compute canada

CVMFS

Ryan Taylor
on behalf of
Compute Canada CVMFS National Team
Compute Canada

- 4 regional consortia
- 35 member institutions
- ~200 technical staff
- ~15,000 user accounts
  - 20% growth per year

- All research disciplines supported
- Free access for any researcher at a Canadian institution
Compute Canada

- 5 major national systems, ~15 legacy
  - 200K cores, 22 PF, 70 PB disk
- Nearly all research software on CVMFS
  - CVMFS National Team: 8 members
  - 45 staff members can publish software

<table>
<thead>
<tr>
<th>System</th>
<th>Type</th>
<th>Network</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbutus</td>
<td>Cloud</td>
<td>10 GbE</td>
<td>2016 H2</td>
</tr>
<tr>
<td>Cedar</td>
<td>General</td>
<td>OPA</td>
<td>2017 H1</td>
</tr>
<tr>
<td>Graham</td>
<td>General</td>
<td>EDR IB</td>
<td>2017 H1</td>
</tr>
<tr>
<td>Niagara</td>
<td>Large MPI</td>
<td>EDR IB</td>
<td>2018 H1</td>
</tr>
<tr>
<td>Béluga</td>
<td>General</td>
<td>EDR IB</td>
<td>2019 H1</td>
</tr>
</tbody>
</table>
Compute Canada Software Stack

Available software

600+ scientific applications
4,000+ permutations of version/arch/toolchain

<table>
<thead>
<tr>
<th>Type</th>
<th>Modules</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>5</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>239</td>
</tr>
<tr>
<td>Chemistry</td>
<td>63</td>
</tr>
<tr>
<td>Data</td>
<td>19</td>
</tr>
<tr>
<td>Geo/Earth</td>
<td>23</td>
</tr>
<tr>
<td>Mathematics</td>
<td>82</td>
</tr>
<tr>
<td>MPI libraries</td>
<td>7</td>
</tr>
<tr>
<td>Physics</td>
<td>48</td>
</tr>
<tr>
<td>Various tools</td>
<td>176</td>
</tr>
<tr>
<td>Visualisation</td>
<td>28</td>
</tr>
<tr>
<td>Misc</td>
<td>38</td>
</tr>
</tbody>
</table>

- Two major new clusters with Skylake CPUs
- Built new modules with AVX512 for most packages
- High deduplication
- Further details
Open access to CC repos

- For users’ laptops/desktops, and external organizations
  - software environment has no OS dependence (just Linux)
- Need openness and protection
- Strategies for resilience
  - tiered local disk cache + preloaded alien cache
    - independent access mechanism for major sites
    - sheltered from DDoS or loss of WAN
  - stratum 1 on Cloudflare (openhtc.io)
    - mitigation against DDoS
    - pursuing peering with CANARIE
- Public access: bookmark this page!
Genomics data repositories

- Partnership with C3G
- 15 TB of data to start
  - mix of compressible and incompressible
- Refactoring to integrate with data.galaxyproject.org
  - deployed Canadian stratum 1 replica
- Evaluating gateway for publisher access
- Tiered local disk + writeable alien cache works well
- Server deployment facilitated by Ansible role
  - manages pubkeys, replication, storage, firewalls, etc.
Repositories start out simple...
... and grow over time ...
... increasing complexity.
Config repos provide central manageability ...
... but can’t be shared.
Possible Solution

• Global config repo?

• Gateway could handle publishing access
  • each organization authorized to update certain subpaths
  • possibly some modifications required

• Which stratum 1 servers would replicate it?
  • addresses need to be distributed by RPM

• More unified but also more risk
Pod Security Policy
---
kind: PodSecurityPolicy
apiVersion: policy/v1beta1
metadata:
  name: restricted-cvmfs
spec:
  privileged: false
  allowPrivilegeEscalation: false
  volumes:
  - hostPath
  - secret
  allowedHostPaths:
  - pathPrefix: '/cvmfs'
    readOnly: true

Role
---
kind: Role
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  name: test-role
  namespace: test
spec:
  rules:
  - apiGroups: ['policy']
    resources: ['podsecuritypolicies']
    verbs: ['use']
    resourceNames:
    - 'restricted-cvmfs'

Pod
---
apiVersion: v1
kind: Pod
metadata:
  name: cvmfs-test
spec:
  volumes:
  - name: cvmfsvol
    hostPath:
      path: /cvmfs
      type: Directory
    containers:
    - name: fedora
      image: fedora
      command: [ "sh", "-c", "sleep 1h" ]
      volumeMounts:
      - name: cvmfsvol
        mountPath: /cvmfs
        readOnly: true

- Experimental k8s cluster for ATLAS
  - CVMFS client installed on kubelet nodes
  - Apply role binding, enable PSP admission controller
  - Unprivileged container can mount /cvmfs only

Compute Canada CVMFS - CernVM Workshop 2019
Kubernetes
fetching container images

- Already have CVMFS client on kubelet node
- Thin image distribution via CVMFS
  - efficient, instant container start
  - CVMFS Graph Driver plugin for Docker
- The ideal
  - one CVMFS client on host for both images and software
  - container runtime starts images directly from filesystem
- Or just run local container registry
Discussion