

# 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:

	2013				2014												
	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
<b>Project Infrastructure</b>																	
1. Group meeting and project organization	█																100%
2. Hire and train new staff	█	█															100%
3. Testbed design, purchase and commissioning	█	█	█	█													100%
<b>Requirements gathering, analysis and design</b>																	
4. ATLAS application requirements	█	█															100%
5. Distributed VM repository review and design	█	█	█	█	█	█	█										100%
6. WebDAV federator review and design	█	█	█	█	█	█	█										100%
7. FAX review	█	█	█	█													100%
8. Streaming/staging data options review	█	█	█	█													100%
9. Shoal/Squid requirements	█				█	█	█										100%
10. MicroVM evaluation					█	█	█										100%
11. CloudScheduler – adding OpenStack API		█	█	█													100%
12. Requirements document		█	█	█													100%
<b>Implementation and testing</b>																	
13. Storage federation testing								█	█	█	█	█	█	█			83%
14. Shoal/Squid testing								█	█	█	█	█	█				100%
15. Shoal integration with MicroVM								█	█	█	█	█	█				100%
16. MicroVM testing					█	█	█	█	█	█	█	█	█				100%
17. CloudScheduler – test new OpenStack API					█	█	█	█	█	█	█	█	█				100%
18. CloudScheduler – integration with new repository								█	█	█	█	█	█				100%
19. CloudScheduler – monitoring and diagnostics						█	█	█	█	█	█	█	█				70%
20. Distributed VM repository testing								█	█	█	█	█	█				100%
21. System design document Version 1							█										100%
22. System design document Version 2										█	█						100%
<b>Beta testing</b>																	
23. Integrated testing of all services												█	█	█	█		17%
24. System design document Version 3												█	█	█			0%
<b>Pre-production sting and deployment</b>																	
25. Integrated testing and production deployment															█	█	0%
26. Deploy services on DAIR														█	█		17%
27. Project demonstration meeting														█	█		0%
28. Final project documentation														█	█		0%
<b>Maintenance</b>																	
																	(74%)
																	<b>78%</b>

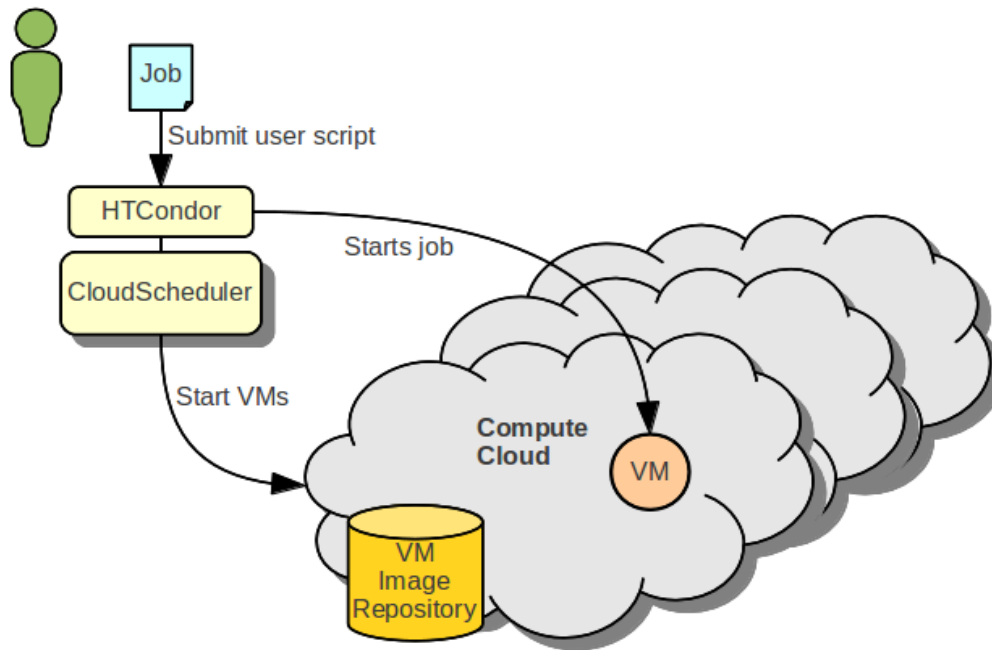
# NEP-101 HEP Data-Intensive Distributed Cloud Computing

## 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

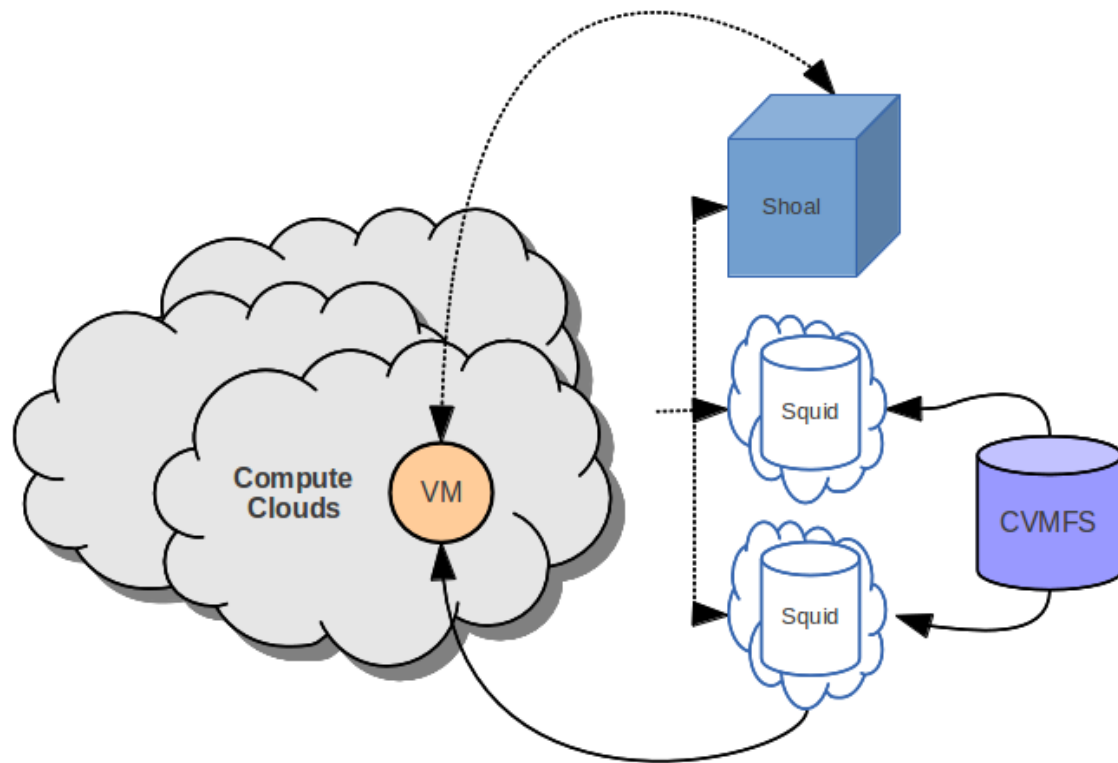
\* In production

- Monitoring & Diagnostics

\* Work continues

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

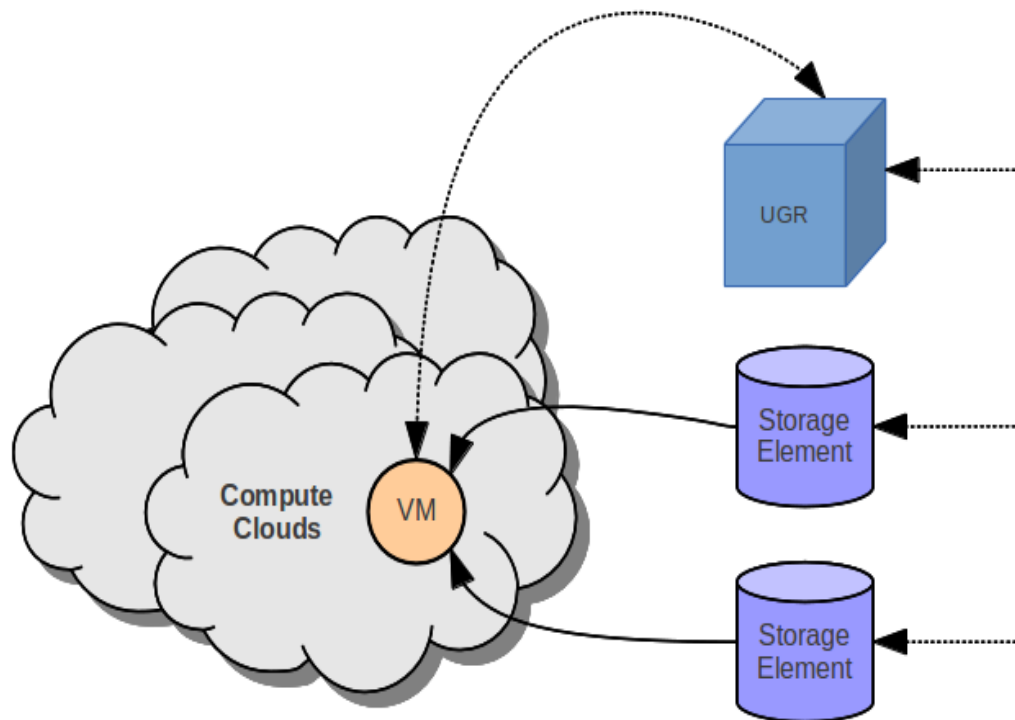
## 2.2. Project Progress Discussion, Software Distribution:



- 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:

The image shows two overlapping windows. The top window is a Mozilla Firefox browser displaying a directory listing for the path `/myfed/AOD/mc11_7TeV.105011.J2_pythia_jetjet.merge.AOD.e815_s1273_s1274_r2731_r2780_tid541566_00/`. The listing includes columns for Mode, UID, GID, Size, Modified, and Name, with entries for various pool.root.1 files.

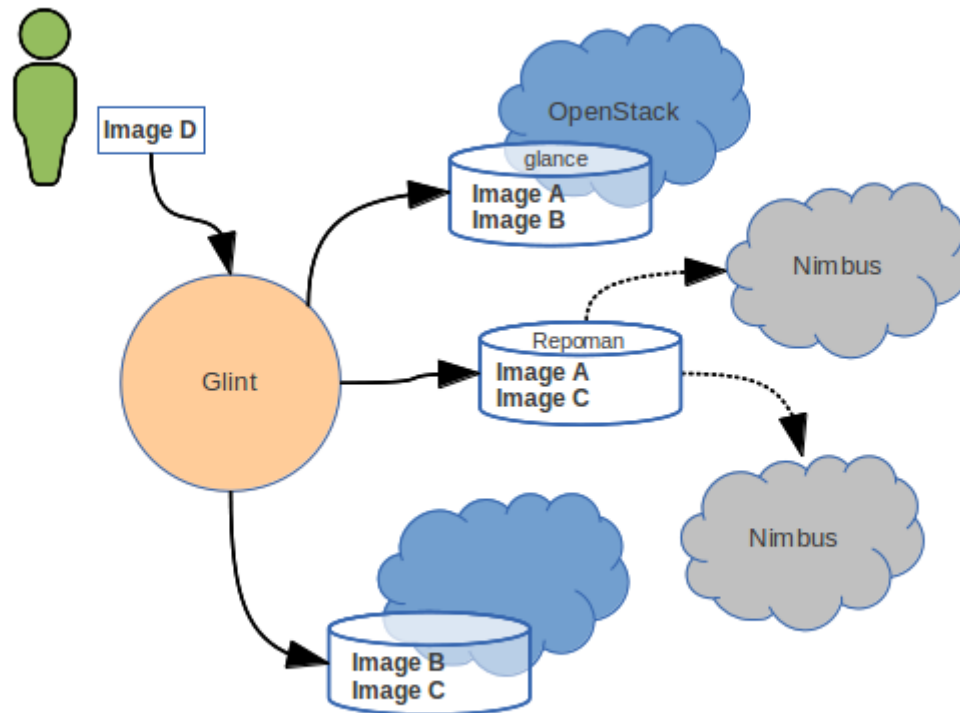
The bottom window is a gedit text editor editing a file named `_AOD_mc-5.metalink`. The content is an XML metalink file with the following structure:

```
<?xml version="1.0" encoding="UTF-8"?>
<metalink version="3.0" xmlns="http://www.metalinker.org/" xmlns:lcmdm="LCGDM:"
generator="lcmdm-dav" pubdate="Thu, 01 Jan 1970 00:00:00 GMT">
<files>
<file name="/AOD/mc">
  <size>1077899833</size>
  <resources>
    <url type="https">http://elephant08.heprc.uvic.ca/atlas/AOD/
mc11_7TeV.105011.J2_pythia_jetjet.merge.AOD.e815_s1273_s1274_r2731_r2780_tid541566_00
AOD.541566._000010.pool.root.1</url>
    <url type="https">http://elephant10.heprc.uvic.ca/atlas/AOD/
mc11_7TeV.105011.J2_pythia_jetjet.merge.AOD.e815_s1273_s1274_r2731_r2780_tid541566_00
AOD.541566._000010.pool.root.1</url>
  </resources>
</file>
</files>
</metalink>
```



# 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:

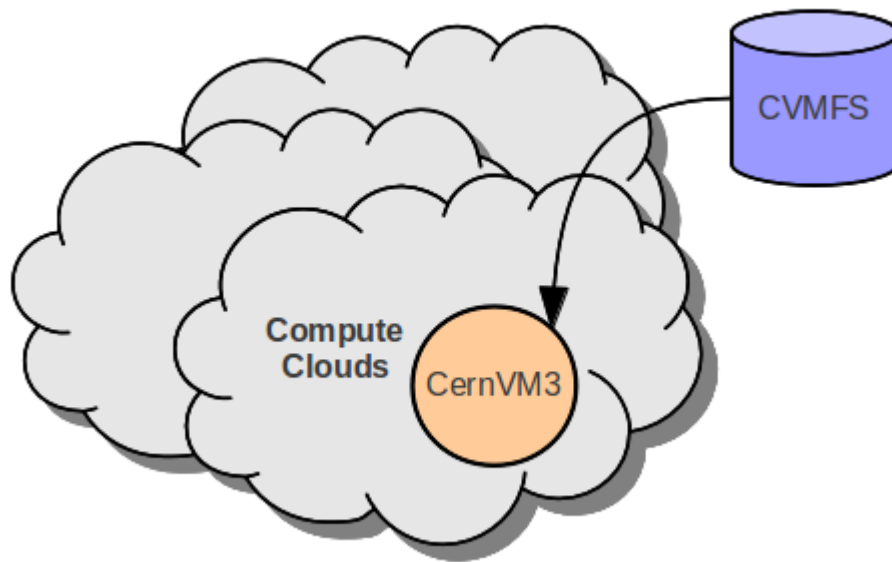
The top screenshot shows the OpenStack Images dashboard for the 'hep' project. The 'Images' tab is active, displaying a list of images with columns for Image Name, Type, Status, Public, Protected, Format, and Actions. The images listed are: atlas-node-si64, ossos\_v2\_test, y, sobie-belle, sobie-belle2, ucemvm-prod.1.17-11, and fedora-image.

The bottom screenshot shows the 'Image Distribution' tab. A table displays the distribution status of images across different environments. The table has columns for the image name and four distribution targets: 'hep on Rat01', 'HEP on Mouse', 'hep on ALTO', and 'NEP-HEPnet on DAIR Cloud'. The 'Search' button is highlighted.

	hep on Rat01	HEP on Mouse	hep on ALTO	NEP-HEPnet on DAIR Cloud
test-worker-si64_2.qcow2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
b_si6_april24	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
atlas-node-si64	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ossos_v2_test	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
y	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sobie-belle	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
sobie-belle2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ucemvm-prod.1.17-11	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
to_share_with_ron	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
fedora-image	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fedora-x86_64-19-20130627-sda.qcow2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

# 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.