



Cloud Overview

- ▼ Adam Huffman
- ▼ Imperial College
- ▼ GridPP 32 Collaboration Meeting

Agenda

- ▼ GridPP Cloud at Imperial
- ▼ Cloud Activity at other GridPP Sites
- ▼ Experiment Activities
- ▼ CERN Private Cloud
- ▼ Common Themes
- ▼ Plans

GridPP Cloud at Imperial College

GridPP Cloud at Imperial College

▼ OpenStack Havana

- Deployed with Packstack (from Red Hat RDO)

▼ Hardware:

- 1 controller, 6 compute nodes
- 176 cores, 448GB RAM
- 1.5TB instance storage currently (3TB in reserve)
- 40 core/256GB compute node ready to add
- 1 node for testing next OpenStack release (RDO Test Days)

GridPP Cloud at Imperial College - Networking

▼ Tyranny of choice:

- Nova-network
 - Stable, heading for deprecation
 - Simple to setup and use
- Neutron (was Quantum)
 - Very flexible, very complex
 - Fragmented documentation
 - Will test with Icehouse upgrade

▼ External access

- Wish to keep grid and cloud VLANs separate
- Public IPs “controversial”
- Login instance workaround

GridPP Cloud at Imperial College - Storage

- ▼ Storage types:
 - Instance (Nova), volume (Cinder), image (Glance)
 - Local instance storage at first
- ▼ Shared storage
 - Expansibility
 - Flexibility
 - Migrations

GridPP Cloud at Imperial College – GlusterFS/Ceph

▼ GlusterFS experience

- Good:
 - Setup and basic operations simple
- Bad:
 - FUSE performance
 - Fragile upgrades
 - Permissions bug with migration
 - IPv6 (?)

▼ Plans:

- Expand Ceph test cluster (~10TB)
- Interest in CephFS
- Replace GlusterFS?

GridPP Cloud at Imperial College - Usage

- ▼ Usage for 2014 so far:
 - gridpp tenant (primarily ATLAS)
 - **12,850.74** vCPU hours
 - **1,938,725.76** disk GB hours
 - ichep-test tenant (CMS Stealth Cloud)
 - **88,054.28** vCPU hours
 - **4,324,985.66** disk GB hours
 - Measures usage, not load

OpenStack Dashboard Interface

The screenshot shows the OpenStack Dashboard interface in a Mozilla Firefox browser window. The browser title is "Hypervisors - GridPP Cloud, hosted at Imperial College". The address bar shows the URL: <https://gridppcl02.grid.hep.ph.ic.ac.uk/dashboard/admin/hypervisors/>. The user is logged in as "admin".

The dashboard features a sidebar with navigation options: Project, Admin, System Panel (Overview, Resource Usage, Hypervisors, Instances, Volumes, Flavours, Images, Defaults, System Info), and Identity Panel (Projects, Users). The main content area is titled "All Hypervisors" and includes a "Hypervisor Summary" section with three pie charts:

- VCPU Usage:** Used 169 of 176
- Memory Usage:** Used 341GB of 440GB
- Disk Usage:** Used 3.7TB of 9.0TB

Below the summary is a table titled "Hypervisors" with the following data:

Hostname	Type	VCPUs (total)	VCPUs (used)	RAM (total)	RAM (used)	Storage (total)	Storage (used)	Instances
gridppcl01.grid.hep.ph.ic.ac.uk	QEMU	24	17	94GB	34GB	1.5TB	380.0GB	3
gridppcl00.grid.hep.ph.ic.ac.uk	QEMU	24	24	94GB	48GB	1.5TB	540.0GB	3
wh01.grid.hep.ph.ic.ac.uk	QEMU	32	32	62GB	64GB	1.5TB	720.0GB	4
wh00.grid.hep.ph.ic.ac.uk	QEMU	32	32	62GB	64GB	1.5TB	720.0GB	4
wh03.grid.hep.ph.ic.ac.uk	QEMU	32	32	62GB	64GB	1.5TB	720.0GB	4
wh02.grid.hep.ph.ic.ac.uk	QEMU	32	32	62GB	64GB	1.5TB	720.0GB	4

At the bottom of the table, it says "Displaying 6 items".

Cloud activity at other GridPP sites

Oxford

- ▼ Setup by Kashif Mohammad
- ▼ OpenStack Havana
 - Deployed with Packstack
- ▼ 1 controller, 2 compute nodes
 - 32 cores
- ▼ Floating IPs available for debugging
- ▼ Successfully running ATLAS jobs (see later)
- ▼ Accounting/funding concerns limiting expansion

- ▼ StratusLab prototype private IaaS
 - 800 cores, 1.4TB RAM, 14TB shared storage
 - Internal development, testbeds, HTCondor bursting (see Andrew Lahiff's talk)
- ▼ Migrate to production-quality OpenNebula IaaS over next year
 - 900 vCPUs, 3.5TB RAM and ~1PB raw storage for Ceph
 - Integrate with Tier 1 batch system
 - Virtualize all Tier 1 worker nodes?
 - Offer cloud endpoints for LHC VOs

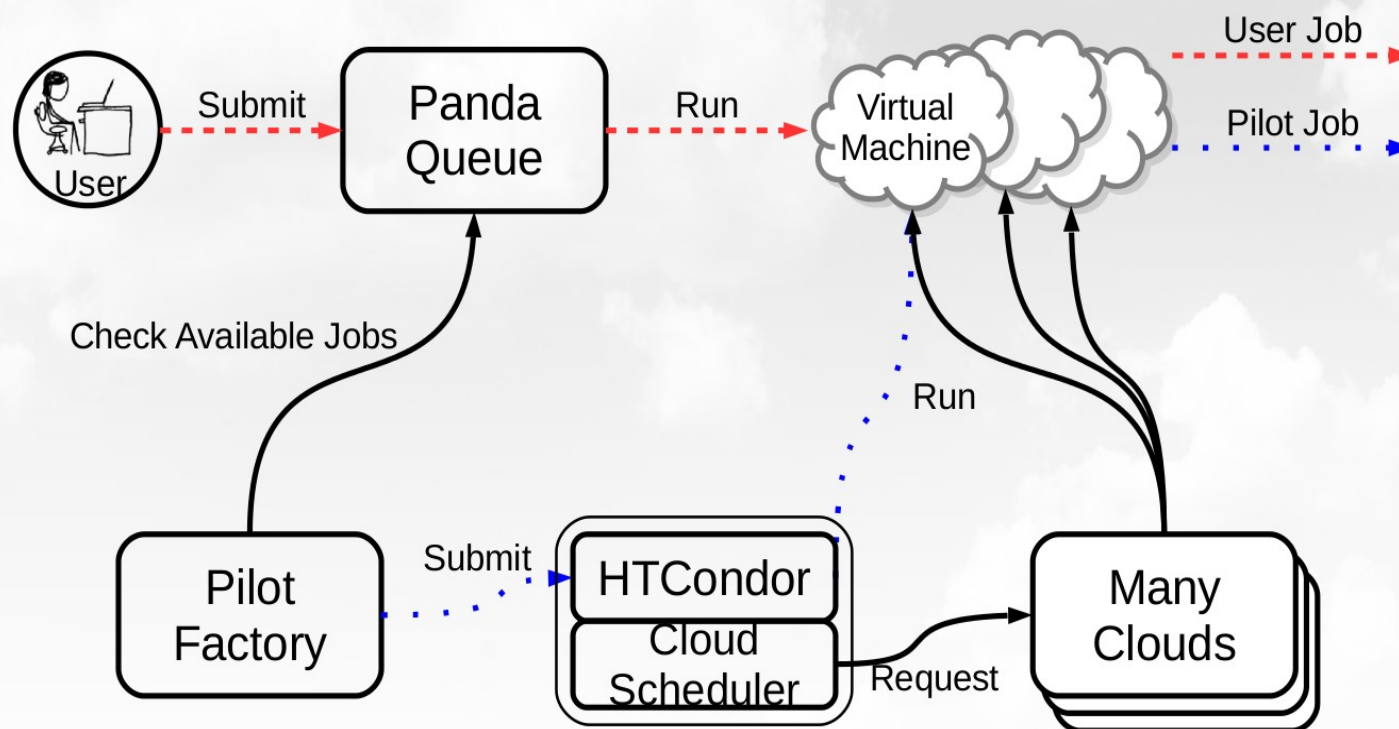
LHC VO Cloud Activity

ATLAS

- ▼ Frank Berghaus (UVic), Peter Love (Lancaster) & others
- ▼ Aims:
 - Minimise local (site) dependencies
 - More flexible resource allocation
- ▼ Cloud Scheduler and APF
- ▼ Images based on SL 6.4, looking at uCernVM

ATLAS Cloud architecture

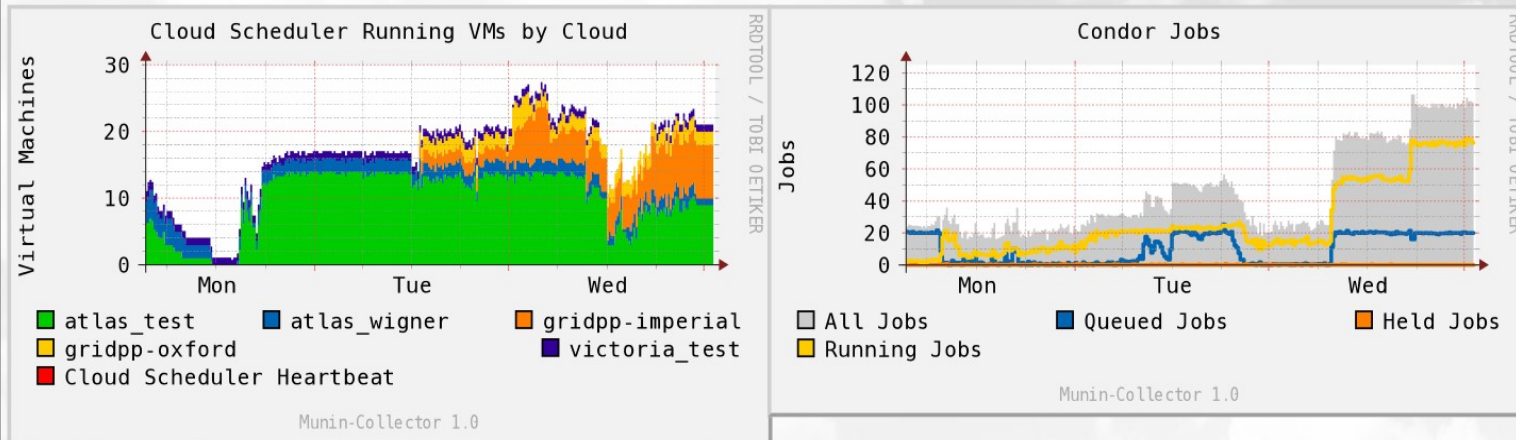
Cloud Job Flow (on the Grid)



- Easy to connect and use many clouds

GridPP Clouds running ATLAS jobs

CERN Cloud Scheduler



- Running ATLAS jobs on CERN and GridPP clouds
- Panda queues:
 - OPENSTACK_CLOUDSCHEDULER
 - GRIDPP_CLOUD
 - Sean is moving the Australian Nectar clouds over

ATLAS Status and Plans

▼ Status:

- GridPP running single-core production
- Ready to use cloud resources for ATLAS production
- Also running production jobs on HLT
 - 15,000 cores (see [SIM@P1](#) talks)

▼ Plans:

- Scale-up
- Add analysis & multi-core production jobs
- Accounting
 - Separate Panda queues per country/organization
 - What information is needed and how to present it?

CMS

- ▼ Stealth Cloud (Daniela Bauer & Simon Fayer, Imperial)
 - Runs grid jobs transparently on a cloud, using glideinWMS
- ▼ CMSooooCloud (large team led by David Colling, Imperial)
 - OpenStack installed alongside HLT software
 - 13,000 cores, ~5,000 jobs, saturating 40Gb link, now 60Gb
 - Investigating inter-fill running
- ▼ UK testing (Andrew Lahiff, RAL)
 - Successfully run jobs at Imperial and Oxford
- ▼ Interest in opportunistic resources

CMS

▼ CERN Private Cloud

- Up to 140 instances, 1,120 vCPU, 2.1TB RAM
- Massimo Sgaravatto (INFN) setup glideinWMS
- Image based on SLC6.5
- Worked around some wrinkles e.g. flavours, HTCondor bugs

▼ CERN Tier-0

- Currently up to 50 instances, 400 vCPU, 800GB RAM (will be expanded soon to 7,000 cores)
- Very similar image to Private Cloud

LHCb

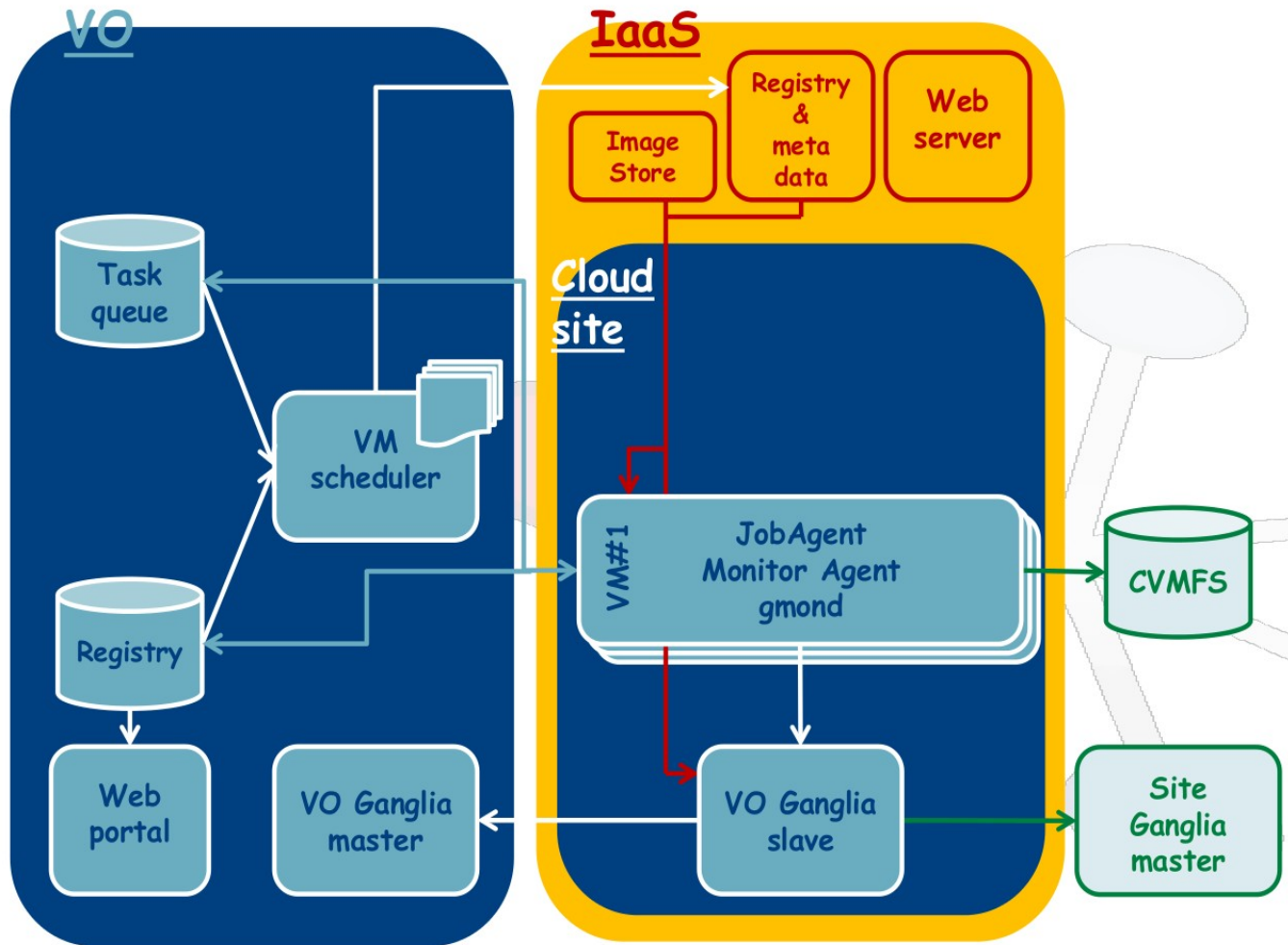
- ▼ Cloud use cases:
 - Cloud sites
 - Testbeds
 - Continuous Integration
 - LHCbDIRAC
- ▼ Successfully run jobs on CERN Private Cloud
 - Change of staff since then
- ▼ Most UK effort on contextualization and Vac (see Andrew's talk)

LHCb Cloud Architecture



Architecture: Goodies and more

Integration .. Cloud .. LHCb



CERN Private Cloud

CERN Private Cloud

- ▼ Currently 50,000+ cores, 35,000 more in a few months at Wigner
- ▼ Migrating 100 servers/week as new OpenStack compute nodes
- ▼ 2 Cells at CERN, 1 at Wigner
- ▼ Successful migration to OpenStack Havana
- ▼ High CERN engagement in upstream project

Common Themes

Accounting
Monitoring
Security

Accounting

- ▼ Benchmarking complicated
 - Compute node may be running various workloads normally
 - Otherwise quiet node may activate Intel Turbo, skewing results
- ▼ Cloud/VM flexibility comes at a potential cost to efficiency
 - Need to match flavours carefully to available hardware
 - Effect of overcommit?
- ▼ EGI work on accounting
 - See Andrew McNab's talk for more discussion

Accounting/Allocation

▼ Allocation

- Want hardware to be busy
- How to avoid a monopoly?

EGI Federated Cloud Accounting Test Portal

Cloud View --> Production

Data to graph:	Computation Monetary Cost ▾ Computation Monetary Cost			
Period:	Start year: 2013 ▾	Start month: 1 ▾	End year: 2014 ▾	End month: 3 ▾
Groupings:	Show data for: SITE ▾	as a function of: DATE ▾		

Refresh

by SITE and DATE.

The following table shows the distribution of grouped by SITE and DATE.

Problems to calculate the percentage

by SITE and DATE

SITE	Feb 2013	Mar 2013	Apr 2013	May 2013	Jun 2013	Jul 2013	Aug 2013	Sep 2013	Oct 2013	Nov 2013	Dec 2013	Jan 2014	Feb 2014	Mar 2014	Total	%
100IT	0	0	0	550,967	854,566	0	153,297	1,155,869	0	0	5,577,117	0	0	7,808	8,299,624	75.86%
CESGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
IFCA-LCG2	0	0	0	0	0	0	195,041	641,516	373,709	0	0	0	153,949	80,280	1,444,495	13.20%
HSAS-FedCloud	0	0	0	0	0	0	0	0	507,923	0	0	519,892	6,238	162,528	1,196,582	10.94%
Total	0	0	0	550,967	854,566	0	348,338	1,797,385	881,631	0	5,577,117	519,892	160,188	250,616	10,940,701	
Percentage	0.00%	0.00%	0.00%	5.04%	7.81%	0.00%	3.18%	16.43%	8.06%	0.00%	50.98%	4.75%	1.46%	2.29%		

[Click here for XML encoded data](#)

Monitoring

- ▼ Who should provide the monitoring?
 - Sites?
 - VOs?
- ▼ How to deal with ephemerality of cloud instances?
 - CERN Private Cloud
 - Keep records in Ganglia for 7 days, for CMS
 - Useful for debugging
 - BUT: usability problems when scaled up?

Security

- ▼ Who controls cloud/VM image creation?
- ▼ Who monitors that images are kept up to date?
- ▼ EGI draft security questionnaire

Plans

- ▼ Ceilometer
- ▼ Containerization
- ▼ Device access (GPUs etc.)
- ▼ TripleO deployment
- ▼ EGI Federated Cloud
- ▼ Object Storage/Ceph
- ▼ LHCb

Thank you ...

- ▼ Material from:
 - ▼ Peter Love / Frank Berghaus
 - ▼ Ian Collier
 - ▼ Andrew McNab / Mario Ubeda Garcia
 - ▼ Kashif Mohammad



All text and image content in this document is licensed under the [Creative Commons Attribution-Share Alike 3.0 License](#) (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these therefore is subject to the [trademark policy](#).