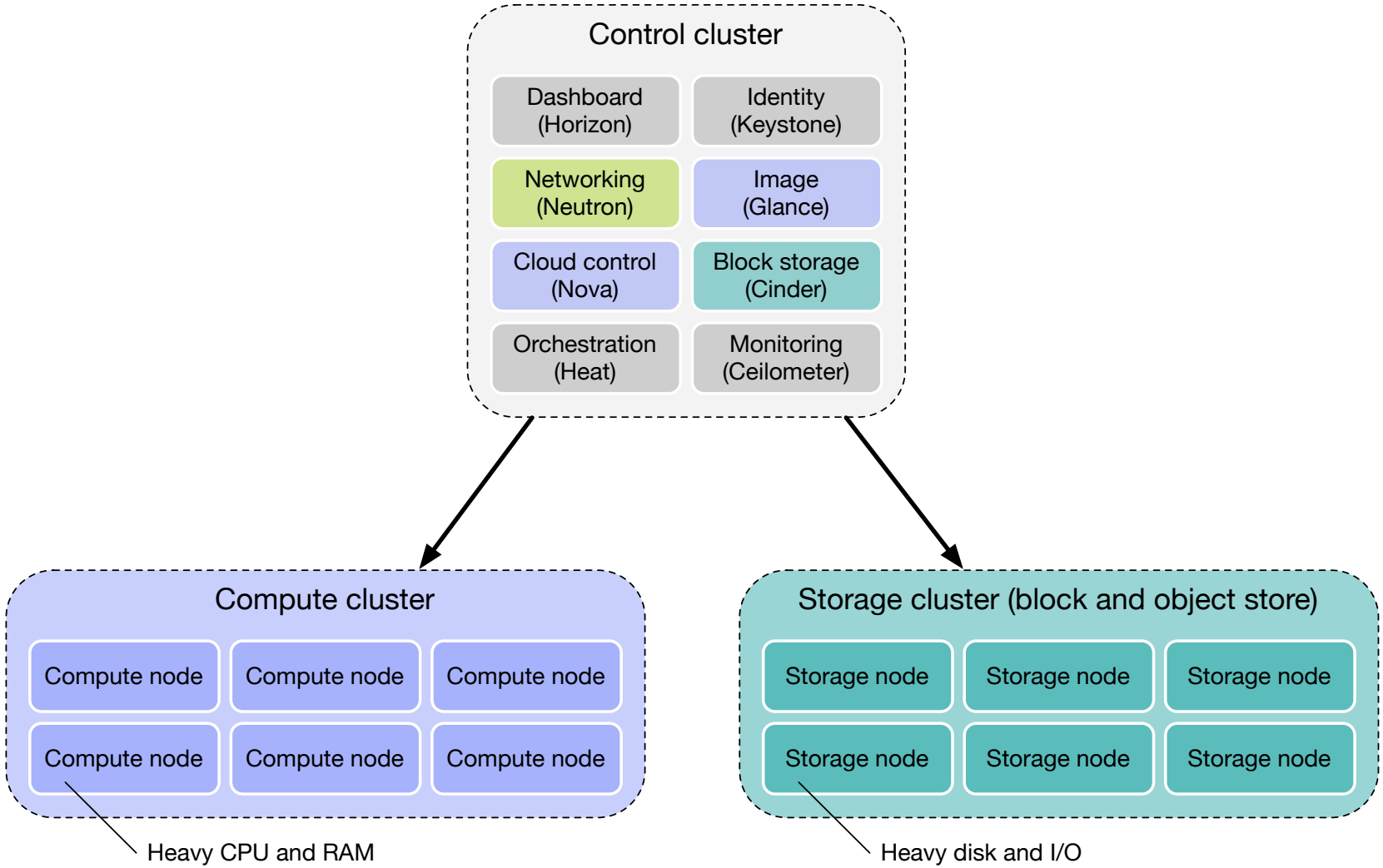
OpenStack
Object Storage — Swift

- Service for storing “objects"
  - Files without structure
  - Directories are not files
- Distributed architecture
  - Deployed on a cluster of servers
  - Three server roles
    - **Proxy server**: Receives user requests and coordinates storage servers.
    - **Storage servers**: Offer disk space. Three types of storage servers:
      - Account server: Stores user accounts
      - Container server: Stores containers
      - Object server: Stores objects
    - **Consistency server**: Responsible for looking for and finding errors (hardware or software failures) and correcting them.



Source: Creationline http://www.creationline.com/lab/772

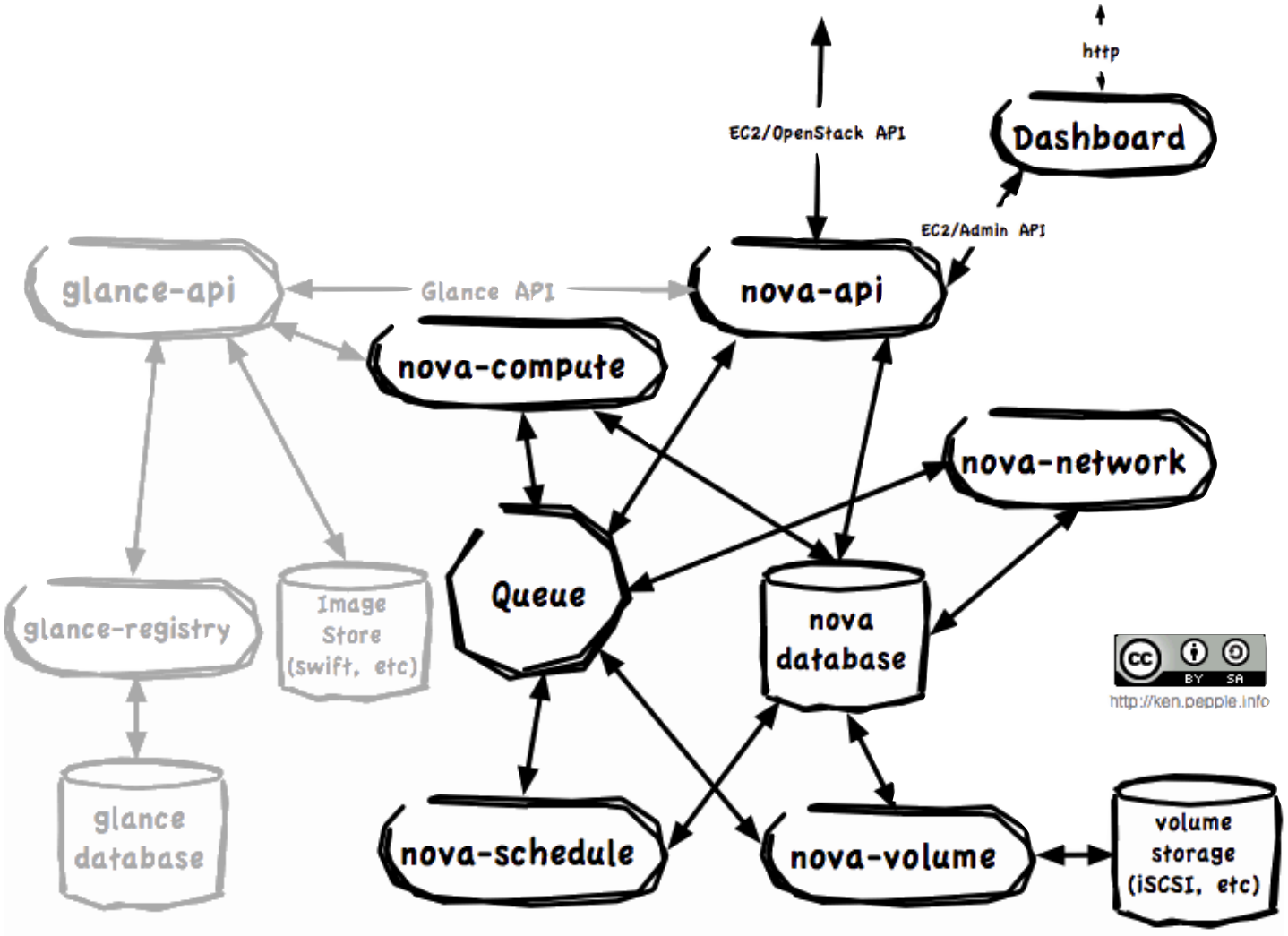
OpenStack
Logical deployment topology



Based on: Mirantis – OpenStack Overview and History, http://www.slideshare.net/mirantis/overview-43176920

OpenStack architecture

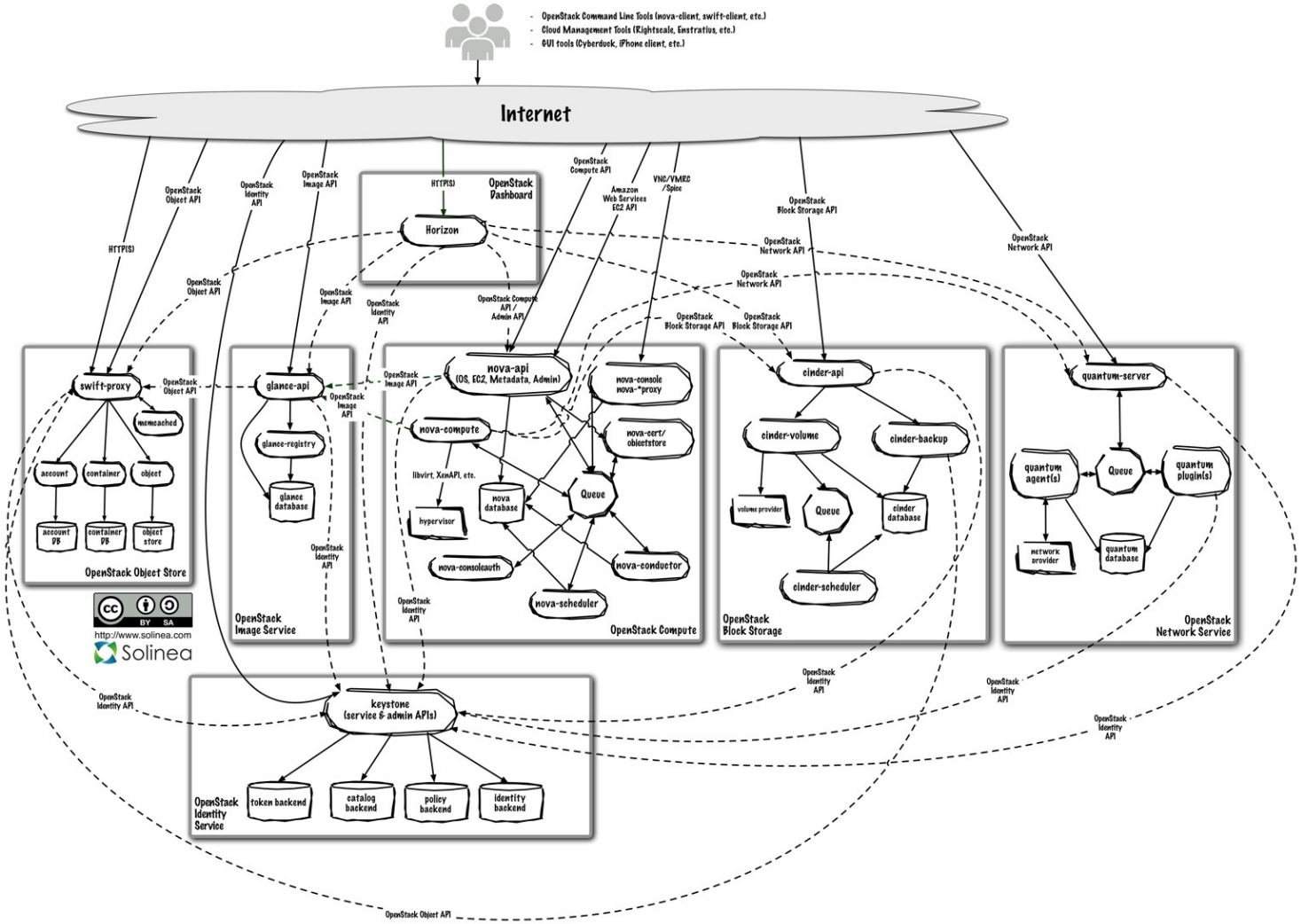
Beginning (2011)



Source: Ken Pepple, http://ken.pepple.info/openstack/2011/04/22/openstack-nova-architecture/

OpenStack architecture

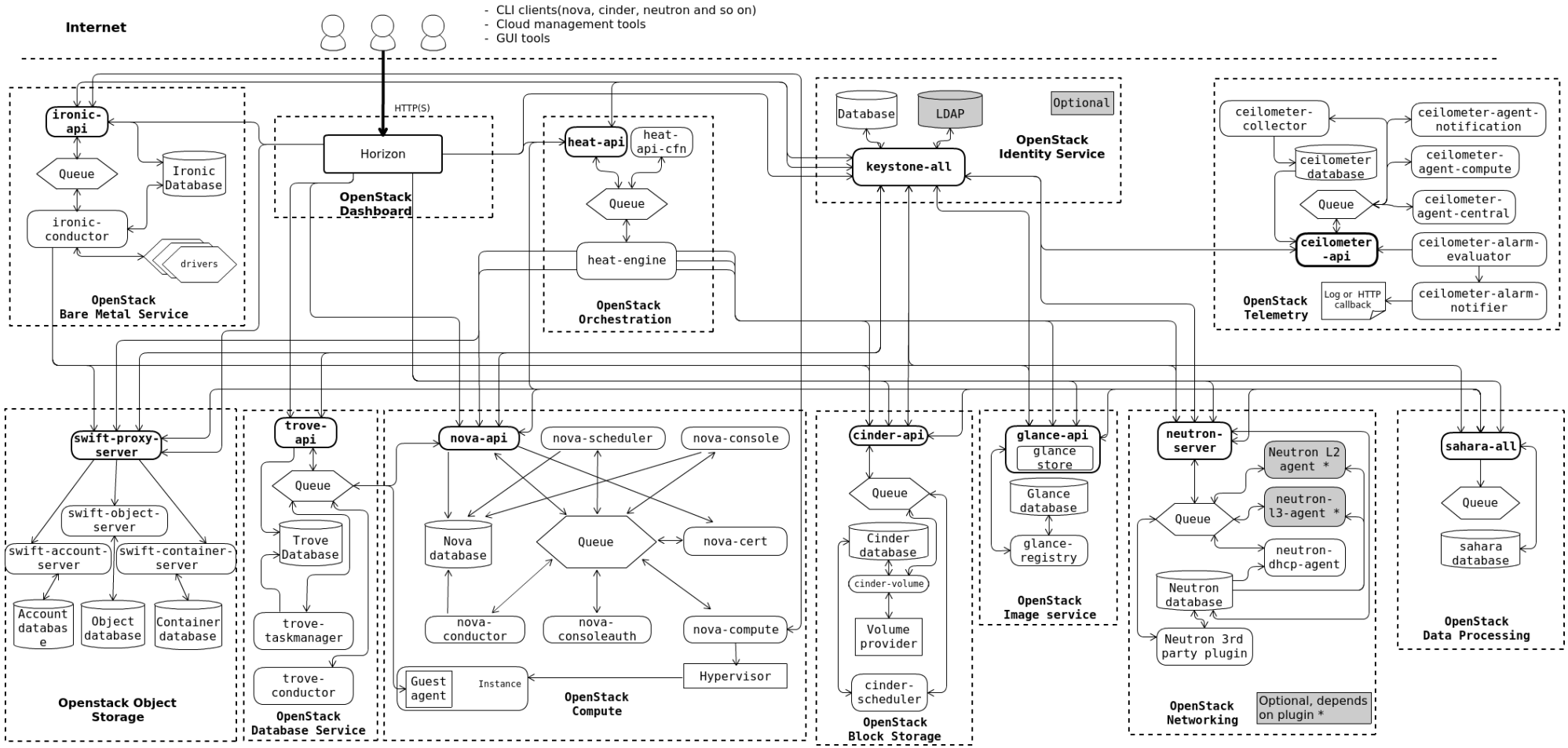
Two years later (2013)



Source: OpenStack Grizzly logical architecture

OpenStack architecture

Today (2016)



Source: OpenStack Liberty logical architecture, http://docs.openstack.org/admin-guide-cloud/common/get_started_logical_architecture.html

OpenStack
Release schedule

- OpenStack is developed and released around 6-month cycles.
  - Every 6 months a new release
  - Each release gets security updates for only 12 months, then it is declared end-of-life
- Much shorter support timespan than Linux distributions (3 - 10 years). Hard to keep up!

Series	Status
Ocata	*Future*
Newton	*Future*
Mitaka	*Under Development*
Liberty	Current stable release, security-supported (EOL: 2016-11-17)
Kilo	Security-supported (EOL: 2016-05-02)
Juno	EOL: 2015-12-07
Icehouse	EOL: 2015-07-02
Havana	EOL: 2014-09-30
Grizzly	EOL: 2014-03-29
Folsom	EOL: 2013-11-19
Essex	EOL: 2013-05-06
Diablo	EOL: 2013-05-06
Cactus	Deprecated
Bexar	Deprecated
Austin	Deprecated

OpenStack
Automated installation options

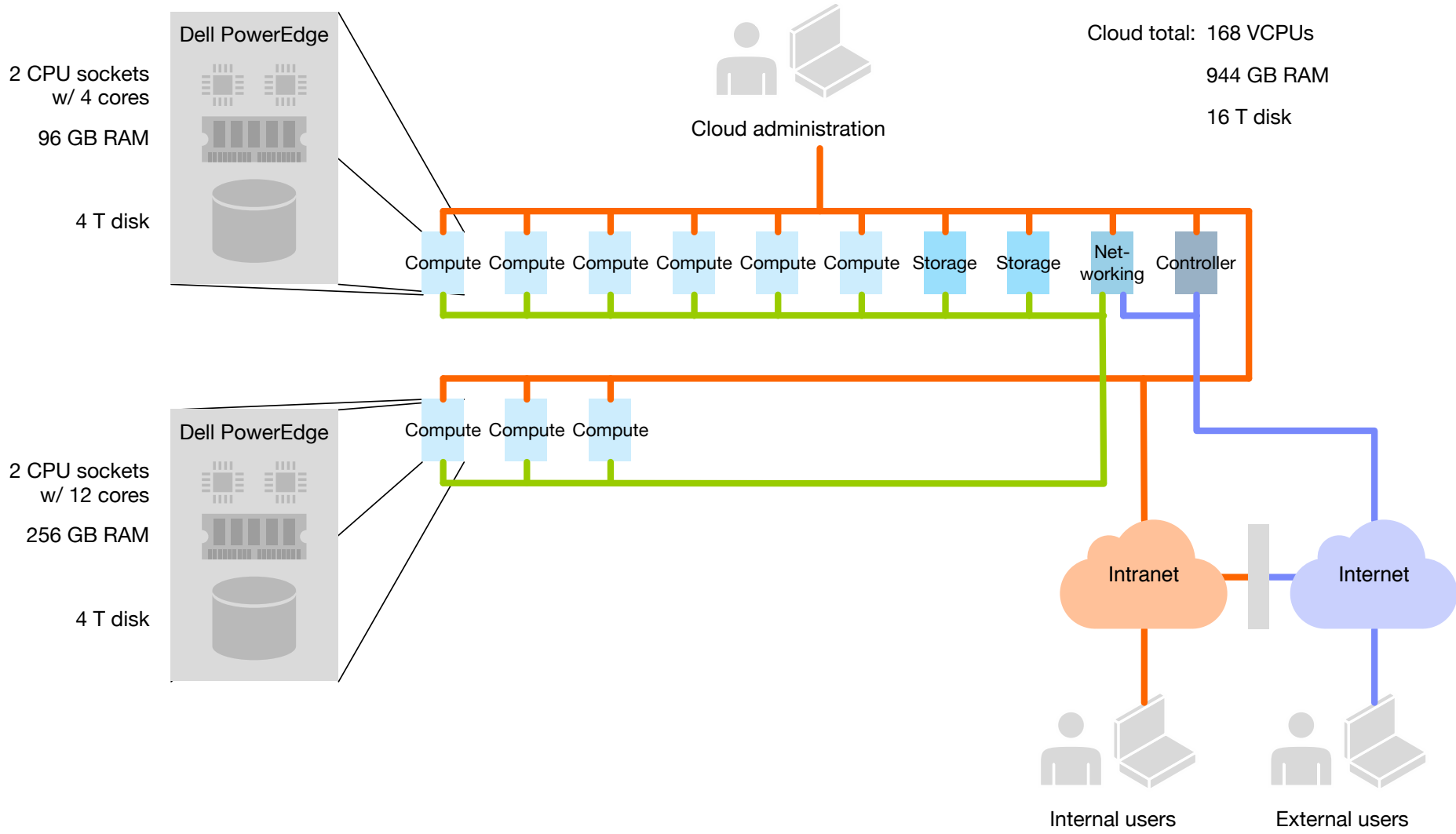
- Red Hat
  - RDO — Install proof of concept on single node, extend to more nodes later
  - TripleO — Deploy a production cloud
  - TryStack — Play with OpenStack in a sandbox. Sandbox is provided as a cloud service
- Ubuntu
  - OpenStack Autopilot — Canonical OpenStack installer for production and test environments
  - BootStack — Installation service by Canonical
- SUSE
  - SUSE OpenStack Cloud Admin
- Mirantis
  - Fuel — Deployment and management tool for OpenStack
- Rackspace
  - Openstack-Ansible — Ansible scripts and extensions for deploying production clouds
- …

HEIG-Cloud

- Goal: Deploy a private OpenStack cloud at HEIG-VD for research and teaching
  - Want to run clusters for Big Data analysis (Hadoop, Spark)
  - Want to deploy a private PaaS (CloudFoundry)
  - ...
- Hardware: 13 Dell PowerEdge servers
- Had previously installed Havanna
- In August 2015 started to install Kilo

HEIG-Cloud

Deployment architecture

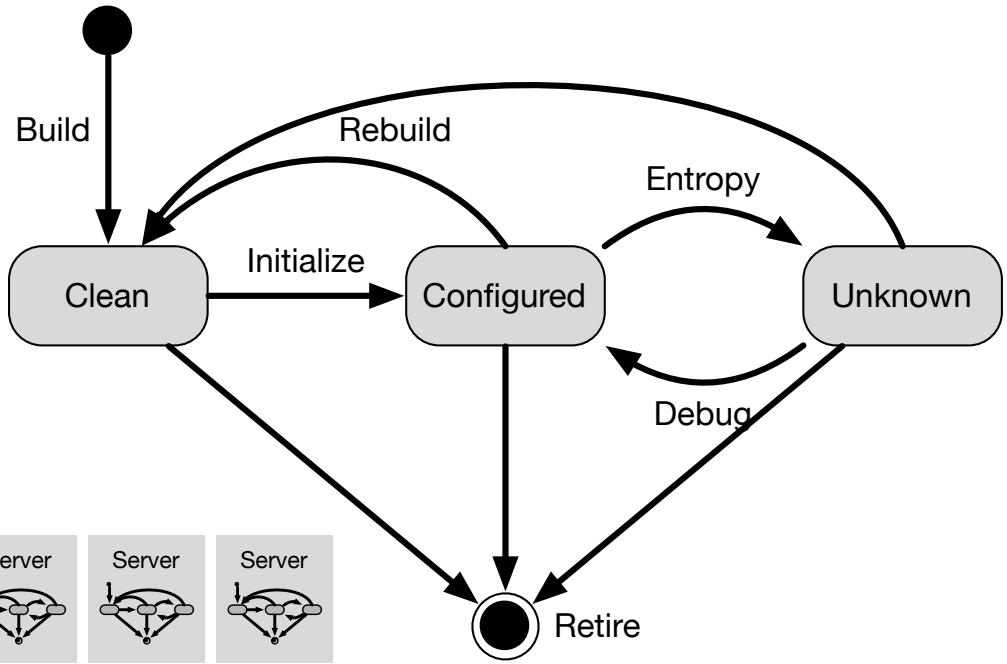


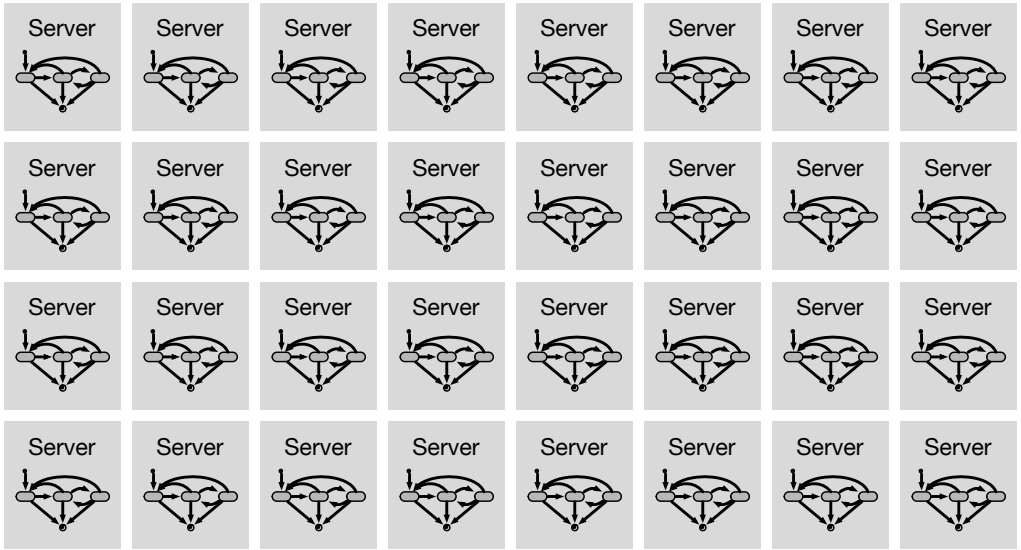
Configuration management

Introduction

- When the number of managed machines exceeds a handful, manual configuration becomes too cumbersome.
- Configuration management tools automate the installation and configuration of software.

Machine life cycle





Configuration management
Tools

- Some popular configuration management tools:
  - CFEngine
    - Started in 1993 by Mark Burgess at Oslo University
    - CFEngine 3 released 2009

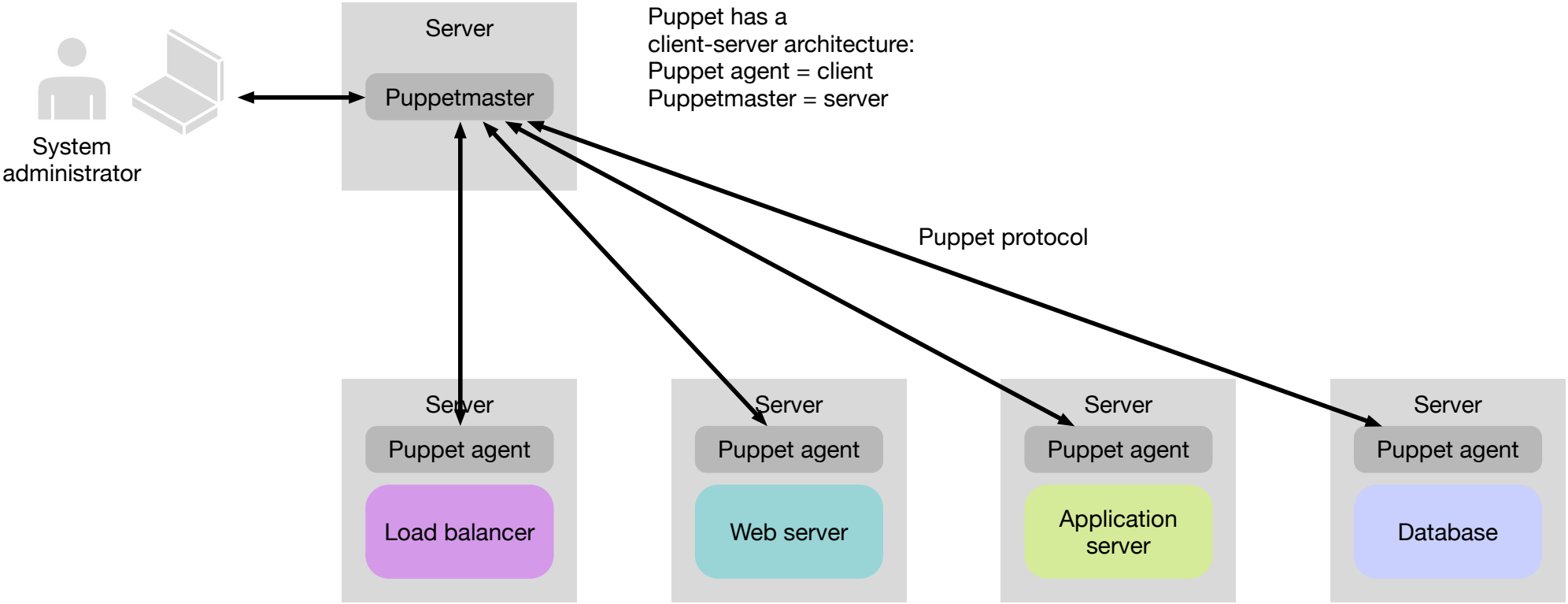
  - Puppet
    - Started in 2005 by Luke Kanies
    - Written in Ruby, uses Ruby domain-specific language

  - Chef
    - Started in 2009 by Adam Jacob and people from Amazon
    - Written in Ruby and Erlang, uses Ruby domain-specific language

  - Ansible
    - Started in 2012 by Michael DeHaan
    - Written in Python

Configuration management

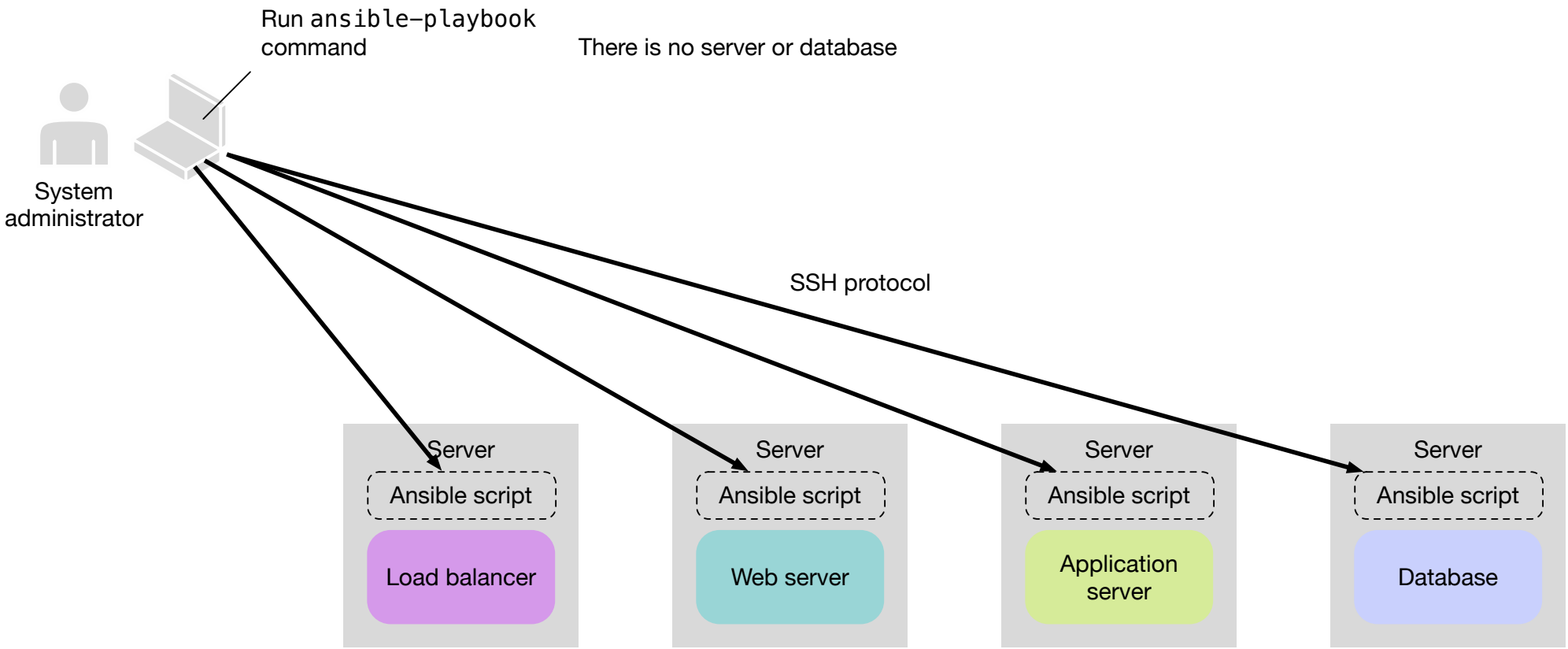
Client-server architecture (for example Puppet)



Managed machines must be prepared by installing Puppet agent.

Configuration management

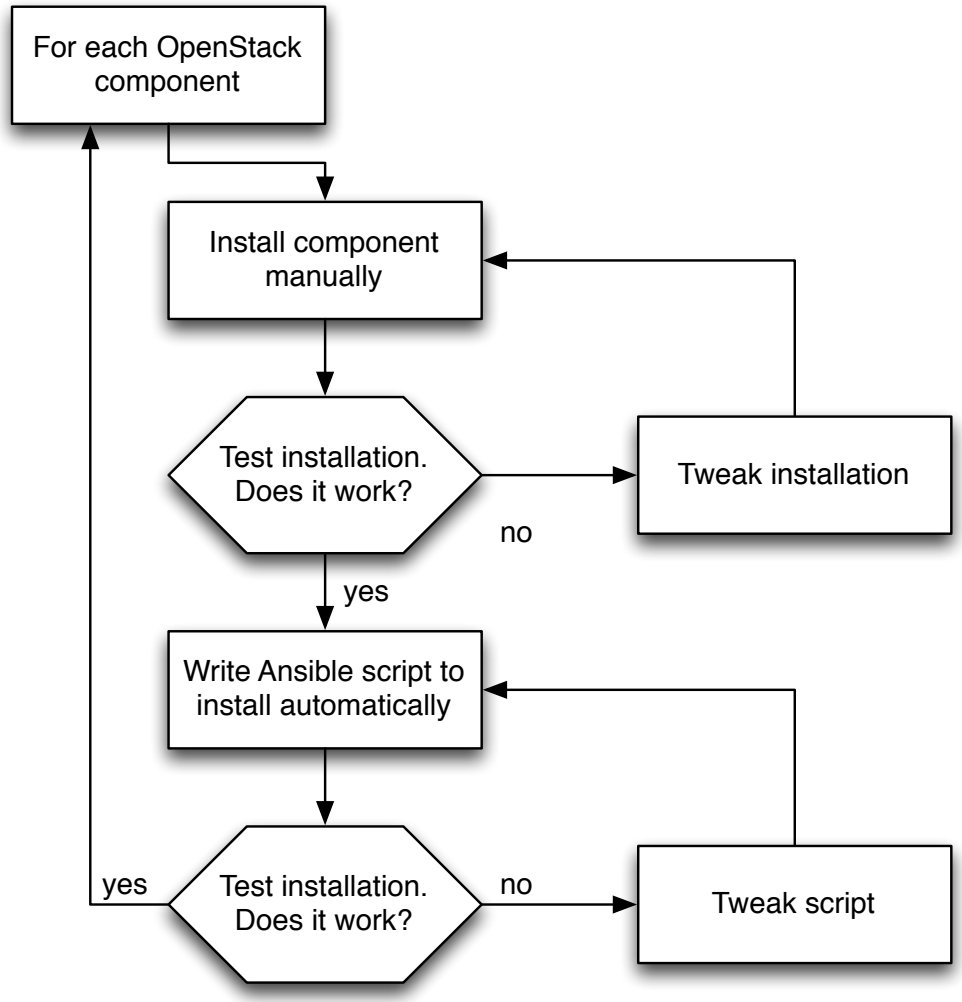
Serverless architecture (Ansible)

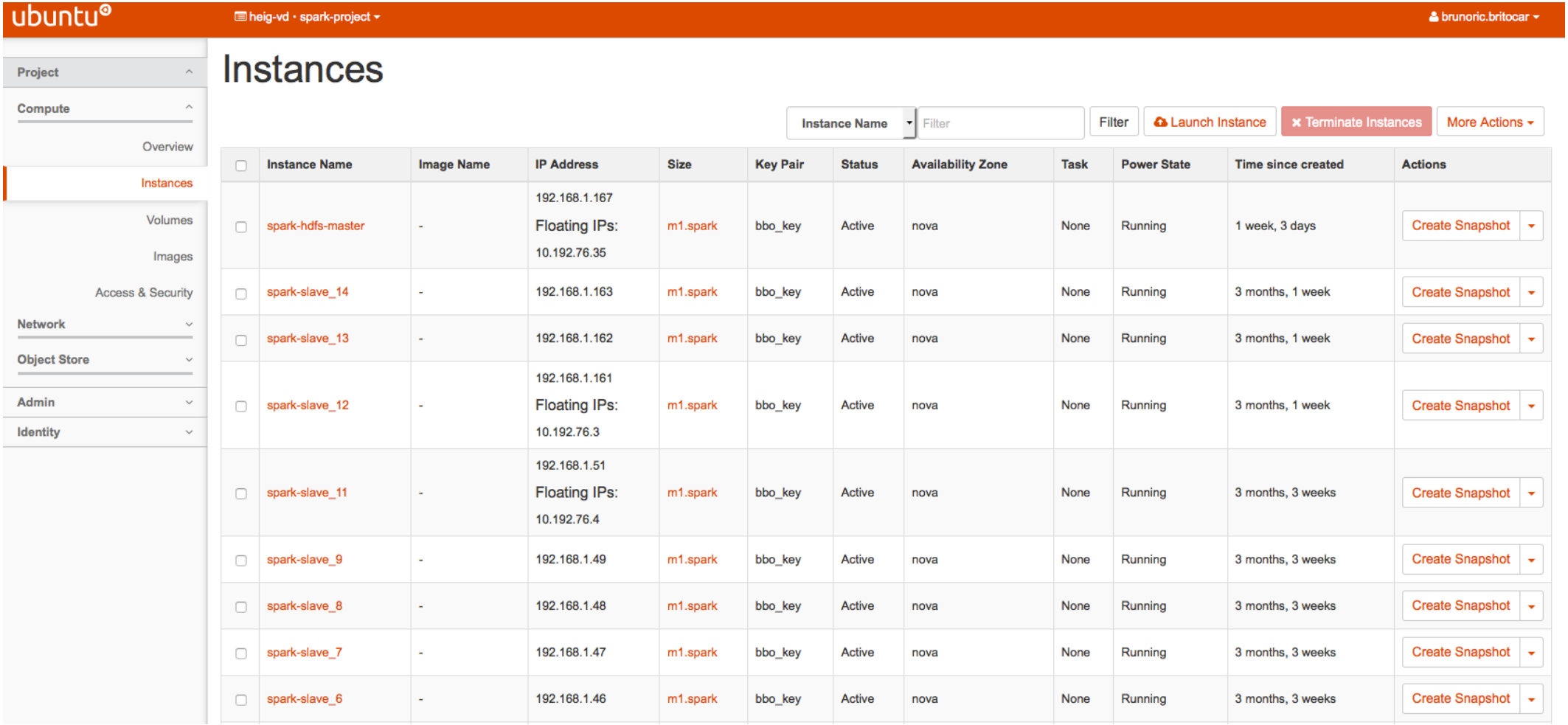


No special software must be installed on managed machines. No agents need to run.
Uses existing Python interpreter and SSH.

Installing OpenStack with Ansible

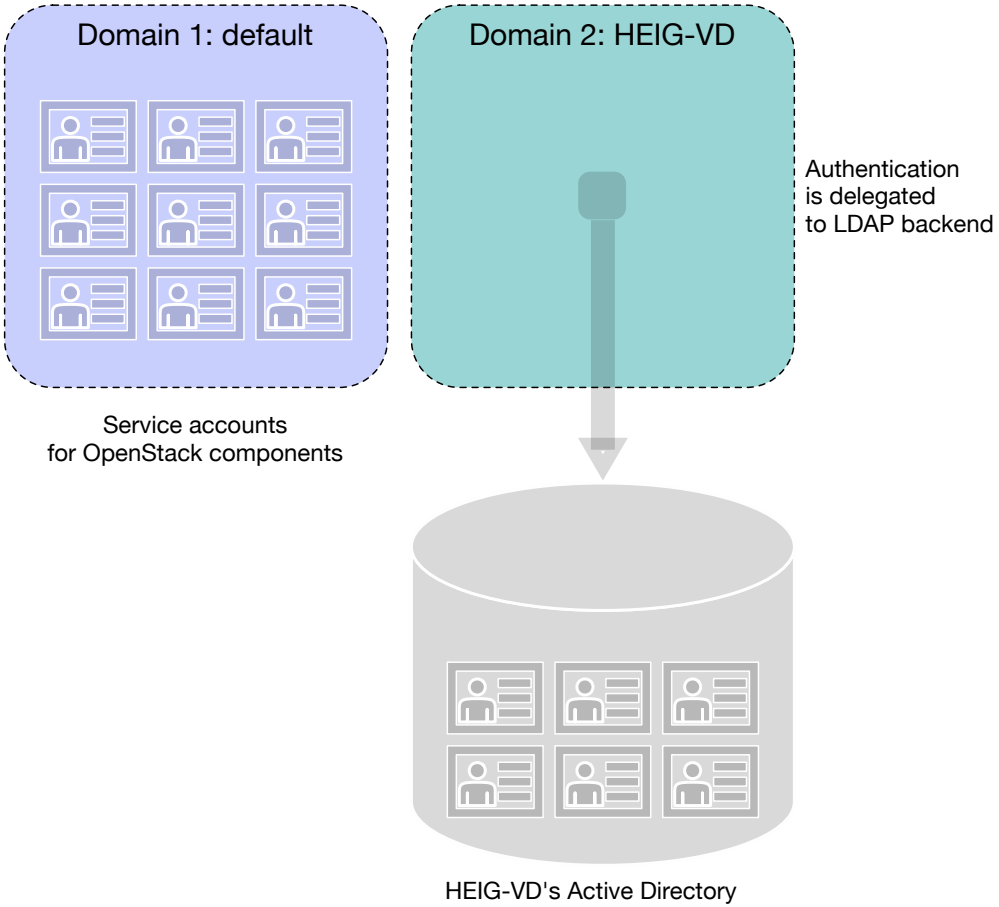
- Unfortunately the installation procedure changes significantly with each new release of OpenStack
- Difficult to re-use Ansible scripts for previous versions
  - Developed new scripts from scratch




ubuntu
heig-vd · spark-project
brunoric.britocar
Project
Compute
Overview
Instances
Volumes
Images
Access & Security
Network
Object Store
Admin
Identity
Instances
Instance Name
Filter
Filter
Launch Instance
Terminate Instances
More Actions
Instance Name | Image Name | IP Address | Size | Key Pair | Status | Availability Zone | Task | Power State | Time since created | Actions
spark-hdfs-master | - | 192.168.1.167 Floating IPs: 10.192.76.35 | m1.spark | bbo_key | Active | nova | None | Running | 1 week, 3 days | Create Snapshot
spark-slave_14 | - | 192.168.1.163 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 1 week | Create Snapshot
spark-slave_13 | - | 192.168.1.162 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 1 week | Create Snapshot
spark-slave_12 | - | 192.168.1.161 Floating IPs: 10.192.76.3 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 1 week | Create Snapshot
spark-slave_11 | - | 192.168.1.51 Floating IPs: 10.192.76.4 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 3 weeks | Create Snapshot
spark-slave_9 | - | 192.168.1.49 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 3 weeks | Create Snapshot
spark-slave_8 | - | 192.168.1.48 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 3 weeks | Create Snapshot
spark-slave_7 | - | 192.168.1.47 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 3 weeks | Create Snapshot
spark-slave_6 | - | 192.168.1.46 | m1.spark | bbo_key | Active | nova | None | Running | 3 months, 3 weeks | Create Snapshot

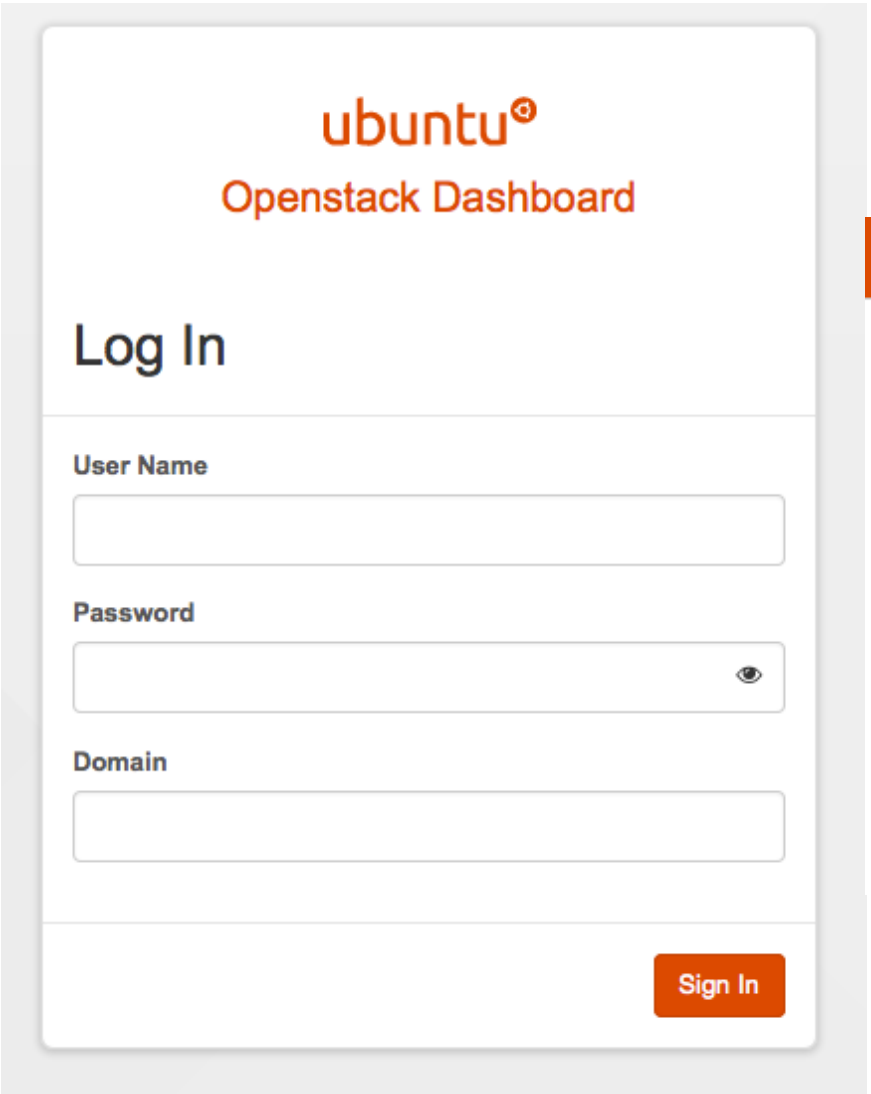
Multi-domain Keystone

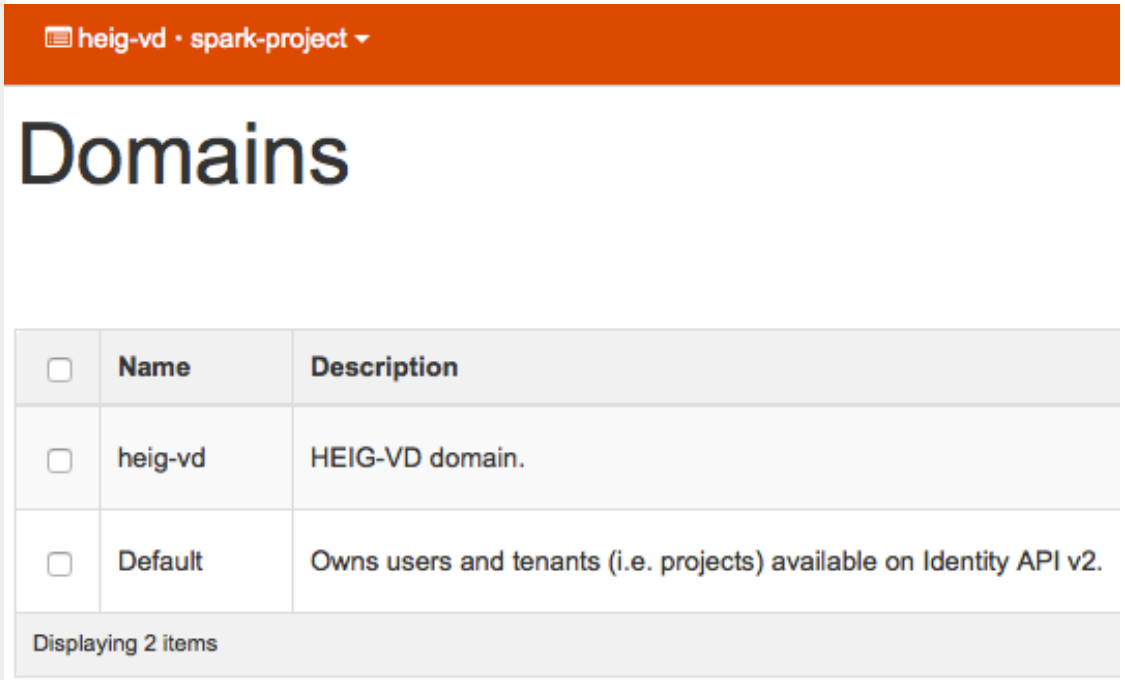
- Keystone performs authentication and authorization of users.
- For authentication the user population can be divided into different domains.
  - Cloud resources in different domains are completely separated from each other.
  - Each domain can be configured differently.
  - Authentication can be delegated to a backend, for example an LDAP server.



- Multiple domains are available in Keystone v3 API
  - **Not all OpenStack projects support the v3 API, or they support it only partially!**

Multi-domain Keystone

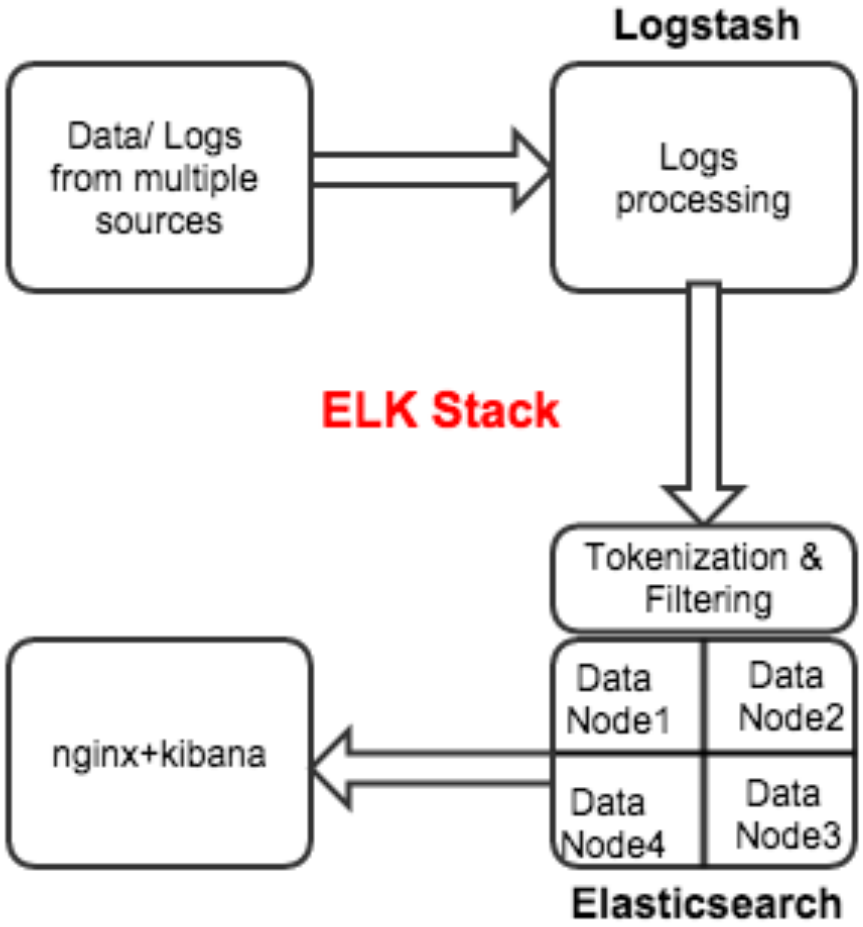




Troubleshooting tools

Elasticsearch + LogStash + Kibana (ELK)

- **Elasticsearch** — Search server based on Lucene
- **Logstash** — General-purpose log management tool to gather logs from multiple sources, process/parse them to a required format and push them to multiple outputs
- **Kibana** - Data visualization plug-in for Elasticsearch

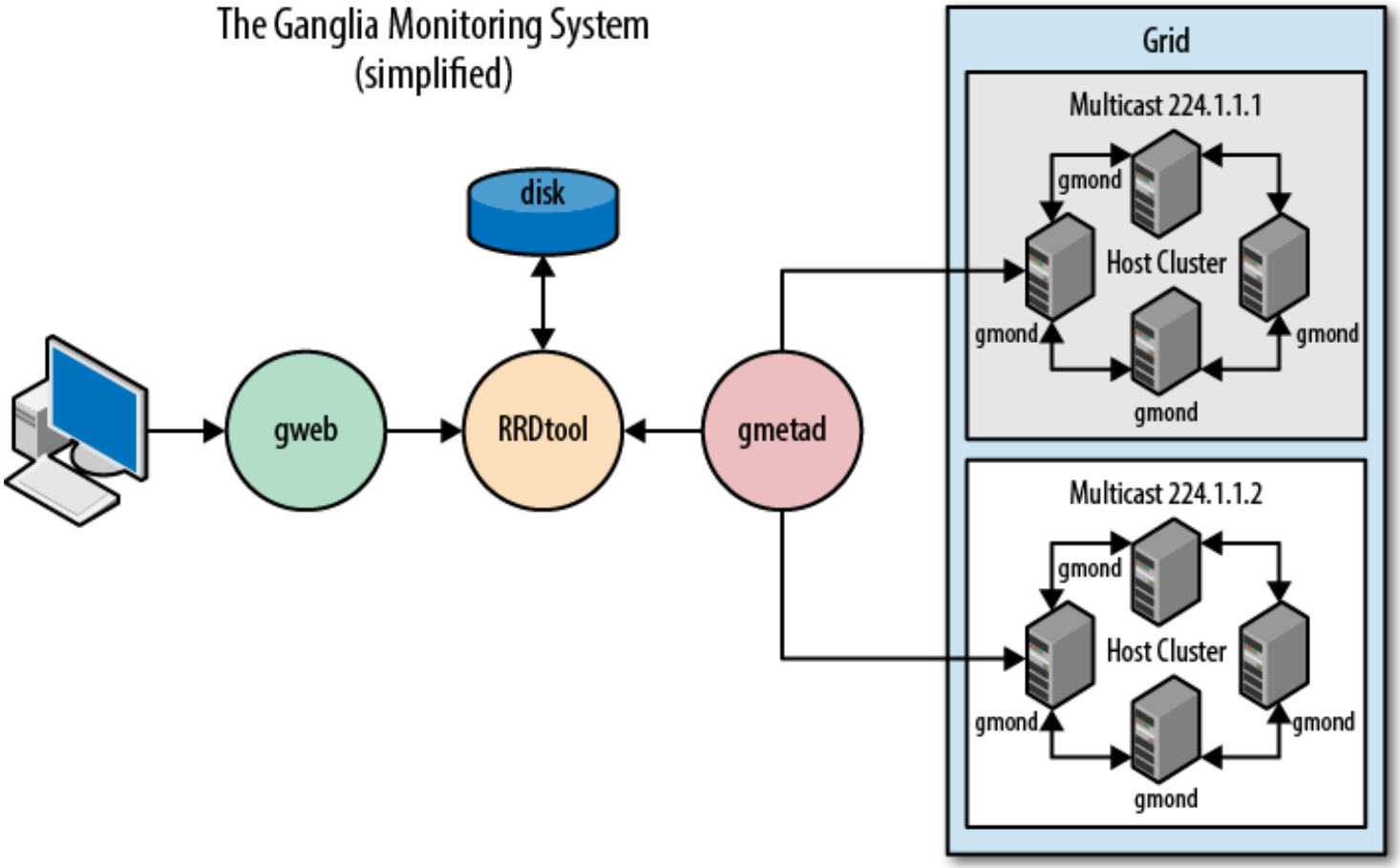


Source: Amit Balode, http://balodeamit.blogspot.ch/2014/12/elk-elasticsearch-logstash-and-kibana.html


error
logstash-*
16 hits
Selected Fields
? _source
Available Fields
@timestamp
t @version
t _id
t _index
_score
t _type
t beat.hostname
t beat.name
count
? fields
t host
t input_type
t message
offset
t source
t syslog_facility
syslog_facility_code
t syslog_severity
syslog_severity_code
t tags
t type
March 17th 2016, 00:00:00.000 - March 17th 2016, 23:59:59.999 — by 30 minutes
Count
15
10
5
0
01:00
04:00
07:00
10:00
13:00
16:00
19:00
22:00
@timestamp per 30 minutes
Time
_source
March 17th 2016, 06:46:07.078
message: 2016-03-17 06:44:47.436 48801 ERROR nova.servicegroup.drivers.db [req-18621f1b-4424-4990-8486-00ec5f7e7c0e - - - - -] Recovered model server connection! @version: 1 @timestamp: March 17th 2016, 06:46:07.078 beat.hostname: IICT-SV3295 beat.name: IICT-SV3295 count: 1 fields: - input_type: log offset: 1,535,375 source: /var/log/nova/nova-compute.log type: syslog host: IICT-SV3295 tags: beats_input_codec_plain_applied, _grokparsefailure syslog_severity_code: 5 syslog_facility_code: 1 syslog_facility: user-level syslog_severity: notice _id: AVODHCti-tAnSUR5i4fy _type: syslog _index: logstash-2016.03.17 _score:
March 17th 2016, 06:45:04.157
message: 2016-03-17 06:44:52.421 37743 ERROR nova.servicegroup.drivers.db [req-bdaeb032-06c2-44ec-b5c6-881d1de24683 - - - - -] Recovered model server connection! @version: 1 @timestamp: March 17th 2016, 06:45:04.157 beat.hostname: IICT-SV3296 beat.name: IICT-SV3296 count: 1 fields: - input_type: log offset: 882 source: /var/log/nova/nova-compute.log type: syslog host: IICT-SV3296 tags: beats_input_codec_plain_applied, _grokparsefailure syslog_severity_code: 5 syslog_facility_code: 1 syslog_facility: user-level syslog_severity: notice _id: AVODGzVX-tAnSUR5i4R6 _type: syslog _index: logstash-2016.03.17 _score:
March 17th 2016, 06:44:55.334
message: 2016-03-17 06:44:52.421 8898 ERROR nova.servicegroup.drivers.db [req-32388d92-315f-4e52-8f14-bed54ee089f5 - - - - -] Recovered model server connection! @version: 1 @timestamp: March 17th 2016, 06:44:55.334 beat.hostname: IICT-SV3294 beat.name: IICT-SV3294 count: 1 fields: - input_type: log offset: 12,394 source: /var/log/nova/nova-compute.log type: syslog host: IICT-SV3294 tags: beats_input_codec_plain_applied, _grokparsefailure syslog_severity_code: 5 syslog_facility_code: 1 syslog_facility: user-level syslog_severity: notice _id: AVODGw-n-tAnSUR5i4P- _type: syslog _index: logstash-2016.03.17 _score:

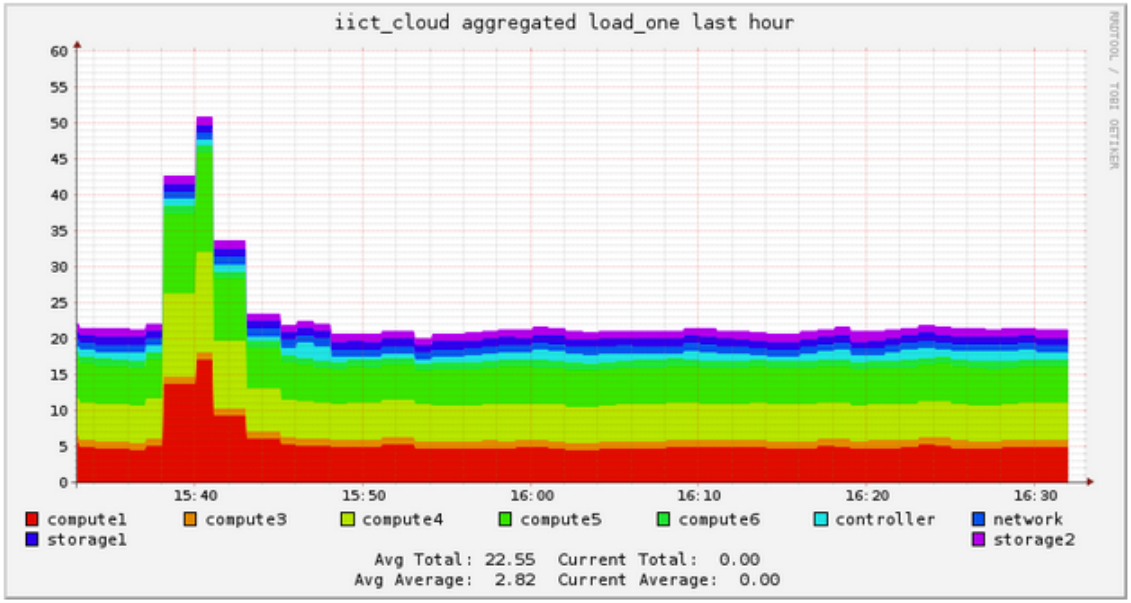
Monitoring
Ganglia

- Ganglia is a scalable distributed monitoring system for clusters
- Components:
  - **gmond** — Monitorig daemon installed on every server to be monitored
  - **gmetad** — Daemon on the master node that collects data from all the gmond daemons
  - **RDDtool** — (Round-robin database tool) Creates a database with circular buffer
  - **gweb** — Web-based user interface



Source: Matt Massie — Monitoring with Ganglia — O'Reilly Media

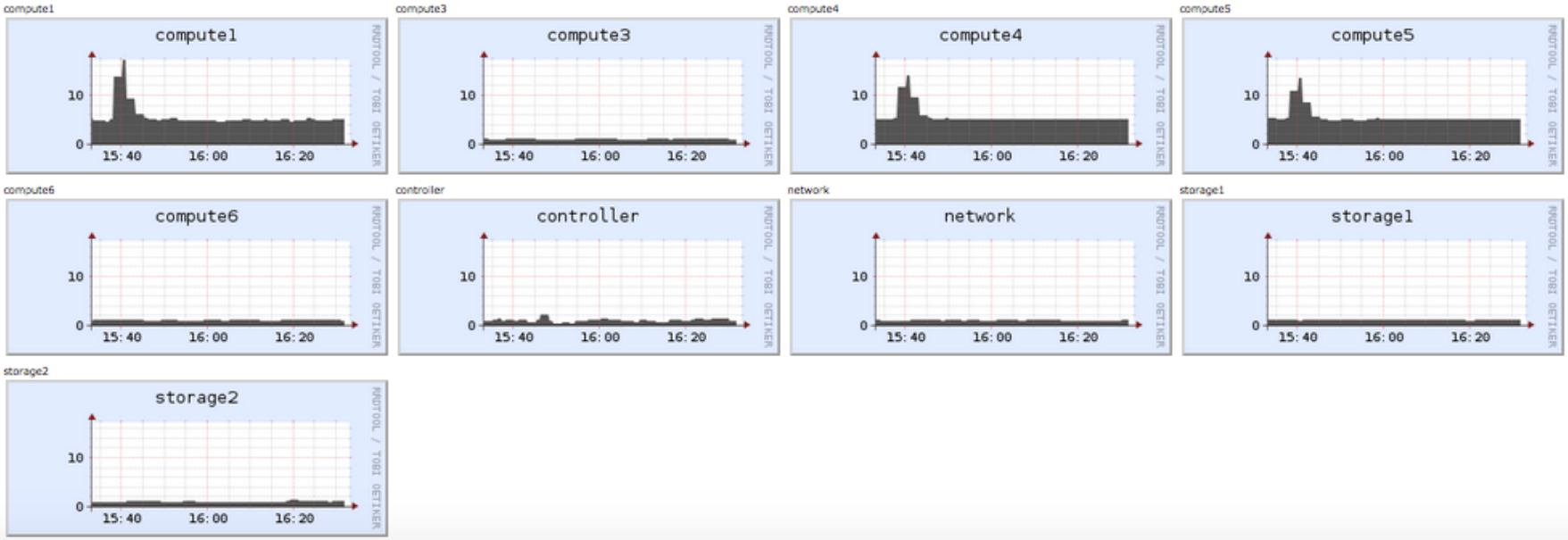
Stacked Graph - load_one



iict_cloud **load_one** last **hour** sorted **by name**

Metric load_one Show Hosts Scaled: Auto Same None Size small Columns 4 (0 = metric + reports)

Show only nodes matching Filter Max graphs to show all Sorted ascending descending by name



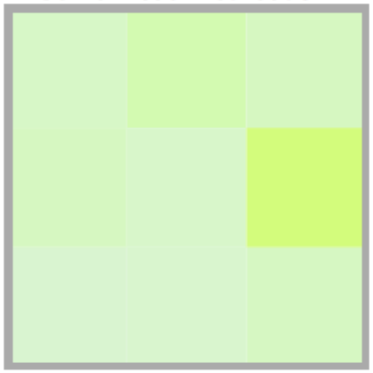
Grid > iict_cloud > --Choose a Node

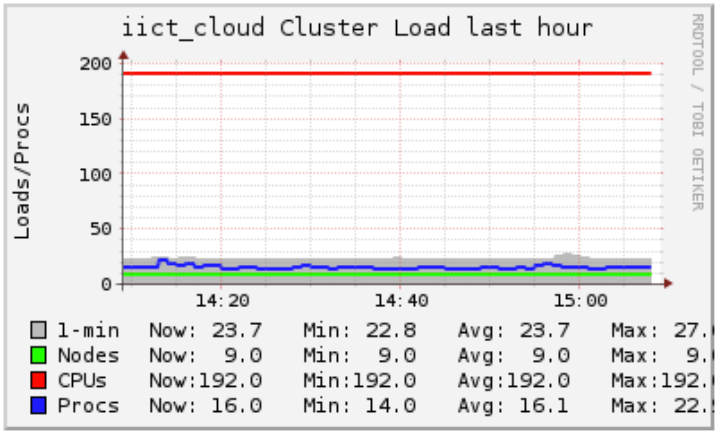
Overview of iict_cloud @ 2015-12-21 15:08

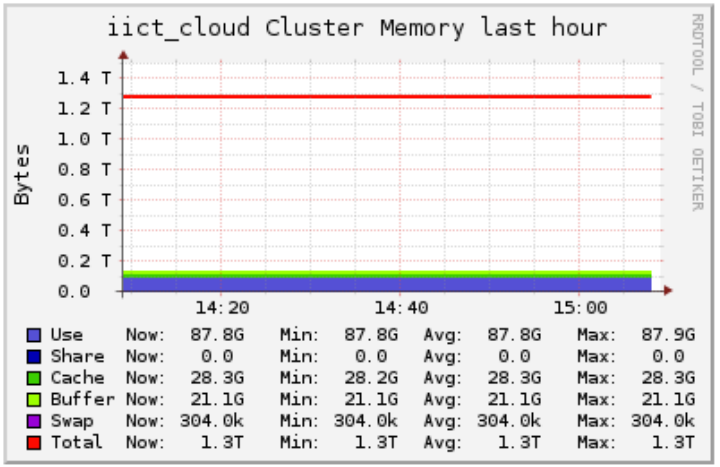
CPUs Total: **192**
Hosts up: **9**
Hosts down: **0**

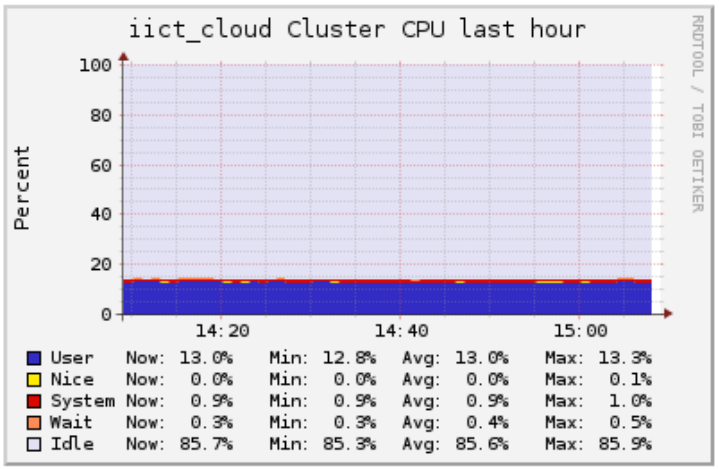
Current Load Avg (15, 5, 1m):
13%, 12%, 12%
Avg Utilization (last hour):
12%

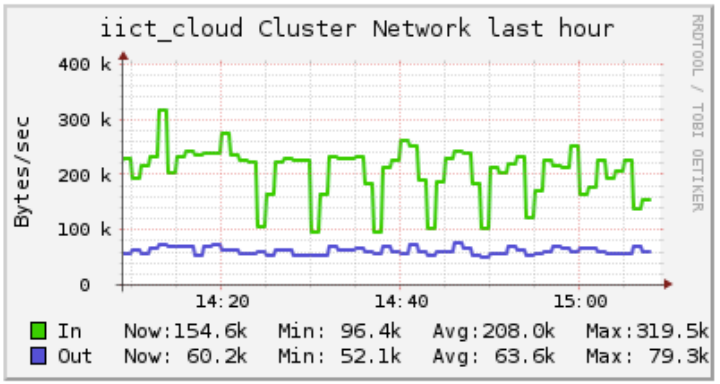
Server Load Distribution



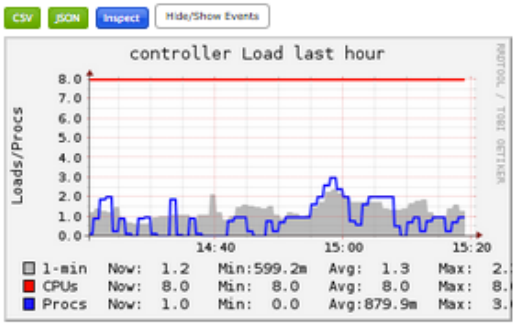


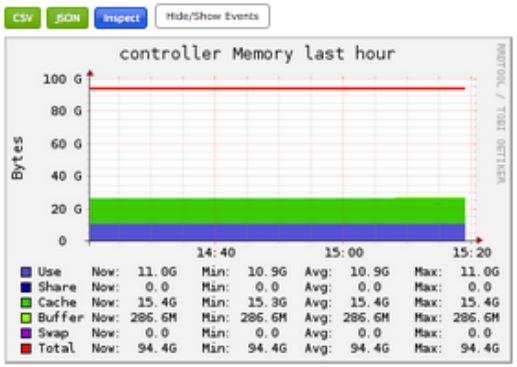


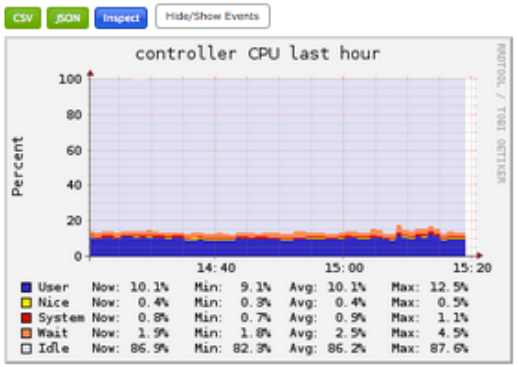


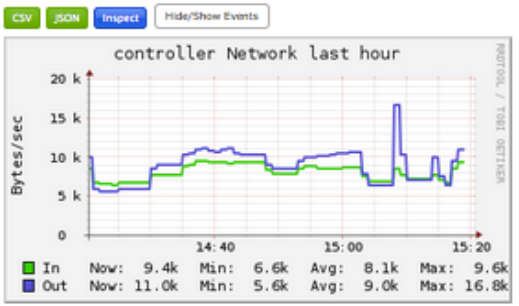


Host Overview









controller **graphs** (23) last **hour** sorted **by name** Columns 2 Size small

Expand All Metric Groups | Collapse All Metric Groups | Timeshift Overlay | Jump To Metric Group...

cpu metrics (6)

Thank you for your time

heig-vd

MARCEL GRAF
Dr.-Ing.
Professeur

INSTITUT DES TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION (IICT)

Haute École d'Ingénierie et de Gestion du Canton de Vaud
Route de Cheseaux 1 - CP 521 - 1401 Yverdon-les-Bains

Prof.: +41 24 557 27 70
marcel.graf@heig-vd.ch - www.heig-vd.ch

Hes·so
Haute Ecole Spécialisée de Suisse occidentale

heig-vd

BRUNO BRITO CARVALHO
Ing. HES
Collaborateur scientifique

INSTITUT DES TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION (IICT)

Haute École d'Ingénierie et de Gestion du Canton de Vaud
Route de Cheseaux 1 - CP 521 - 1401 Yverdon-les-Bains

bruno.britocarvalho@heig-vd.ch
www.heig-vd.ch

Hes·so
Haute Ecole Spécialisée de Suisse occidentale

More information at the HEIG-Cloud blog at http://heig-cloud.github.io/