Cloud computing for HEP applications

Randall Sobie

Institute of Particle Physics
University of Victoria

Motivation
Design and operation of distributed computing cloud
Unique services and software
Data federation
Context-aware cloud computing
Future plans and summary

Why do we want to migrate our computing to the cloud ?

Why do we want to migrate our computing to the cloud ?

Separates system administration from the application software

Utilize opportunistic (private and commercial) clouds

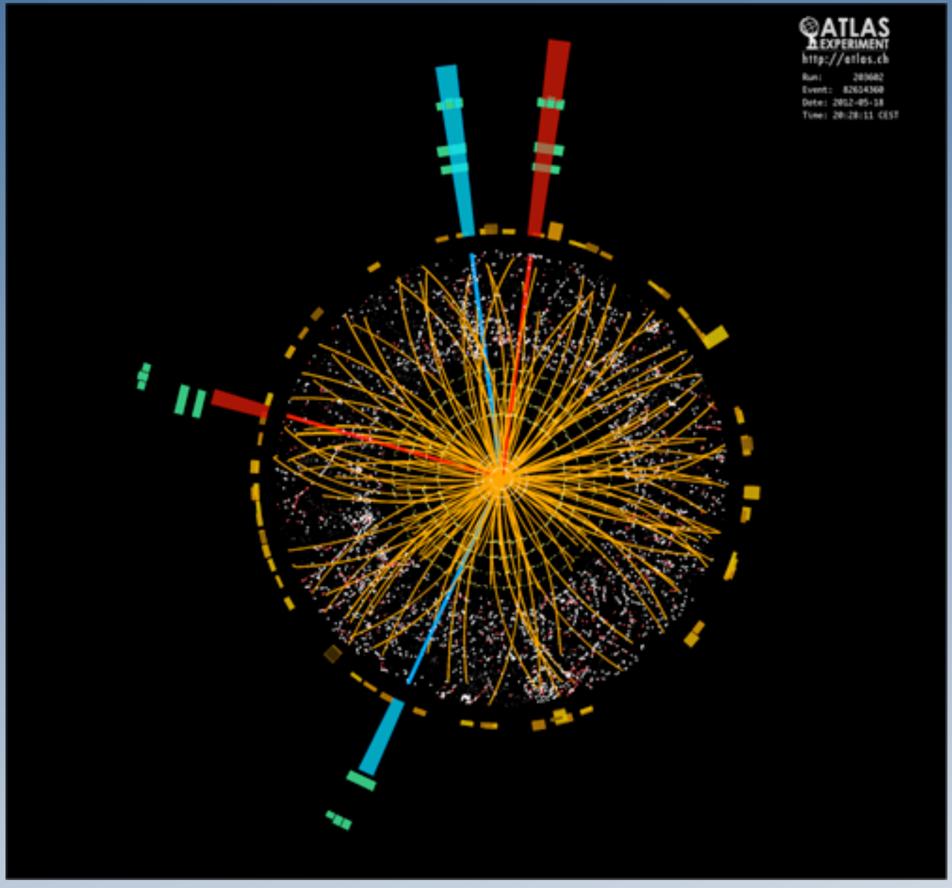
Many centres are becoming clouds as they are easier to manage

Reduces manpower and costs

Makes more effective use of our limited resources

Global effort to develop cloud computing technologies

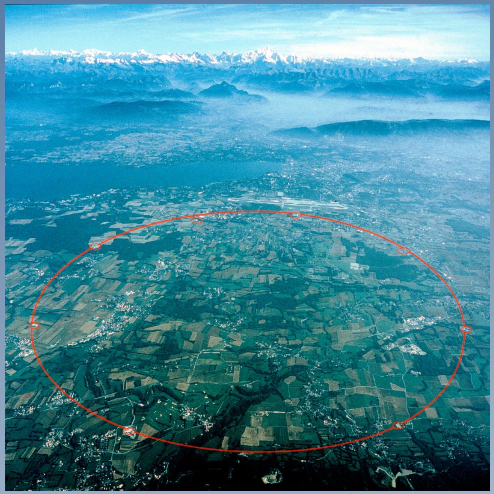
The computing challenge of HEP



Embarrassingly parallel tasks

Well suited to a cloud infrastructure

ATLAS and Belle II Projects

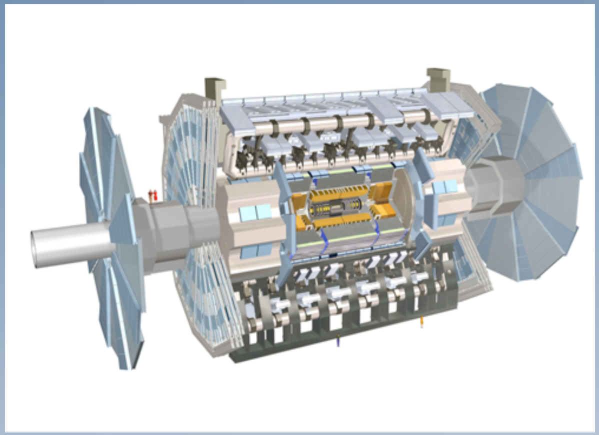


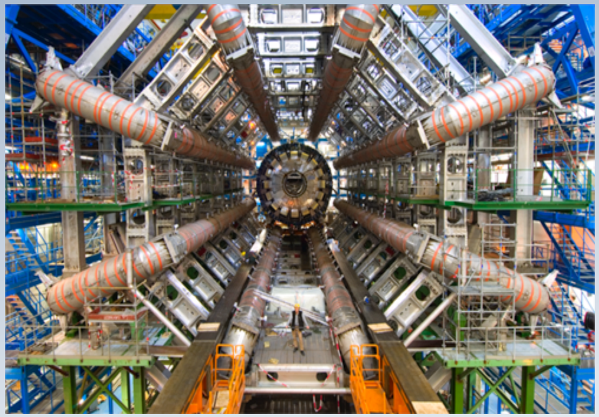
ATLAS

High-energy proton collisions
Large, general purpose detector
3000 researchers, 35 countries
Higgs discovery

100 Centres
200 PB data

Started again in 2016
Searching for “new physics”



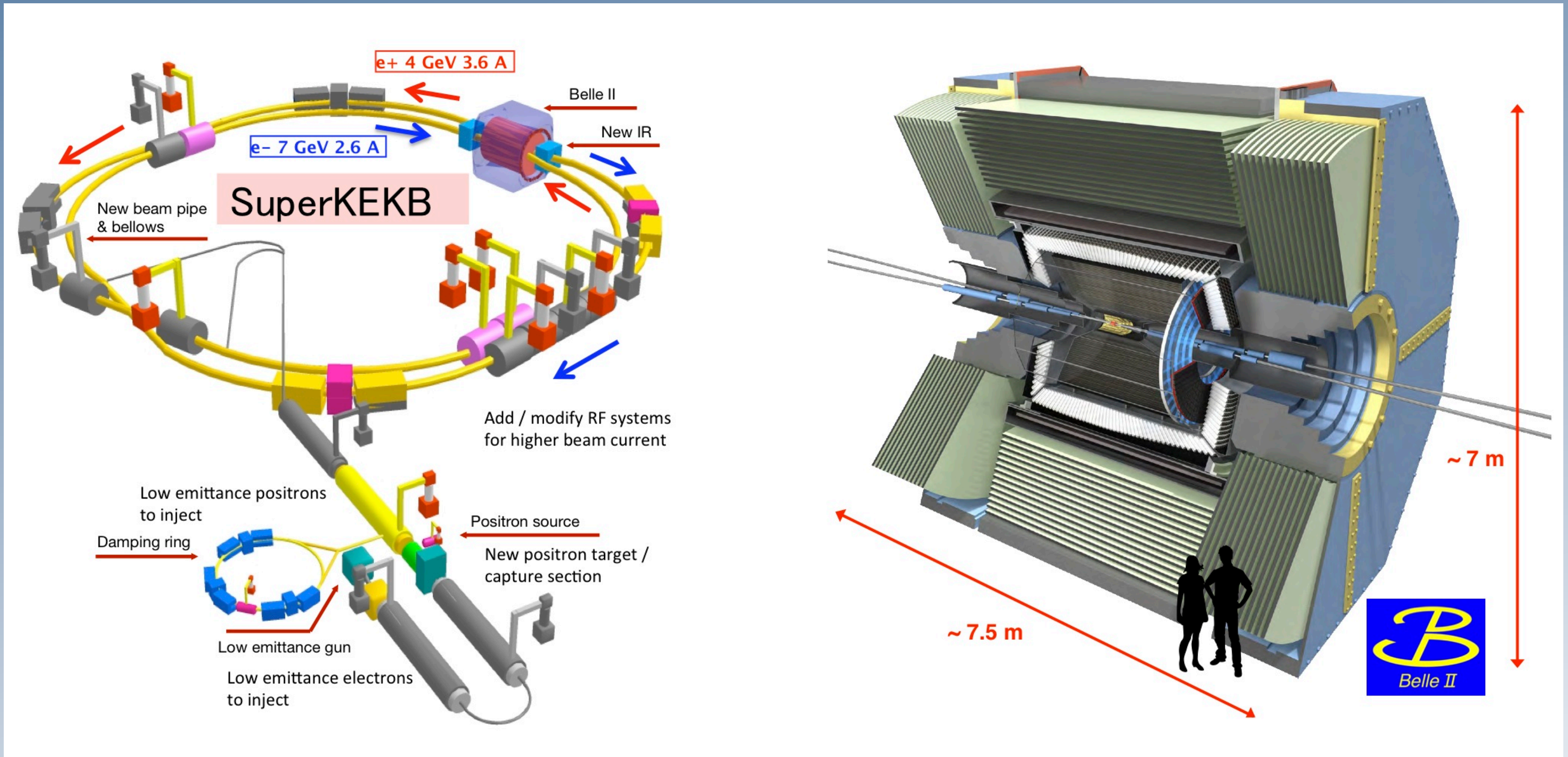


Belle II project at KEK Laboratory in Japan

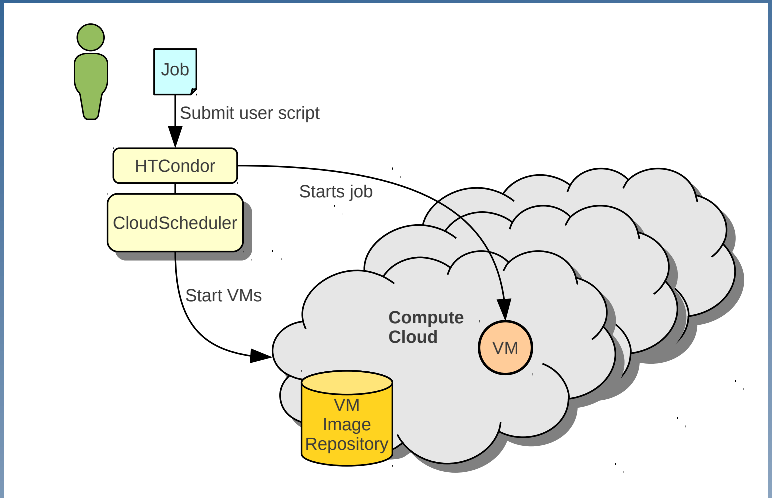
High-intensity frontier

Accelerator commissioning in 2016

Data taking starts in 2017



Search for the origin of CP asymmetry and new physics



Our goal is use distributed clouds as a single system for HEP workloads

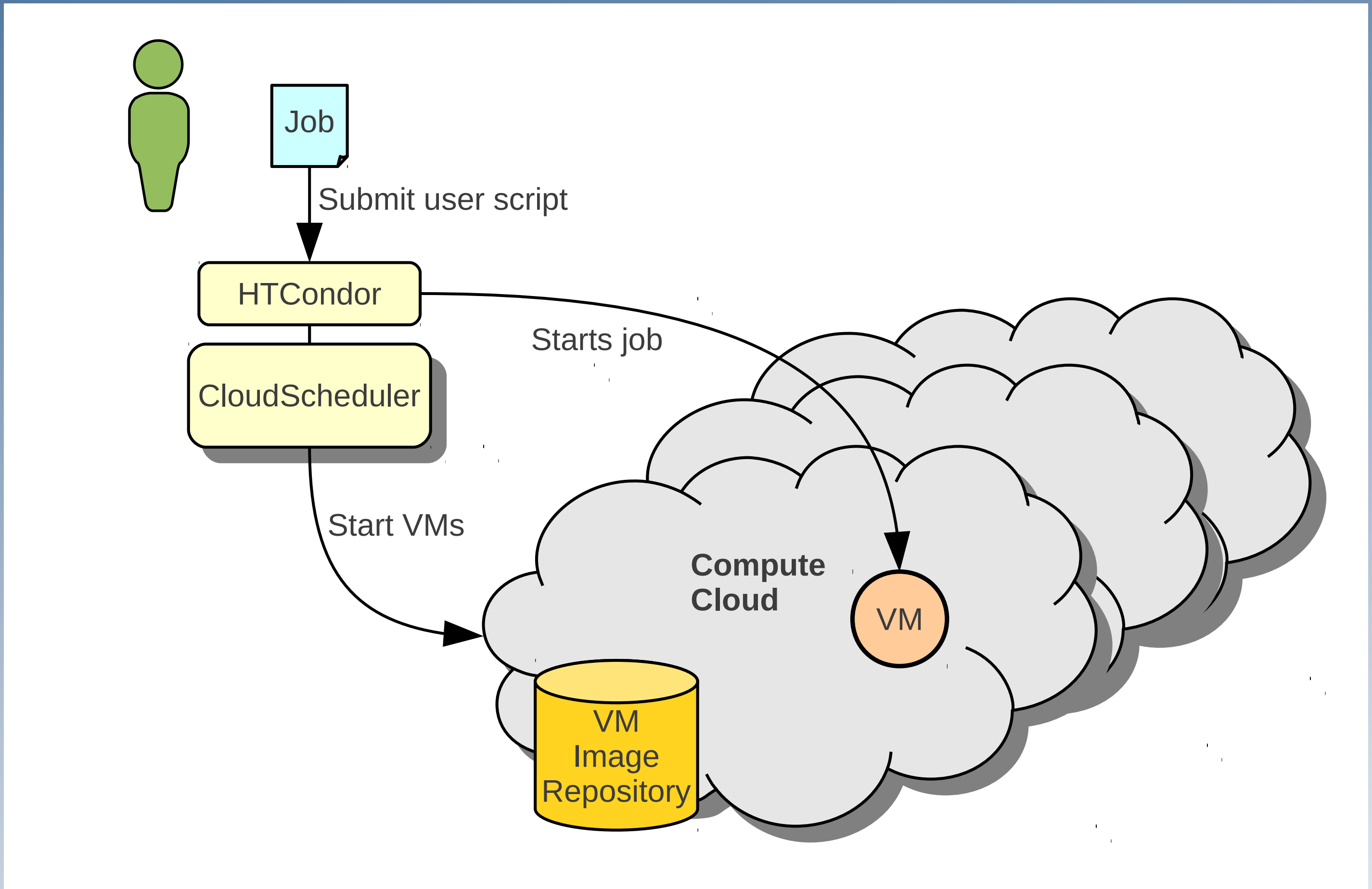
Utilize dedicated clouds (operated by HEP) and opportunistic clouds (private or commercial)

Integrate them into the project grids (e.g. WLCG) or computing infrastructure

Utilize existing software and services and only develop the missing components

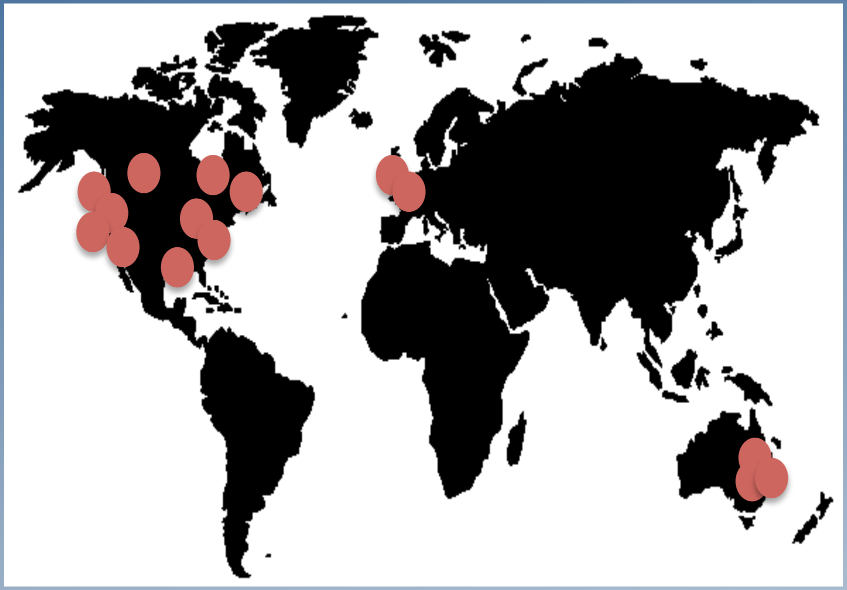
Build a system that can be used by any HEP project or other research communities

Distributed batch cloud computing

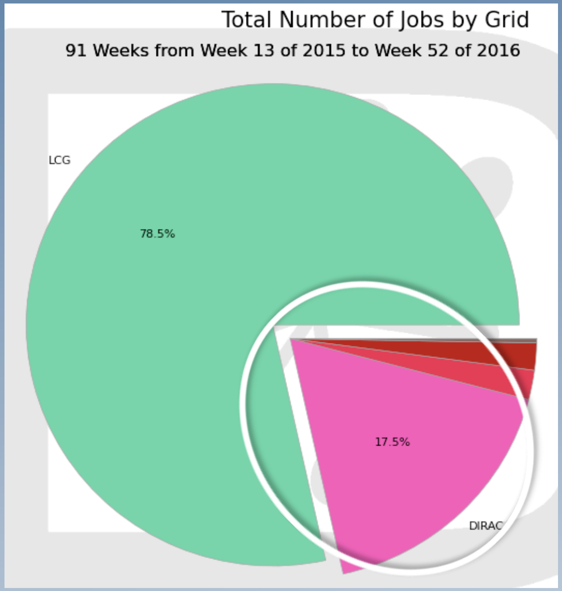


Design conceived 2008 and CloudScheduler first deployed in 2009

Distributed batch cloud system



Dedicated and opportunistic resources
(ATLAS and BelleII)



17% of BelleII computing in 2015 used clouds

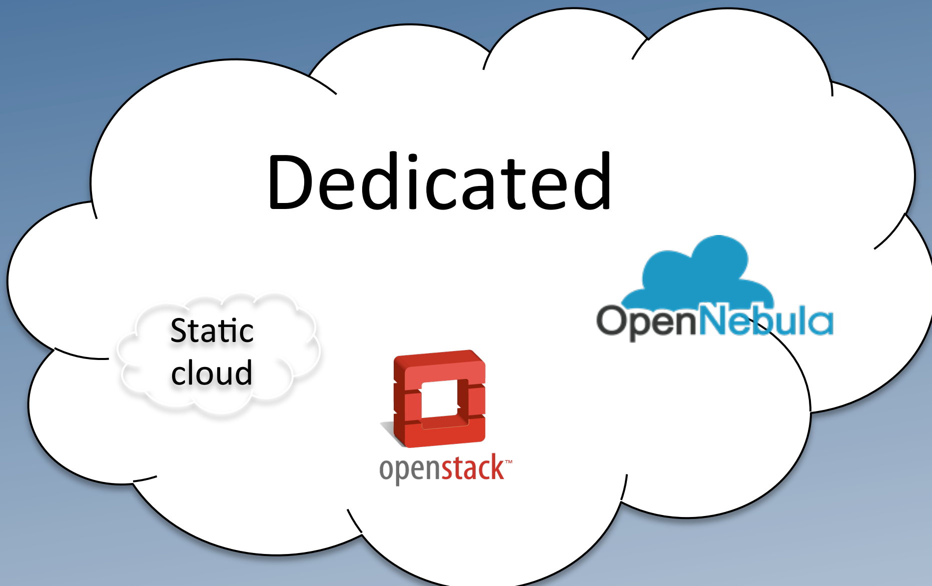
Production use of clouds for many years with gradual increase in utilization

Technology is still young and rapidly evolving

Integrating new technologies into a production system is challenging

(e.g. OpenStack cloud software only a few years old)

Cloud computing in HEP



"Dedicated" clouds
(Owned by HEP)

"Opportunistic" clouds
(private and commercial)

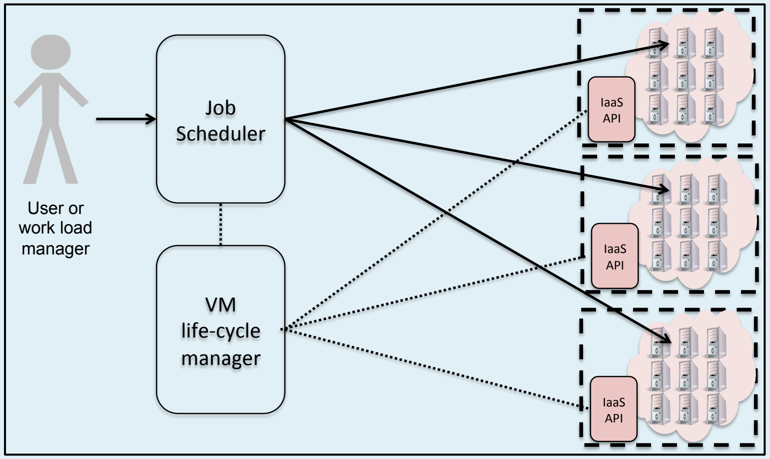
Software and services

Integration of many existing, open-source components
(Only develop missing elements)

Panda, DIRAC, HTCondor-client
Client job submission

HTCondor
Batch job system

CloudScheduler
VM provisioning and management



OpenStack
Amazon EC2
Microsoft Azure
(GCE)

microCernVM (cloud_init)

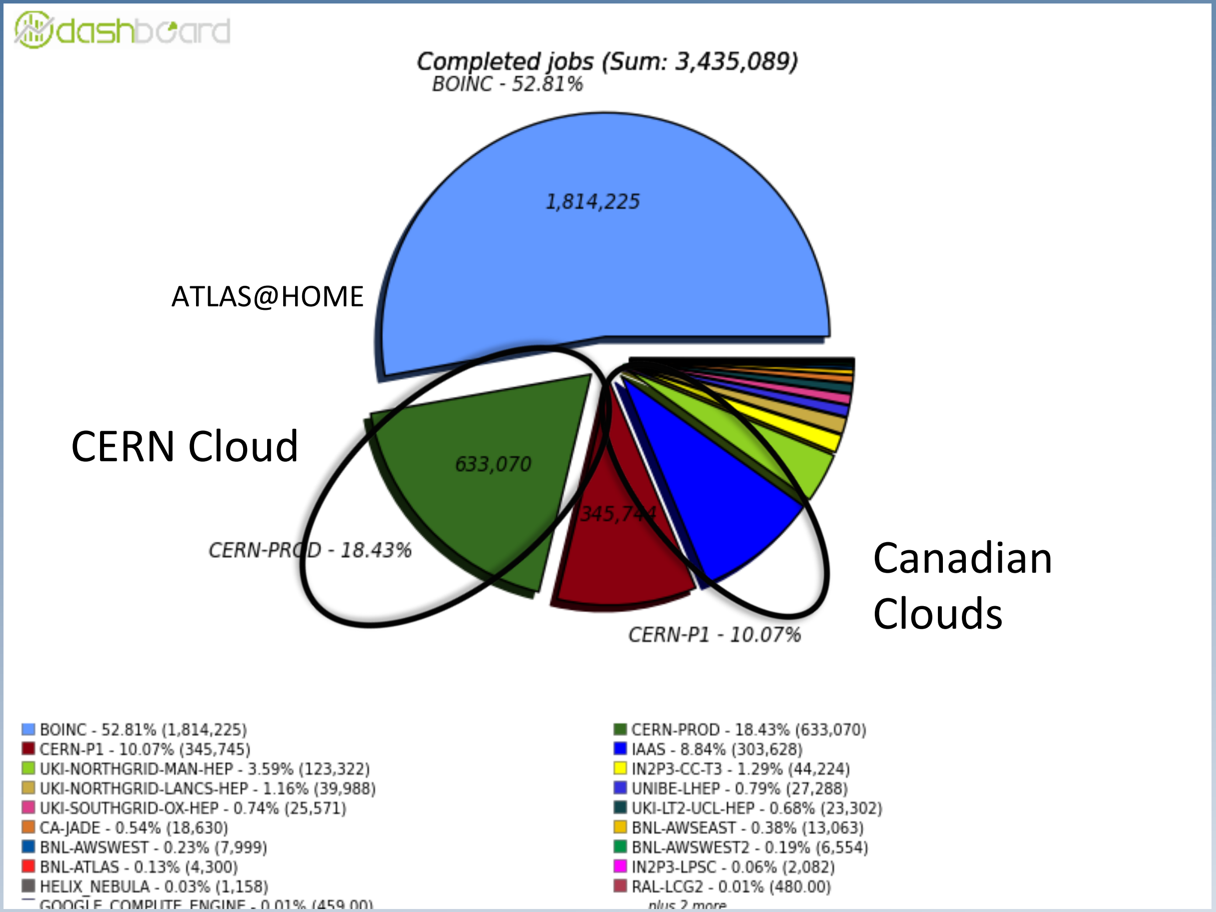
Glint
VM distribution over remote clouds

Shoal/Squids/CVMFS
Squid cache discovery service

Munin/Ganglia/Grafana/…
Monitoring systems

Production system for many years
On-going development to manage technology changes, improve reliability and adding new capabilities

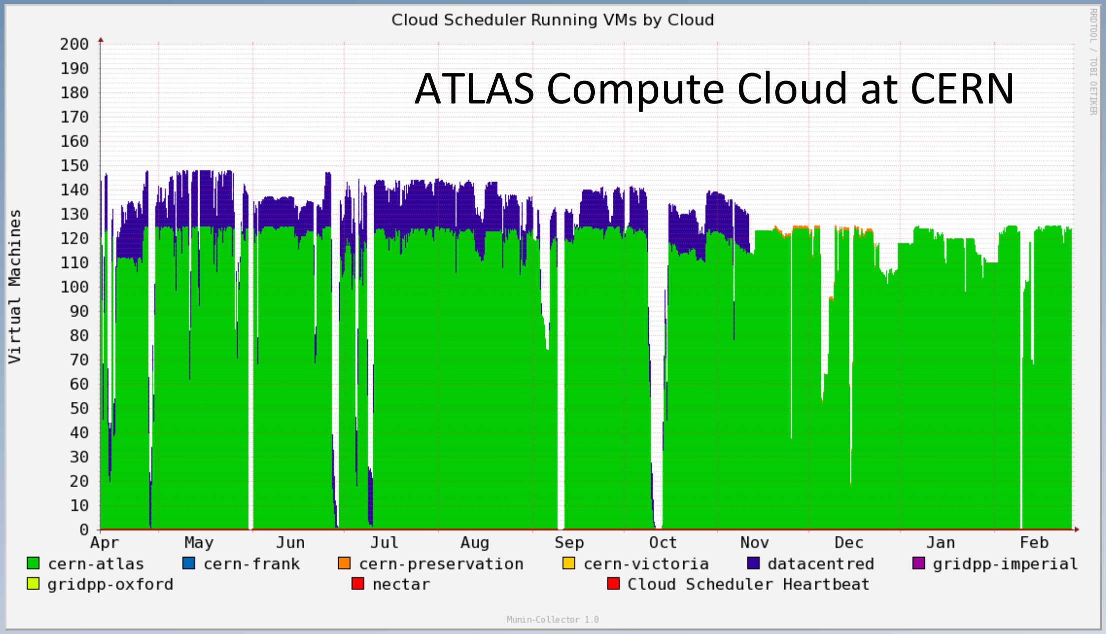
ATLAS Cloud use in 2016



UVIC group manages cloud resources for ATLAS at CERN and in Canada (2000-3000 cores)

Reliable systems

Clouds are as reliable as our traditional centres

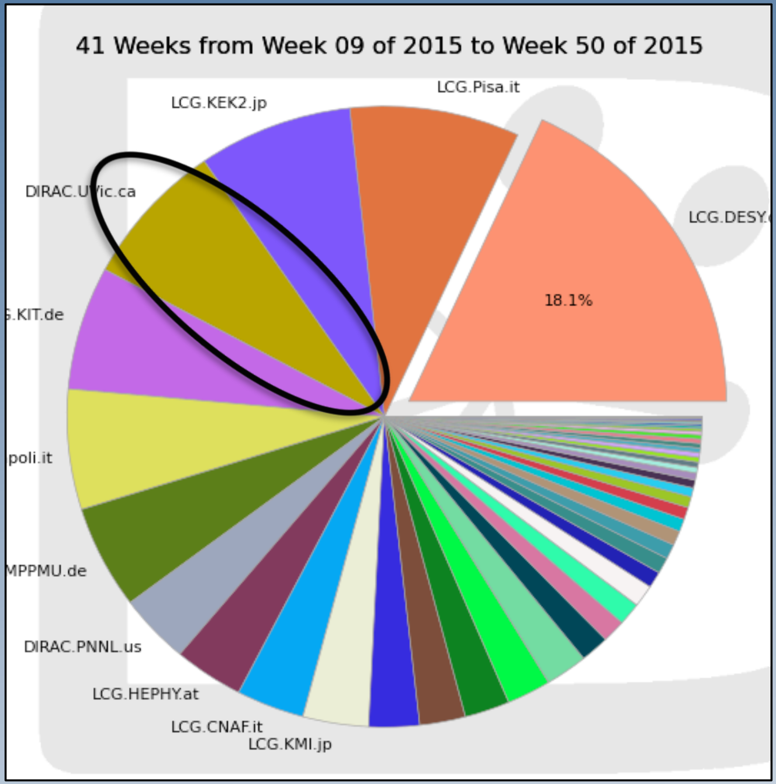


Number of virtual machines running over the 10 months

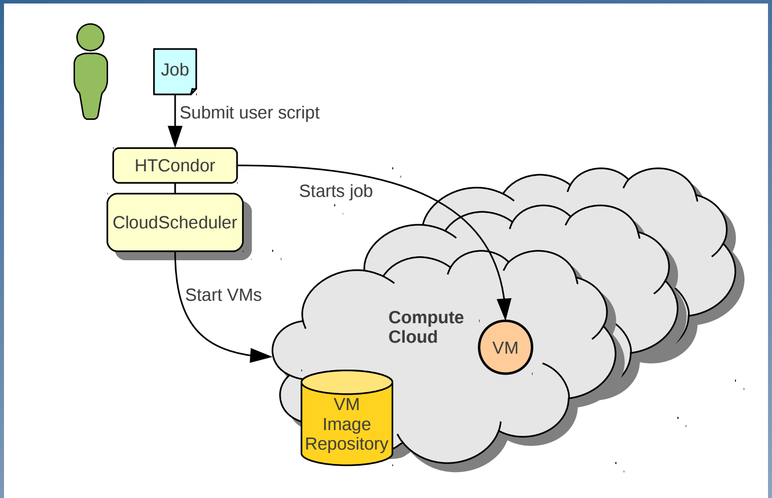
Fewer outages
(scheduled and unscheduled)

Running all low I/O, high I/O and high-memory jobs

Belle II Cloud Computing



In 2015, Canada cloud computing accounted for 10% of the total resources (2000-3000 cores)



Areas of interest and development

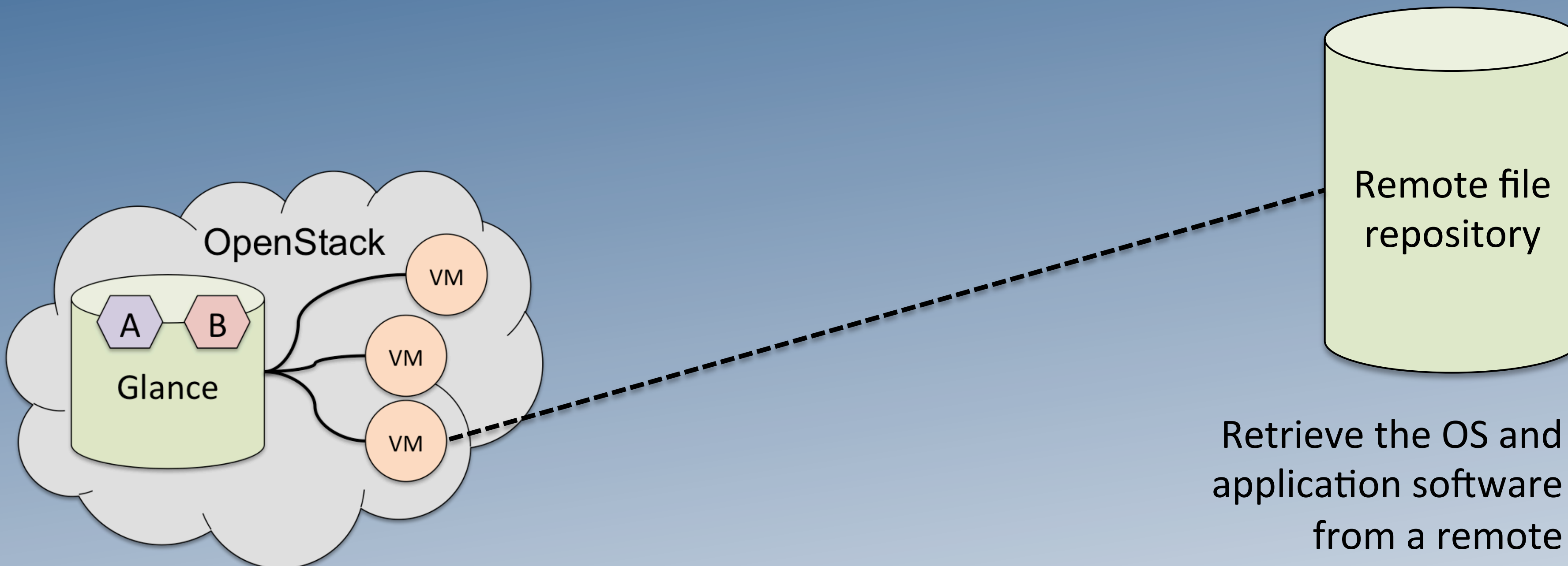
Dynamic software management

VM image management in a distributed cloud system

Data federations

Monitoring and context-aware cloud computing

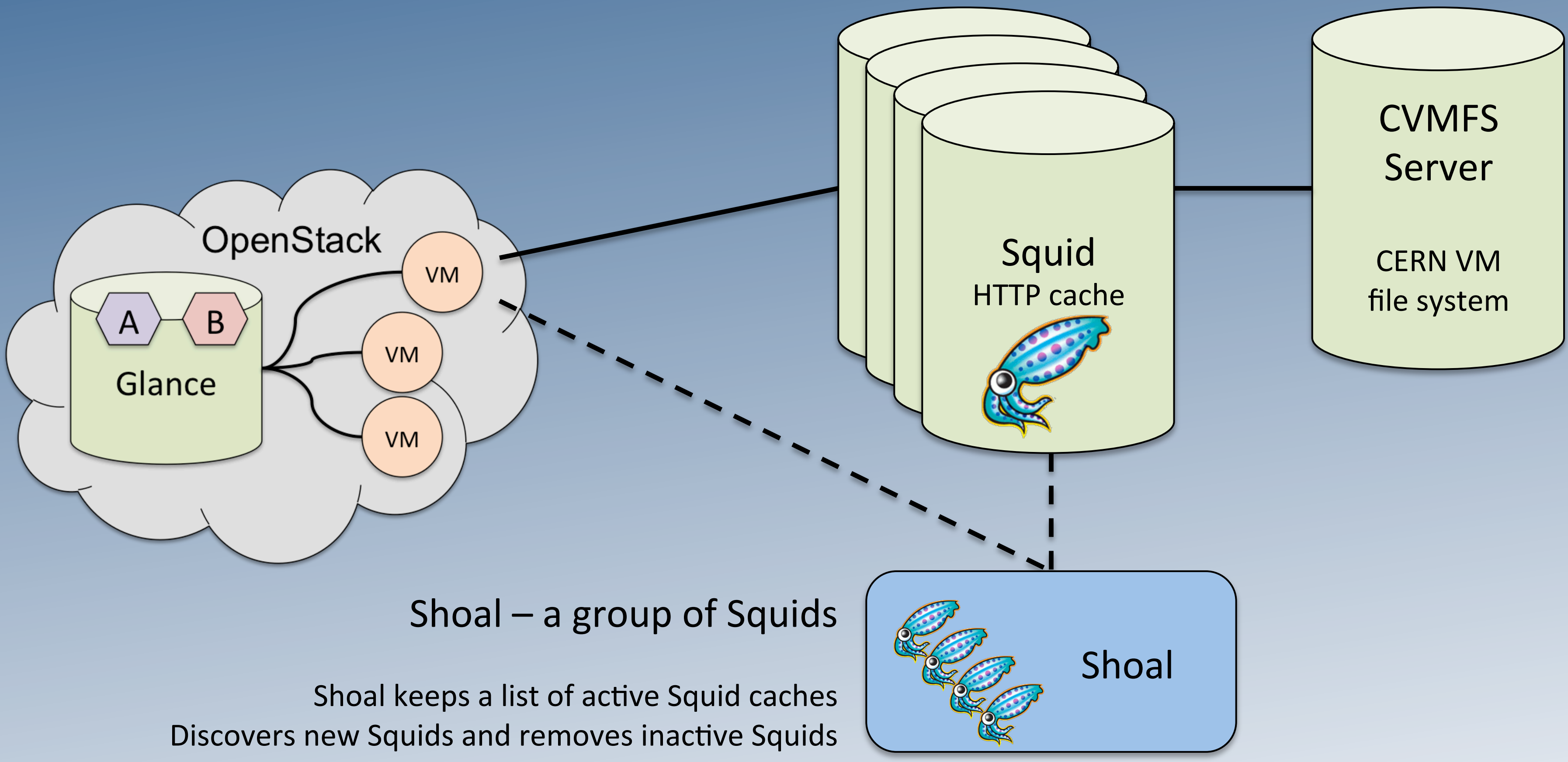
Micro-images and remote software



Retrieve the OS and application software from a remote repository

Micro-CernVM images
(only enough OS to boot the image)

OS and application software distribution

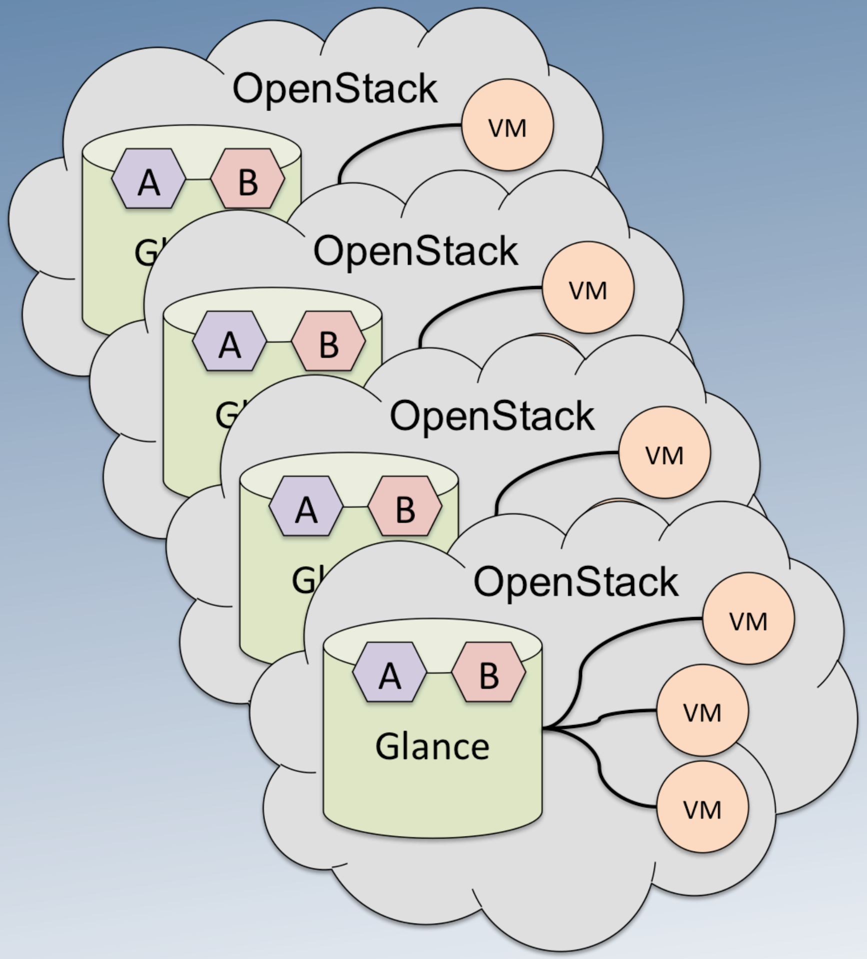


Shoal – a group of Squids

Shoal keeps a list of active Squid caches
Discovers new Squids and removes inactive Squids
VM queries Shoal for the location of the nearest Squid
Shoal: github.com/hep-gc/shoal

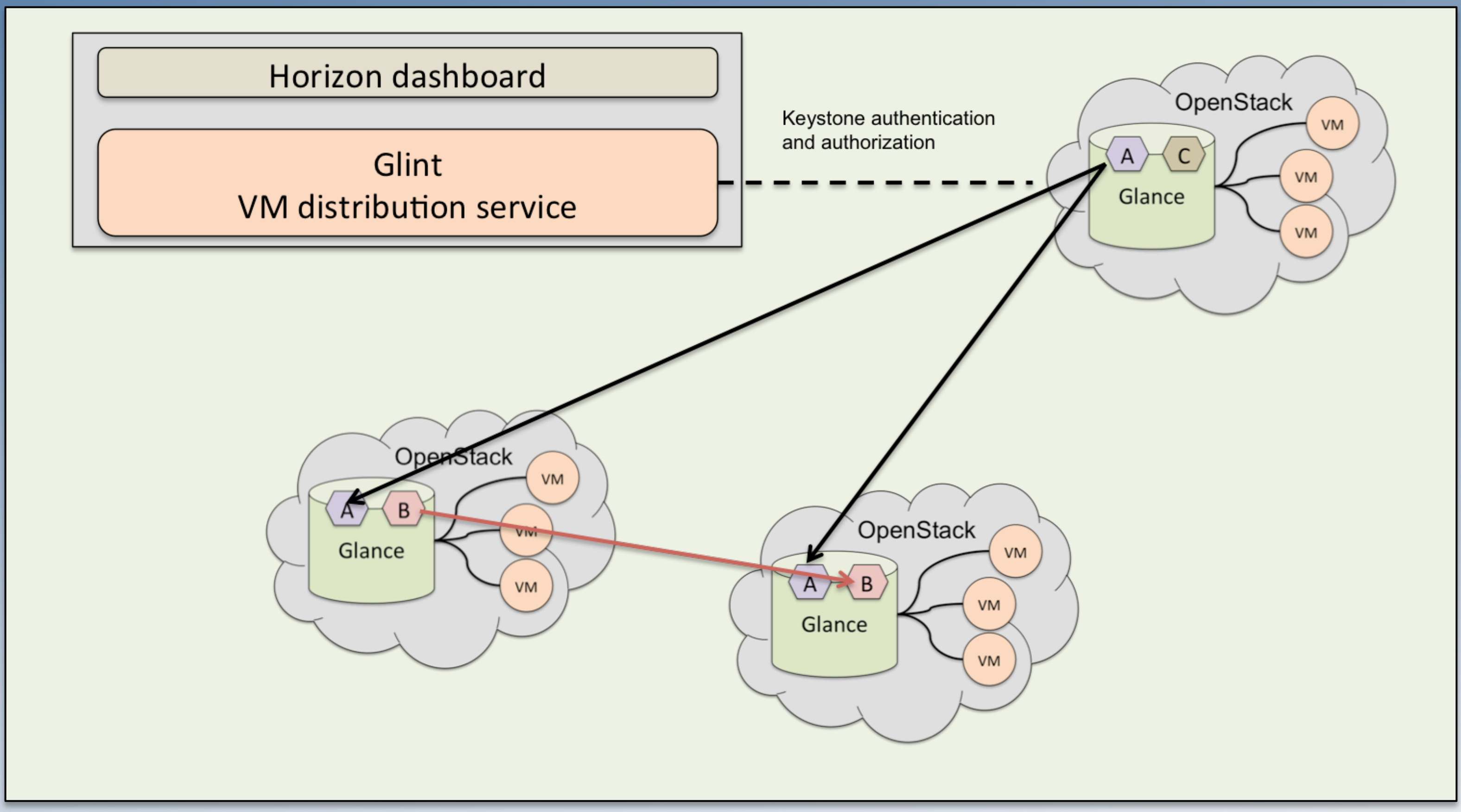
VM image management

How do we ensure the VM images are consistent on all clouds?



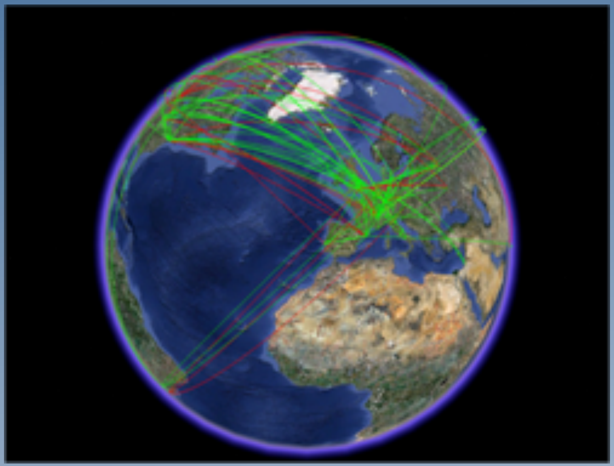
Glint – a VM distribution service

Glint provides the user the ability to manage their images on multiple OpenStack clouds



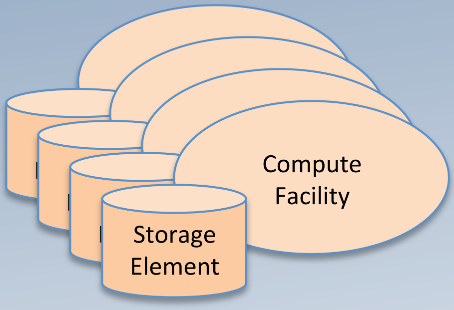
Integrated in the OpenStack Glance GUI

Data Management

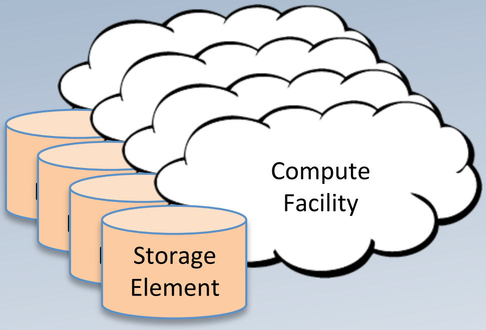


Approximately 100 computing sites
Each HEP site has CPU and storage
10-100G network

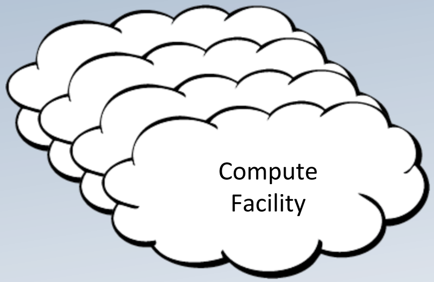
We want to use the data on all clouds



Non-cloud HEP

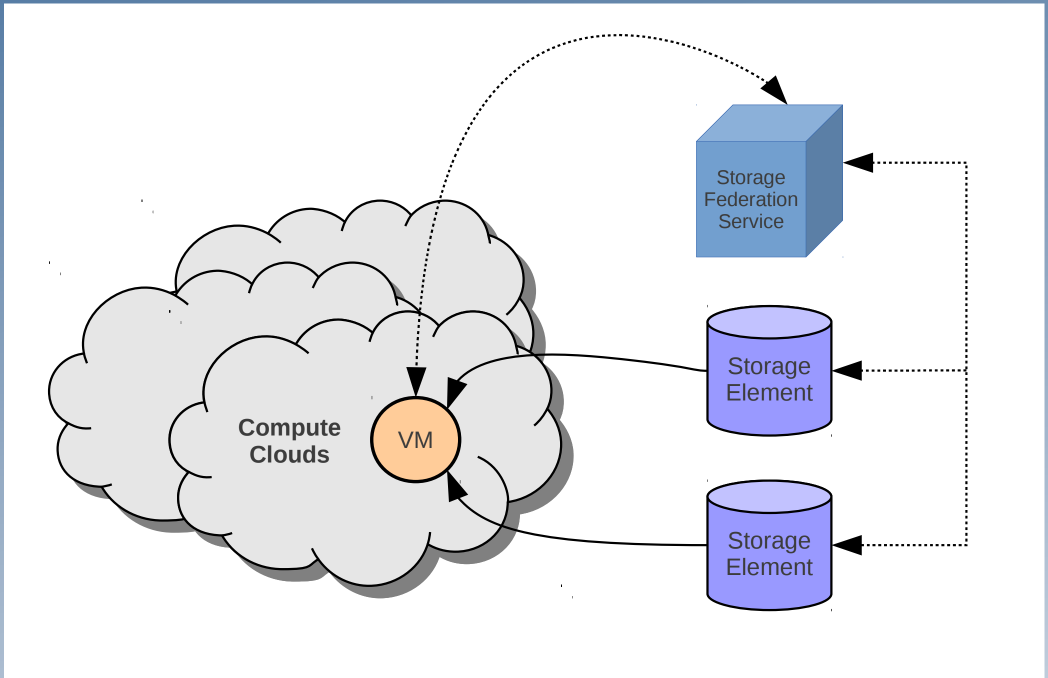


Cloud HEP



Non-HEP cloud

Data Federation



We use WebDAV for file access
Open-source HTTP protocol standard

Generic Data Federator
Developed at CERN

Presents a unified view of the data file tree to the application

Dynamic discovery

Directs the VM to the nearest Storage Element with the input data

GeoIP information to pick the closest site

Context-aware systems in mobile communication

How can we use the information (context) of a user to their benefit?

Location
Direction
Time
Health
Social

Context-aware systems in mobile communication

The information can meet on-demand request or be proactively sent to the user

Directions
Food and lodging
Financial
Commercial
Health

Collecting information

Challenging area – developers and physicists look at numbers differently

General status page – state of HTCondor, Cloudscheduler, clouds, jobs

Cloud Monitor Last 1 hour | Export | Refresh 53

ATLAS-Cern Monitor CS HTCondor

cern-atlas (125)

Cloud	CloudScheduler VMs Starting	Running	Retiring	Error	Condor Slots 1	2	3	4	5	6	7	8	Idle VMs
cern-atlas cern-worker	0	53	0	0	52	52	53	4	0	0			0
cern-atlas cern-mcore-worker	0	71	0	0	71								0

Jobs	Condor Job Status Total	Idle	Running	Completed	Held
All	431	199	232	0	0
Analy	2	0	2	0	0
Himem	258	99	159	0	0
MCore	171	100	71	0	0

Belle-II Monitor CS HTCondor

amazon (30) amazon2 (30) azure (2) cc-east (60) chameleon (2) cybera-c (6) dair-ab (3) dair-qc (3) ecloud (5) mouse (5) nefos (80)

Cloud	CloudScheduler VMs Starting	Running	Retiring	Error	Condor Slots 1	2	3	4	5	6	7	8	Idle VMs
azure belle-worker	0	2	0	0	2								0
chameleon belle-worker	0	2	0	0	2	2	2						0
mouse belle-worker	0	0	1	0	0	1	1	0	1	0	0	0	0

Jobs	Condor Job Status Total	Idle	Running	Completed	Held
All	321	260	11	0	50

IAAS Monitor CS HTCondor

cc-east (100) cc-west (75) dair-ab (4) dair-qc (6)

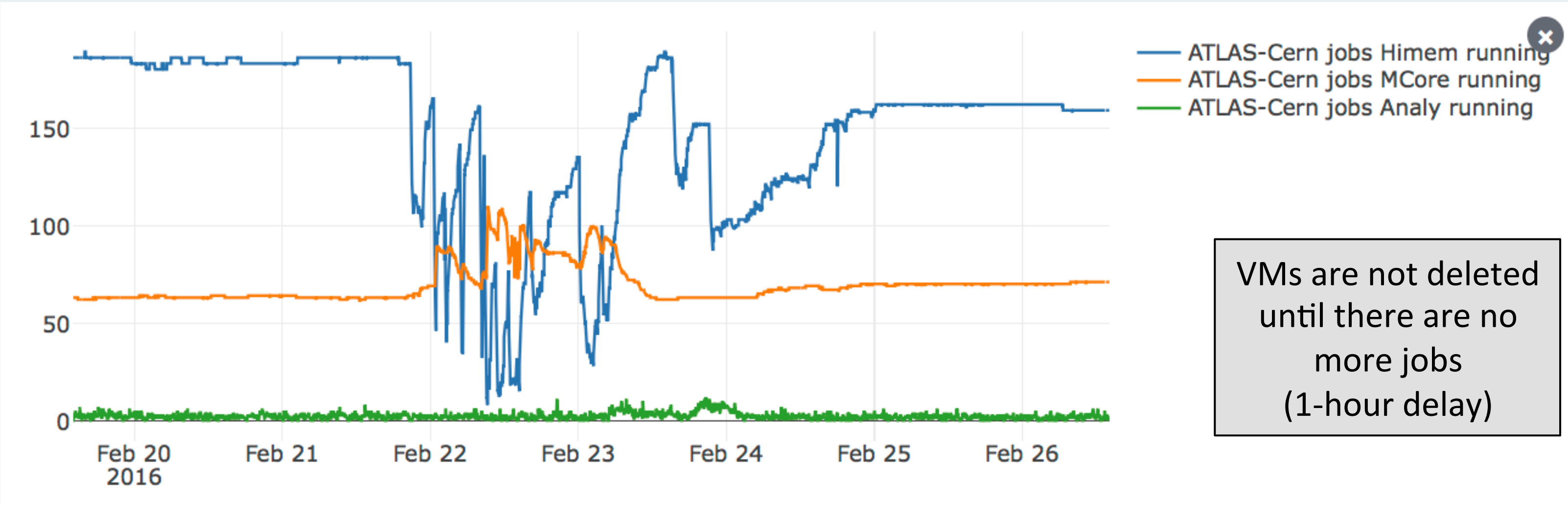
Cloud	CloudScheduler VMs Starting	Running	Retiring	Error	Condor Slots 1	2	3	4	5	6	7	8	Idle VMs
cc-east atlas-worker	0	52	28	0	63	60	64	69	61	64	63	64	0
cc-east atlas-mcore-worker	0	20	1	0	21								0
cc-west atlas-worker	0	37	0	0	37	37	37	37	37	37	37	37	0
cc-west atlas-mcore-worker	0	32	6	0	38								0
dair-ab atlas-worker	0	4	0	0	4	4	4	4	4	4	4	4	0
dair-qc atlas-worker	0	1	0	0	1	1	1	1	1	1	1	1	0
dair-qc atlas-mcore-worker	0	5	0	0	5								0
Jade jade-worker	0	0	0	2	1	1	1	0	1				0

Jobs	Condor Job Status Total	Idle	Running	Completed	Held
All	972	51	920	1	0
1 Core	879	27	852	0	0
8 Core	89	24	64	1	0
Alberta	4	0	4	0	0
Analy	0	0	0	0	0
Himem	0	0	0	0	0
MCore	0	0	0	0	0

We can plot any number as a function of time

Running jobs on ATLAS Compute Cloud for Feb 20-26 (all, High memory, analysis, multi-core)

Cloud Monitor Last 7 days | Export | Refresh 34



VMs are not deleted until there are no more jobs
(1-hour delay)

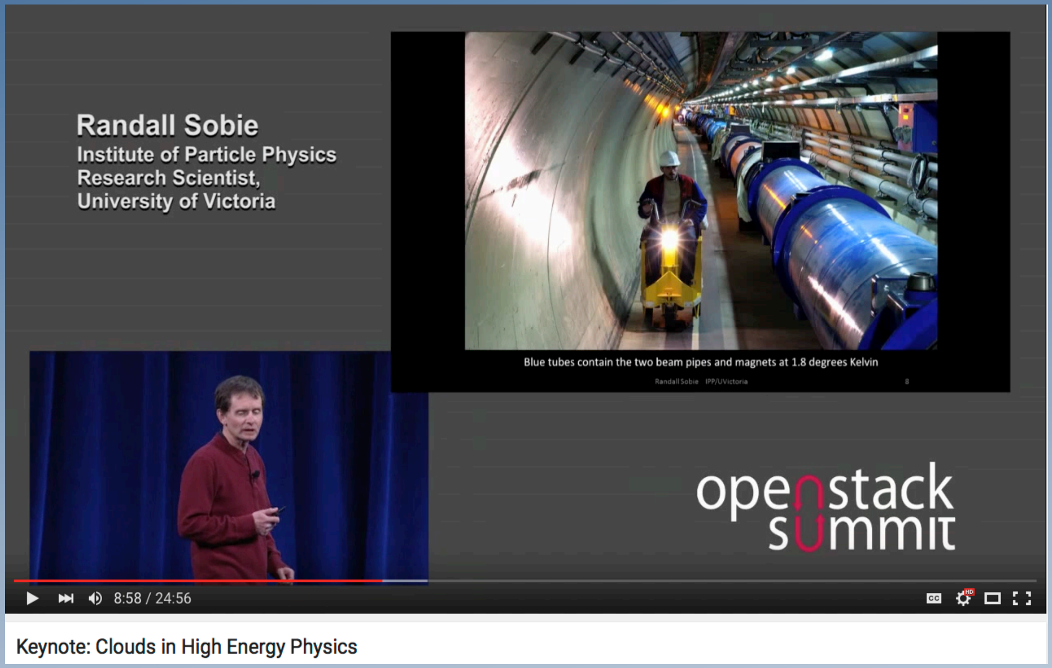
ATLAS-Cern | Monitor | CS | HTCondor

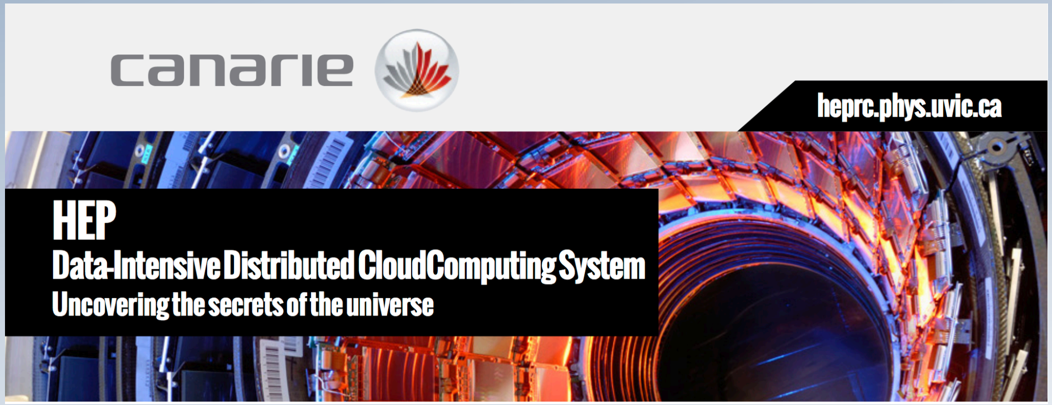
cern-atlas (125)

Cloud	CloudScheduler VMs Starting	Running	Retiring	Error	Condor Slots 1	2	3	4	5	6	7	8	Idle VMs
cern-atlas cern-worker	0	53	0	0	52	52	53	3	0	0			0
cern-atlas cern-mcore-worker	0	71	0	0	71								0

Jobs	Condor Job Status Total	Idle	Running	Completed	Held
All	431	200	231	0	0
Analy	1	0	1	0	0
Himem	259	100	159	0	0
MCore	171	100	71	0	0

Cloud computing impact





Non-traditional support ($2M)

In-kind grants and awards
(Amazon, Microsoft, Google)

Many conference presentations and papers (and videos)

Attractive for undergraduate (coop) engineering and science students
(50 students on cloud and network projects)

Application to CFI Cyberinfrastructure competition ($2M) for manpower for cloud computing and big data for HEP

HQP

Computing and networks in HEP provide an excellent training ground for HQP

Staff typically stay 1-3 years before transitioning to industry in Canada and abroad

Two staff members in CERN-IT department

Recent staff moved to Silicon Valley startup, Ottawa cloud company, DELL.

Over 50 undergraduate computer engineering and science students employed in 4-8 work terms

Funding from NSERC, Google and others

HEPNET technical manager moved to DELL in 2015

Summary

- Developed a unique system to utilize distributed and independent clouds into a batch computing infrastructure
  - Leveraged existing components and services
  - Uses OpenStack, Amazon EC2 and Microsoft Azure clouds

- Expanding the functionality to run data-intensive application

- Improving the reliability and robustness with context-aware technologies

- Significant impact in HEP cloud computing