NEP-101 HEP Data-Intensive Distributed Cloud Computing

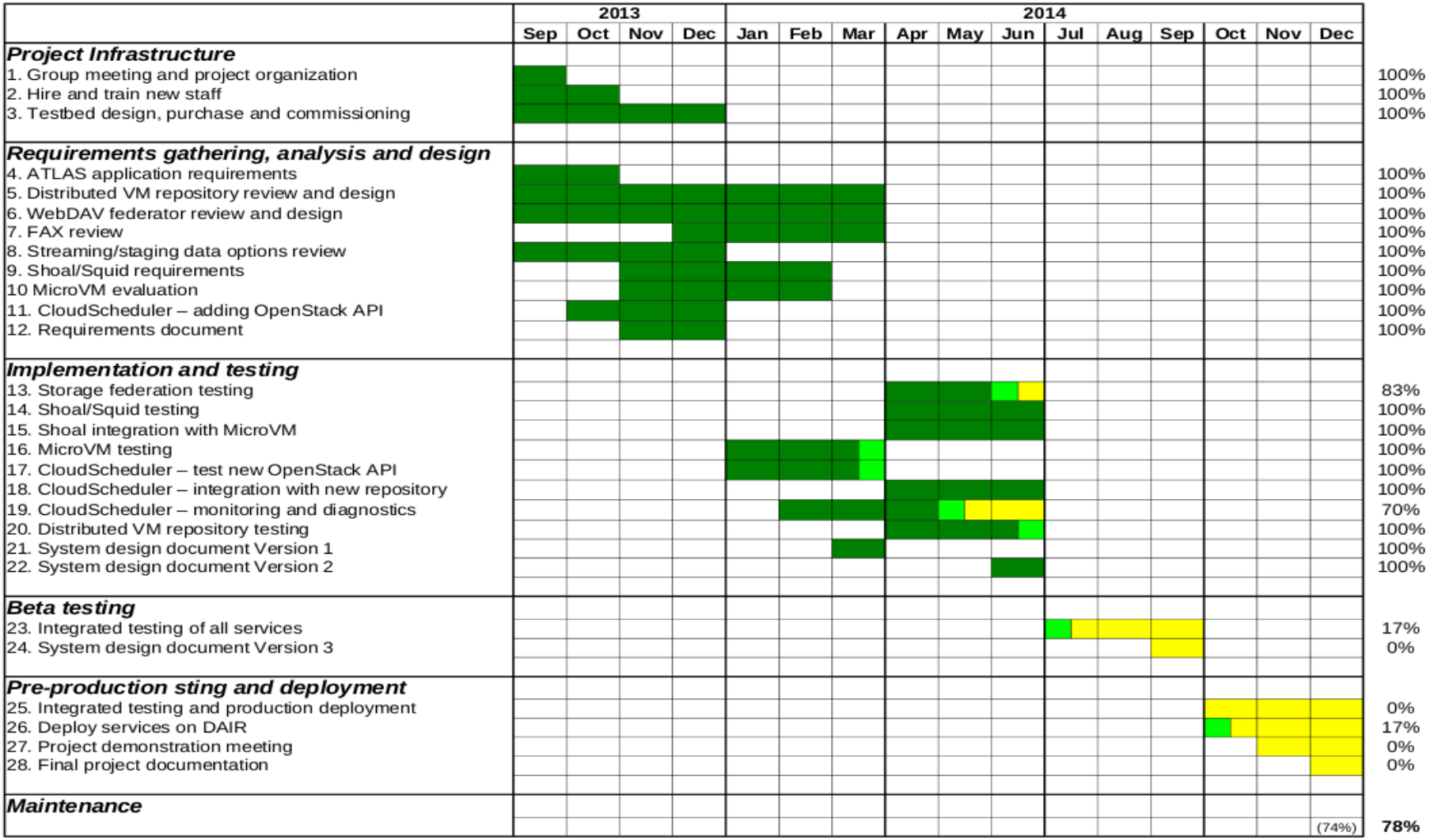
Technical Review

NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Agenda:
  1. Percent Completion
  2. Project Progress Discussion
  3. Collaboration Opportunities
  4. Summary

NEP-101 HEP Data-Intensive Distributed Cloud Computing

1. Percent Completion:

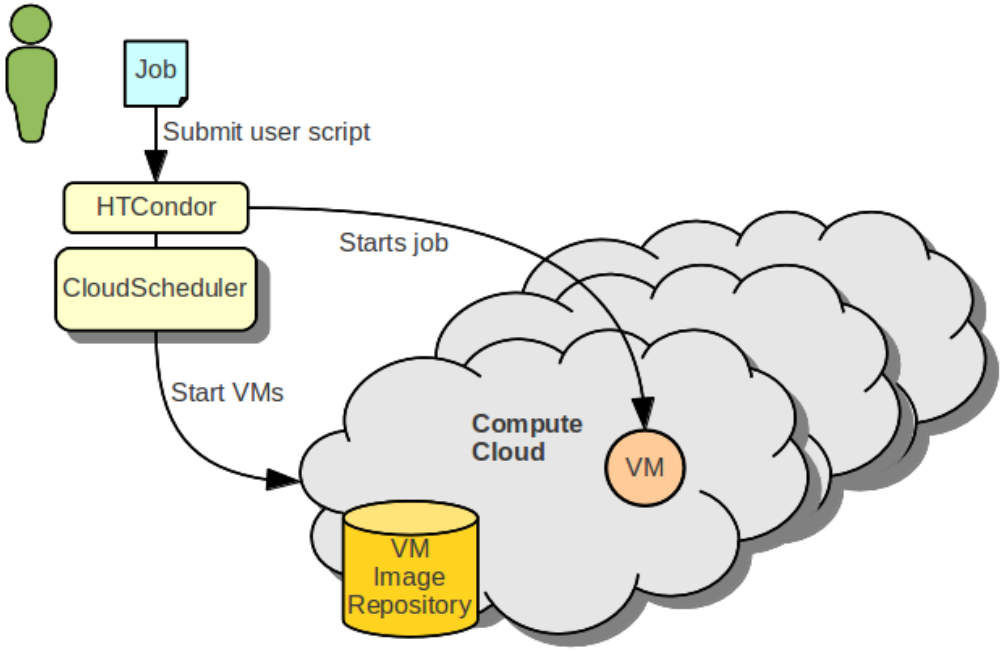


NEP-101 HEP Data-Intensive Distributed Cloud Computing

2. Project Progress Discussion:
   - Batch Services (CloudScheduler)
   - Software Distribution (CVMFS, Shoal, Squid)
   - Storage Federation (UGR)
   - VM Image Distribution (Glint)
   - VM Image Optimization (CernVM3)

NEP-101 HEP Data-Intensive Distributed Cloud Computing

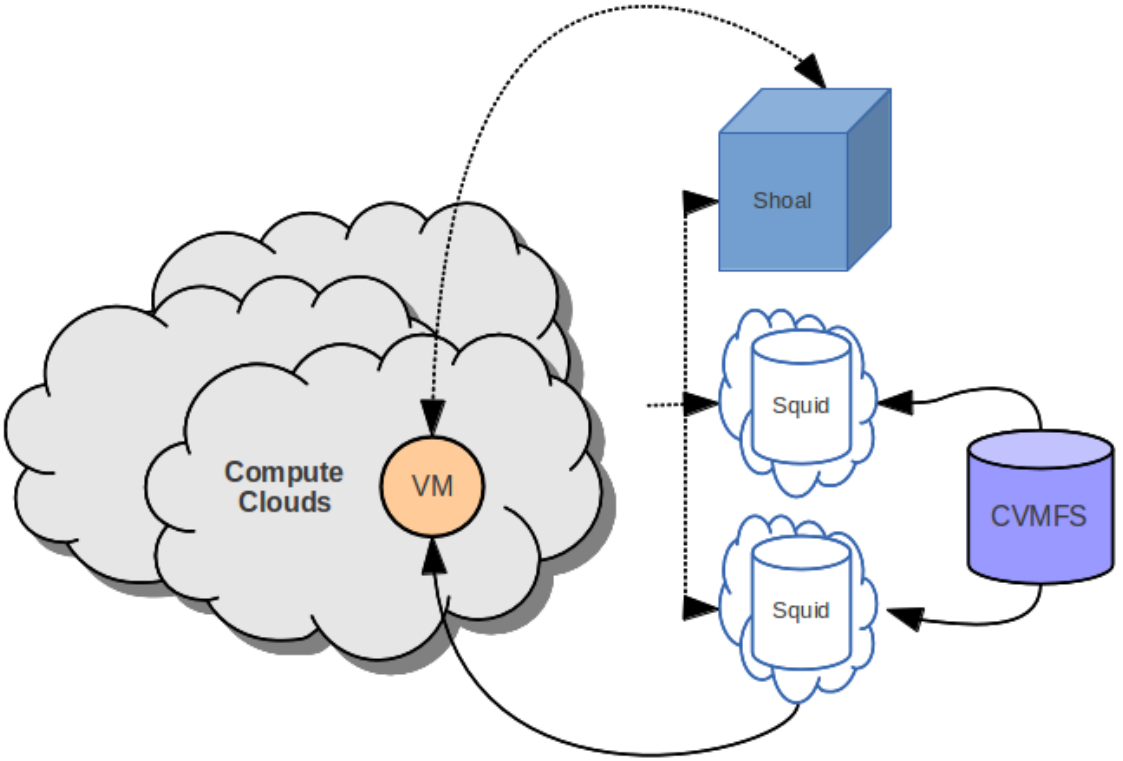
2.1. Project Progress Discussion, Batch Services:



- OpenStack API
  * In production
- New Repository Integration
  * In production
- Monitoring & Diagnostics
  * Work continues

NEP-101 HEP Data-Intensive Distributed Cloud Computing

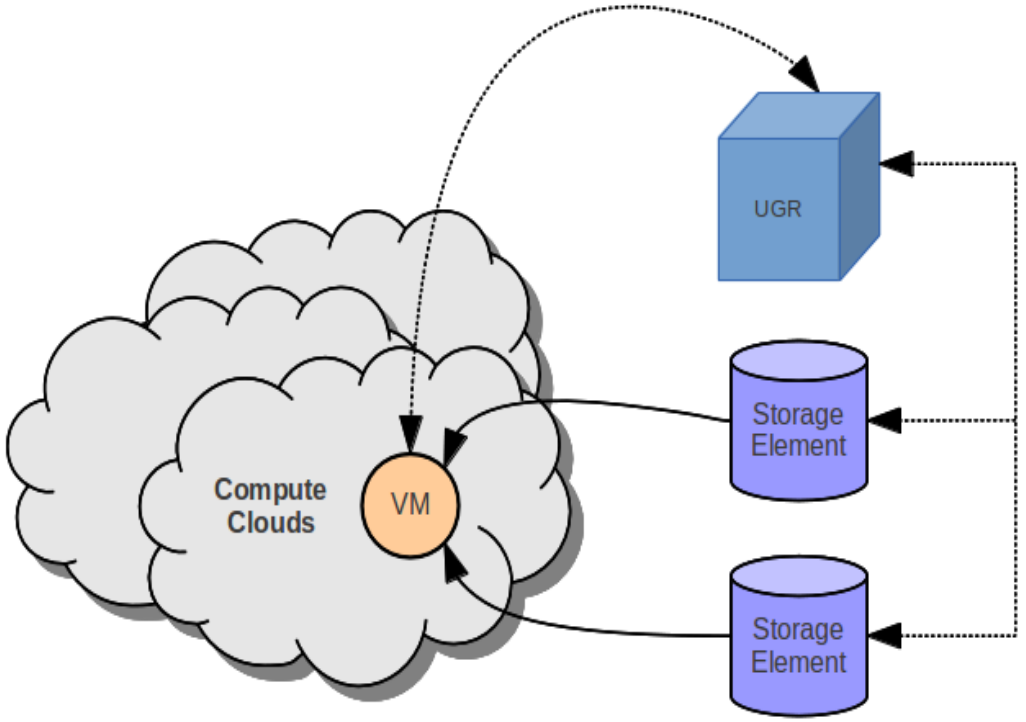
2.2. Project Progress Discussion, Software Distribution:



- In production for 8 months
- Multiple geographically distributed ATLAS squid caches
- Adopted by CERN, Oxford, & others.
- Included in CernVM3

NEP-101 HEP Data-Intensive Distributed Cloud Computing

2.3. Project Progress Discussion, Storage Federation:



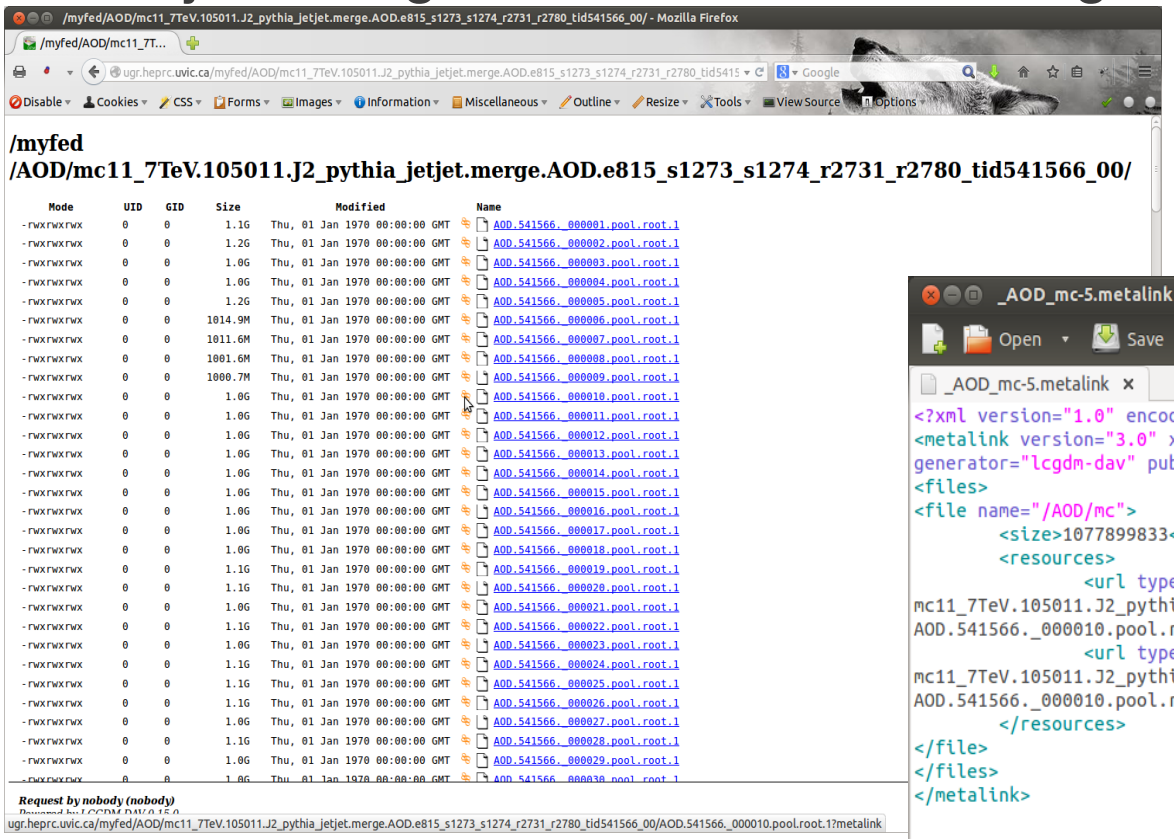
- Supports WebDAV servers and ATLAS storage elements (SE)

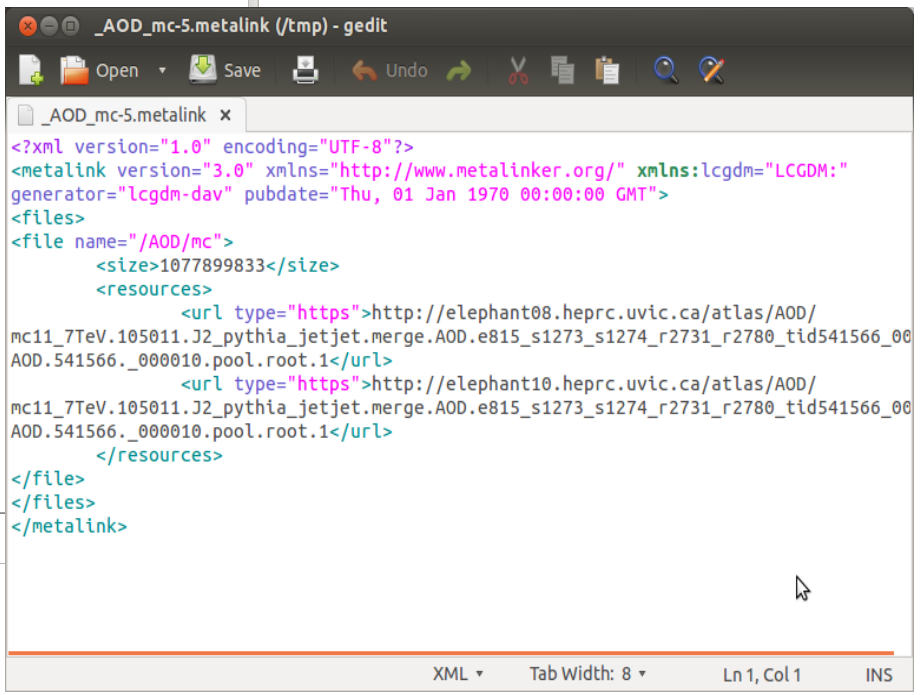
- ATLAS SE authentication via VOMS proxy – Tested interactively

- Takes advantage of pre-existing ATLAS data distribution – Canadian sites configured, testing to commence

NEP-101 HEP Data-Intensive Distributed Cloud Computing

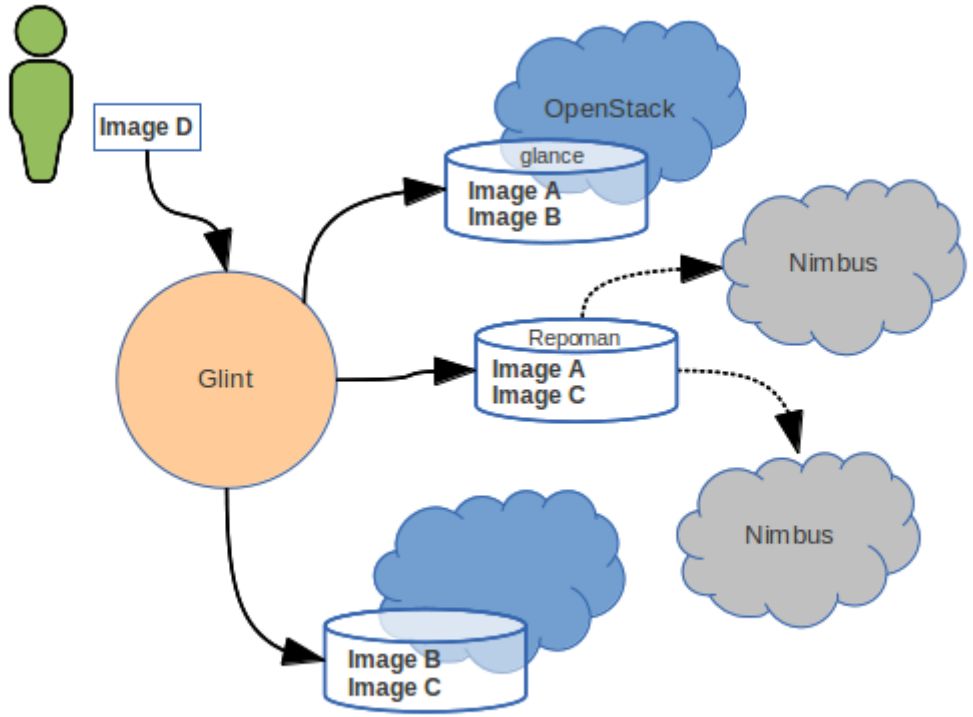
2.3. Project Progress Discussion, Storage Federation:





NEP-101 HEP Data-Intensive Distributed Cloud Computing

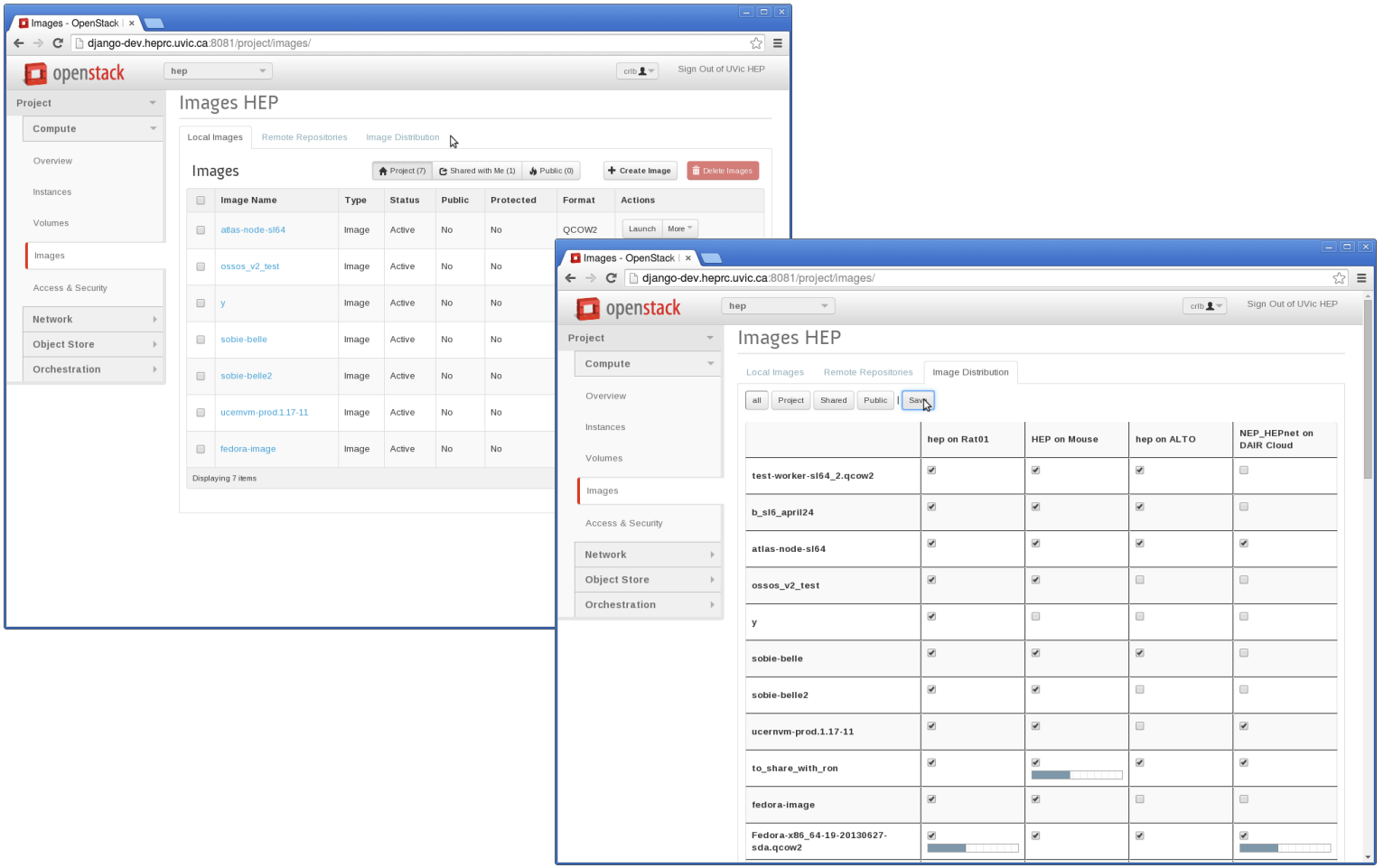
2.4. Project Progress Discussion, Image Distribution:



- Integrated with OpenStack Dashboard
  - * Adding HTTPs & Branding
  - * Moving to production
  - * Presenting Glint to OpenStack Summit in November
- Uses OpenStack development architecture, and keystone authentication
- Supports Glance, EC2, & GCE

NEP-101 HEP Data-Intensive Distributed Cloud Computing

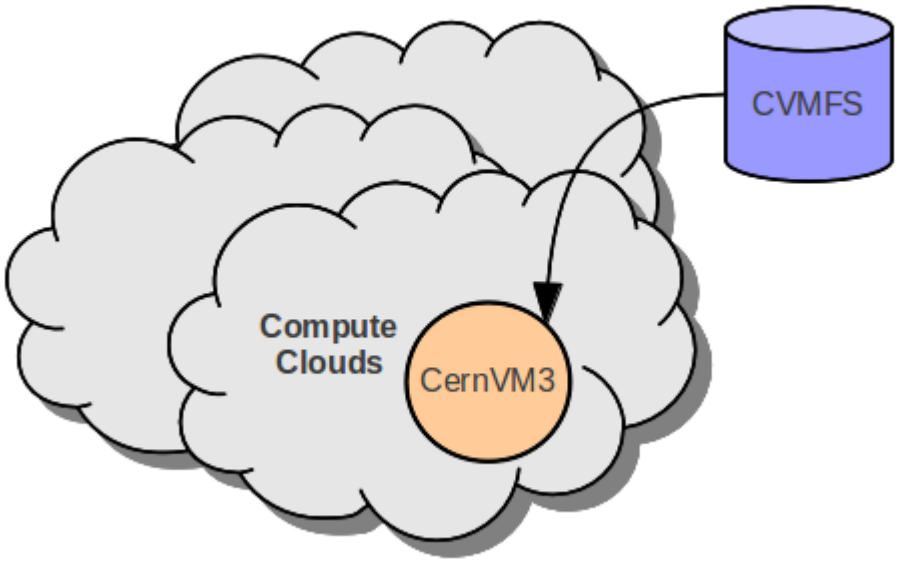
2.4. Project Progress Discussion, Image Distribution:



NEP-101 HEP Data-Intensive Distributed Cloud Computing

2.5. Project Progress Discussion, Image Optimization:

- CernVM3 uses a micro kernel



- Additional kernel modules are loaded from CVMFS as Required
- Developed by CERN
- In production

* Work continues on contextualization using cloud-init

* Installed on DAIR but not yet public

NEP-101 HEP Data-Intensive Distributed Cloud Computing

3. Collaboration Opportunities:
   - Latest CloudScheduler, Glint, and UGR will be available on DAIR
   - All source code developed by the project is on Github
   - Seeking to have Glint installed at CERN, on Westgrid, and other third party sites
   - Seeking to have Glint included as an OpenStack project
   - Using code provided by CERN, OpenStack, and the open source community

NEP-101 HEP Data-Intensive Distributed Cloud Computing

4. Summary:

- Project is on track and making good progress.
- Many of the pieces are already in production by NEP-101 project group, CERN/ATLAS, Belle II, and CANFAR.