Compute Canada ARC
ARC in Canada

- It started with supercomputers!
- Now we run
  - Supercomputers
  - IaaS
  - BaaS
  - Big Data
  - VDI
  - And much more..
Organizations

- Acenet: NB, NS, PEI, NL
- Compute Ontario
- Calcul Quebec
- Westgrid: AB, SK, MB, BC
- Compute Canada
Canadian ARC in numbers

- **4 Supercomputers** (#59, #190 & #271 in Top 500 list)
  - ~190,000 physical CPU cores
- **5 major Cloud deployments**
  - ~15,000 physical CPU cores
  - ~7 PetaBytes CEPH storage
  - ~1.4 Million VMs created since commissioned
- ~60 PetaBytes of online storage,
- ~60 PetaBytes of nearline and backups
Adoption of Openstack

- Realizing the need!
- Started in 2014 with two Clouds
- Icehouse release
- Services added gradually
- Upgrades ... Or Migrations
- Now 5 Major Clouds
National cloud platform

- x86 Clouds
- Accelerators: GPUs/FPGA
- CEPH backends
- 10/25GbE interconnect
- 100Gb Backbone links
National Cloud Team

- Responsible for:
  - Architecture, deployment and operations of the national cloud platform
  - User support
- Distributed Team
- Meet the team:
Evolution of Deployment - West Cloud Phase 1

Limited hardware footprint:
- 6 control plane nodes
- 13 storage nodes
- 40 compute nodes

Installation of OpenStack and Ceph:
- All packages were downloaded and installed to each machine
- Manually configured on each host individually
## West Cloud - Phase 1

<table>
<thead>
<tr>
<th></th>
<th>2014 West Cloud - P1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Cores</td>
<td>640</td>
</tr>
<tr>
<td>Compute Nodes</td>
<td>40</td>
</tr>
<tr>
<td>Storage (Raw)</td>
<td>500 TB</td>
</tr>
<tr>
<td>Storage (Usable)</td>
<td>180 TB</td>
</tr>
</tbody>
</table>
Evolution of Deployment - West Cloud Phase 2

Many more pieces of hardware to configure:
- 8 control plane nodes
- 18 storage nodes
- 280 compute nodes

Significantly more complex:
- Many more nodes to configure
- Started with self generated Bash Scripts
- Moved on to using in-house developed Ansible plays
West Cloud - Phase 2

<table>
<thead>
<tr>
<th></th>
<th>2014 West Cloud - P1</th>
<th>2016 West Cloud - P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Cores</td>
<td>640</td>
<td>7600</td>
</tr>
<tr>
<td>Compute Nodes</td>
<td>40</td>
<td>280</td>
</tr>
<tr>
<td>Storage (Raw)</td>
<td>500 TB</td>
<td>2.2 PB</td>
</tr>
<tr>
<td>Storage (Usable)</td>
<td>180 TB</td>
<td>750 TB</td>
</tr>
</tbody>
</table>
Evolution of Deployment - Arbutus Cloud

Design Decisions:

- Brand new cloud on brand new hardware
- Migration of West Cloud projects to Arbutus Cloud
- Migration of hardware resources
Evolution of Deployment - Arbutus Cloud

Many more pieces of hardware to configure... again:
- 14 control plane nodes
- 44 storage nodes
- 350 compute nodes

Significantly more complex:
- OpenStack Ansible Project
- Ceph Ansible Project
# Our Current Installation - Arbutus Cloud

<table>
<thead>
<tr>
<th></th>
<th>2014 West Cloud - P1</th>
<th>2016 West Cloud - P2</th>
<th>2018 Arbutus Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Cores</td>
<td>640</td>
<td>7600</td>
<td>9000</td>
</tr>
<tr>
<td>Compute Nodes</td>
<td>40</td>
<td>280</td>
<td>350</td>
</tr>
<tr>
<td>Storage (Raw)</td>
<td>500 TB</td>
<td>2.2 PB</td>
<td>5.8 PB</td>
</tr>
<tr>
<td>Storage (Usable)</td>
<td>180 TB</td>
<td>750 TB</td>
<td>4.4 PB</td>
</tr>
</tbody>
</table>
Evolution of Deployment - Arbutus Cloud

Not only did the amount of hardware change the underlying technology changed as well.

Storage changed from a triple replicated storage scheme to an erasure coded scheme.

- Usable storage in triple replication was $\frac{1}{3}$ of raw capacity (2.2 PB meant 750 TB usable).
- Usable storage with erasure coding is roughly $\frac{3}{4}$ of raw capacity (5.8 PB means 4.4 PB usable).

Introduction of new compute nodes enables higher performance and more features for the end user with each upgrade cycle.
Deployment Tools

Physical Provisioning: **xCAT**

- IPMI BMC Control
- Remote console
- Remote power control
- Remote rebuild

OpenStack Installation and Orchestration: **OpenStack Ansible**

- Stages, deploys, and configures OpenStack
- Required additional glue code and site-specific configuration plays
Development Tools and Process

Revision Control: GitLab

- The single source of truth
- Separate feature branches for every issue in progress
- Protected Deployment branches for dev, prod, remote sites
- Continuous Integration testing automatically on every commit

Stated Goal: The entire deployment should be reproducible at any time with just the data in Git

- Proven with continuous ongoing rebuilds of dev cloud from its deployment branch
- Removes the “iceberg of unknown configuration” that plagues many cloud deployments
Observability

Metric Collection: Prometheus

- Lightweight pull-only exporters written in Go
- Hyper-efficient time-series data store server
- Durable storage into InfluxDB

Visualization: Grafana

- Community-sourced and custom-built dashboards
- Native Prometheus integration

Alerting: Prometheus / FLARE
National Collaboration

Cloud sites across Canada operate using these same techniques.

Architectural decisions are made by consensus across the Cloud National Team.

Cloud National Team members are active contributors to OpenStack projects, including OpenStack Ansible, and serve as members of OpenStack governance committees.

Commitment to revision control shows how cloud sites differ while ensuring a common base.
Thanks for Your Time!
Any Questions?

cloud@computecanada.ca