Enabling multi-cloud resources at CERN within the Helix Nebula project

D. Giordano

(CERN IT-SDC)

HEPiX Spring 2014 Workshop
23 May 2014

Outline

- The Helix Nebula Initiative
- CERN Flagship experience
- Conclusions

Timeline

Strategic Plan for a Scientific Cloud Computing infrastructure for Europe

- http://cds.cern.ch/record/1374172/files/CERN-OPEN-2011-036.pdf
- 8th August 2011

CERN-OPEN-2011-036
08/08/2011

Contacts
Dr. Maryline Lengert
ESA - European Space Agency
Senior Advisor
Maryline.Lengert@esa.int
Tel +39 06 941 80430

Dr. Bob Jones
CERN – European Organization for Nuclear Research
IT department
Bob.Jones@cern.ch
Tel. +41 22 767 14 82

2011

- Endorse the Common **Strategy**
- Agree on the **Partnership**
- Select **flagships** use cases
- Define **governance** model

2012-2013

- **Pilot** Phase
- **Deploy** flagships,
- **Analysis** of functionality, performance & financial model

2014 …

Towards an **open market for Science**

Public Event 2014

- Showcasing the Helix Nebula production platform
- Future directions and roadmap of the overall Helix Nebula Initiative

http://indico.cern.ch/e/Helix_Nebula_Cloud_Productive

Helix Nebula
Public & Private Partnership

HELIX NEBULA THESCIENCECLOUD

Strategic Plan

- Establish multi-tenant, multi-provider cloud infrastructure
- Identify and adopt policies for trust, security and privacy
- Create governance structure
- Define funding schemes

CERN

To support the computing capacity needs for the LHC experiments

EMBL

Setting up a new service to simplify analysis of large genomes, for a deeper insight into evolution and biodiversity

esa DLR cnes

To create an Earth Observation platform, focusing on earthquake and volcano research

PIC port d'informació científica

To improve the speed and quality of research for finding surrogate biomarkers based on brain images

Suppliers

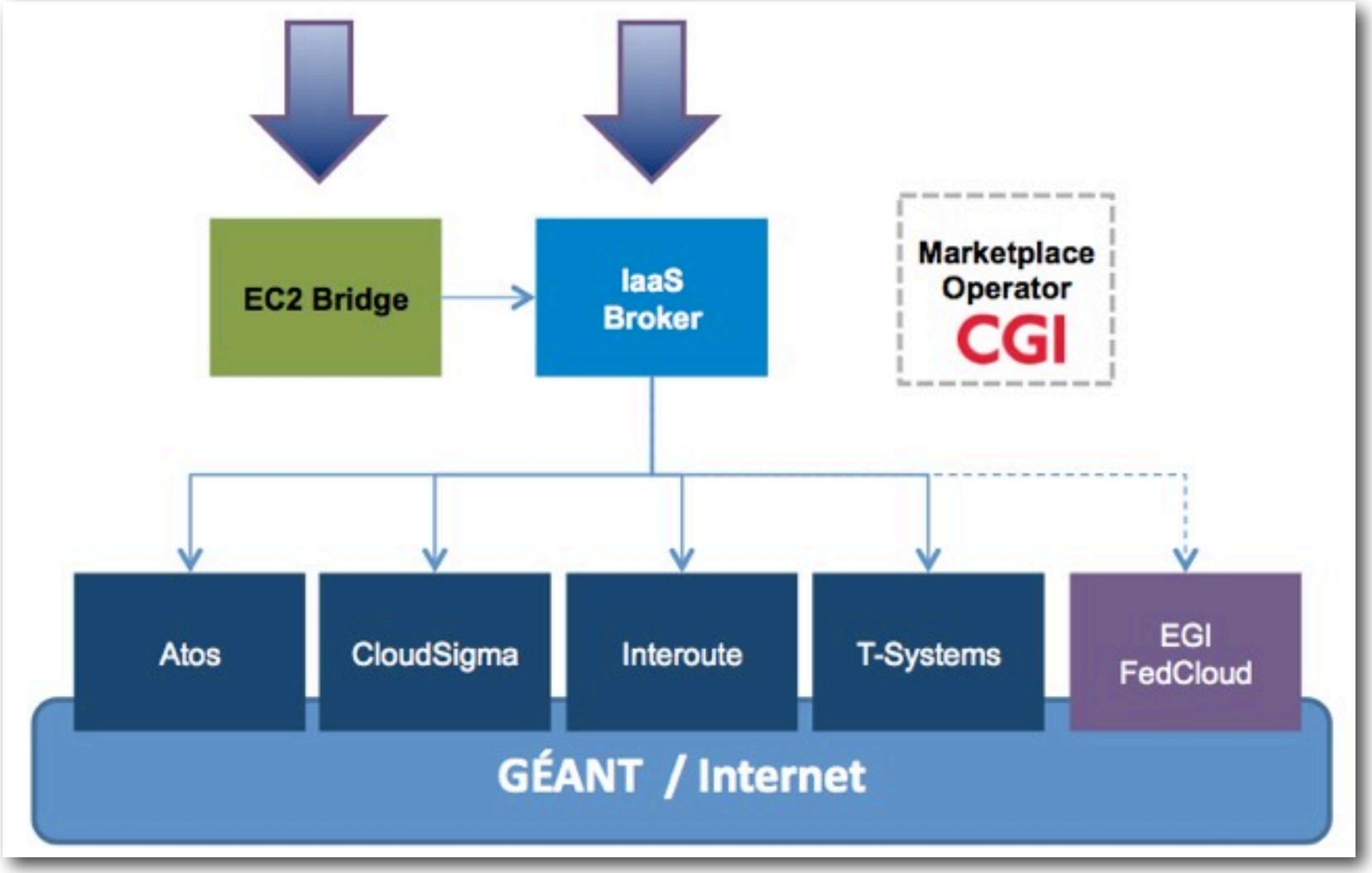
Atos CGI CloudSigma CSA cloud security alliance DANTE www.dante.net egi indra interoute MEMSET HOSTING SAP sixsq SWITCH T-Systems terradue 2.0 the SERVER LABS the IT architects ultimum technologies Yandex

Adopters

AWST Capgemini cnrs ECMWF Ifremer orange Business Services Telefónica THALES Trust-IT

Helix Nebula Platform

- Initial four commercial cloud providers integrated
- IaaS Broker: SlipStream BlueBox
- Amazon EC2 Bridge for compatibility with third party tools, such as StarCluster or any EC2-compatible tool
- Integration with the EGI FedCloud on the 2014 roadmap



Broker

SlipStream®: "1-click multi-cloud application deployment platform."

- Open Source PaaS
- Supports all mainstream cloud solutions
- Independency from the target cloud provider
- Concurrent deployment across several clouds

http://sixsq.com/products/slipstream.html

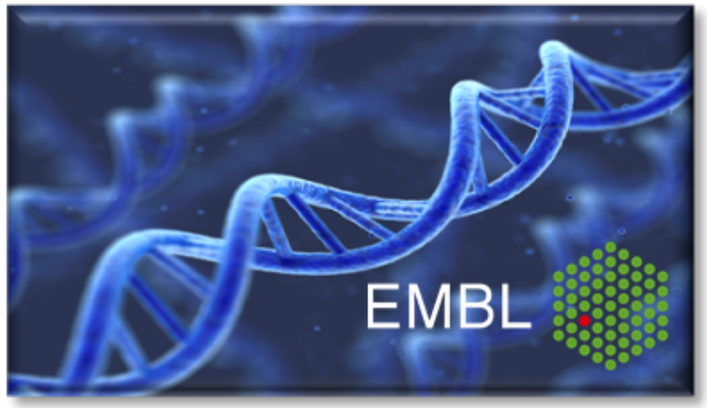
Flagship Use Cases

Stretch what is possible with the cloud today for multidisciplinary data intensive science

ATLAS High Energy Physics Cloud Use

To support the computing capacity needs for the ATLAS experiment

Genomic Assembly in the Cloud

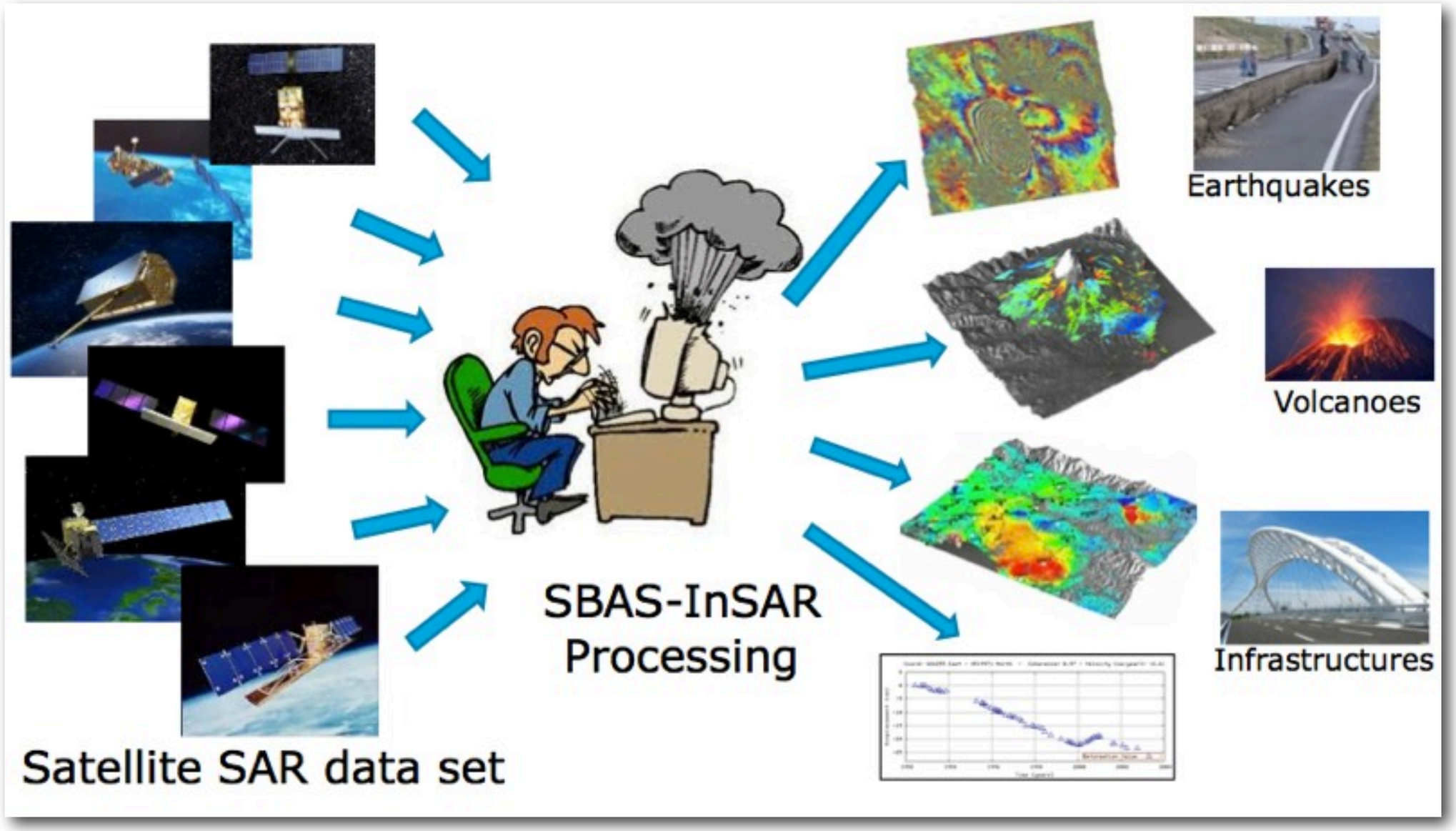


A new service to simplify large scale genome analysis; for a deeper insight into evolution and biodiversity

SuperSites Exploitation Platform

To create an Earth Observation platform, focusing on earthquake and volcano research

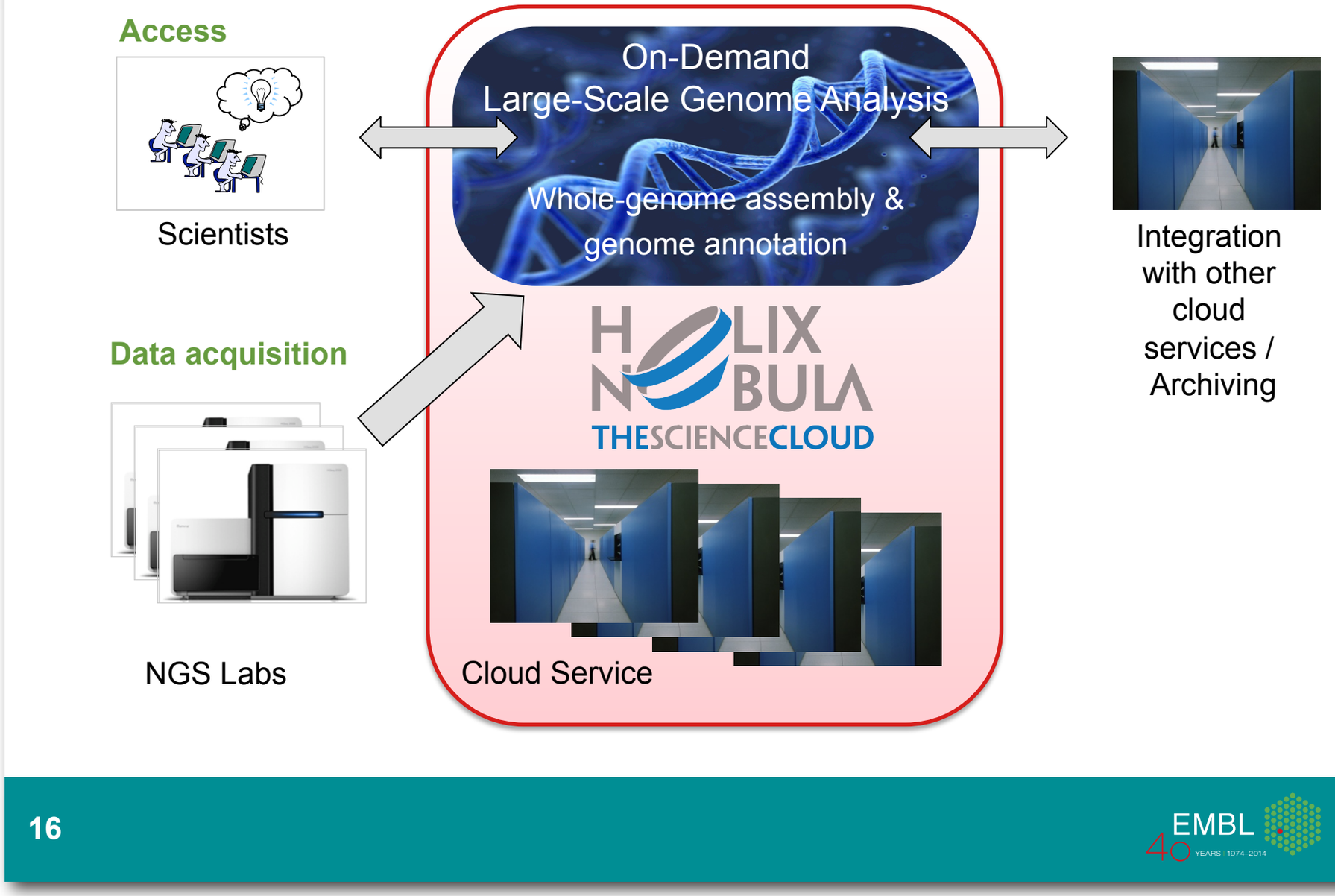
ESA Helix Nebula Flagship Use Case



Assessment of geohazards and monitoring the dynamic and complex solid-Earth system.

- Enable the correlation and processing of observation data for supersites monitoring.

EMBL Helix Nebula Flagship Use Case

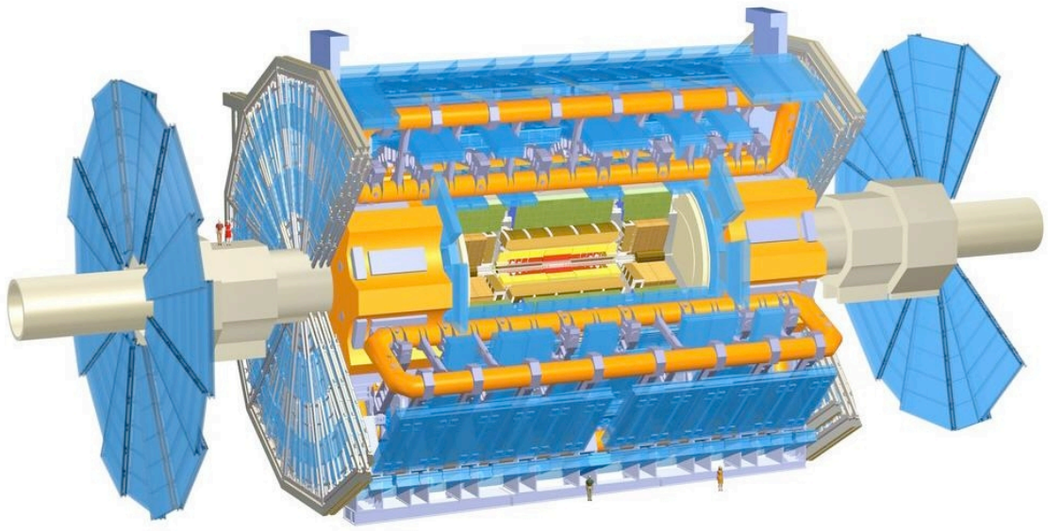


- Successful end-to-end tests of bioinformatics pipelines
- Using real world large genome sequencing data
- Mix of quick parallel jobs and long running serial jobs

Helix Nebula: CERN Flagship Use Case

Aim

- Evaluating the use of cloud technologies for LHC data processing
- Transparent integration of cloud computing resources with ATLAS distributed computing software and services
- Evaluation of financial costs of processing, data transfer and data storage
- Service Level Agreements and Governance model

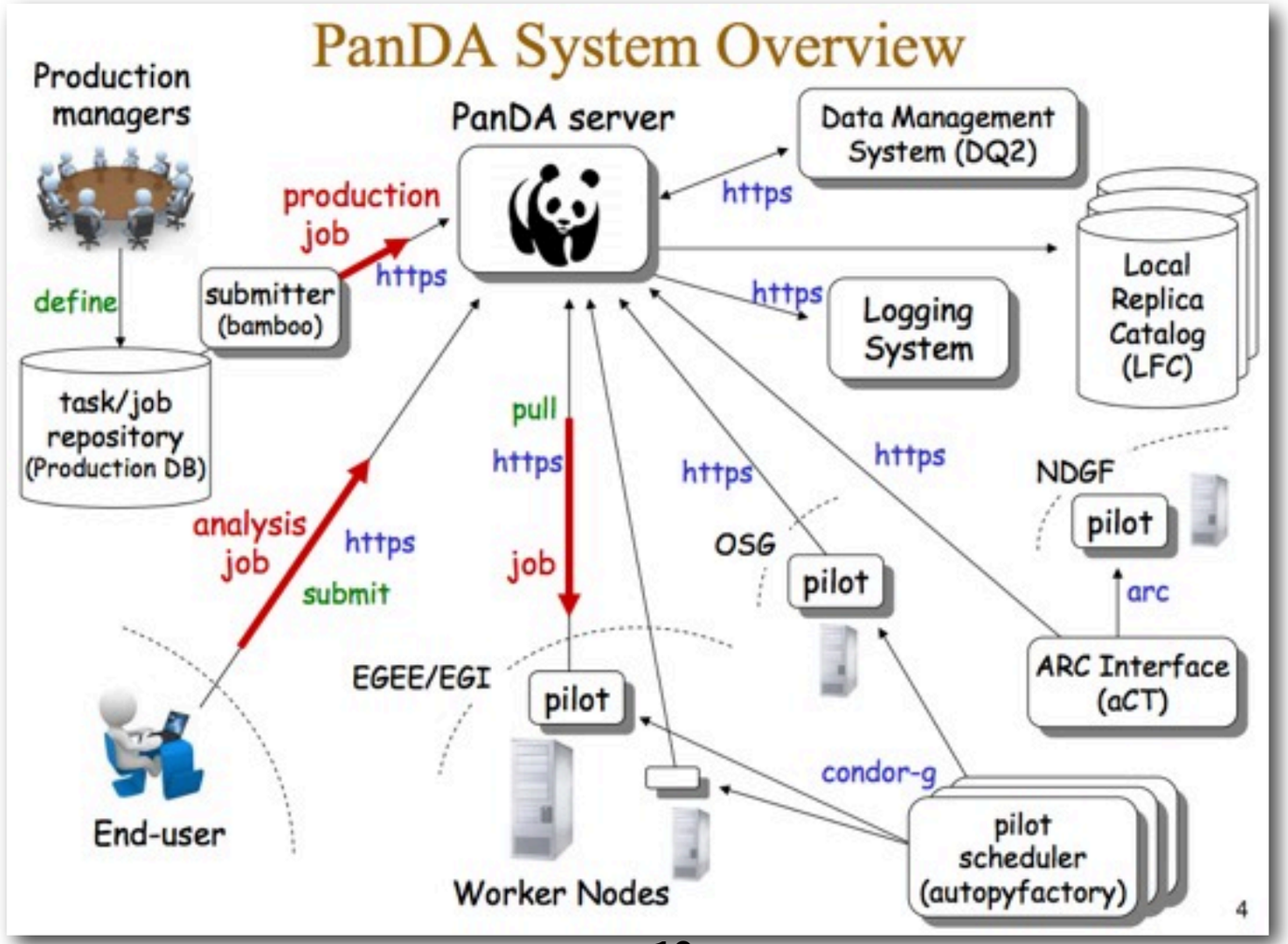


ATLAS detector

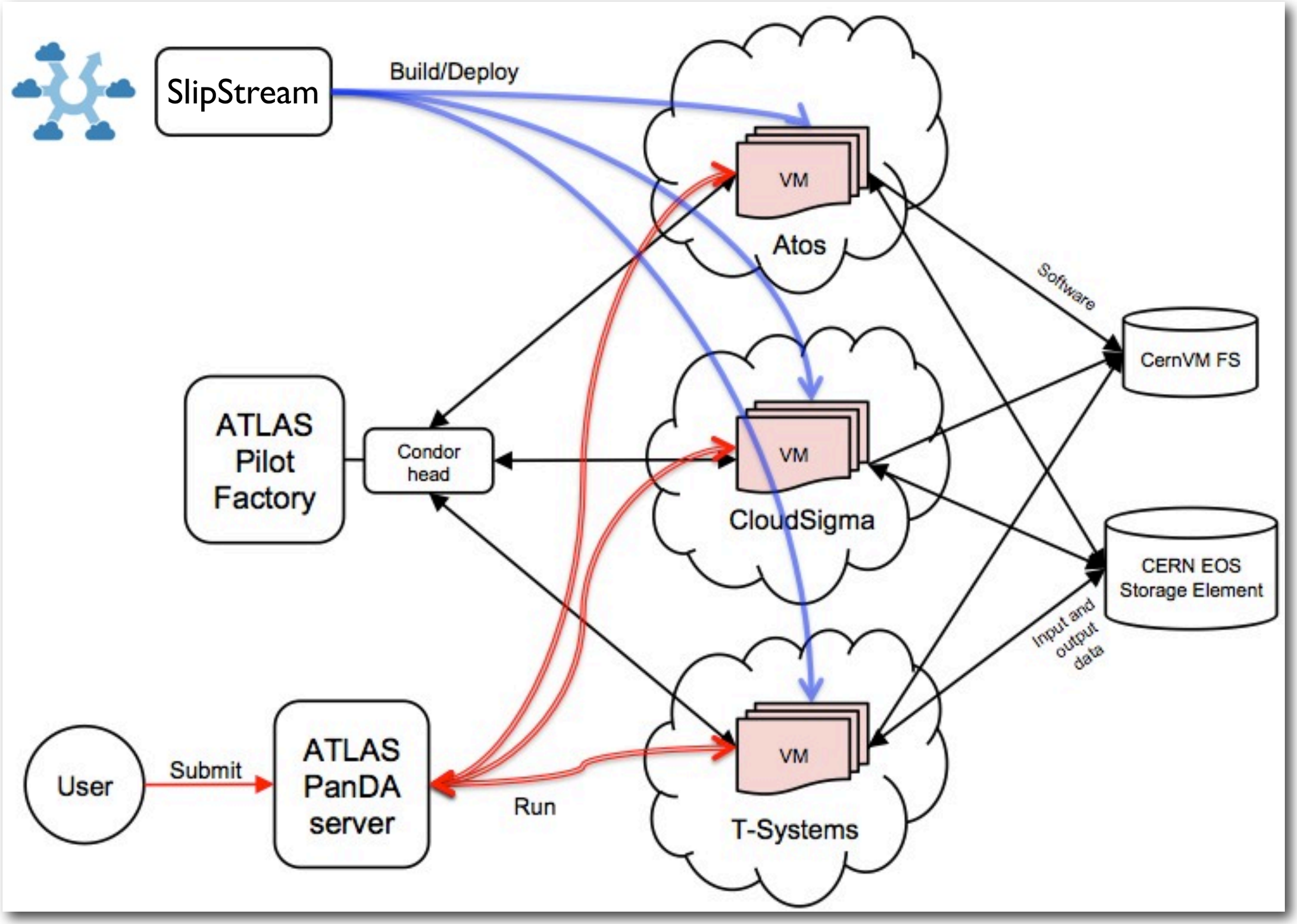
Experiment Workflow

ATLAS Production ANd Distributed Analysis (PanDA) system

- A homogeneous processing system layered over heterogeneous resources
- Use of Condor for job submission
- Use of pilot jobs for acquisition of processing resources.
- Support for both managed production and analysis



Cloud Job Flow



Setup

Homogeneous configuration across suppliers

VM Configuration

- Based on CentOS-6
- 1 vCPU, 2 GB RAM, HD size ~10 GB (instance type m1.small)
- Additional disk 20 GB needed for ATLAS jobs

Image Creation

- Install RPMs for repository configuration, condor, ATLAS worker node, EMI

Contextualization

- Additional disk partitioning/mounting
- Configure CVMFS, Condor, Ganglia
  - Define for each provider a specific PanDA & Ganglia resource: Helix_Nebula_*

Network

- VMs with public IPs, but for T-Systems (private network with gateway)

Tests

Experiment workflow tested with Monte Carlo jobs

- Geant4 based simulation of the particles propagation through the ATLAS detector

Long (~4h), very intensive CPU usage, low I/O usage.

- Input: MC generator 4 vector files. Output: ~50 MB/file of 50 events

Monitoring

Large variety of complementary monitoring views in order to track, log, cross-check, debug

- VM side (Ganglia), WMS side (PanDA, APF), SlipStream dashboard

Jobs:

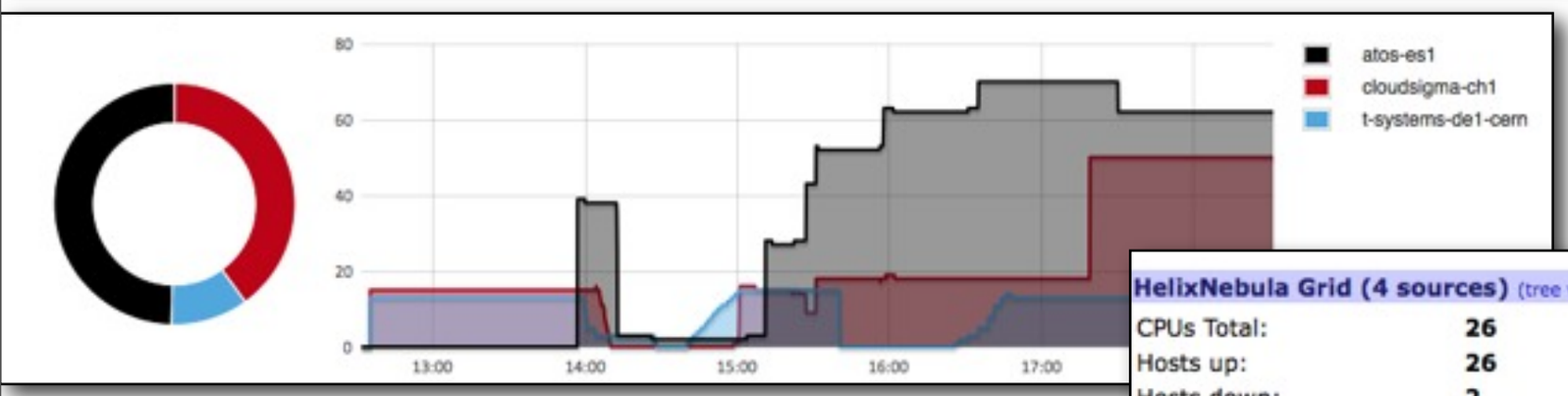
PandaID, Owner, Working group	Job	Status	Created	Time to start	Duration	Ended/ Modified	Cloud/Site, Type	Priority
2162262985 gangarbt	trans=AtlasG4_trf.py, pkg=AtlasProduction/17.2.2.2	finished	2014-05-11 10:51	1:48:55	6:15:36	05-11 18:56	CERN/CERN_HELIX_NEBULA_CloudSigma, prod_test	10000
	In: mc12_8TeV.107651.AlpgenJimmy_AUET2CTEQ6L1_ZeeNp1.evgen.EVNT.e1571_tid01001789_00 Out: hc_test.gangarbt.hc20035895.HELIX_NEBULA_CloudSigma.24							
2162262981 gangarbt	trans=AtlasG4_trf.py, pkg=AtlasProduction/17.2.2.2	finished	2014-05-11 10:51	0:11:06	6:17:37	05-11 17:20	CERN/CERN_HELIX_NEBULA_CloudSigma, prod_test	10000
	In: mc12_8TeV.107651.AlpgenJimmy_AUET2CTEQ6L1_ZeeNp1.evgen.EVNT.e1571_tid01001789_00 Out: hc_test.gangarbt.hc20035895.HELIX_NEBULA_CloudSigma.24							
2162262980 gangarbt	trans=AtlasG4_trf.py, pkg=AtlasProduction/17.2.2.2	finished	2014-05-11 10:51	0:10:19	6:58:51	05-11 18:00	CERN/CERN_HELIX_NEBULA_CloudSigma, prod_test	10000
	In: mc12_8TeV.107651.AlpgenJimmy_AUET2CTEQ6L1_ZeeNp1.evgen.EVNT.e1571_tid01001789_00 Out: hc_test.gangarbt.hc20035895.HELIX_NEBULA_CloudSigma.24							
2162262979	trans=AtlasG4_trf.py, pkg=AtlasProduction/17.2.2.2	finished	2014-05-11 10:51	0:08:03	5:02:24	05-11 16:02	CERN/CERN_HELIX_NEBULA_CloudSigma, prod_test	10000

BatchQueue (pandaq) view

Batch queue	HELIX_NEBULA_ATOS
WMS queue	HELIX_NEBULA_ATOS
Site	HELIX_NEBULA
State	test
Links	agis pandamon ssb

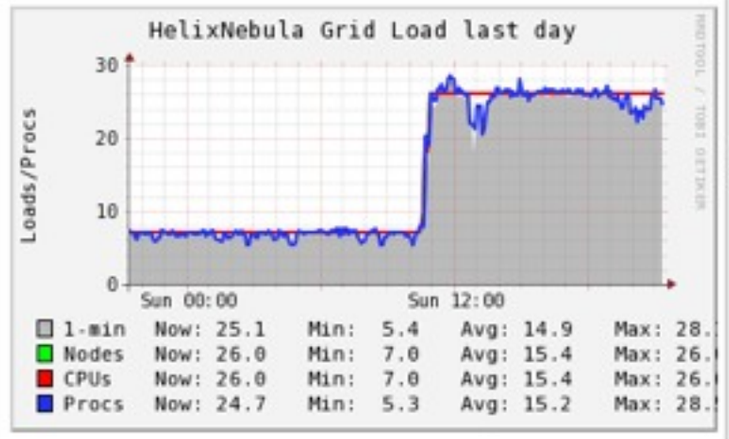
Label	factory	created	running	exiting	done	miss
HELIX_NEBULA_ATOS	aipanda002	165	7	0	0	0
HELIX_NEBULA_ATOS	aipanda009	1673	7	0	5	0
HELIX_NEBULA_ATOS	aipanda013	151	7	0	2	0

Factory	job	state	payload?	created
aipanda009	520291.0	created	-	seconds ago
aipanda002	884469.0	created	-	seconds ago
aipanda013	542345.0	created	-	seconds ago
aipanda009	520285.0	created	-	1 min ago
aipanda002	884463.0	created	-	1 min ago



HelixNebula Grid (4 sources) (tree view)

CPUs Total:	26
Hosts up:	26
Hosts down:	2



Note: Sorting by multiple columns at the same time can be activated by 'shift' clicking on the column headers which they want to add to the sort. Hovering mouse over the column headers to get descriptions of table columns.

Search:

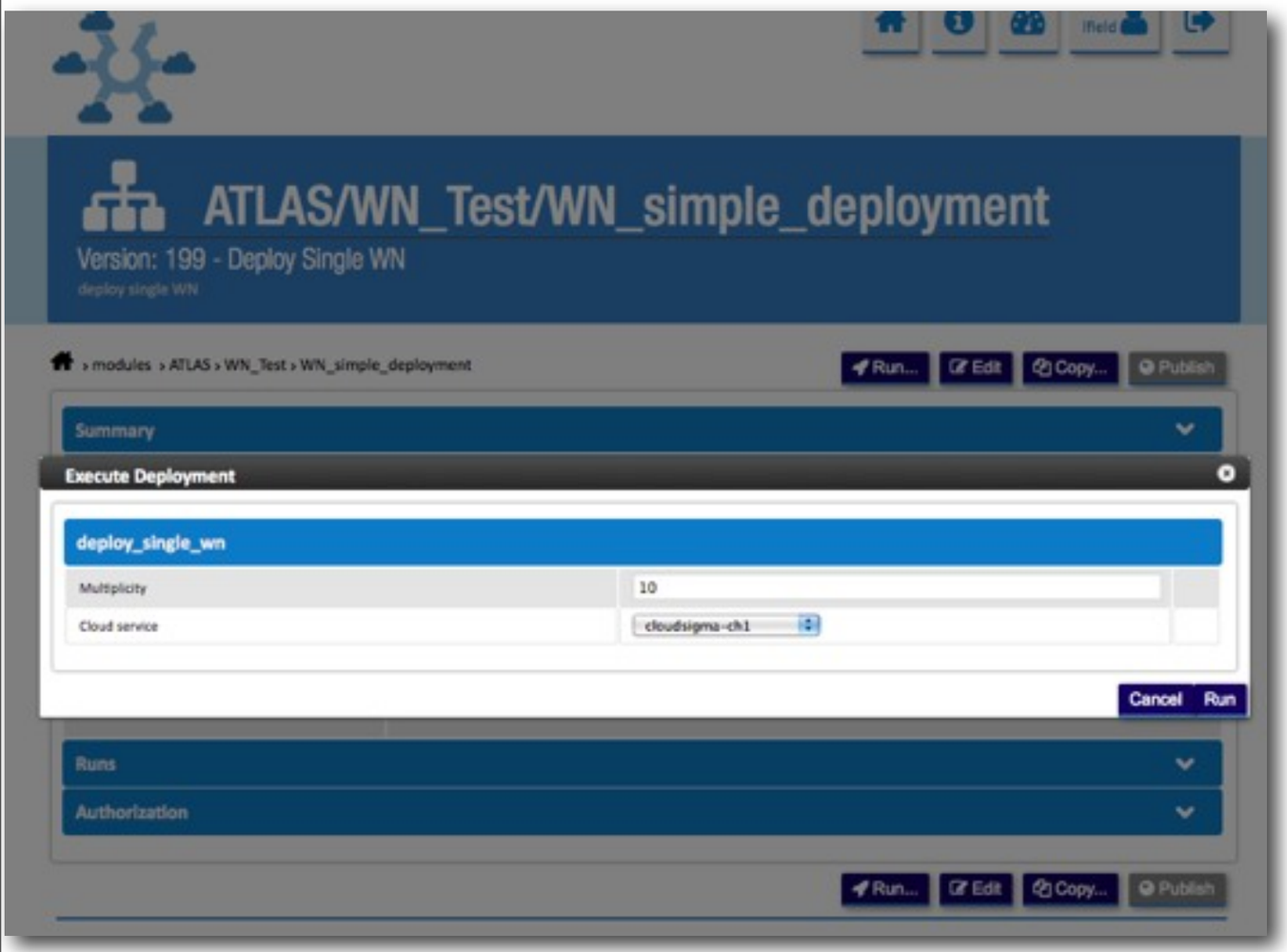
Site	Wall Duration	CPU Duration	CPU Count	Network inbound(Gb/sg)	Network outbound(Gb/sg)	Memory(GB)	Disk(GB)	Cloud Type
HELIX_NEBULA_ATOS	96.00	71.18	3.93	0.00	0.00	7.71	0	OpenNebula
HELIX_NEBULA_CloudSigma	0.00	0.00	0.00	0.00	0.00	0.00	0	OpenNebula
HELIX_NEBULA_TSystems	24.00	2.09	1.00	0.00	0.00	1.83	0	OpenNebula
Total:	120.00	73.27	4.93	0.00	0.00	9.54	0.00	

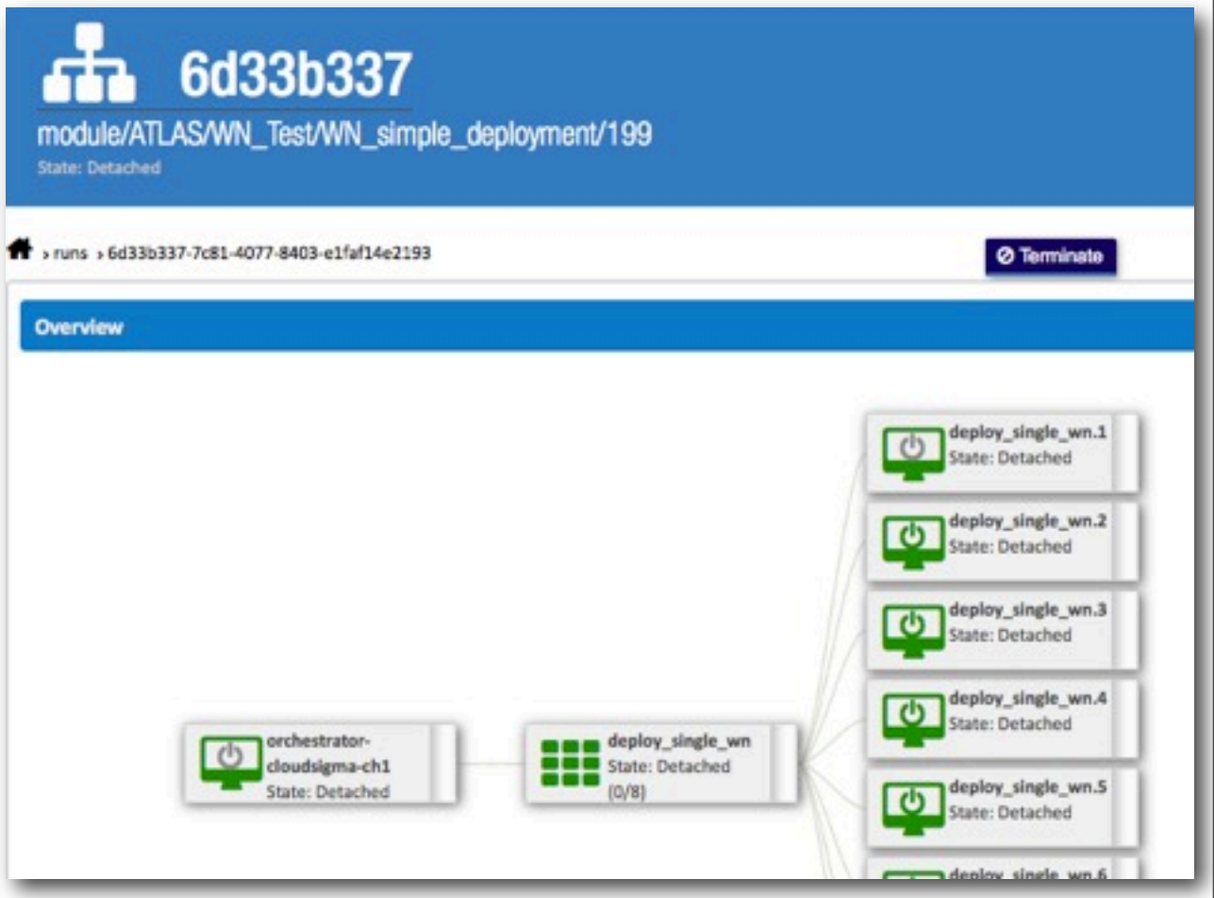
Showing 1 to 3 of 3 entries

Deployment

Multiple VMs started in a single deployment

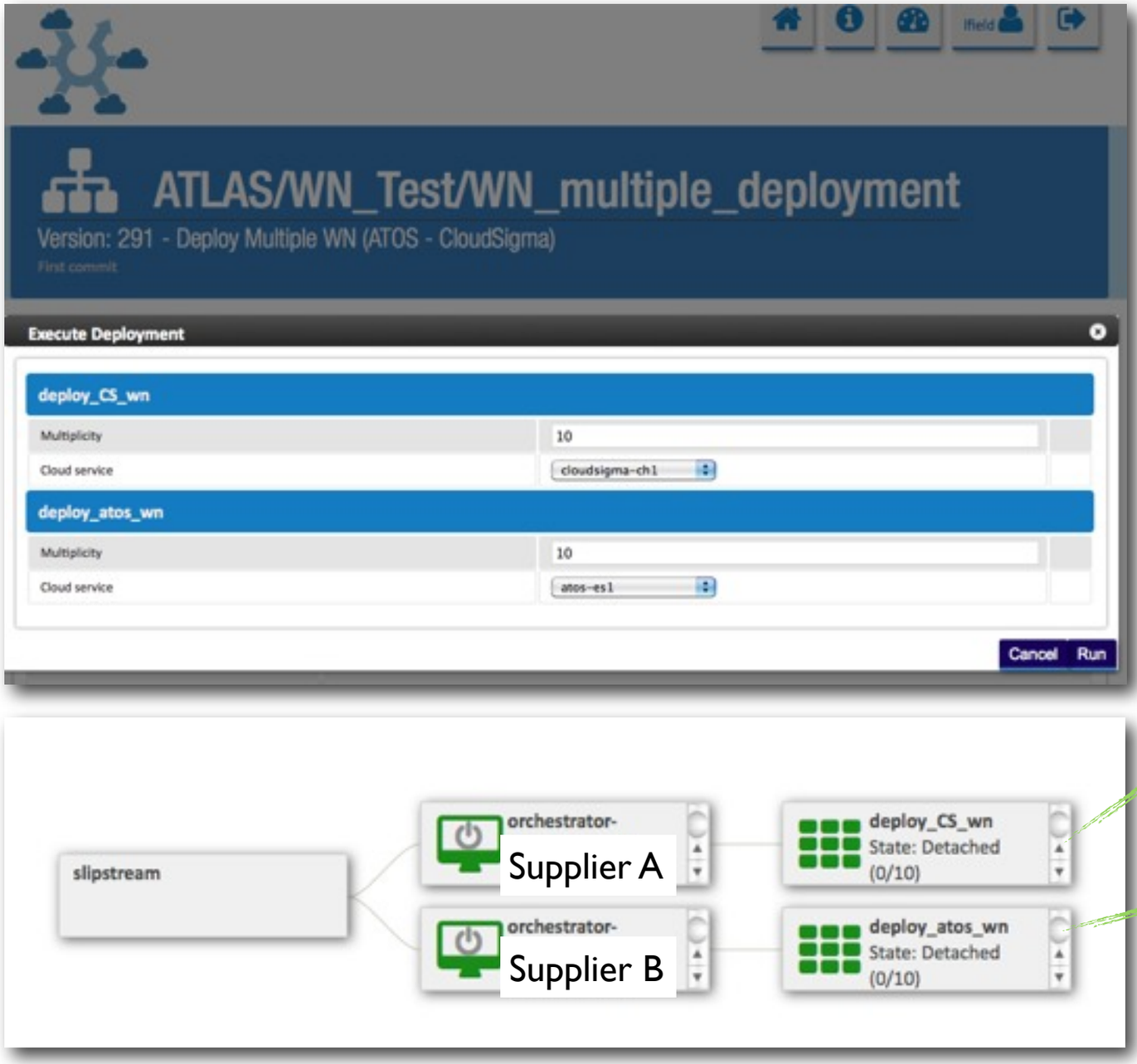
- Submitted up to 25 VMs per deployment

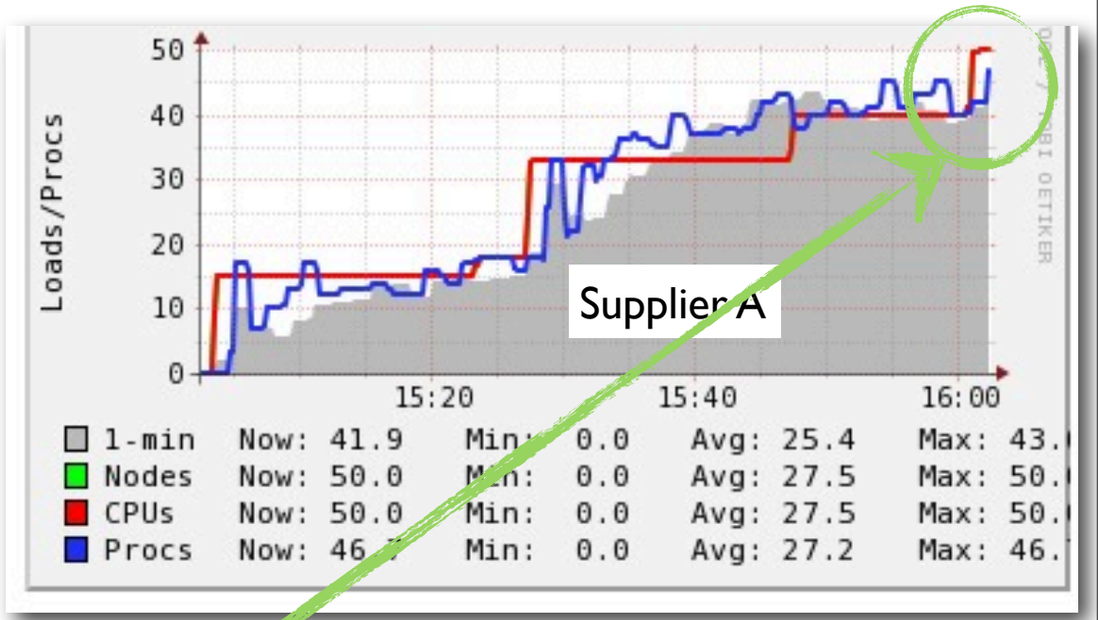


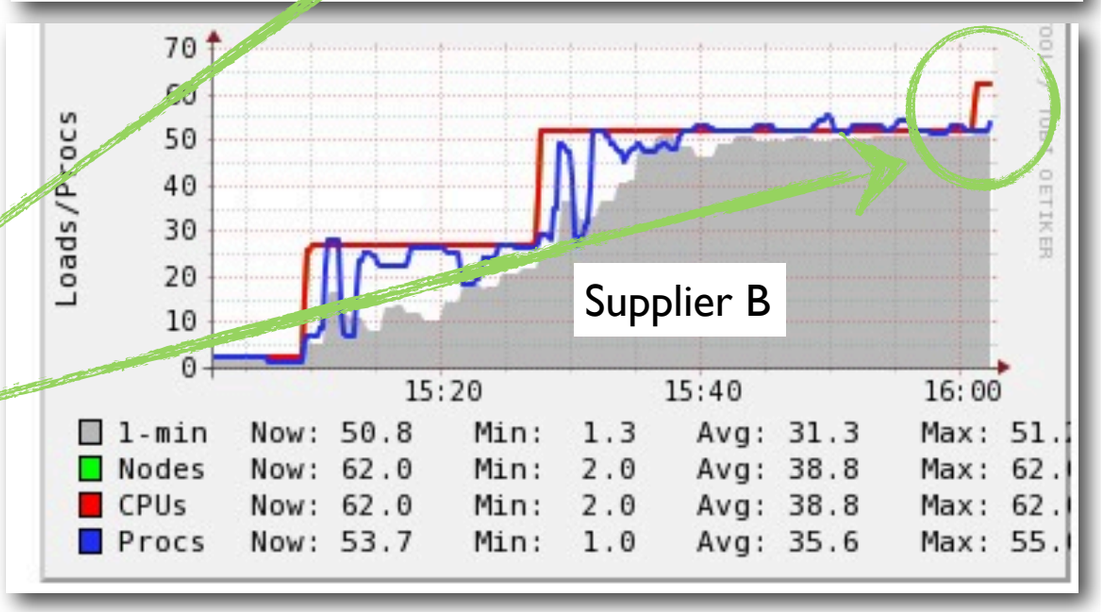


Deployment on multiple suppliers

Successfully deployed VMs in different suppliers through a single deployment

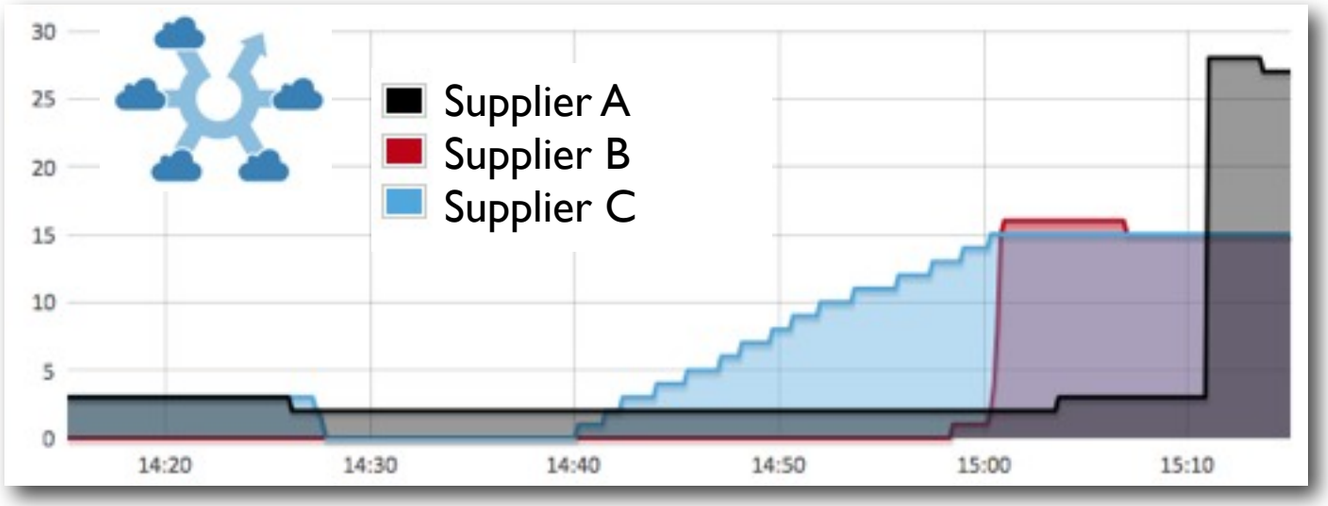


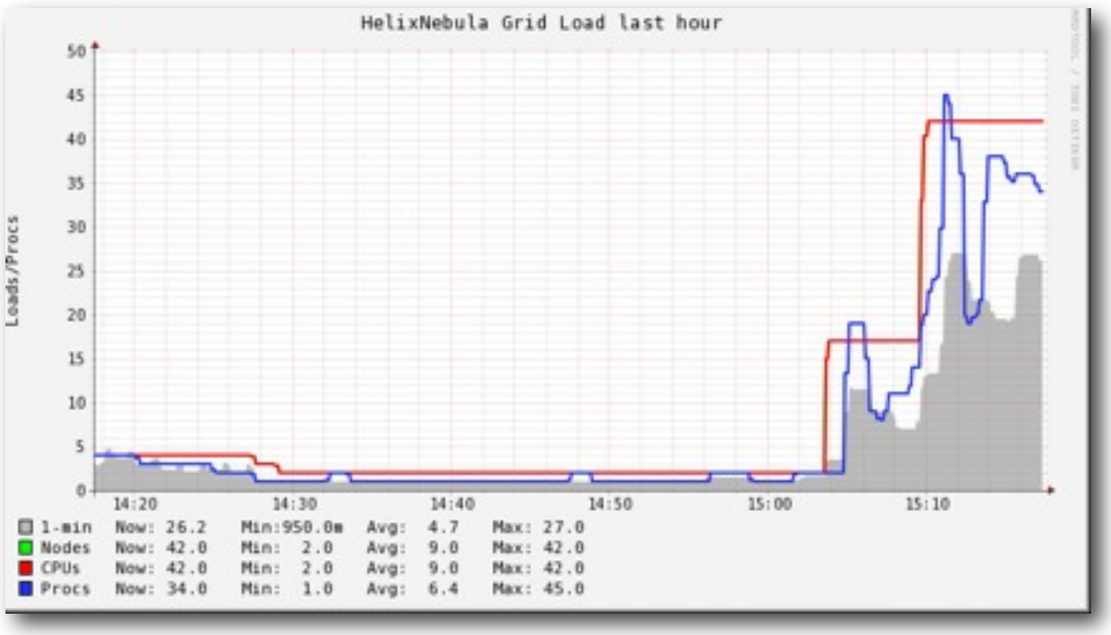




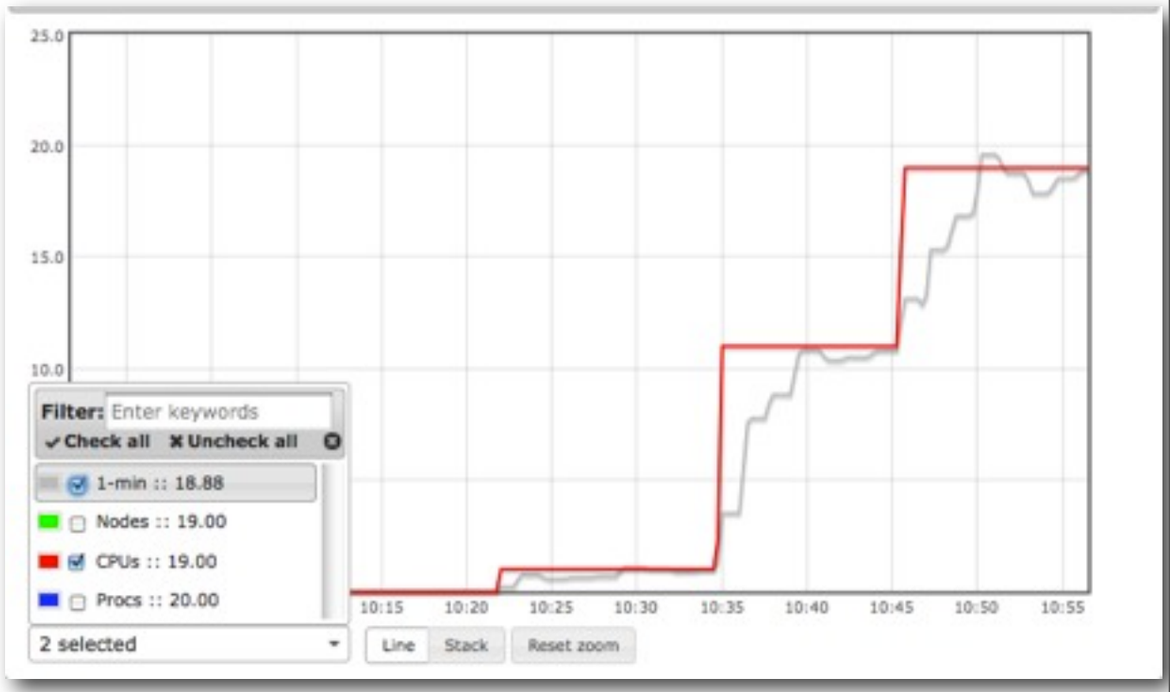
Ramping up

- Time to have VMs running from beginning of deployment depends on supplier
  - ~ 5′ - 25′

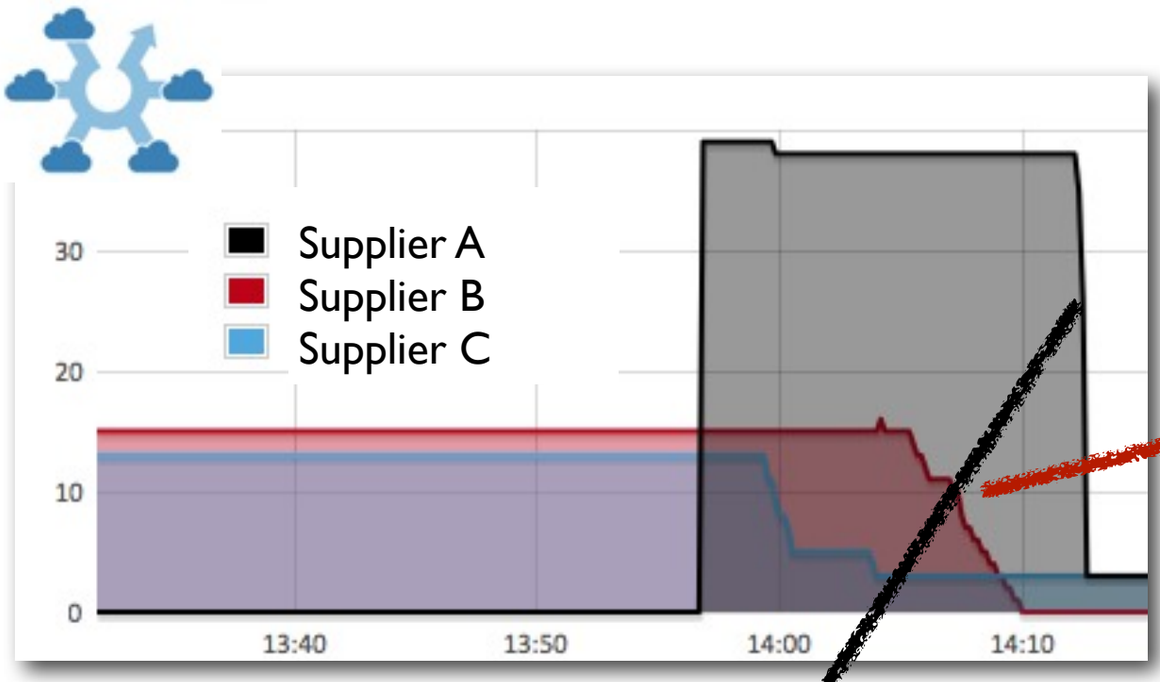


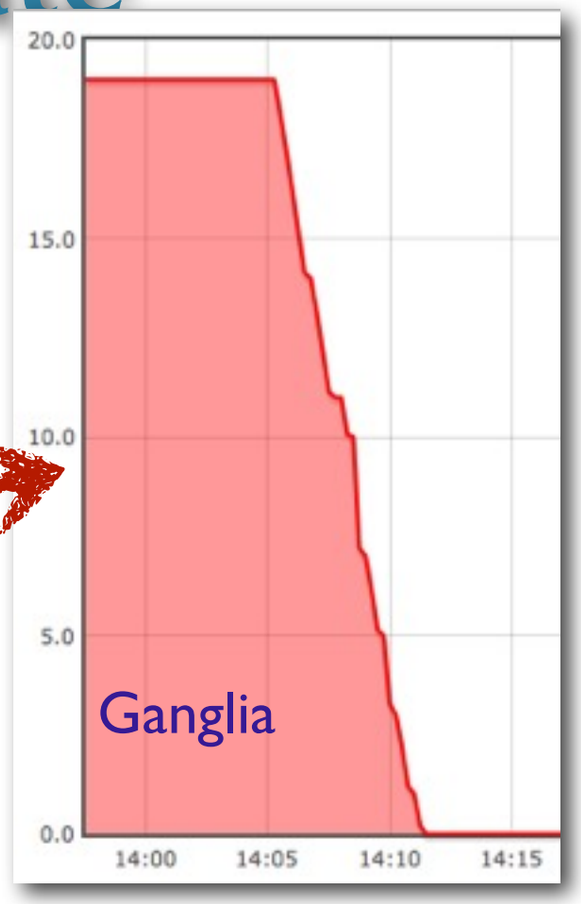


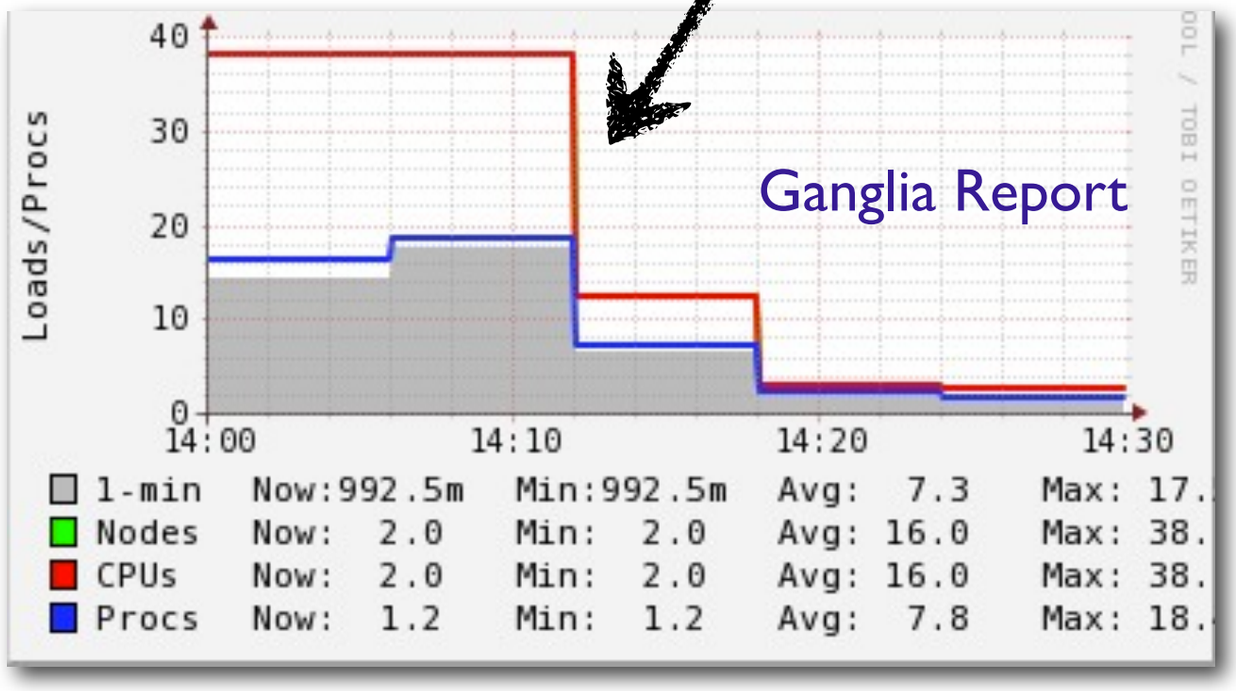
- Experiment jobs starts to run in O(5′) from the VM start



Terminate



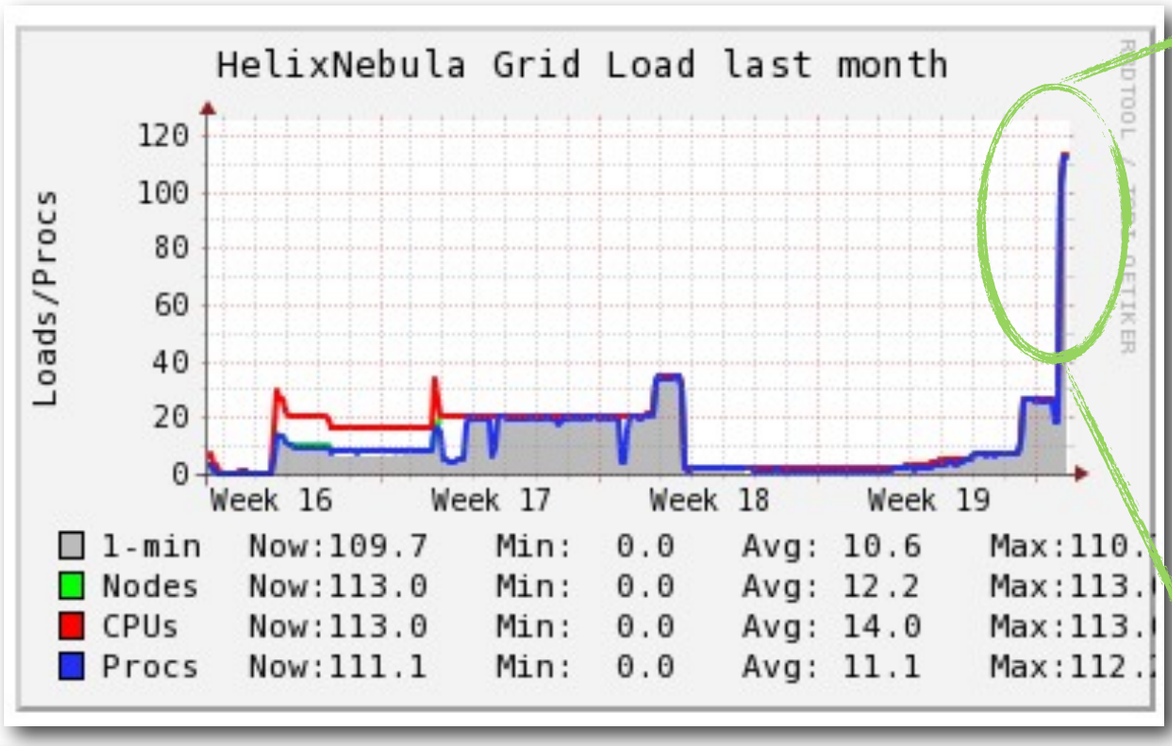


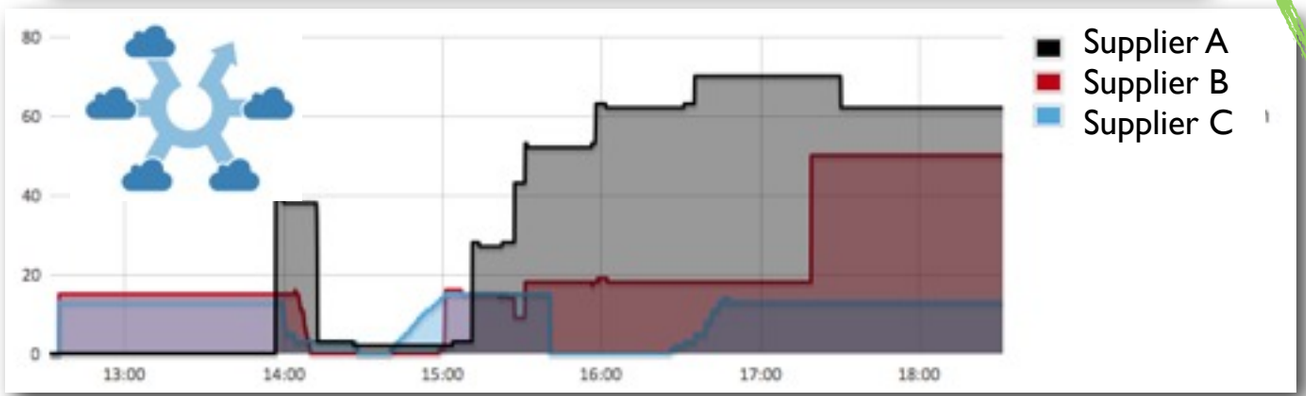


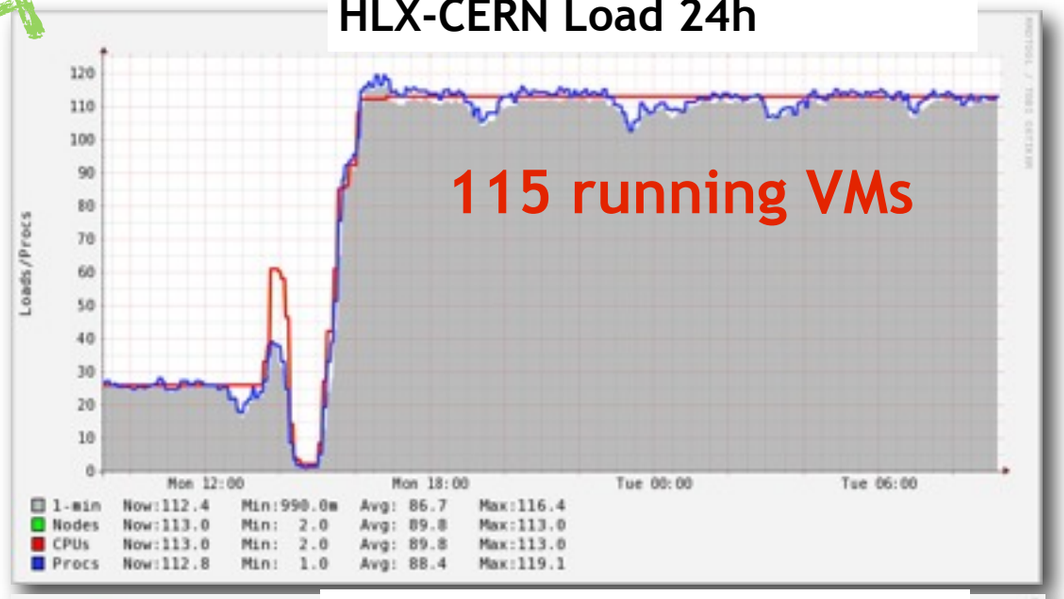
- Fast termination of machines in O(60″) from the “Terminate” command

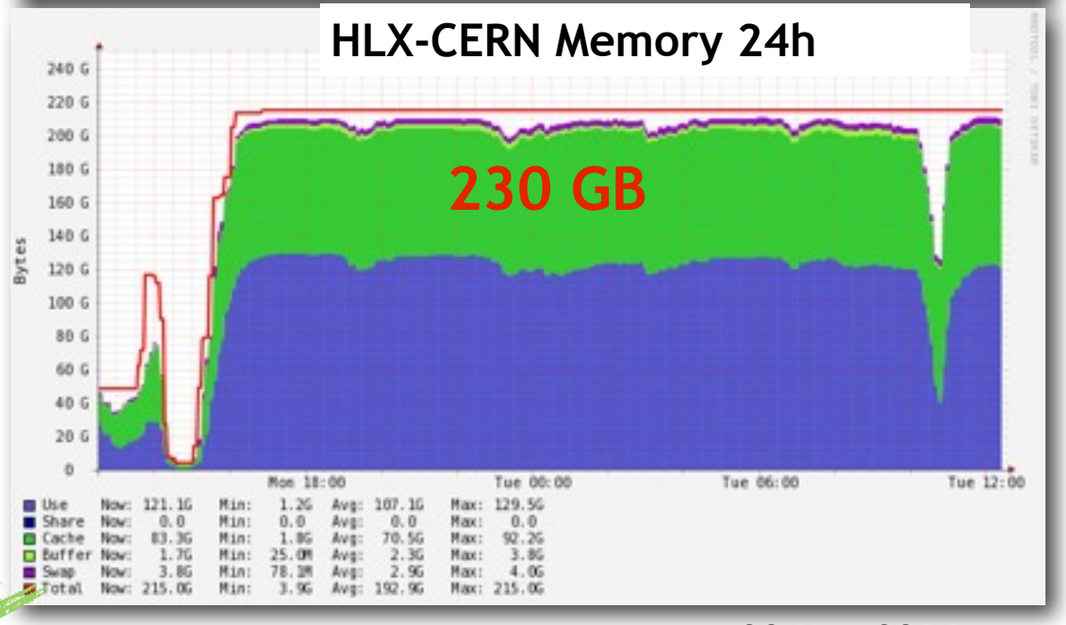
Scale tests

- Deployments show long time stability after startup
  - VMs left running for several weeks, running ATLAS functional tests
- Able to rapidly scale up to use available resources





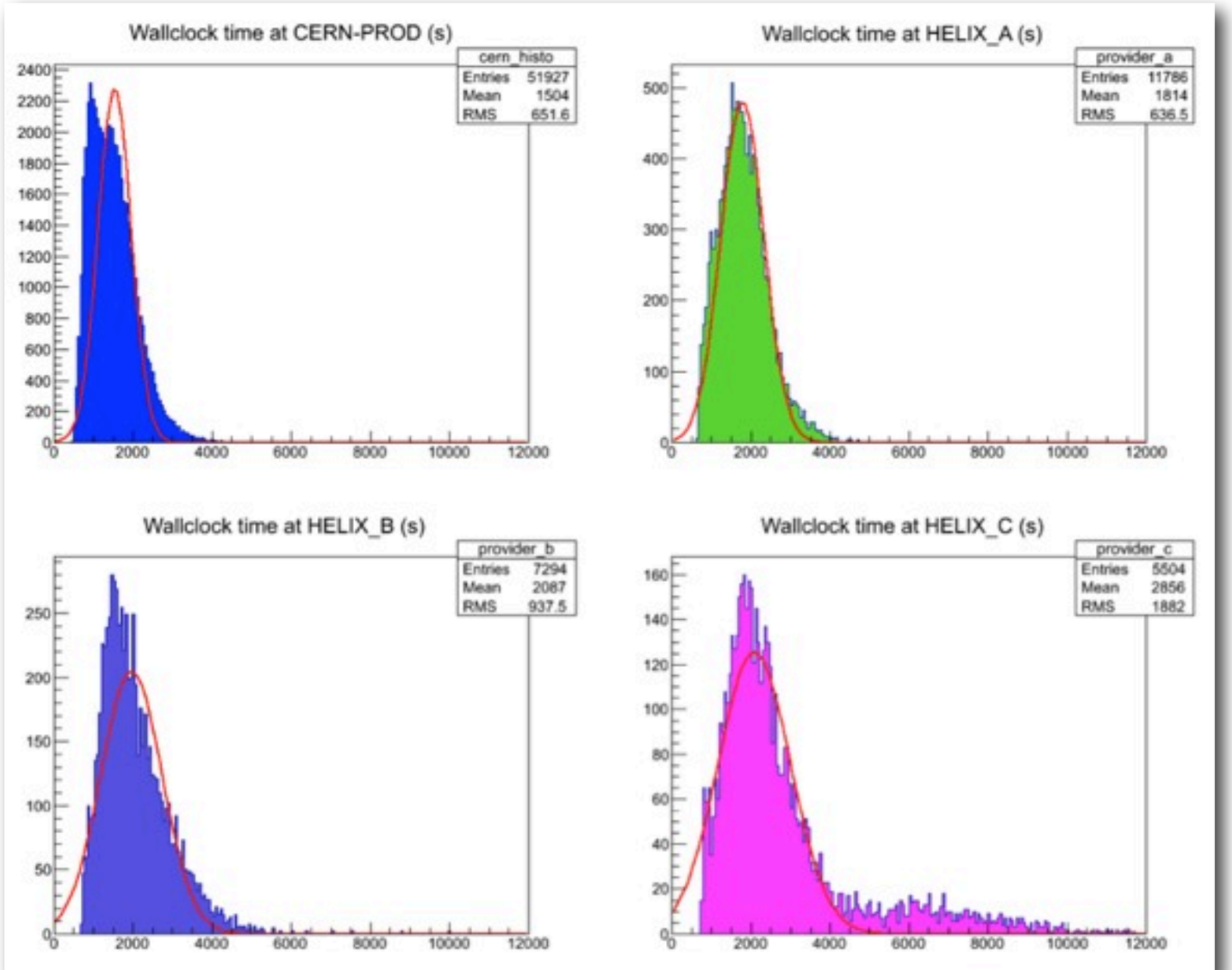




Past experience

Scale tests already performed in the past phases of the CERN flagship tests

- ~40k CPU days of processing during the pilot phase
  - Tests performed in 2013 connecting directly to each single provider (SlipStream BlueBox was still not in the picture)



Lessons learnt

Building and deployments still require efforts

- Sometimes base images not available / removed from the cloud provider
- Not possible to upload a user image (CernVM)

Service reliability needs to improve

- Aborted / Failed deployments could leave zombie VMs
- Monitoring and metering still limited

Capacity: be able to scale seamless.

- Limitation of 25-50 VMs for deployment
  - Future auto-scale feature looks interesting

Service Catalog and pricing reports are currently very basic

Conclusions

CERN flagship deployed the ATLAS experiment workflow on commercial clouds

- Successfully tested primary functionalities: start/stop/status and medium scale deployments
- Still heterogeneity seen among providers
- Costs still high/undefined to compete with in house resources

Helix Nebula Initiative over the last two years has allowed to

- Enable a federation of European public - private commercial cloud service providers
  - Creation of the Helix Nebula Market Place (HNX)