Clouds at other sites

T2-type computing

Randall Sobie

University of Victoria

Overview

- Clouds are used in a variety of ways for Tier-2 type computing
  - MC simulation, production and analysis
  - Commercial/private, in-house/distributed
- Motivation for using clouds
  - Ease of use, reduced manpower costs, resource sharing
  - Separation of application and system administration
  - Leverage software development by commercial world
- How are clouds being used?
  - VM provisioning, job management, benchmarks, storage, networking, monitoring

Cloud computing in HEP

Cloud computing in HEP is typically providing 5-20% of the processing of current projects

"Dedicated" clouds
(Owned by HEP)

"Opportunistic" clouds
(private and commercial)

Cloud deployments

Traditional
bare-metal

Specific purpose cloud
(e.g.. LTDA BaBar, HLT clouds)

Standalone/private
cloud
(e.g. PNNL, NorduGrid)

Distributed clouds
(e.g. UK, Canada,
Australia, INFN Clouds)

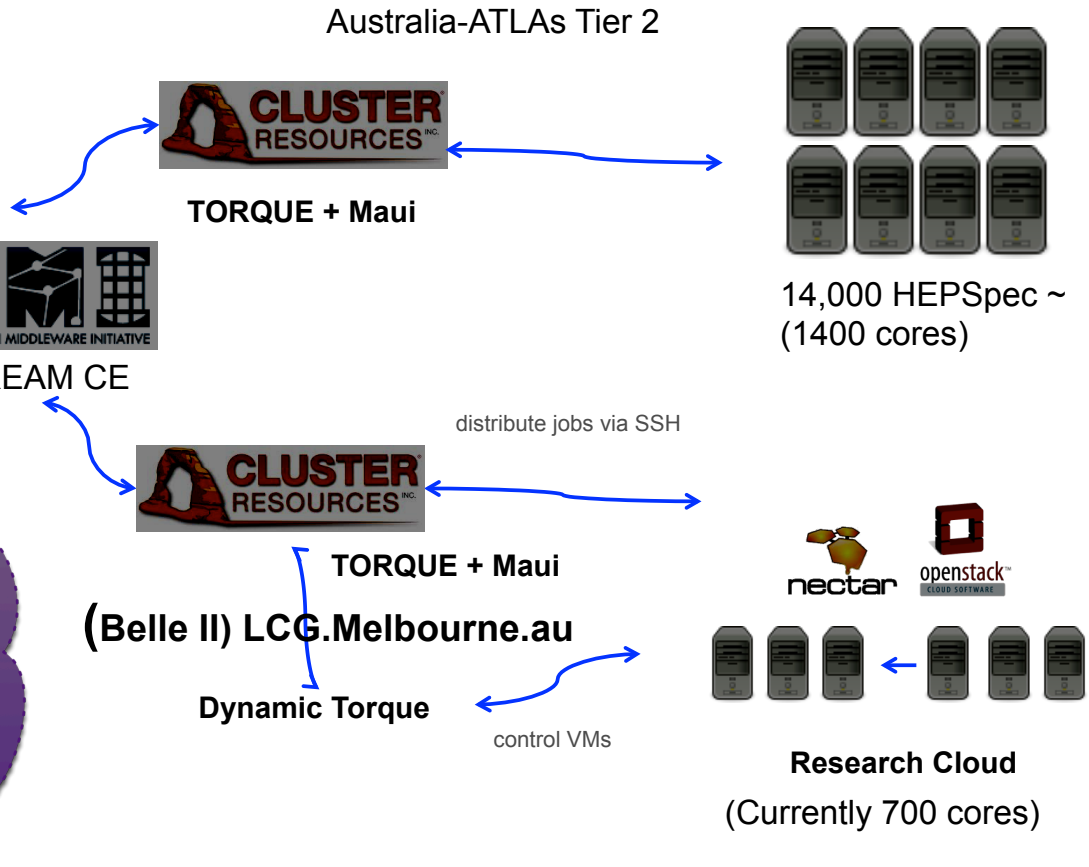
Bare-metal or in-house cloud with external cloud
(e.g.. CERN, BNL)

Examples of cloud deployments

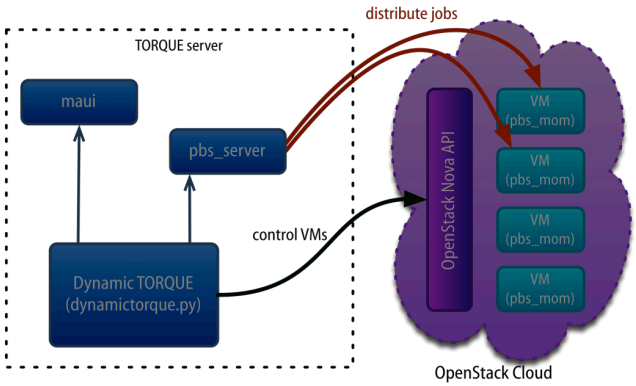
(meant to illustrate our use of clouds)

Australian Belle II Grid Site

Single CREAM CE services
ATLAS Tier-2 (Torque)
and
Belle II site (Dynamic Torque)



Dynamic Torque



Why private cloud?

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Chosen for flexibility, efficient use of compute resources for services
- Provides easy load-balancing and availability features
- Provides templating features
- Easy re-use of templates to test and instantiate new server instances
- Non-systems staff can provision their own instances of services
- Software Defined Networking is more malleable than physical networking, encourages better networking practices, including security

Lessons learned

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- VM's and/or containers provide needed flexibility to support multiple collaborations and different user needs
- Ceph storage is *very* robust and flexible
- VM's impose a 15%-20% performance penalty on HEP compute workload without careful tuning
  - Move to containers on bare metal planned
- OpenStack features do not help us make sure a certain number of instances are up and healthy and consistent
  - Kubernetes looks appealing in this respect

GridPP (P.Love/A.McNab)

University Openstack instances

- Clouds at HEP institutions (Oxford/Imperial).
- ECDF cloud in Edinburgh has recently made available to the HEP

UK Vacuum deployments

- Key to our light-weight Tier-2 strategy where we operate with minimal manpower at the site (<1000 cores).

Datacentred commercial Openstack

- Scale of a Tier-2 facility.
- Free access to the their system (ATLAS) whilst they were commissioning things; paid for access when funds available.
- Network connectivity to the UK academic network is only 1Gbit but they have plans to upgrade

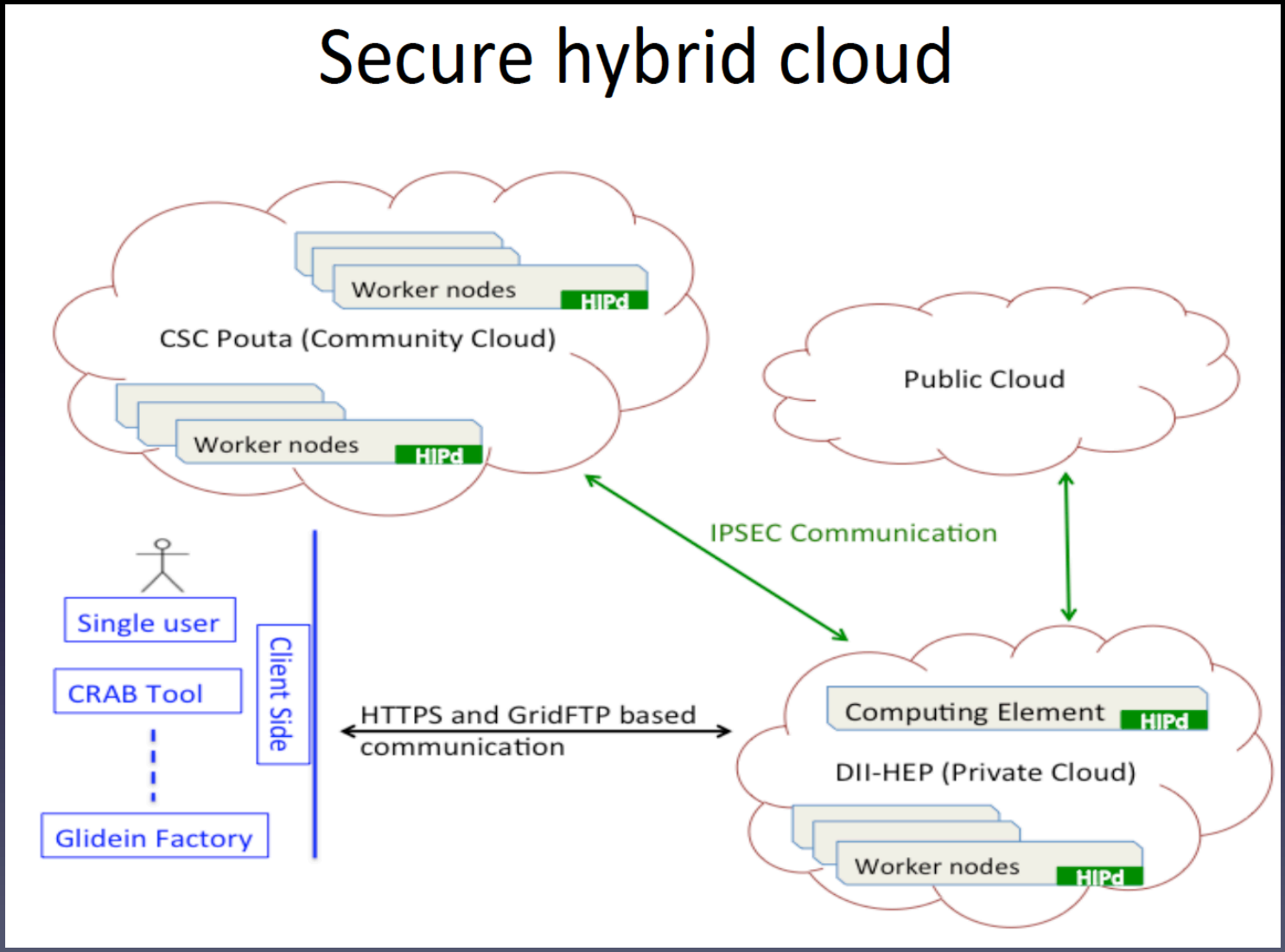
Italy (INFN; Massimo Sgaravatto et al)

PrivateOpenStack Cloud (Padova-Legnaro) called CLOUD AREA PADOVANA
Used by ~ 25 user groups/project that financially contributed for the resources

Batch processing

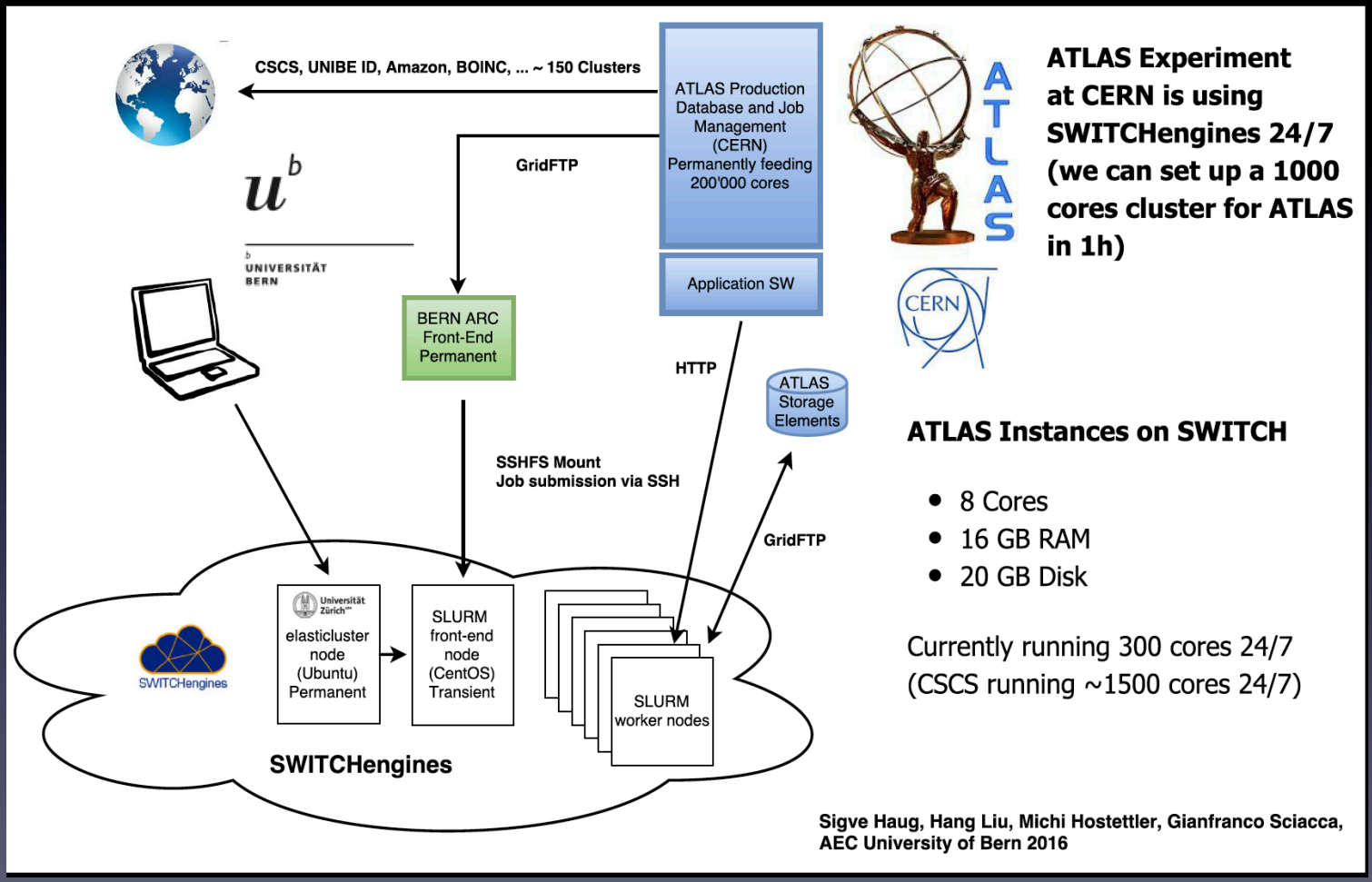
- Relying on the elastiq framework, HTCondor batch clusters are instantiated.
- These batch clusters are 'dynamic': new worker nodes are automatically added or are removed depending on load.
- CMS Cloud project is integrated with the local Tier-2.
  - E.g. CMS VMs can access the T2 storage (dcache) using the same local protocol (dCAP) used by the T2 WNs.
- Plans to deploy the Synergy service, which allows to manage the resource allocation using a fair-share approach, without a static partitioning of such resources among the relevant user communities.

NorduGrid


Secure hybrid cloud
Worker nodes
HIPd
CSC Pouta (Community Cloud)
Worker nodes
HIPd
Public Cloud
IPSEC Communication
Single user
CRAB Tool
Glidein Factory
Client Side
HTTPS and GridFTP based communication
Computing Element
HIPd
DII-HEP (Private Cloud)
Worker nodes
HIPd

Bern Switzerland

SWITCHengines – Swiss NREN commercial cloud (OpenStack)
(free during development phase)



Canada

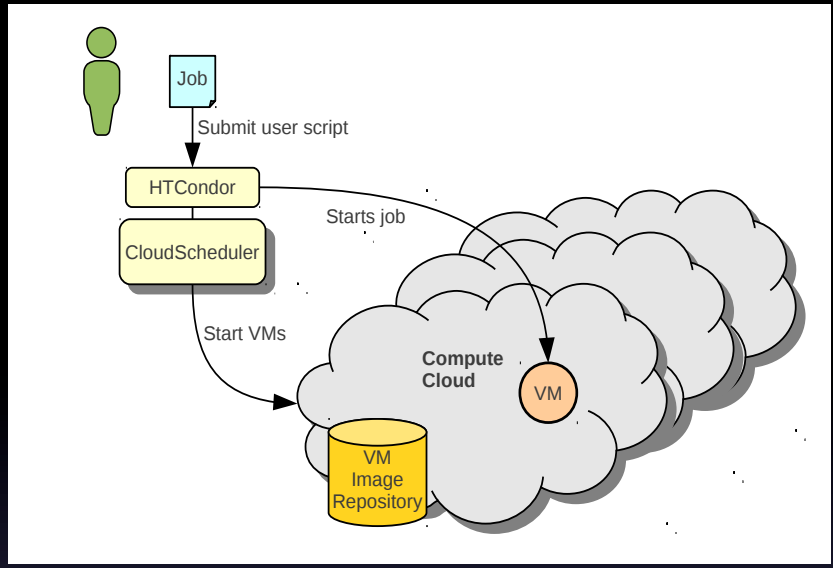
Distributed cloud system for ATLAS and Belle II

- Integrated into Panda/DIRAC
- In production for 3-4 years
- Also used by Canadian astronomy

- uCernVM, CVMFS, Squid-discovery (Shoal)
- Distributed VM image repository
- Data written to local storage and transferred
- Benchmarks run at VM boot
- VM time measurements for accounting
- Reasonable monitoring

- Updating system for Open Nebula
- Studying data federations (e.g. Dynafed)
- Context-awareness

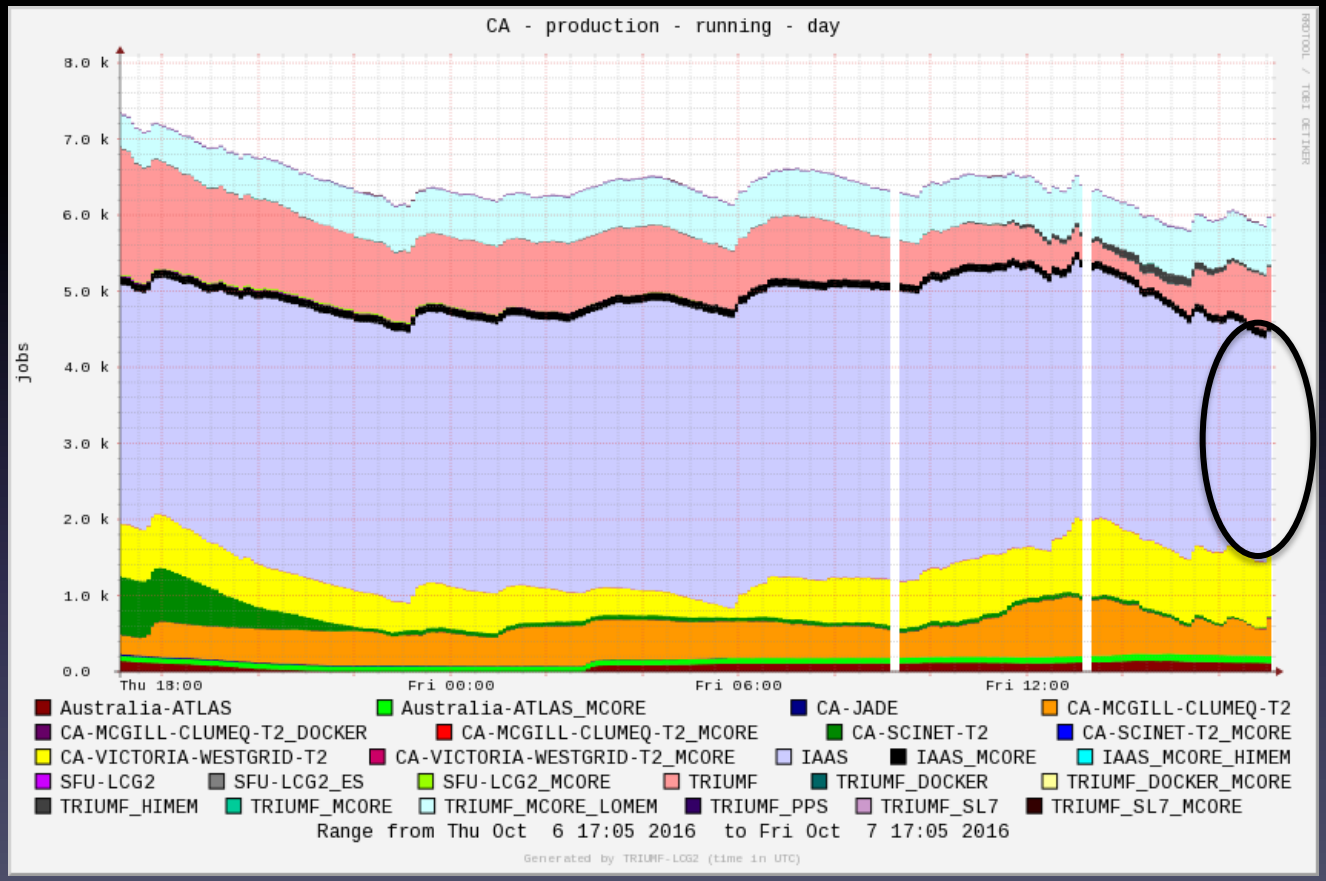
- Challenges include managing resources across many administrative domains



10-15 clouds managed by HTCondor/ CloudScheduler (4000-5000 cores)

800-1000 cores (each) EC2/Azure
(Egress fees waived)

Canadian WLCG "cloud" – includes Australian T2
Friday October 6



<u>Cloud resources</u>
10 clouds
4300 cores

Job scheduling/VM provisioning

- Variety of methods for running HEP workloads on clouds
  - VM-DIRAC (LHCb and Belle II)
  - VAC/Vcycle (UK)
  - HTCondor/CloudScheduler (Canada)
  - HTC/GlideinWMS (FNAL), HTC/VM (PNNL), HTC/APR (BNL)
  - Dynamic-Torque (Australia)
  - Cloud Area Padovana (INFN)
  - ARC (NorduGrid)
- Each method has its own merits and often was designed to integrated clouds into an existing infrastructure (e.g. local, WLCG and experiment)

Commercial and private clouds

- Commercial cloud use
  - Primarily Amazon EC2 and Microsoft Azure (with grants)
  - ATLAS discussing use of GCE
  - Other commercial OpenStack clouds
    - DataCentred (UK), SWITCHengines (Switzerland)
  - CERN commercial cloud procurement

- Private clouds
  - OpenStack and OpenNebula research-funded clouds but not involved in HEP

Network connectivity

- Amazon and Microsoft clouds are connected to the research networks in North America (probably GCE as well)
  - Egress charges can be waived upon request
- Trans-border or trans-ocean traffic can be an issue
  - Become an important discussion topic in the LHCONE meetings
- Private opportunistic clouds
  - traffic flows over research network but not LHCONE network

CPU Benchmarks

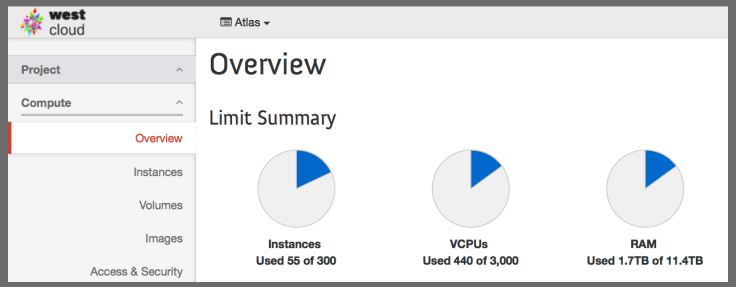
New suite of "fast" benchmarks

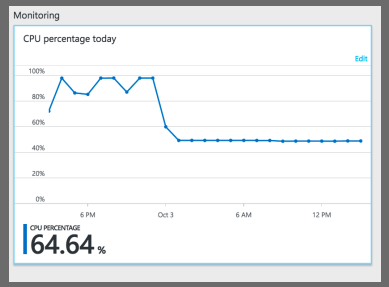
- HEPiX Benchmark Working Group
- Suite available includes "fast HS" (LHCb) and Whetstone benchmarks
  - Write to ElasticSearch DB
- Run benchmarks in the pilot job or during the boot of the VM

Data storage

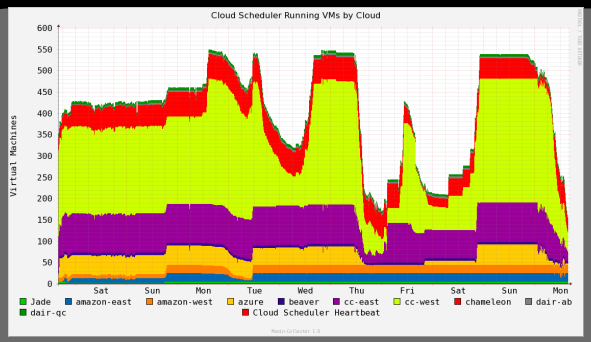
- Data written to local storage on node and then transferred to selected SE
- UK group has done some work integrating their object store with ATLAS
- BNL using S_3 storage on EC_2 for T2-SE

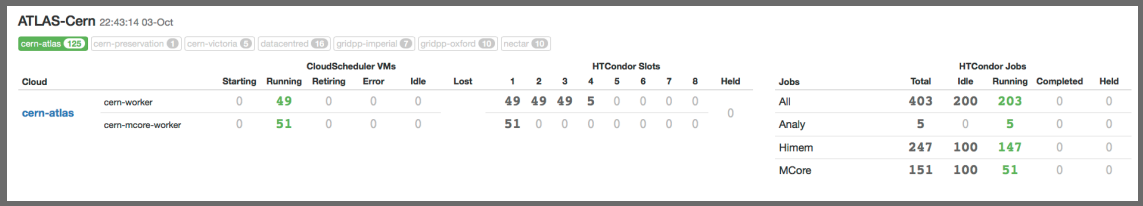
Monitoring





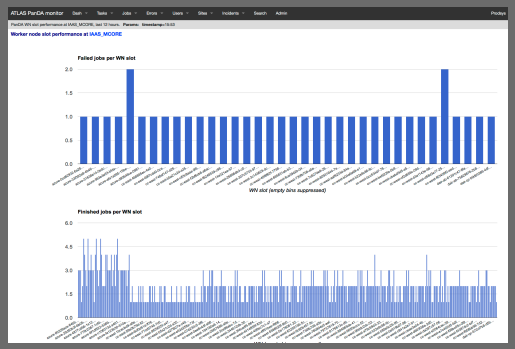
Cloud or site monitor





Cloud System monitor

Sensu, Munin, RabbitMQ, Mongo-DB, Ganglia



Application monitor

Panda monitoring

MONTH

Cloud	#	Bmk	User	Total
beaver	18	15.6	49.5	60.5
cc-west	1610	19.0	3228.4	4205.4
cc-east	296	14.5	931.1	1129.1
chameleon	136	21.1	976.1	1269.5
dair-ab	2	12.7	30.6	32.9
dair-qc	7	12.3	85.3	92.2
azure	126	21.4	1518.2	1800.3
ec2	106	9.8	36.4	184.3
Total			6855.6	8774.2

Monday October 03 15:00:01

Benchmarks and accounting

ElasticSearch DB

Summary

- Clouds at HEP sites
  - Typically integrated into an existing infrastructure
  - Seen as a way to better manage multi-user resources
  - Cloud R&D funding opportunities
- Opportunistic research clouds
  - Easy way to utilize clouds at non-HEP research computing facilities
  - No requirement for on-site application specialists or complex software
- Commercial clouds
  - EC2/Azure/GCE dominate but other OpenStack clouds
  - Grant and some contracted resources
  - Trans-border network connectivity being addressed