

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

## Project Status

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Outline
  - Time line
  - Activities
  - Deliverables
  - Work Plan
  - Summary

	2013				2014		
	Sep	Oct	Nov	Dec	Jan	Feb	Mar
<b>Project infrastructure</b>							
Group meeting and project organization	■						
Hire and train new staff	■	■					
Testbed design, purchase and commissioning	■	■	■	■			
<b>Requirement gathering, analysis and design</b>							
ATLAS application requirements	■	■					
Distributed VM repository review and design	■	■	■	■	■	■	■
Webdav Federator review and design	■	■	■	■	■	■	■
FAX review			■	■	■	■	■
Streaming/staging data options review	■	■	■	■			
Shoal/Squid requirements			■	■	■	■	
MicroVM evaluation			■	■	■	■	
Cloudscheduler - adding OpenStack API		■	■	■			
Requirements document			■	■			
<b>Implementation and testing</b>							
Storage federation testing							
Shoal/Squid testing							
Shoal integration with MicroVM							
MicroVM testing					■	■	■
Cloud scheduler - test new OpenStack API					■	■	■
Cloud scheduler - integration with VM repository							
Cloud scheduler - monitoring and diagnostics						■	■
Distributed VM repository testing							
System design document Version 1							■
System design document Version 2							

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Time line:
  - Project approved 4 October 2013.
  - Contract signed 4 December 2013.
  - Project setup (staff & equipment) extended into February.
  - Development time is very short.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities by sub-projects:
  - Setup
  - Batch services
  - Software distribution
  - Storage federation
  - VM image distribution
  - VM image optimization

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities (1)
  - Setup - hardware acquisition
    - 4 nodes: 16 cores, 64GB, 10Gib each
    - Compute for OpenStack cloud
    - Supplemented with an elephant node
    - Drivers for Intel 57800 not available
    - Substitued BCM5720 + Intel X520 DA
    - Network performance to be resolved

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities (2)
  - Batch Services
    - Work started on the OpenStack API.
    - Job Description Language (JDL) enhancements:
      - Cloud aliases, +TargetClouds, +VMAMI, +VMInstanceType
    - Amazon EC2, spot pricing.
    - Google Compute Engine (GCE).

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities (3)
  - Software Distribution
    - Web cache broker (Shoal): server, agent, and client.
    - Single server running in a VM.
    - Web caches (Squid + Shoal agent); 3 manually deployed.
    - New VM images: SL6 with Shoal client.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities (4)
  - Storage Federation
    - Prototype federation service hosted in a VM.
    - Testing Unified Generic Redirector (UGR, from CERN).
    - Federating ATLAS Storage Element (SE) plus two WebDAV servers.
    - All federated storage servers are currently in Victoria.
    - System needs tuning and testing.



# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Activities (5)
  - VM Image Distribution
    - Prototype created.
    - Production & development hosted in VMs.
    - Django web service (as is OpenStack dashboard).
    - Currently works with Glance.
    - Other types through plugin architecture.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Deliverables

- Three documents; all publicly available:

- Use cases: <https://wiki.heprc.uvic.ca/twiki/pub/Main/NEP101FirstTechnicalReview/UseCases.pdf>
    - Requirements: <https://wiki.heprc.uvic.ca/twiki/pub/Main/NEP101FirstTechnicalReview/Requirements.pdf>
    - Design: <https://wiki.heprc.uvic.ca/twiki/pub/Main/NEP101FirstTechnicalReview/Design.pdf>

- All documents are working documents

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Work Plan – involves all sub-projects

	2014					
	Apr	May	Jun	Jul	Aug	Sep
<b>Implementation and testing</b>						
Storage federation testing						
Shoal/Squid testing						
Shoal intergration with MicroVM						
MicroVM testing						
Cloud scheduler - test new OpenStack API						
Cloud scheduler - integration with VM repository						
Cloud scheduler - monitoring and diagnostics						
Distributed VM repository testing						
System design document Version 1						
System design document Version 2						
<b>Beta-Testing</b>						
Integrated testing of all services						
System design document Version 2						

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Batch Services
  - Develop OpenStack API support.
  - Additional JDL enhancements.
    - VM image names, +VMType
  - Diagnostics and monitoring.
  - Amazon EC2, spot pricing enhancements.
  - Google Compute Engine (GCE).
    - Maintenance and functionality testing.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Software Distribution
  - Functionality and performance testing.
  - Additional manual deployments of squid caches.
  - Auto-deployment of squid caches.
  - Work with ATLAS and Belle II projects.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Storage Federation
  - Initial federation performance tuning.
  - Build geographically dispersed functional testbed.
  - Functional and performance testing.
  - Integrate production storage elements.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- VM Image Distribution
  - Build user base, testing.
  - VM Name/Type management.
  - Distributed image integrity checking.
  - Tighter integration with keystone.
  - Support other required cloud types.
  - Rolling upgrades.

# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- MicroVM
  - Work will be focused around CernVM 3.
  - Image customization:
    - Shoal/Squid, UGR, ATLAS, Belle II
- Documentation



# NEP-101 HEP Data-Intensive Distributed Cloud Computing

- Summary
  - Late start; still playing catch-up.
  - Good progress to date on all other aspects.
  - Work plan will keep us on track to achieve project goals.