IBM OpenStack

Justin Anderson

Cesar Ramirez

Jack Nagel

John Rensberger

Company & Sponsor

- International Business Machines Corporation
  - Hardware and Software Solutions
  - Mainframe Computers to Nanotechnology
  - Platinum Sponsor for OpenStack
- Sponsor: Jeremy Berg
- Mentors: Michael Fork, Lance Bragstad, Mathew Odden, Adam Reznechek

What is OpenStack?

- OpenStack is an Infrastructure as a Service (IaaS) application that controls pools of:
  - Compute Nodes
  - Storage Nodes
  - Networking Resources
- Users control and deploy these resources through a web portal
- Similar to AWS, Google Compute, and Azure
- Current release is Grizzly

Background

- Expansion of last year's project
- Goal was to increase awareness and promote use by:
  - Updating current environment
  - Reaching out to faculty and staff for use cases
  - Implementing and delivering use cases

Limitations

- There are plenty of resources but they are only managed by a few people
- Researchers and students have very specific requirements that make the allocation of their resources very time consuming
- Professors need to over-allocate in order to be able to handle peaks

Opportunity

- Improve existing OpenStack infrastructure
- Provide scalable solutions and on-demand resources for current C.S. department computing needs
- Ease the use and management of resources
- Cater the environment for the users

Use Cases & Benefits

- Student and Researcher VMs
  - C.S. Systems Administrator
- Hadoop
  - Dr. Ludwig and Dr. Denton
- Education
  - Linux Course and Cloud Course
- Thin-clients and Resource Management
  - College of Engineering Technical Support
- Documentation

How We learned

- Online reading and research
- Discussions with Dr. Denton and Dr. Ludwig
- Practicing Hadoop installations and deployment on Amazon EC2
- Playing around with DevStack (scripted installation of OpenStack)
- Reading thousands of lines of error logs
- Pulling our hair out at times

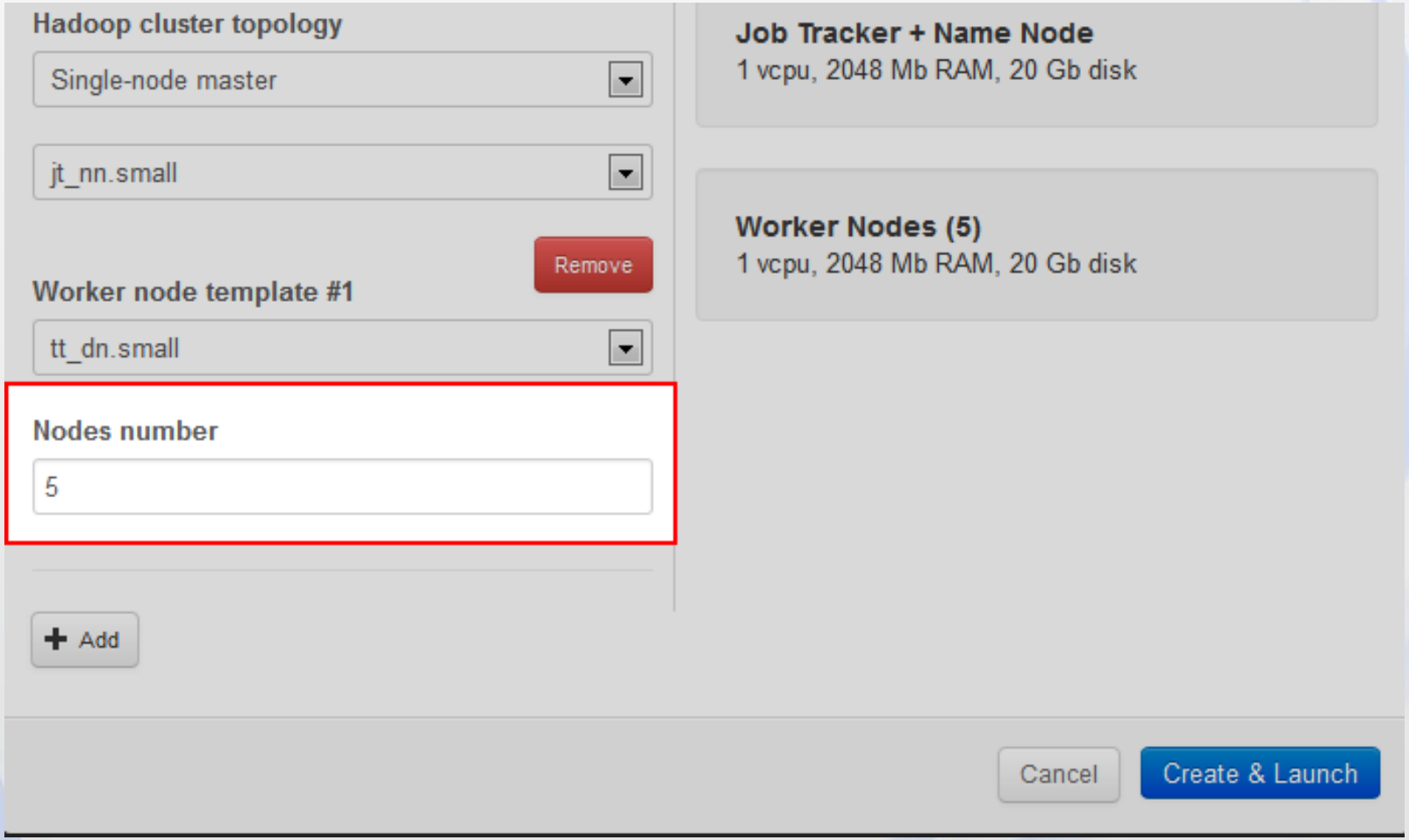
Hadoop in C.S. Department

- Current solution
  - Hadoop cluster utilizing IACC 244 Linux lab
  - Must be installed by sysadmin
  - Available on a first-come, first-serve basis
  - Heavy impact on performance for end users while jobs are running

Hadoop & Savanna

- OpenStack Savanna
  - Elastic Hadoop implementation for OpenStack
  - Provision, deploy, and destroy clusters on-demand
  - Control over node configuration using templates
  - Integration with Horizon (OpenStack dashboard)
  - REST API
  - Planned features: integration with Swift (object storage)

Hadoop & Savanna

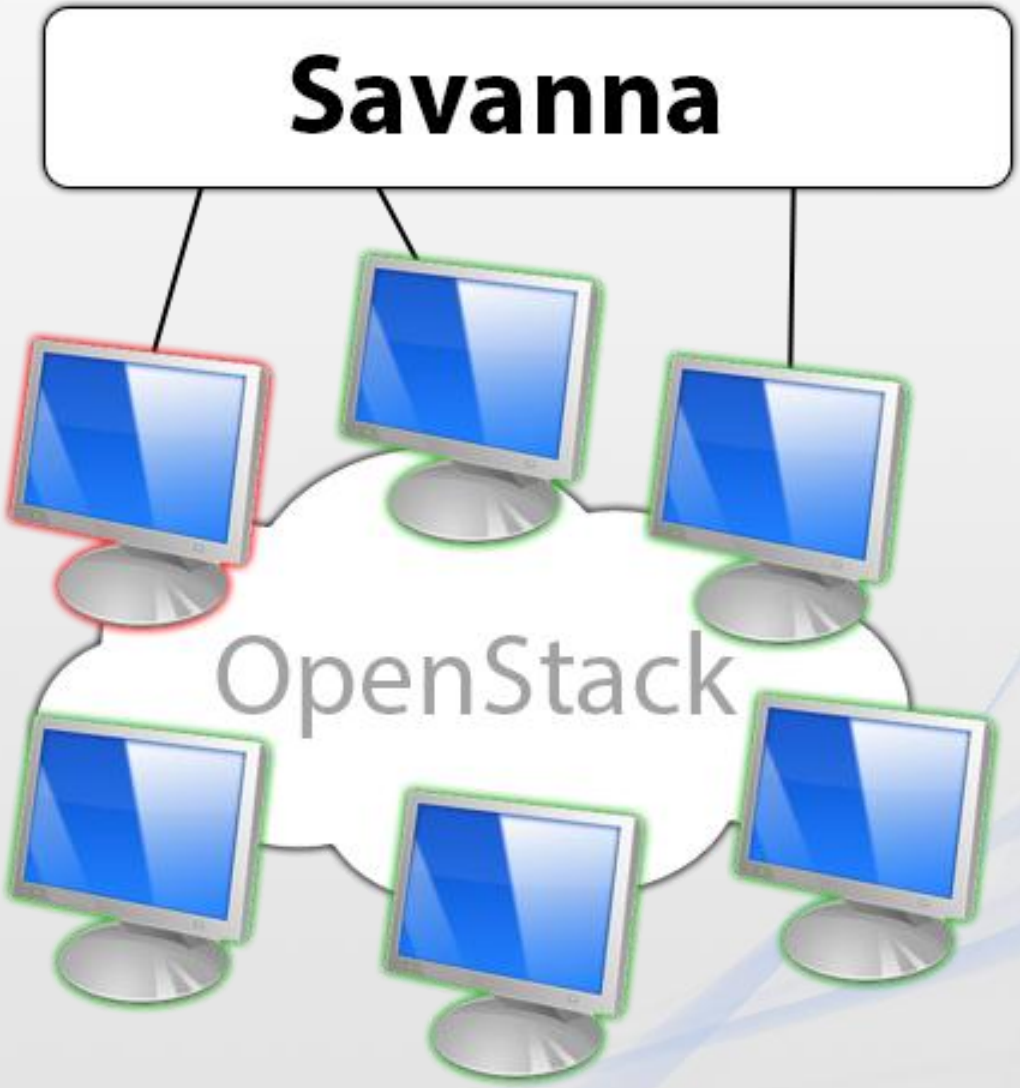


User Sends Request to Savanna

Savanna

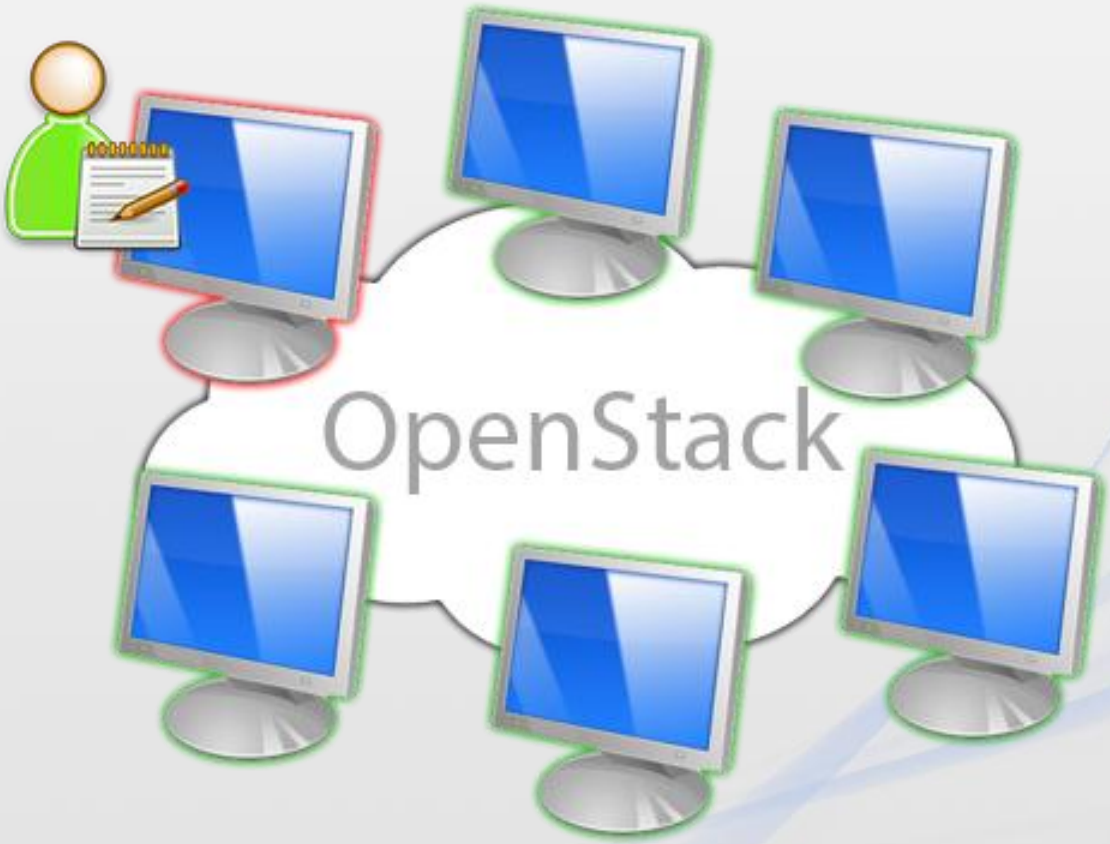
OpenStack

Savanna spawns VM's for master and worker nodes



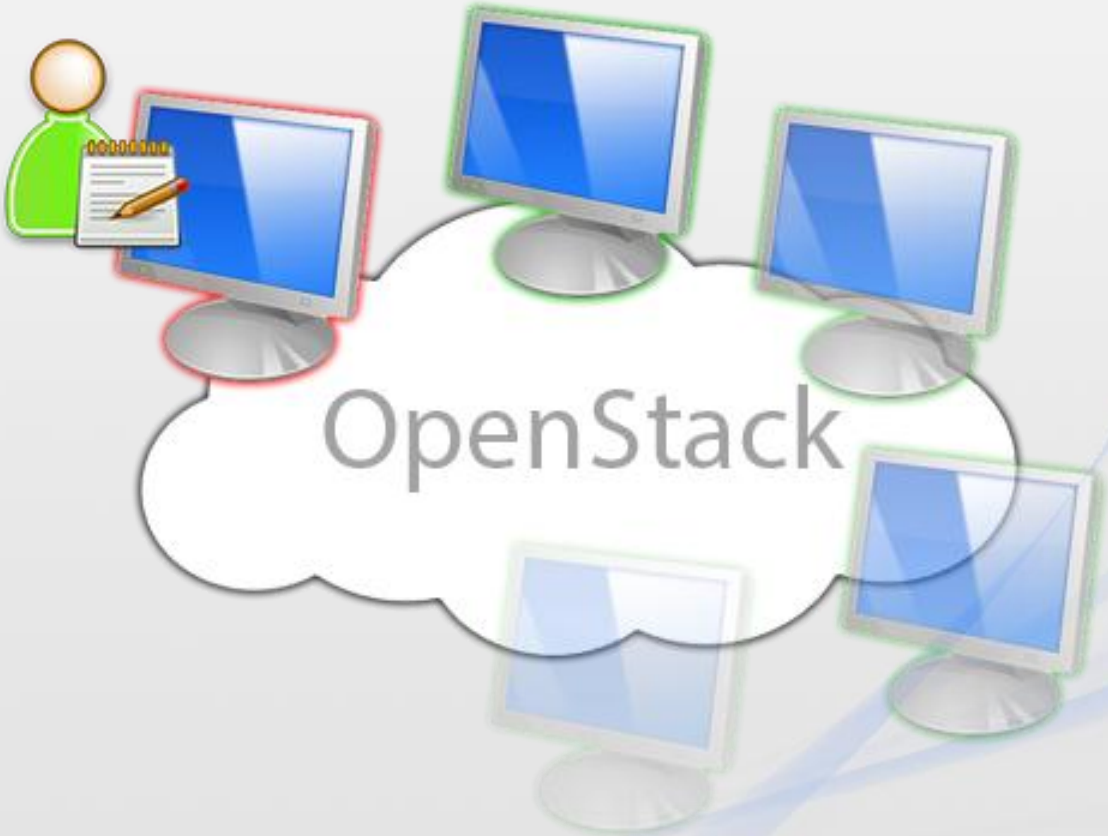
User logs into master and runs a job

Savanna



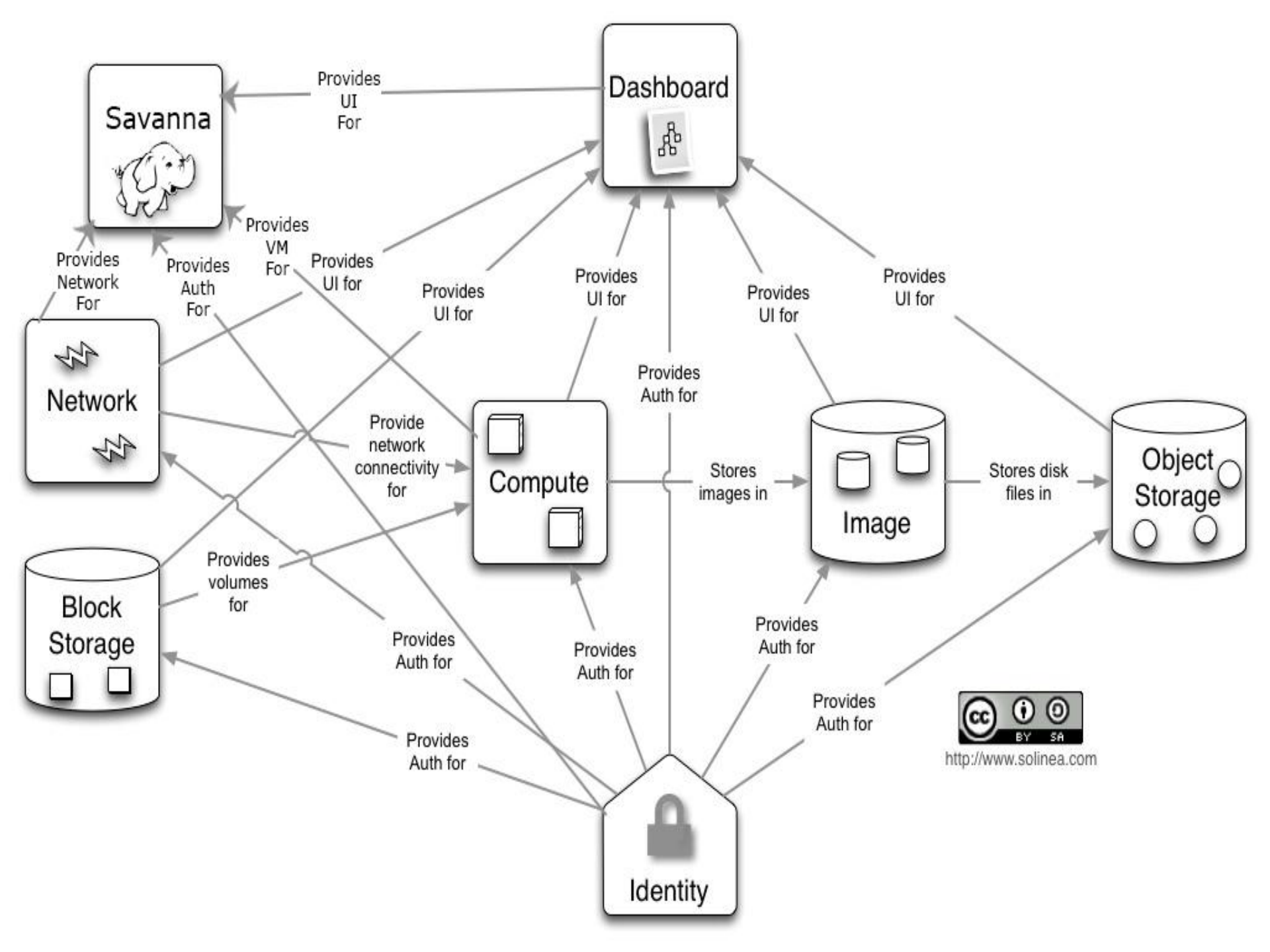
After user gets results, the user can delete the cluster.

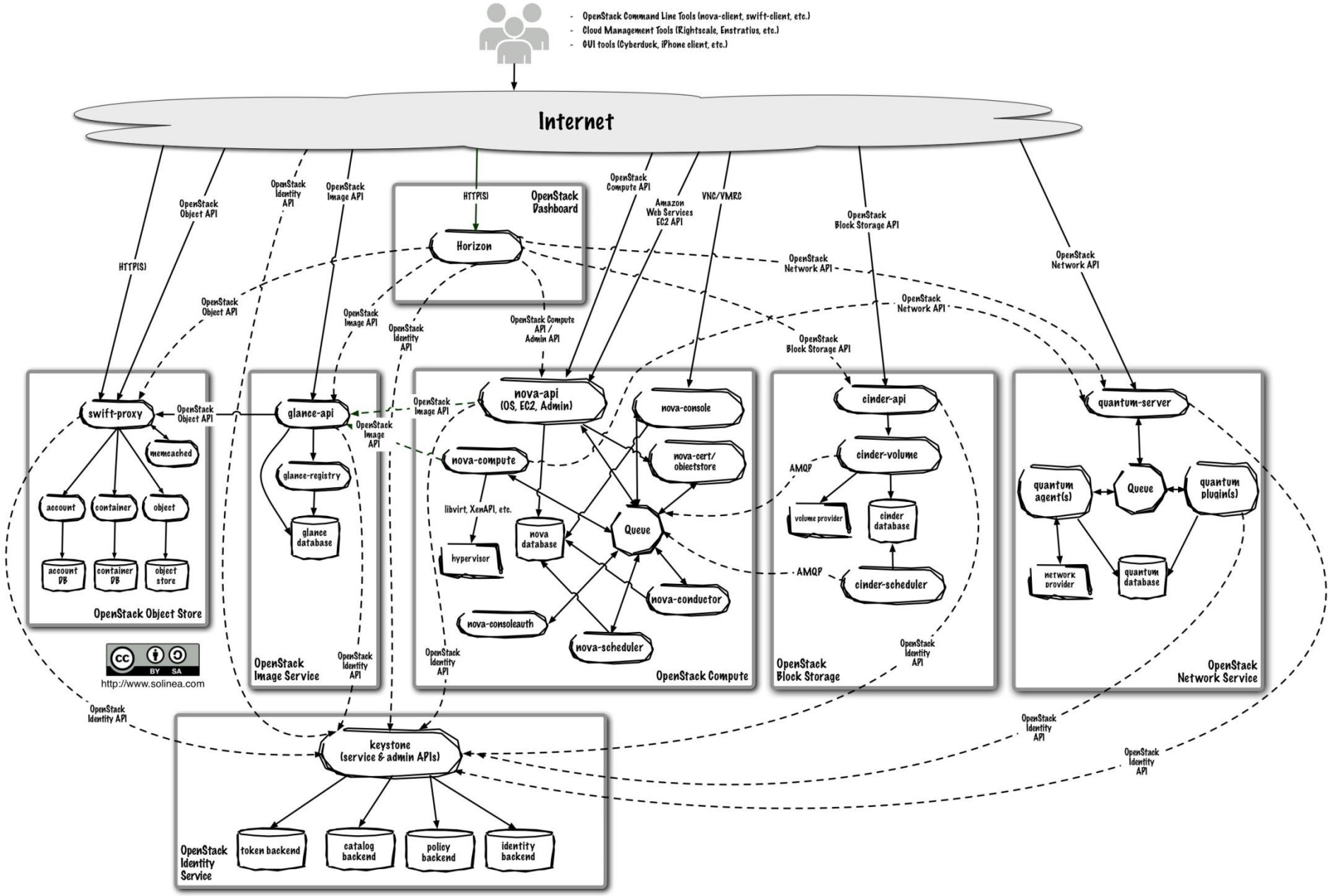
Savanna



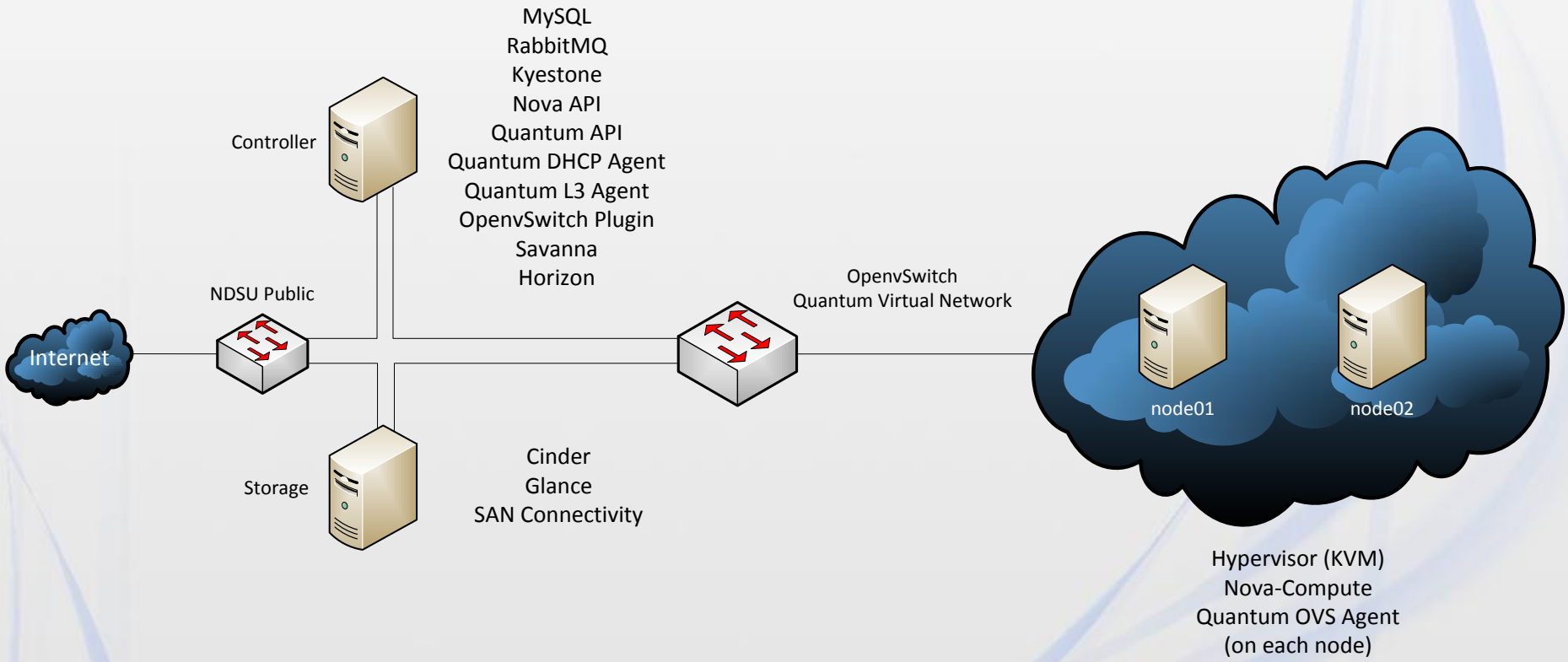
Hadoop and Savanna

- How does this help?
  - Fast setup and teardown of clusters (elastic)
  - Run multiple Hadoop jobs at the same time
  - Per-tenant resource quotas
  - Isolate clusters with per-tenant networks


Savanna
Provides UI For
Dashboard
Provides Network For
Provides Auth For
Provides VM For
Provides UI for
Provides UI for
Provides UI for
Provides UI for
Provides UI for
Network
Provide network connectivity for
Compute
Provides Auth for
Stores images in
Image
Stores disk files in
Object Storage
Block Storage
Provides volumes for
Provides Auth for
Provides Auth for
Provides Auth for
Provides Auth for
Provides Auth for
Identity
http://www.solinea.com


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

Our Deployment



Currently

- Our current deployment has 20 VCPUs, 45 GB of RAM and 4TB of storage
- Resources are managed on-demand
- Virtual networking and floating IPs
- Web-based dashboard

We are not lying to you!

DEMO

Issues & Risks

- Core Issue: Integrating all the services and getting all features to work together
  - Basic features were achieved quickly but more advanced ones took significantly more time
  - Lack of documentation and occasional bugs
- Core Risk: Updates and Bugs
  - OpenStack is still young and is rapidly developed including significant new features
  - There are frequent updates for bug fixes

Lessons

- OpenStack should be installed from source
  - More effort to install, but it's educational and easier to maintain
- There's never too much documentation
  - It needs to be organized
- Understand the use-cases from companies that run OpenStack

Next Steps

- Finish LDAP integration for authentication
- Keep up with OpenStack's rapid updates
- Follow development of Savanna
- Improve awareness of OpenStack and its benefits within the department

Ask us questions!

Q & A

Extra Material!

BACK UP

Schedule

- Learn about OpenStack
- Use-cases and requirements
- Deploy OpenStack
  - Deploy service
  - Debugging
  - Integration
  - Documentation

Status Update

- Things that we finished
  - All OpenStack components are working
  - Deploy Hadoop clusters using Savanna
  - Documented our process for future usage
- Things that we did not finish
  - C.S. LDAP integration with AD