

Helix Nebula – The Science Cloud

Title: D5.1 Evaluation of Initial Flagship Deployments

Editor: Phil Evans

Work Package: 5

Submission Date: 17th June 2013

Distribution: Public

Nature: Report



Log Table

Issue	Date	Description	Author/Partner
V0.1	17 June 2013	Initial draft	Phil Evans (Logica)
V0.2	17 June 2013	Provided comments	Robert Jones (CERN)
V0.3	18 June 2013	Provided comments	Mick Symonds (Atos)
V0.4	18 June 2013	Comments received from Atos and CERN addressed	Phil Evans (Logica)
V0.5	19 June 2013	Comments provided by Interoute & T-Systems	Jonathan Graham (Interoute), Bernd Shirkpe (T-Systems)
V0.6	19 June 2013	Addressed comments around clarity of when environments were used	Phil Evans (Logica)

Contents

Executive Summary.....	5
1 Introduction.....	7
2 Flagship Overview.....	8
3 Deployment Profiles/Overview	9
3.1 CERN.....	9
3.2 ESA.....	9
3.3 EMBL	10
4 Cloud Provider Overview	13
5 Results.....	14
5.1 CERN.....	14
5.1.1 Time to First Instance	14
5.1.2 Scaling	14
5.1.3 Performance.....	14
5.1.4 Cost.....	15
5.1.5 Support	15
5.1.6 General.....	15
5.1.7 Conclusion	15
5.2 EMBL	15
5.2.1 Time to First Instance	15
5.2.2 Scaling	16
5.2.3 Performance.....	16
5.2.4 Cost.....	16
5.2.5 Support	16
5.2.6 General.....	17
5.2.7 Conclusion	17
5.3 ESA.....	17

5.3.1	Scaling	18
5.3.2	Performance.....	18
5.3.3	Cost.....	18
5.3.4	Support	18
5.3.5	General.....	18
5.3.6	Summary	18
6	Lessons Learnt.....	20
7	Recommendations.....	21

Executive Summary

During the first half of 2012, three demand-side flagship applications were ported and deployed to the cloud infrastructure platforms of the Helix Nebula Cloud Service Providers infrastructure as part of the first Helix Nebula Proof of Concept (PoC) phase. This paper presents an Initial Evaluation of those deployments.

Although many of the demand-side organisations had experience of deploying to cloud infrastructure in the past, the Proof of Concept phase gave the demand-side the opportunity to understand, amongst others aspects, the technologies, resource availability and complexities involved in deploying their flagship application to multiple cloud providers all of whom offer quite different platforms and setups.

Additionally, the Proof of Concept deployments gave the supply-side the opportunity to gain a better understanding of the requirements and wide ranging technical complexities of the flagship applications, each of which presented their own unique challenge for the suppliers.

The Proof of Concept deployments were a success with each flagship application being deployed to at least two cloud providers' infrastructure. The evaluations of the deployments presented in this paper demonstrate some clear themes that need to be addressed in Helix Nebula's roadmap going forward. In summary these are:

- The introduction of a common API to remove the difficulties encountered of each provider having a different API;
- To identify a way of removing the overhead of having to port the server images to the various cloud providers' platforms;
- Move to a scheme of transparent billing based on actual usage costs and away from the flat-rate fee;
- To ensure dedicated resources and an increased time-scale on future Proof of Concept deployments;
- That the cloud providers ensure that they are able to better handle large-scale ad-hoc resource requests whilst providing a stable platform that handles large scale long running jobs.

- To remove/simplify the difficulties experienced with “Private Cloud” solutions regarding the effort taken to gain network connectivity.

In many cases, the evaluations presented in this document do not refer to by name the cloud providers with which various issues were experienced. In addition, many of the Cloud Service Providers’ infrastructure will have evolved since the running of the Proof of Concept deployments in the first half of 2012. The aim of this report is not to provide a comparison of the cloud providers but to report the lessons learnt and to help formulate the roadmap to achieve a successful federated cloud platform.

1 Introduction

Following the Helix Nebula General Assembly in Heidelberg on the 26 – 27 October 2011, a roadmap was agreed and the Helix Nebula team started activities to implement the “Strategic Plan for a Scientific Cloud Computing infrastructure for Europe”.

Since November 2011, the Helix Nebula consortium, composed of Demand-Side, Supply-Side, Management Team and Technical-Architecture Team, has been working on a number of activities including the implementation of a Proof of Concept (PoC) stage for the three flagship projects accepted by the Supply-Side for the Helix Nebula two-year pilot phase.

The three flagship projects are:

- CERN: ATLAS data management and processing.
- EMBL: Large scale analysis of genomic sequence data
- ESA, CNES, DLR, & CNR: Processing of ERS and Envisat data to derive information of relevance for the earthquake and volcano analysis

This report provides an Initial Evaluation of the three Proof of Concept deployments, held in the first half of 2012, based on the Success Criteria defined as part of Work Package 5.

2 Flagship Overview

The three flagship applications selected for the Proof of Concept covered a varying number of Use Cases. Each flagship presented differing challenges for the Cloud Service providers during the migration of the application to the cloud infrastructure. The table below provides a high-level overview of these Use Cases against each of the flagship applications:

Table 1: Use Case Attributes of the Flagship Applications

	ATLAS H.E.P Cloud Use (CERN)	Genomic Assembly in the Cloud (EMBL)	SuperSites Exploitation Platform (ESA/CNES/DLR)
Scientific goal/Society Impact/Photogenic	✓	✓	✓
Scale of Resources Used	✓	✓	✓
Federation/Aggregation of Datasets		✓	✓
Long-Term Archiving of Data			✓
On-Demand Processing	✓	✓	✓
Impact on Community & Benefits	✓	✓	✓
Potential Increase of Users	✓	✓	✓
Interoperability	✓	✓	✓
Data Security	✓	✓	✓
Maturity	✓	✓	✓
Access to License-Controlled Software			✓

A full presentation detailing each of the Flagship applications was presented at the Second Helix Nebula General Assembly in January 2013. Videos for each of the presentations can be found at the following locations:

- CERN: <http://www.youtube.com/watch?v=Fd-VKk0nA8c>
- EMBL: http://www.youtube.com/watch?v=xVaQLlc_u1k
- ESA: <http://www.youtube.com/watch?v=CsyUYw1GQq8>

3 Deployment Profiles/Overview

This section provides a brief overview of the flagship application deployment profiles and resource requirements collated prior to the kick-off of the Proof of Concept deployments. These requirements represent the target deployments that the Demand-Side organisations aimed to achieve within the Proof of Concept phase.

3.1 CERN

The aim of the CERN PoC was to provision an environment of around 100 nodes that could be temporarily integrated with ATLAS to run production jobs. The CERN flagship uses the CERN grid management system which is based on Panda/Condor.

Requirement	Specification
Local Storage	0.5TB
Cores Per Node	12 Cores
Preferred Nodes per Core	4 Cores
RAM Per Node	2 GB Minimum, 3GB More efficient
Total Initial Node Requirement	100
Total Eventual Node Requirement	∞

3.2 ESA

The ESA Proof of Concept consisted of two applications that have been running for several years on the ESA GRID infrastructure. The CNR application (based on Envisat data only) has been available for a number of years on GPOD.

The enhanced version, specifically designed to take full advantage of Cloud Computing capacities, was deployed to Helix Nebula during the Proof of Concept phase.

Requirement	Specification
Local Storage	500 GB
Cores Per Node	1 to 12 Cores
Preferred Nodes per Core	4 Cores
RAM Per Node	4 to 24 GB
Total Initial Node Requirement	100-200

Total Eventual Node Requirement	1000
---------------------------------	------

3.3 EMBL

Initially the EMBL flagship deployment consisted of three major steps with increasing complexity (including one “optional” step). However this was simplified in order to allow a better focus on the additional software components required for the EMBL Proof of Concept. The Proof of Concept aims to test the three major software components of the flagship:

- Assembly pipeline
- Annotation pipeline
- Star Cluster

The EMBL Cloud Management System is based on StarCluster. EMBL used a stepwise approach to implement the test: The minimal tests involve:

- Step 0: Image deployment, setup of shared file system (based on GlusterFS) and initial test of EMBL software using small input data sets;
- Step 2: StarCluster functionality test and large-scale tests of EMBL software using a cloud HPC cluster (provisioned by StarCluster) to analyse large input data sets.

Step 0 Assembly Requirement	Specification
Local Storage	10 GB
Cores Per Node	8
Preferred Nodes per Core	8
RAM Per Node	32
Total Initial Node Requirement	1
Total Eventual Node Requirement	1

Step 0 Annotation Requirement	Specification
Local Storage	10
Cores Per Node	8

Preferred Nodes per Core	8
RAM Per Node	32
Total Initial Node Requirement	4
Total Eventual Node Requirement	4

Step 2 Assembly Requirement	Specification
Local Storage	10
Cores Per Node	8
Preferred Nodes per Core	8
RAM Per Node	32
Total Initial Node Requirement	49
Total Eventual Node Requirement	49

Step 2 Assembly Requirement	Specification
Local Storage	10
Cores Per Node	8
Preferred Nodes per Core	8
RAM Per Node	64
Total Initial Node Requirement	1
Total Eventual Node Requirement	1

Step 2 Assembly Requirement	Specification
Local Storage	10
Cores Per Node	8
Preferred Nodes per Core	8
RAM Per Node	64
Total Initial Node Requirement	49
Total Eventual Node Requirement	49

GlusterFS Requirements	Specification
Local Storage	600
Cores Per Node	8
Preferred Nodes per Core	8
RAM Per Node	32
Total Initial Node Requirement	4
Total Eventual Node Requirement	4

4 Cloud Provider Overview

Four Cloud Service Providers (CSPs) participated in the initial flagship deployments. The table below details which providers participated and an overview of the technologies that were, at the time of the initial flagship Proof of Concepts, deployed at the CSPs.

Table 2: Technologies Deployed by Each Service Provider

Provider	Cloud Stack	Availability	Hypervisor	API	Network
Atos ¹	StratusLab	Custom	KVM	OpenNebula	Specific
CloudSigma	Proprietary	General	KVM	Proprietary	Open
Interoute	Abiquo	General	KVM/ESX	N/A ²	Open
T-Systems	Zimory	Custom	VMWare	Zimory	Specific

It should be noted that the technologies now deployed by the Cloud Service Providers have evolved since the first half of 2012. The pilot in 2013 will use improved technologies.

Due to various constraints such as timing, not all flagship deployments were completed or attempted on all service providers. The table below presents which flagship deployments were deployed on which cloud infrastructures.

Table 3: Cloud Supply/Flagship Deployment Overview

Provider	CERN	EMBL	ESA
Atos	✓	✓	✓
CloudSigma	✓	✓	✓
Interoute	-	-	✓
T-Systems	✓	-	✓

¹ The technologies have evolved since the first half of 2012. The pilot in 2013 will use improved technologies.

² Subsequent to the 2012 Q1 pilots (and as used in the 2013 pilots) Interoute has released the jclouds API

5 Results

Following is an evaluation of the flagship deployments performed during the Proof of Concept phase. The evaluations have been performed at a high level against the Success Criteria that was defined in M11: Documented criteria and metrics. A full analysis against the criteria metrics was not possible due to limited data available from the PoC deployments.

5.1 CERN

The CERN Proof of Concept was supported by CloudSigma, T-Systems and ATOS. All deployments were successfully and eventually integrated into the ATLAS simulation production environment on a temporary basis.

5.1.1 Time to First Instance

Significant effort was required for configuration around the VPN requirements where the Cloud Service Provider offered a “Private cloud” based solution. This resulted in the installation of the Atlas application taking longer since the security set-up and configuration of the VPN required resources of all parties.

Porting of the image to the various Cloud Providers platforms also caused some delays to the deployment.

5.1.2 Scaling

During the Proof of Concept with CloudSigma, it was possible to Scale to around 100 nodes and run production jobs successfully. The deployment was also successfully integrated with the ATLAS production environment on a temporary basis.

At the time, some issues of how quickly new nodes could be cloned/provisioned were experienced. However, CloudSigma are considering tuning their system to better accommodate burst requests.

On some providers it was necessary to scale via the web-console rather than an API which does not seem best suited for responding to the burst-requirements required by the CERN flagship and does not lend itself well to ad-hoc large scale resource usage.

5.1.3 Performance

The Proof of Concept continued to run for a number of weeks. However, some failure rates were observed, the reason for which is under investigation. There are some hints that this

was a problem with proxy renewals in Condor. In total, around 90 to 100 concurrent jobs were run for several weeks.

5.1.4 Cost

For the Proof of Concept phase a flat fee was agreed with each of the providers.

5.1.5 Support

Support was provided on a per-supplier basis.

The overall installation with CloudSigma was straightforward and accompanied with a good working relationship ensuring short communication lines with people performing the work.

5.1.6 General

Suppliers had a different APIs which required some learning and adjustments for deployments which slowed down the initial activities. In some cases, it was found that even the definition of what a “Cloud” is differed.

5.1.7 Conclusion

Going forward, common interfaces are important/essential. The deployments performed during the PoC were not of the same scale compared to Amazon deployments that had been made previously.

The set-up and configuration of VMs with each of the suppliers’ systems was time consuming, limiting the amount of testing that could be performed. The experience so far has shown the importance of common interfaces across-suppliers, which still needs to be worked on.

5.2 EMBL

The EMBL PoC was supported by CloudSigma and ATOS/SixSq.

5.2.1 Time to First Instance

Providers which implement a “private cloud” solution posed some challenges to open access. It was found that “private cloud” infrastructures tend to be protected by a tight security setup which could pose a challenge at a later stage during the pilot when organisations/customers on the public network need to upload data and access it on the future Helix Nebula service.

Public providers were found not to be as difficult. In particular, with CloudSigma, it was possible to integrate StarCluster with their Cloud API very quickly allowing us to complete Step 0 in around 1.5 weeks. However with most providers it was found that the initial image deployment was time consuming due to the varying technologies deployed by the Cloud Service Providers.

5.2.2 Scaling

Both providers tested were able to achieve the deployment scale required. However there was a clear difference in performance compared to running the jobs in-house and on the cloud infrastructure.

5.2.3 Performance

Assembly & annotation applications were challenging for some of the CSP's infrastructure particularly around maintaining GlusterFS stability under high I/O levels. Long running jobs require a stable environment. A run of 20,000 annotation jobs per hour on 50 nodes was achieved on the suppliers tested.

Occasional connectivity issues on one of the CSPs' internal cloud network infrastructure were observed.

CloudSigma's strategy of cloning cloud instances has room for improvement, in particular when fast and on-demand provisioning of a larger number of instances is required. During the PoC this was addressed using a work around. CloudSigma's support to improve the performance of the EMBL code was very much appreciated.

Overall it was found that running the same application on the EMBL internal production system takes around 1/3 of the computing time than that experienced on the HN suppliers' infrastructure.

5.2.4 Cost

For the Proof of Concept phase a flat fee was agreed with each of the providers.

5.2.5 Support

Good to Excellent support was received from all Cloud Service Providers in assisting with the porting of the flagship applications. However, EMBL would recommend that future PoCs have dedicated resources and more time.

Support efforts seemed to suffer with one provider when more than one PoC was being handled at a time.

CloudSigma's staff were found to be exceptionally competent and helpful and managed to implement the StarCluster integration (including other additional supporting tools) very quickly during the PoC.

5.2.6 General

As the EMBL Flagship application relies on the EC2 API it was recommended by the Technical Architect Committee to implement a bridge that would handle the conversion from the StarCluster tool to the Cloud Provider's API. As each Cloud Provider has a different API it was necessary to implement the bridge for each provider.

On one provider's infrastructure that was OpenNebula based and already implemented part of the Amazon EC2 interface, an integration of StarCluster with OpenNebula/StratusLab was not any more straightforward.

5.2.7 Conclusion

On the whole, the cloud infrastructures at the time of testing did not offer the same scale or ease of use compared to the tests previously run on Amazon infrastructure. In addition, the performance of our in-house computing infrastructure is significantly higher than the one experienced throughout the PoC. Helix Nebula needs to provide the advantage of covering burst requirements.

The efforts required on both the supply and demand-side to run a PoC should not be underestimated and an extension of the duration for an individual PoC to, at a minimum, 2 months seems adequate. The earlier HN is able to implement the initial elements of a future federated cloud, the better for all parties involved in any future PoCs.

5.3 ESA

The ESA PoC aimed to learn whether HN cloud computing can serve ESA Earth Observation (EO) processing ICT needs and whether ESA could deploy an end-to-end platform for EO exploitation on HN in addition to exploring whether an ecosystem of value added service providers could be developed around such a platform. The ESA PoC was supported by Atos, CloudSigma, Interoute and T-Systems.

5.3.1 Scaling

It was not possible for all providers to meet the desired number of nodes, and in one case it was necessary to move the PoC to a different data centre in order to meet the burst processing requirements. This resulted in a couple of days delay whilst the migration was performed.

5.3.2 Performance

With one provider issues were encountered during the uploading/staging-in of data and programs. About 50% of the time processing could not be performed due to the connectivity issues. In addition, some service down time was experienced.

5.3.3 Cost

For the Proof of Concept phase a flat fee was agreed with each of the providers.

5.3.4 Support

Excellent support was experienced by all providers. In particular, the installation at Atos was straight forward and the team (including SixSq) were very active in resolving problems particularly around enabling automatic cloning of working nodes.

5.3.5 General

SixSq provided three interfaces to support provisioning in the cloud, which went beyond expectations, outlining an excellent service.

5.3.6 Summary

It is felt that the areas requiring improvement should be addressed during the following project phases. The large differences among providers, such as the differences in supported image types and APIs would need to be addressed.

In all cases the execution of the PoCs was feasible and the testing scripts provided by Logica allowed objective/comparable performance measurements. Generally, the PoC performances are suitable for the continuation of the SSEP flagship project but are not yet at the level expected for an operational service.

A further assessment needs to be completed considering the service terms and conditions. In addition, moving towards actual infrastructure costs must be considered, rather than the approach of using a flat rate as used during the PoCs.

A detailed analysis of the costs would provide more information on the benefits of Helix Nebula. From the technical view point, the HN offerings have the maturity to move the PoC activities to the next phase.

6 Lessons Learnt

The manpower resources on the demand- and supply-side as well as the time needed to successfully complete a PoC were underestimated. Two-man months of specialist people being dedicated to the PoC with a two-month duration with full access to the necessary Cloud assets would be more adequate to complete a PoC efficiently.

It was not the purpose of the PoC to validate the interoperability of HN cloud assets. However, given that every provider tested so far implements a different cloud interface, the need for a federated cloud in terms of homogeneous interfaces, accounting and billing must be addressed as soon as possible.

It would have been helpful to have up front a document/matrix outlining the demand-supplier side interfaces, on-demand elasticity handling, federated identity management and accounting of the supplier side. Furthermore to ease access to the HN offerings, a brokerage system would need to be considered.

The PoC has provided good information about the technical availability, capabilities and “fitness” of the cloud of mainly 3 to 4 HN Cloud providers; however more focus is needed on the capacity of responding quickly to burst- and on-demand requirements, which is considered a key advantage of using the cloud and common interfaces.

7 Recommendations

Following recommendations have been elaborated:

- 1 Considering that all of the current demand-side and in addition most of the future demand-side organisations own high performance IT infrastructure, the Tech Arch team should envisage an architecture that supports Hybrid Clouds. In particular for commercial or IP sensitive data owners would be reluctant to put all of their data in a public cloud.
- 2 The initial aspects of a federated environment such as a common API should be explored in further details before further tests are performed.
- 3 In order to make the HN an even more attractive working environment software packages that are commonly used in science and research should be deployed and used as SaaS or PaaS. Potential candidate software packages include:
 - Dropbox style service for file/data transfer between research labs
 - ROOT (Proof-on-demand <http://pod.gsi.de/http://root.cern.ch/drupal/content/proof>)
 - MS Office tools (MS 365)
 - Matlab
 - IDL
 - Oracle
 - Cloud based Condor service to be made available in a similar way as has been done by Cycle Computing (<http://blog.cyclecomputing.com/condor/>)
 - Collaboration tools such as Alfresco
 - Desktop and video conferencing tools (e.g. Adobe Connect)
 - Bibliographic tools such as Endnote or Reference Manager
- 4 Accommodating fast responses to burst requests needs to be emphasised.
- 5 Considering the effort on all sides to identify the “cloud capacity” of the suppliers, the demand-side recommends that new suppliers joining Helix Nebula as a Cloud provider

pass the 3 PoCs. For this installation procedure, -kits, and automated test scripts would need to be established validating the “Cloud capacity” and that the suppliers are conforming to security requirements defined by Helix Nebula. In the long run, a Helix Nebula “certification” would be needed to ensure a high level of “..aaS” performance and security compliance.

- 6 Accounting and invoicing based on used resources needs to be clear for the usage of HN assets. Best would be a standard system where the customers can easily see the expected cost and also have control of the cost during the system usage. Invoices must provide a clear cost breakdown (storage, computing, traffic, fees, etc.).
- 7 The demand side, providing unique data of a large scale to a global scientific community, encourages all HN partners (industry, SMES, demand side) to discuss and agree on a business model ensuring that the HN infrastructure becomes a long-term sustainable infrastructure.