Helix Nebula – The Science Cloud

Helix Nebula Architecture

Abstract: This document captures the current knowledge of the Helix Nebula Technology and Architecture Group regarding the need of a federated framework to simplify discovery, access, usage and management of a federated cloud system. Alongside this objective, we aim at providing an integration framework, where current and future suppliers (i.e. cloud service providers) can easily interface their system in order to attract and receive cloud workload.

Document Information Summary	
Deliverable title:	*Helix Nebula Architecture*
Editor:	*Marc-Elian Bégin*
Contributing Authors:	*Marc-Elian Bégin (Sections 1, 3, 6)* *Paul Parsons (Section 3,4)* *Phil Evans (Section 5)* *Hervé Caumont (Sections 2)* *Micheal Higgins (Section 4)* *Gunnar Billing (Overall structure)* *Stephan Riebach (Sections 3 and 4)* *Note: all authors have had active and numerous inputs into the resulting architecture.*
Distribution:	*Public*
Version/Revision:	*1.5*
Keywords:	*Architecture*

Disclaimer

Helix Nebula targets scientific communities, decision makers from SMEs, industry and government as well as European Union policy makers. This document contains information on Helix Nebula core activities, as prepared by the Helix Nebula Technology and Architecture Group. Any reference to content in this document should clearly indicate the authors, source, organisation and publication date.

This document is work in progress. The authors reserve the right to significantly change the underlying assumptions, system and specification described in this document.

Any redistribution or reproduction of part or all of the contents of this document in any form is prohibited without explicit permission from the Helix Nebula Consortium and its partners.

Log Table

Issue	Date	Description	Author/Partner
V1.0	02/07/2012	Final Version	Tech Arch
V1.2	11/09/2012	Public Version	Tech Arch
V1.3	13/09/2012	Addition of T-Systems environment	Tech Arch
V1.4	14/09/2012	Addition of CloudSigma environment	Tech Arch
V1.5	14/09/2012	Minor changes	Tech Arch

Table of Contents

1. Introduction

This document captures the current knowledge of the Helix Nebula Technology and Architecture Group regarding the need of a federated framework to simplify discovery, access, usage and management of a federated cloud system. Alongside this objective, we aim at providing an integration framework, where current and future suppliers (i.e. cloud service providers) can easily interface their system in order to attract and receive cloud workload.

This document is work in progress and only reflects the views and vision of its authors. The system proposed in this document builds on the knowledge of the participants, but also integrates the significant lessons learned from executing the Helix Nebula Proof of Concept performed to date.

Still significant work is left to do, as well as convergence with the demand-side, such that the proposed system and the roadmap (including the priorities put forward) represents the wider possible views and reaches large agreement.

The document is structured as follows: Section 2 Cloud Vision introduces our common vision, understanding and scope of the cloud in the Helix Nebula context. From this common understanding, Section 3 Requirements presents the result of our requirements analysis and presents it in the form of user stories. Section 4 Helix Nebula Service Enabling Framework provides the foundation of the Helix Nebula system. In order to turn this ambitious system into reality, Section 5 Roadmap proposes a roadmap where incremental releases of the Helix Nebula system will be created, with added features for each version. The roadmap also maps several functional elements to user stories (see Section 3) for easier cross-referencing. Section 6 concludes the document with a short summary.

2. Cloud Vision

This section described the cloud vision Helix Nebula shares and promotes.

2.1. Shared goals

The Helix Nebula Strategic Plan for a Scientific Cloud Computing infrastructure for Europe defines the vision of an industrial strategy for a federated cloud framework to be implemented by 2020. The main Helix Nebula goals to reach this vision are as follows:

- Establish a Cloud Computing Infrastructure for the European Research Area and the Space Agencies, serving as a platform for innovation and evolution of the overall federated cloud framework.
- Identify and adopt suitable policies for trust, security and privacy on a European-level.
- Create a lightweight governance structure that involves all the stakeholders and can evolve over time as the infrastructure, services and user-base grows.
- Define a funding scheme involving all the stakeholder groups (service suppliers, users, EC and national funding agencies) into a Public-Private-Partnership model, that delivers a sustainable and profitable business environment adhering to European-level policies.

A set of functional requirements and technical properties are defined in order to meet the defined goals, and bring the Helix Nebula vision into reality.

2.2. The Science Cloud properties

The main properties Helix Nebula aims at exhibiting are:

- Multi-demand
- Multi-supply
- Easy selection between providers
- Faster response to stakeholders' needs

These properties define overall a 'trusted' domain with transparency for invoicing and other accounting over cloud resources provisioning, so consumers using the European Cloud Computing Infrastructure are not entering a vendor "lock-in" situation. The overall resource scalability and offering of diversity and quality of service are aimed at ensuring consumers a faster and wider than ever interdisciplinary environment to support science compared to any single supplier.

2.3. Services to be provided

Helix Nebula promotes a model for enabling on-demand network access to a shared pool of configurable computing resources:

- high-bandwidth network connectivity,
- large sets of commodity servers,
- federated storage with secure data layers,
- easy to deploy cloud software appliances for user applications,
- and infrastructure management and monitoring services.

Scientists and developers, especially SMEs looking for an "innovation engine" for the mass-market, need an operational e-infrastructure, supporting access to multidisciplinary data on a very large scale, as well as providing highly elastic resources.

3. Requirements

This section attempts at assembling the bulk of the requirements guiding the definition of the Helix Nebula Architecture. We start by identifying the actors involved in the system, followed by user stories, in order to identify the usage patterns and features the system must deliver to the users.

3.1. Actors

Several actors, both human and software systems are involved in the Helix Nebula system. The following is a list of the main actors. These actors are then used in the rest of this document when needing to identify who or what is performing a given action, or is the intended recipient of a response.

1. Supply-side - Providers (aka system administrator in the following user stories)

- provides IAAS
- bills for resources
- accepts federated access
- provides information to the catalogue of services with capabilities
- supports a unified/common API interface (e.g. StratusLab, OpenNebula, Open Stack, etc.)

2. Demand-side - Consumers (aka user in the following user stories)

- discovers capabilities via the catalogue of services
- provisions resources and initiates jobs/work
- monitors and controls its resources
- pays bills

3. Broker - Middle men (aka system under construction in the following user stories)

- provides the blue box system, including the portal (see below)
- hosts a catalogue of services
- enables "the cube" - the ability to select providers based on cost, time, capability, etc.
- unifies billing

4. SME Supplier - Independent Service Providers (aka SME in the following user stories)

- provides value added images
- provides value added datasets

- provides value added system provisioning capabilities

5. SME Consumer - Independent Service Providers (see note below regarding demand-side amalgamation)

- take advantage of 'pay-as-you-go' and 'no-upfront' infrastructure investment
- consume unique and available capabilities (IaaS, PaaS and eventually SaaS) in the Helix Nebula federated cloud

Note: the two SME actors require further thoughts, especially the second one. In the following user stories, the SME Consumer is amalgamated with the Demand-side Consumers, but we suspect that they will have distinctive usage patterns and thought relevant to distinguish them at this early stage.

3.2. Functional Requirements - User Stories

In this section, we capture functional requirements for the Helix Nebula architecture. In order to create a richer dialog between the supply and demand sides, these functional requirements are expressed as user stories in the form:
As a <actor>, I can <action>, such that (or in order to) <benefit>.

3.2.1. Simple Provisioning

1. As a HN consumer, I can provision a single VM, in a single API call, specifying the reference image to start from as well as resources to be allocated (e.g. CPU, RAM), such that I can access the created VM in the cloud of my choice
2. As a HN consumer, I can provision N VMs, in a single API call, specifying the same as the user story #1, such that I can access the N created VMs in the cloud of my choice
3. As a HN consumer, I can create, in a single API call, a new empty disk/volume, such that I can use it to build a new machine or as an additional disk/volume to a VM
4. As a HN consumer, I can provision as for user story #1, with extra disks/volumes, such that I can keep my reference VM small and change available disk space using extra disks/volumes. The extra disks can be: persisted or volatile

Note: All cloud resources identified above (i.e. image, VM/instance, volume) can be managed using CRUD semantic.

3.2.2. Contextualisation

5. As a HN consumer, I can specify when provisioning machines, a public SSH key, such that it is available from within the machine during boot for a script to populate the .ssh/authorized_keys file. This means that I do not require to know username/password of the images I provision in order to access them in a secure way.

6. As a HN consumer, I can specify when provisioning machines, user data (in string/text format), such that it is available from within the machine during boot for a script to act on this information.
7. As a HN consumer, I can provision middleware and applications on VMs, using pre-configured templates

3.2.3. Image Management

8. As a HN consumer, I can list public images, or images belonging to me, such that I can discover what images are available and select the right image to instantiate
9. As a HN consumer, I can delete images belonging to me, such that I can keep my list of images clean and tidy
10. As a HN consumer, I can convert an existing image to the cloud specific format and contextualisation, following clear instructions and available tools, such that I can do this on my own, without requiring specialist support from the cloud provider
11. As a HN consumer, I can upload a new image to the cloud provider, such that I can instantiate this image following user story #1 and #2
12. As a HN consumer, I can save an existing VM as a reference image, such that I can instantiate this image following user story #1 and #2
13. As a HN consumer, I can transfer/convert an existing image from two cloud providers, using API calls, such that I do not have to recreate the image every time I change cloud provider

Note: All cloud resources identified above (i.e. image, VM/instance, volume) can be managed using CRUD semantic.
System Provisioning

"System provisioning" in this context means ability to provision a set of VMs, including the configuration and orchestration of the VMs, such that these VMs, together, can deliver a system that can be used for scientific HPC processing. Examples of system provisioning already identified during the Helix Nebula Proof of Concept phase (PoC) are:

- GlusterFS: EMBL is using a four-node GlusterFS, requiring fast disk access, provisioned together to deliver 7-8 Mb/s
- G-POD: ESA is using a Computing Element (CE) / Working Node (WNs) deployment, along with the User Portal for selection of Data and Processors, on-demand processing, and job monitoring, where specific data handling applications can be seamlessly plugged into the system, and where the number of WN varies depending on the usage scenario.

14. As a HN consumer, I can provision a system (see definition above), in a single API call, such that the system is configured and ready to deliver the intended value. When required, the system provisioning can include parameters that the user must provide with the API call (e.g. external service endpoint, source of software, license information)

3.2.4. Data Transfer and Placement

15. As a HN consumer, I can transfer a large amount of data (>GB?) in and out of a given cloud, such that I can push data into the cloud of my choice, or pull data out of a given cloud.
16. As a HN consumer, I can provision computing resources close to the data I am interested in processing, such that I do not have to transfer large amount of data to process it.
17. As a HN system administrator, I can easily replicate popular datasets in my cloud, such that users do not incur large data transfer costs to crunch the data on my cloud.

Constraint: a dataset (same applies for images) can be tagged as 'seed', such that these are not removed by cleanup procedures. This is in contrast with volatile replicas of datasets, to attract sporadically workload on popular datasets (e.g. following disasters or significant discovery).

Note: can connection to GEANT be expected from each HN supplier, such that it can implicitly be used when transferring data between providers?

3.2.5. Image Marketplace

18. As a HN consumer, I can access a catalogue of private and public images, such that I can easily search and discover them.
19. As a HN consumer, I can refer to an image via a unique identifier, such that I can more easily know if this image is present in a given cloud or not. This means that images are annotated and/or tagged with unique identifier, and maintained across copy and replication processes.
20. As a HN consumer, I can apply filters to my searches, such the following:
    - Architecture: e.g. 32 bit, 64 bit
    - O/S: e.g. Linux(all flavours), Windows, Solaris, AIX
    - Disk layout
    - Purpose, keywords
    - Installed software: e.g. package list
    - Cost
    - Capabilities and limitations

    in order to ease the selection process

21. As a HN system administrator, I can access basic analytics, such that I can see what images are popular and why
22. As a SME (for example), I can register and advertise in the marketplace images pre-installed and configured with added value software, such that I can generate revenue from these images. This implies an 'AWS devpay'-like feature, such that cloud providers can invoice users extra charges that get paid to the image owner.

Note: both image and data marketplaces offer standard search capabilities

3.2.6. Data Marketplace (DataMart)

23. As a HN consumer, I can access a datamart (i.e. dataset marketplace), such that I can discover high-quality scientific data, which I can process to generate high value assets
24. As a HN consumer, I can upload high-quality datasets to the datamart (i.e. dataset marketplace), such that I can share the datasets with the wider scientific community
25. As a HN consumer, I can protect the datasets I share on the HN federated cloud, such that only a restricted set of users (aka Virtual Organisations) have access to the data
26. As a HN consumer, I can refer to a dataset via a unique identifier, such that I can more easily know if this dataset is present in a given cloud or not. This means that datasets are annotated and/or tagged with unique identifier, and maintained across copy and replication processes.
27. As a SME (for example), I can register and advertise in the marketplace added value datasets, such that I can generate revenue from these. This implies an 'AWS devpay'-like feature, such that cloud providers can invoice users extra charges that get paid to the dataset owner.
28. As a HN system administrator, I can access basic analytics, such that I can see what datasets are popular and why

3.2.7. Security

29. As a HN consumer, I can set-up firewall rules for each instantiated VM, such that I can control access of my cloud resources
30. As a HN consumer, I can set up firewall rules between different Cloud Providers, in order to establish transparent networks.
31. As a HN consumer, I can manage my instance keys, such that each have pre-assigned firewall rules, which I can specify when provisioning instances, or modify while the instances are running
32. As a HN consumer, I have a mechanism to rotate login credentials and instance keys, such that different instances are started with different credentials and keys, thus improving protection and security of my VMs
33. As a HN consumer, I can encrypt data in movement as well as persistent data.
34. As a HN consumer, I can use the same login credentials to access all the HN services and Cloud Providers using SSO

3.2.8. Networking

35. As a HN consumer, I can assume that all cloud providers and connected to GEANT, such that I have high-quality and predictable network performance between resources running in different HN cloud providers

3.2.9. Cross Cloud / Federated Provisioning

36. As a HN consumer, I can provision cloud resources on several clouds at the same time, such that I can take advantage of the several clouds (e.g. geographical distribution, specific cloud provider feature and/or capabilities, data proximity)

3.2.10. Hybrid Cloud Deployment

37. As a demand-side system administrator, I can install a Helix Nebula connector, such that my cloud provisioning solution can provision resources on the public HN cloud, alongside local cloud resources
38. As a demand-side system administrator, I can provision HN cloud resources able to connect to existing resources in my private cloud, such that a secure access is granted between the public HN resources and my private resources.

3.2.11. Service Discovery

39. As a HN consumer, I can access the HN service catalogue (e.g. service discovery service), such that I can discover what cloud provider is available, and what capabilities they each offer and support
40. As a HN system administrator, I can populate/interconnect with the HN service catalogue, such that I can advertise my cloud's capabilities, in order to attract users onto my cloud

3.3. IaaS Enabling Frameworks involved in PoC

In this section we provide a high-level description of the different proof of concept environments, followed by an assessment from each supplier, of the potential contribution the solution can have on the identified Helix Nebula Architecture.

3.3.1. Atos

PoC Architecture

The Atos cloud solution is built as a proof of concept environment with dedicated hardware resources, that are provisioned and deployed exclusively for the purpose of this PoC. The deployed hardware capacity included the following:

- 28 dedicated enterprise class servers with 8 cores each and 64 GB of RAM
- 8 TB of dedicated storage
- Dedicated secure network to isolate from the rest of the data center traffic
- Fast internet connection
- 300 Mb/sec dedicated local connectivity

The software solution is based on StratusLab and SlipStream and is deployed in collaboration with SixSq. The StratusLab distribution, an open source solution distributed under the Apache 2.0 license, integrates the OpenNebula 3.2 virtual machine manager toolkit. StratusLab augments OpenNebula with the following services:

- **Federated Marketplace Service**: this service, developed in collaboration with HEPiX, CERN and EGI, provides a catalogue of metadata related to images and datasets. Built as a RESTful web service, it "talks" HTML, XML and JSON. Users can interact with the service using a browser or the command-line client. Each entry on the marketplace is an XML document signed with an X.509 certificate, ensuring the integrity of the entry (e.g. it has not being tampered with since its creation). The other important trust element provided by this service is that the metadata contains hashes (e.g. MD5, SHA) of the resources it represents (e.g. image, dataset). Location information is also present in the metadata, such that a cloud engine can retrieve (and cache) the raw image or dataset over the network.
  Finally, the Marketplace supports the concept of endorsement, where users can 'endorse' existing entries, by signing with their X.509 certificate an existing metadata entry. The Policy Enforcement Engine can then apply white and black lists filtering logic, resulting in only allowing images endorsed by known authors or trusted individuals.
- **Command-line client**: the StratusLab distribution supports copy-on-write VM instantiation (an important addition to the standard OpenNebula capabilities) such that several VMs can be started from a single reference image, in a single API call.
- **Persistent Disk Storage Service**: this service provides persistence functionality to disks attached to VMs, but also offers a caching service, with an upload/download image feature. This means that images can be automatically transferred between StratusLab sites, and cached locally. Integrated with the OpenNebula scheduler in the form of extension drivers, the service provides thin provisioning, including copy-on-write provisioning, resulting in fast start-up times based on reference images.
- **Policy Enforcement Engine**: this policy engine binds together the metadata provided by the Marketplace, the caching engine of the Persistent Disk Storage Service and the filtering policy site administrators configure. This mechanism also supports the ability to block instantiation of revoked or expired images, ensuring that only valid images can be instantiated by users.
- **Pull Contextualisation**: StratusLab supports contextualisation, similar to AWS EC2's, such that users can pass VM specific information (e.g. public SSH keys, key/value pairs, scripts), that can be retrieved by applications running inside the VM. Typically, a script executed during the boot sequence executes this script. This process is now compatible with the CernVM contextualisation mechanism, such that all CernVM are compatible with StratusLab, as was demonstrated during the CERN PoC.

StratusLab ships by default with KVM support, which is the preferred hypervisor. The distribution is an open source community project, including core contributors such as CNRS, GRNet, Trinity College Dublin, SixSq, UCM and Telefonica.

For the Atos PoC environment, StratusLab was deployed in a dedicated management VM. Storage was provided via an iSCSI TGTD software backend. The exception was for the GlusterFS system for EMBL, where 5 VMs were equipped each with a 500 GB disk directly provided by a NetApp backend. This allowed Atos to deliver the required 7-8 Gb/sec throughput.

The Atos cloud environment also includes SlipStream, a system provisioning service and image factory engine. SlipStream was used, for example, to provision a GRID cluster for ESA G-POD processing, as part of the ESA PoC. More details can be found on SlipStream in Section 4.3.

Relevance to Helix Nebula Architecture

In the context of the Helix Nebula proposed system, a number of technologies and services deployed by Atos are relevant and would form a reasonable starting point. Since all of these technologies are open source, it makes shared development simpler.

The following list of existing services and components could be for a reasonable starting point to fulfil identified Helix Nebula functions, as described in Chapters 4 and 6.

StratusLab Marketplace: the StratusLab Marketplace could be used as the foundation for the Image and Dataset Marketplaces. It is a simple RESTful web service, already supports federation, and is cloud implementation and hypervisor agnostic.

StratusLab Contextualisation: already supported by CernVM and the HEPiX collaboration, the StratusLab contextualisation strategy would form a reasonable starting point. It is also compatible with other efforts, such as Cloud-Init (Ubuntu initiative, now supported by Amazon and several other OS vendors). The StratusLab specific solution, inspired by efforts from OpenNebula, is to create, at instantiation time, a CD-ROM populated with pre-defined files containing the contextualisation information.

StratusLab Policy Engine: trust on the cloud, and even more on federated clouds, between users and supplier system administrators is paramount. The ability to specify, filter and control what images and datasets are allowed, coupled with the ability to trace down to individual the provenance of these artefacts is paramount. This component, interfacing with the Marketplace, provides a potential starting point. However, this component would have to be integrated with the provisioning logic of each cloud, unless it is integrated in a per-supplier component in the 'blue box', such that the filtering is done prior to submission.

StratusLab Persistent Storage Service: the interesting elements that this service offers is the ability to transfer, using the HTTP protocol, images between StratusLab clouds, automatically. This service can also be used a caching engine, supporting a number of backends - e.g. iSCSI, distributed file system.

SlipStream Image Factory: a significant effort was and is currently required to create the right images for each user for and by each supplier. Further, in order to deploy production workload, users often require several distinct images. This is in part due to the fact that users come to each supplier with already existing virtual images, requiring them to be converted and adapted for each cloud solution. An alternative, as recommended by Helix Nebula's Technology and Architecture Group, is for each supplier to provide agreed and standardised reference base images, optimised for each provider's solution. Generic construction recipes would then be applied automatically, by such a factory service, resulting in images behaving near identically, taking advantage of each cloud provider (e.g. virtio for KVM based clouds, contextualisation enabled) capabilities.

SlipStream System Provisioning: the system provisioning capability already identified was used by ESA to provision a GRID cluster of variable size for G-POD processing. Although SlipStream currently only supports StratusLab and AWS EC2, it is in principle straightforward to extend to most cloud solution. This brings the ability for users to configure themselves, their own system provisioning and being able to apply it identically to the cloud provider of their choosing. Finally, such self-system-provisioning service would allow SME, for example, to deliver pre-configured systems, end-users can provision themselves, hence contributing to the vibrant eco-system Helix Nebula aims at developing.

3.3.2. CloudSigma

The CloudSigma solution offered to Helix Nebula flagship users is built around our existing public cloud IaaS offering. No new or dedicated hardware or facilities are required to support the provisioning of virtual resources for the proof-of-concept environments.

The CloudSigma cloud is formed around the KVM hypervisor. KVM offers full virtualization with no "false bottom" root passwords and complete protection of user VM's from packet sniffing or disk drive intercepts. KVM is a modern and innovative Linux-based solution for cloud management forming part of the mainstream Linux kernel.

Standard and publicly available virtualized resources are offered at a completely granular level for CPU, RAM and Storage with the only limitations placed on configurations being the maximum size constraints listed here:

- CPU: from 0.25GHz to 40GHz
- RAM: from 0.25Gb to 96Gb
- Disk Storage: from 1Mb to 20Tb
- CPU Cores: from 1 to 22
- VM's: no limit on the number virtual machines

Other infrastructure components available to PoC users via our standard public cloud offering include:

- The ability to change the granular CPU and RAM configurations
- All 10Gbs networking in the physical servers and backbone switches
- Multiple 10Gbs connections to the public Internet
- Local, persistent disk storage on a highly available distributed file system
- Static IP addresses
- Private VLAN configurations (up to 8 per virtual server)
- High performance VirtIO network and disk drivers

- Rapid cloning of Disks and Servers
- Web Console and JSON enabled RESTful API
- Server Group Names for easy control of large server farms
- Over 100 CD and DVD distributions of Linux, Unix and Windows operation systems
- Browser based VNC and RDP console connections to any running virtual machine
- Simple and fast image upload and download service
- Optional Data encryption on all storage media
- Optional high-performance solid-state-disk (SSD) storage

Pricing of these infrastructure resources is available on the CloudSigma public web site. Billing of these resources is done in 5 minute cycles on a pay-per-use basis.
Provisioning and operation of virtual machines on CloudSigma is accomplished using either our public Web Console or our JSON enabled RESTful API. Bridges to EC1/EC2, Jclouds and other public API's are fully supported.
CloudSigma also provides a localized version of MIT's StarCluster cluster management system as required for some of the PoC testing.

Relevance to Helix Nebula Architecture

In the context of the Helix Nebula requirements for a proposed service offering, the elements of CloudSigma's existing public cloud provide a solid basis for implementation of the initial proof-of-concept architectures.
As the architecture and technical requirements for a federated and open public cloud service mature, CloudSigma is committed to providing the interfaces, tools and support required to meet these specifications.
The CloudSigma public cloud offering is completely compatible with the goals and aspirations of the Helix Nebula project.

3.3.3. T-Systems

PoC Architecture

T-Systems PoC was provided based on its portfolio element Dynamic Services for Infrastructure (DSI). DSI was introduced in 2011 and is the extension of T-Systems Dynamic Services portfolio – established in 2005 – with a standardized IaaS service.

With DSI T-Systems delivers virtualized resources from a highly-secure private cloud. A self-service portal allows users to freely configure virtual servers, and to scale capacity up or down – at anytime. And to only pay for actual use.

In T-Systems production environment following resources were provided for the PoC:
- Up to 125 virtual machines (VM)
- VMs with up to 16 cores and 64 GB of RAM
- VMWare Hypervisor
- NetApp NAS storage
- Secure VPN network making use of existing Géant/NREN infrastructure with up to 250 Mbps throughput

The DSI service is based on the innovative cloud management software from Zimory which for example provides a REST-based API, automatic scale-out, supports various types of hypervisors and allows to define the physical location of data. DSI is continuously being developed towards a managed service (DSI 2.0) with PaaS elements and a wider range of interfaces (vCloud Datacenter). And currently the API is being integrated into OpenNebula.

DSI provides the following services:

- Automated provisioning of VMs, NFS storage areas and VLANs
- Cloud Manager User Portal
- Standard and user-defined VM templates
- Rule-based VM adjustments – triggered by time- or load-dependent threshold values
- Grouping mechanisms
- Image upload facility
- Direct console access
- Online reporting of resource usage
- Resource reservation and access management
- API control of all functions
- Online billing and cost centre management

Relevance to Helix Nebula Architecture

In the context of the Helix Nebula proposed system, a number of elements deployed in the T-Systems PoC are relevant and establish further proof the architecture can be implemented with limited investments. The following provides a non-exhaustive list:

- Existing user images could be converted and made to run on DSI, although the process should be further automated
- The API provided a comprehensive functionality and covers user requirements
- Existing Géant/NREN infrastructures can be customized to fulfil performance and security requirements to connect science infrastructures with commercial private cloud services
- Initial setup efforts defined and helped prioritize the Helix Nebula roadmap

3.3.4. Others

Further to the existing cloud environments provided by the supply-side, several technologies developed by Helix Nebula members should be considered when implementing the Helix Nebula architecture. For example, The Server Labs has developed FastTrack and Terradue has developed Cloud Controller, enabled by Opennebula toolkit exposing an OCCI and EC2 interface for IaaS provisioning. More details on these can be found in Section 4.12.

4. Helix Nebula Service Enabling Framework

The core of HelixNebula is the Service Enabling Framework, the so-called “Blue Box”, a complex component providing API services and a Web Portal that will enable users to interact in a central and transparent manner with all the Cloud Providers.

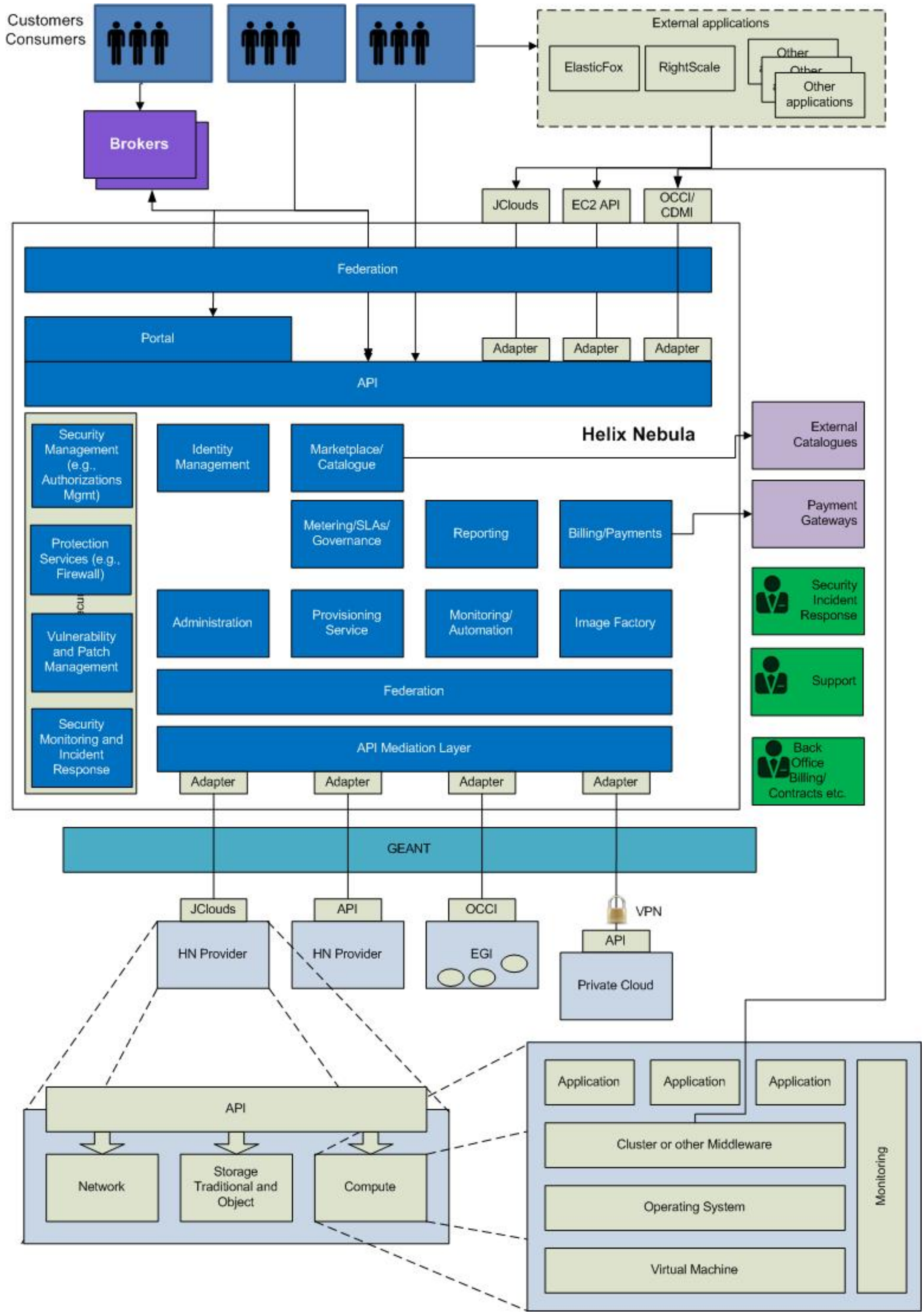


Figure 1: Helix Nebula Service Enabling Framework - overview

4.1. Catalogue / Marketplace

At the core of the HN architecture is a catalogue where the infrastructure provided by each supplier will be centralised so that users will be able to search in terms of price, operating system, architecture, software stacks provided, etc.

The Catalogue/Marketplace will consist of a portal (i.e. HTML application), as well as a REST API, and will allow the user to easily search for images in each of the different Helix Nebula Providers by attributes such as:

- Price
- Granularity of resources (e.g. m1.large or equalizer)
- Architecture. E.g. 32 bit, 64 bit, SPARC,
- O/S . e.g. Linux(all flavours), Windows, Solaris, etc.
- Capabilities, e.g. SSD, GPU
- Purpose
- Keywords
- Software installed

The Marketplace will also contain products and services offered and sold by HN Providers (Cloud vendors and SMEs).

The Marketplace will have an administration capability for HN Providers to provide descriptions of images, applications and services - e.g.:

- Firewalls
- Cluster Management Systems

We envisage that this catalogue will have a decision tree or filtering mechanism so that it will be easy for users to make their choice. It will be important that suppliers not chosen are notified as to why they were excluded from the decision criteria.

4.2. Federation and Identity Management

The Identity Management Service will provide a Single Sign-On (SSO) capability to the user so he/she will be able to use the same credentials irrespective of the Cloud Provider. In addition to cloud management credentials, it is envisaged that this component will be able to centrally manage instance SSH keys.

The Identity Management Service should provide mechanisms to manage and rotate these Keys and Credentials if necessary.

In the future we envisage that the users will be able to logon transparently to Helix Nebula using their corporate login credentials.

4.3. Image Factory

The Image Factory shall be able to interact with each HN Cloud Provider in order to create and upload custom images created by the user. Additionally it will be able to initiate transfers from one Provider to another, including image conversion if and when required.

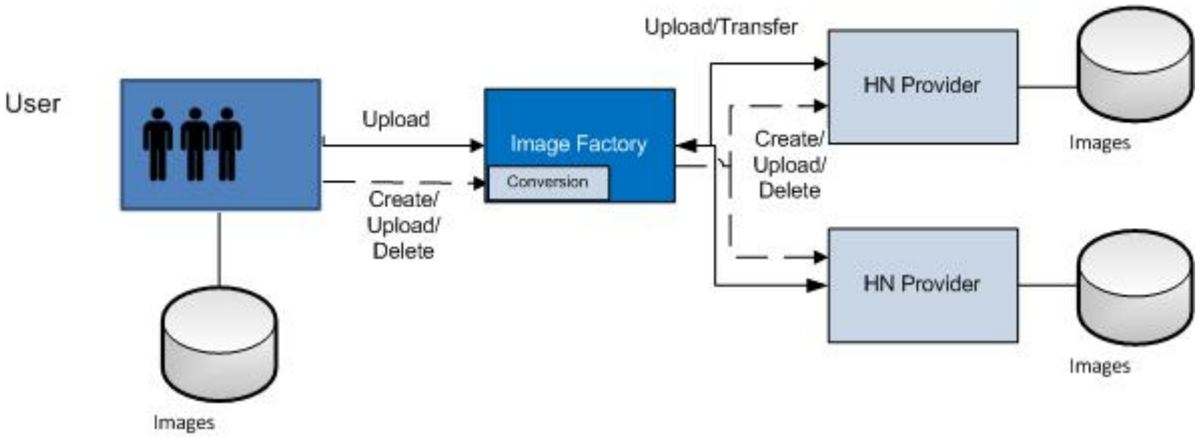


Figure 2: Image Factory architecture

The User will be able to upload and create images using the API or the Portal. Conversion of images from one format to another will be done by a conversion process that will first convert the image from the source format to OVF and then to the destination format.

An alternative mechanism to the upload and conversion workflow described in this section is the ability to re-create images based on existing base reference images and creation recipes.

Each provider might also prefer to use their own converter, which should be interfaced with the overall Helix Nebula image factory functionality.

4.4. Provisioning

The Provisioning Service will provide the facilities to provision at different levels:

- IaaS provisioning (Compute, Storage, Network, Security)
- O/S provisioning (upgrades, security patches, hardening)
- Server Contextualization
- System Provisioning
    - Middleware (e.g. Application Servers, DB's, ESBs).
    - Cluster Provisioning
- Application Provisioning
- Application Contextualization

The cluster management solution must be able to deliver functionality comparable with existing cluster management systems, such as StarCluster, Condor/Panda, GlusterFS, as currently used by the demand-side applications.

Templates stored within Helix Nebula will enable users to provision using the same templates irrespective of the Cloud Provider they are working with.

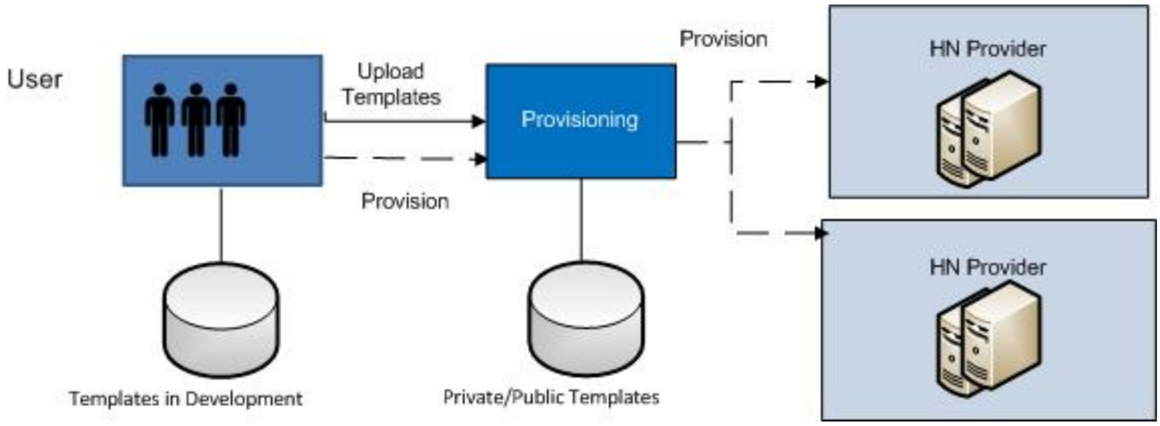


Figure 3: Provisioning architecture

A crucial part of this template concept is that of contextualization, whereby the users ssh keys and software that he wants installing will be added to the instance during the boot process. This process is what makes the public cloud different from traditional virtualization environments.
This template concept could also be applied to image creation and integrated with the image factory service. In such a scenario, providers without a required image could trigger an image transfer/conversion workflow or the creation of the missing image, as part of the provisioning workflow.

4.5. Security

4.5.1. Physical Security

Each cloud within Helix Nebula will have its own security policies and security mechanisms in place to provide physical security.

Helix Nebula will have a Security Policy that will encompass the Security Policies of the individual Cloud Providers. It will be easy to see at a glance the different certifications of each provider.

4.5.2. Perimitral Security

Permitral security is normally made up within each cloud from the following components:

- Routers

- Firewalls
- IDS (Intrusion Detection)
- VPN

The Service Enabling Framework will allow the users to configure these elements within each cloud provider in a consistent manner.

To enable the transfer of data and images between different providers it will be necessary to set up secure networking channels between the providers so that users will be able to see the their network as one.

Additionally users of Helix Nebula may want to configure instances with additional security, hardened images that have firewalls and some level of intrusion detection

Security Monitoring and Incident Response is an important part of any cloud. Helix Nebula will centralise the reporting and the user and provider response to security incidents.

4.5.3. Administration

Helix Nebula will provide a way for users to administer their cloud resources and the systems they have deployed within each cloud provider. We envisage that some of this can be done via API, but there will almost certainly be the need for an administration console.

4.6. Monitoring/Metering

The Helix Nebula API and Portal will provide monitoring of the users' infrastructure and their servers at a Multi-Cloud level. The dashboard must show the state of all the users' servers, regardless of the cloud provider they are running under.

Event Management will be important to allow features such as auto-scaling in the future.

In future versions of Helix Nebula a capability to do Server/Disk Analytics will also be important.

In order to enable accounting and quota enforcement, there must be a common Cost Reporting API together with consumption monitoring.

We aim to make it possible for users to customize their monitoring of their running systems through the installation of monitoring agents in their images or instances.

4.7. Billing/Payments

Although some of the process with regards to billing and payments will be manual, as much of it as possible should be automated within the Helix Nebula. Each supplier will provide the billing information through an API and Helix Nebula will provide the user with a consolidated bill.

In order to enable SMEs to be able to sell products and services within Helix Nebula, some sort of developer payment mechanism will be needed similar to Amazon DevPay that will give them the opportunity to invoice users through the billing system of Helix Nebula.

4.8. Network Federation and Image/Data Movement

Helix Nebula will provide a mechanism for users to import and export their data. The idea is that each provider will be connected to scientific users via GEANT to keep costs down.

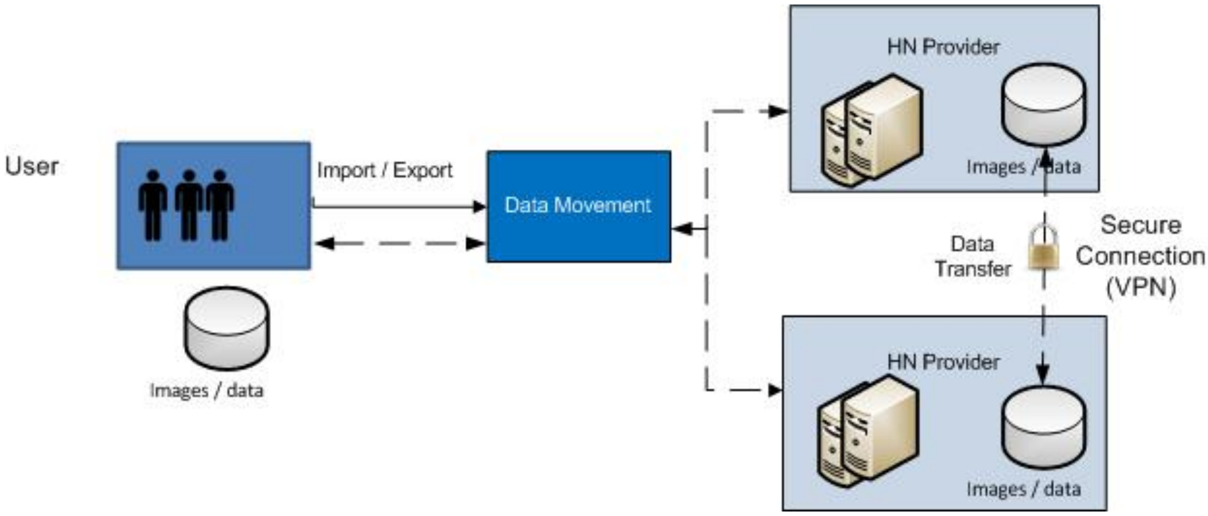


Figure 4: Network federation and image/data movement architecture

We envisage that by Release 3, users will be able to transfer images and data between sites (i.e. between clouds and between their data-source and a cloud site). Where necessary, data will be encrypted at rest and in motion.
In the future Helix Nebula should provide automatic compute load balancing between clouds with transparent data migration

4.9. Termination

The following actions and considerations will have to be taken into account once an account is terminated:

- Account Termination
- Termination Trigger
- Termination Reporting
- Data / Image De-provisioning
- Quarantine
- Employee Termination and Key de-provisioning

4.10. Runtime Infrastructure

The architecture of Helix Nebula will require an infrastructure to run on. Such infrastructure is likely to require the following components as a minimum.

- Database
- LDAP / Active Directory
- Web Service Engine
- Web Server / Portal / Console
- Identity Management System
- Intrusion Detection System

4.11. Technology

In order to construct the architecture of Helix Nebula it will be necessary to build, purchase or use open-source software. The following is a preliminary list of potential candidates technologies, standards and solutions:

- Catalogue (e.g. Sysfera, built.)
- API / provisioning –
    - Standards (e.g. OCCI, CDMI)
    - Mediation Layer (e.g. JClouds, EC2, SlipStream)
- Billing/ Payment Processing - EDI, pdf
- Monitoring (e.g. Ganglia, Nagios)
    - Event Management
- Federation / SSO / A.A.A. (e.g. X.509)
    - Key management, PKI, PCI
- Image formats and Factory - (VMware, Raw, ISO)
- Object Storage (e.g. S3, Willows, etc.)
- Contextualization (e.g. SlipStream, TSL FastTrack, Cloud-init)
- SLA Reporting
- Data movement/migration standards
- Network Transparency
    - Peering
    - Back-haul - private lines
- Support (e.g. Zendesk)

4.12. Service Provider Solutions

This section contains a summary of products from current HN contributors that could be used to build the Helix Nebula architecture.

4.12.1. SixSq SlipStream

The SlipStream service is a provisioning service and an image factory engine. It allows users to automatically create virtual machines based on existing images, in an automated and unattended way. It also allows users to provision systems (i.e. ensemble of virtual machines working together to deliver a system). Examples of systems SlipStream provisioned in the Helix Nebula PoC was a GRID cluster for ESA G-POD processing (as illustrated in Figure 5). Another use case could have been the GlusterFS, although in this case, the single system was provisioned by hand.

SlipStream supports StratusLab and AWS EC2, and a single SlipStream service can be used on different cloud endpoint, based on user preferences.

SlipStream is a SixSq product currently released under a proprietary license. However, SixSq has announced that it will open source SlipStream, probably under the Apache 2.0 license.

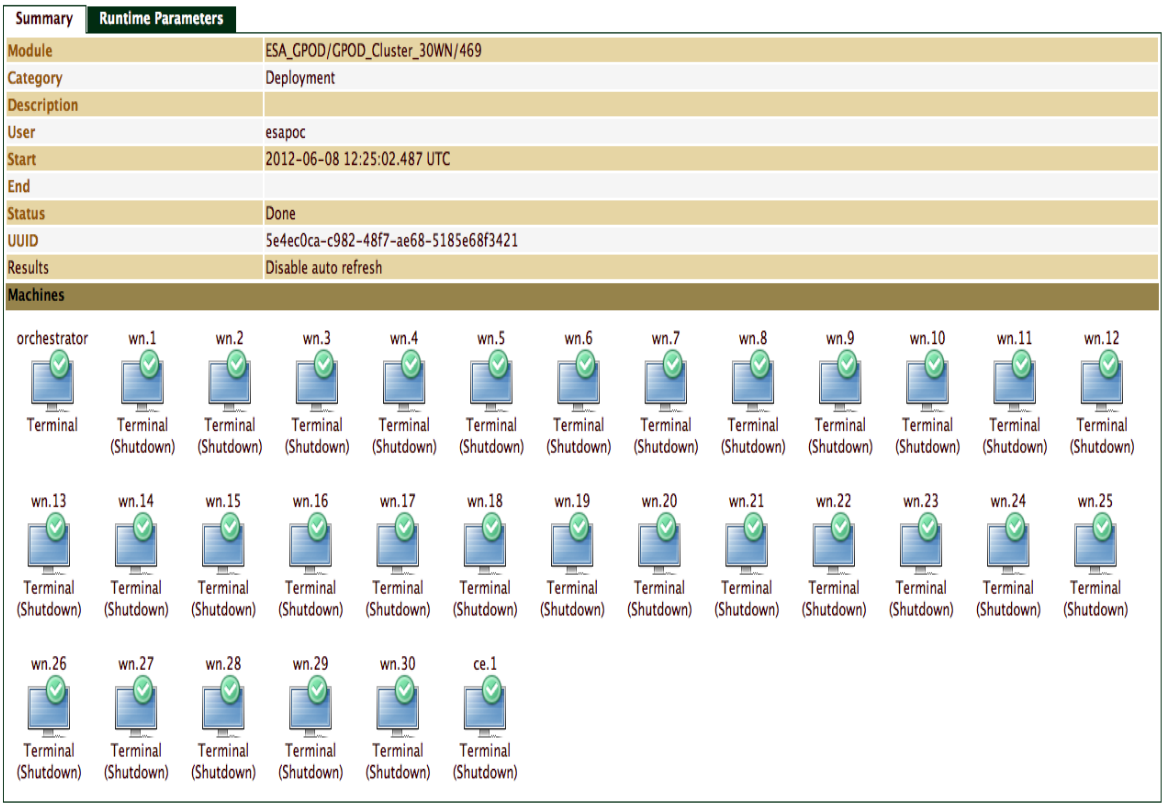


Figure 5: SlipStream provisioning of G-POD 30 WNs GRID cluster for processing as part Atos/ESA PoC

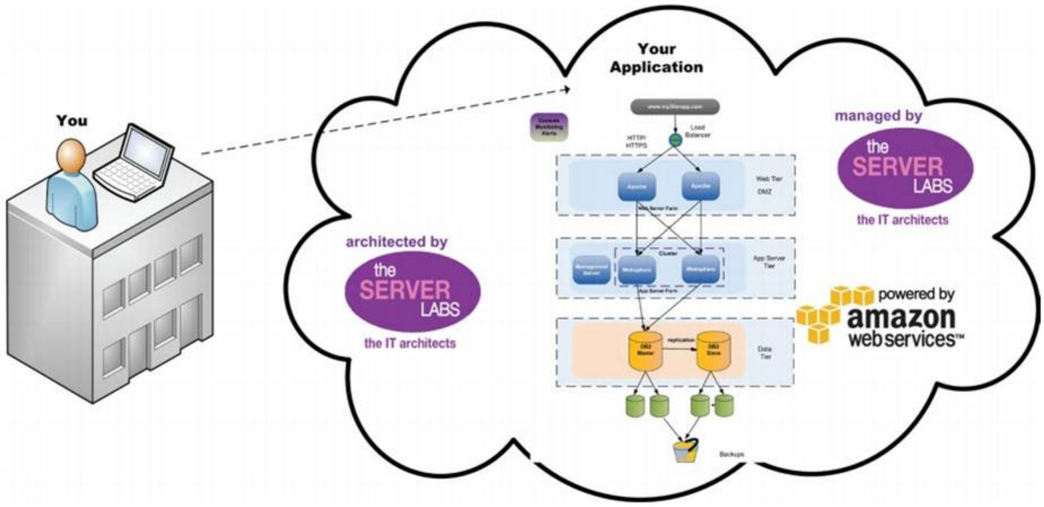
4.12.2. The Server Labs Cloud FastTrack

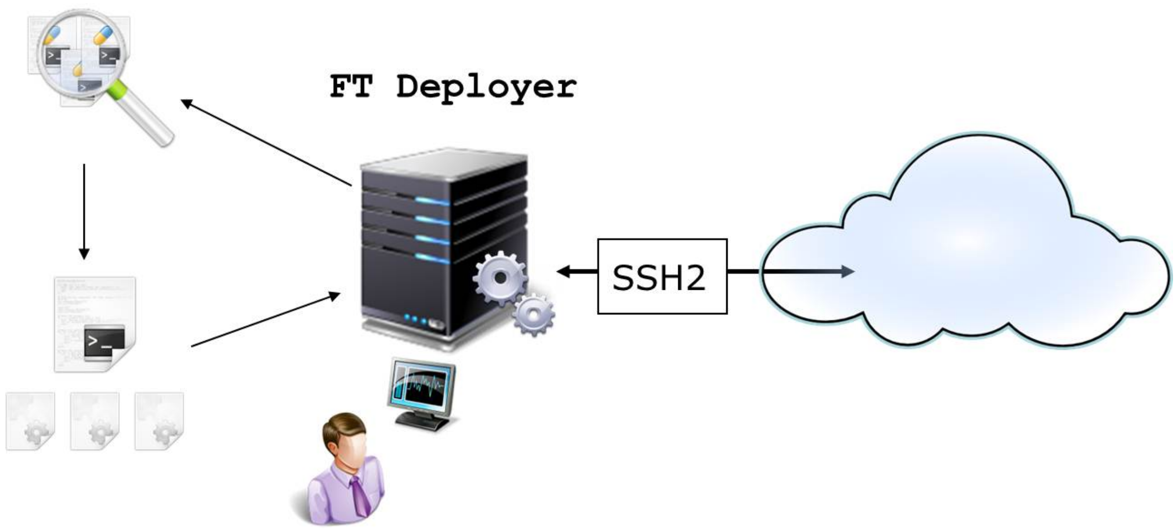
TSL's Cloud Fast Track Service has been designed to provide you with personalised, customised and fully functioning environments in the Cloud in only minutes. We have created this standard off the shelf service with busy IT departments in mind that need to deploy test and development environments in the Cloud and don't have time or expertise to develop their own.

The Server Labs Cloud FastTrack has the following features.

Drag and Drop Web Console. It's very simple to create complex deployments with best practices and security built in.

Multi-cloud. Supports multiple cloud providers
Contextualization. Contextualization is done from outside so you don't need to build special images or store your SSH keys in the cloud.
Template based. You can build on other people's templates and save them
Best practices. All our years of experience are built-in to the way FastTrack creates deployments.
Secure. By default, Fast Track will create your deployment with a DMZ reverse-proxy to provide extra security. It handles all firewall rules to make sure your deployment is secure





The fact that FastTrack does not impose any requirements on the cloud images makes it ideal to be part of the provisioning system within Helix Nebula.

4.12.3. Terradue

Will be provided in the next version of this document.

4.12.4. T-Systems Zimory

The Zimory software that provides T-Systems DSI service is a flexible, interoperable, scalable and standardized cloud management software suite that eliminates vendor lock-in; maximizes value from underlying technologies and delivers a Cloud service that can be tailored to support the Helix Nebula roadmap. Zimory provides:

- Enterprise-grade cloud service enabling secure data and application processing in the Cloud.
- Heterogeneous IaaS clouds utilizing existing infrastructure or dedicated new resources
- Architecture free from business and technical dependencies to deliver maximum value. Supporting technologies include: Xen; Oracle; NetApp; VMware; KVM.
- Reservation technology
- Flexible pricing based on utilization.
- Multitude of delivery models (public, private, community, hybrid and virtual private clouds) through a distributed architecture.

Furthermore, T-Systems provides industry cloud solutions based on vCloud Datacenter, and the open source frameworks OpenStack and Eucalyptus from which expertise and elements can be used to support the Helix Nebula roadmap.

5. Roadmap

This section details the first draft of the Helix Nebula "Blue Box" roadmap.

5.1. Release 1

The first release focuses on providing federated single-sign on functionality across multiple providers, the introduction of common service catalogue and image library and the ability to provision resources via a single common API.

5.1.1. Federation

Full Single Sign On functionality (US33) will be provided within the first release allowing users to be authenticated using common credentials across multiple cloud providers.

5.1.2. Service Catalogue

An extensible, searchable common catalogue format will be defined so that a wide range of attributes of supplier's capabilities can be represented. The catalogue allows users to search and filter the contents based upon attribute such RAM, CPU and disk capacities presenting the user with a list of providers that best match the attributes of their query (US38). A Supplier Upload API will allow Helix Nebula System Administrators to populate/interconnect with the catalogue to allow updating of their cloud's capabilities within the Service Catalogue (US39).

5.1.3. Image Factory & Marketplace

A common image library allows users to provision images based upon public or private images (US7/US8/US10/US11).

The library will be searchable (US17) with each image having associate metadata in a commonly agreed format and through the search functionality users will be able to apply filters to the search (such as Operating System or Architecture) in order to find a suitable image for their requirements (US19). Each image will have an associated unique identifier (US18) allowing users to easily provision VMs based upon particular image.
HN System Administrators will have access to basic analytics allowing them to see what images and popular and why (US20).

Within Release 1, images will be hypervisor specific and no conversion functionality will be available within this release.

5.1.4. Provisioning

A common API will enable users to provision VMs or disks/volumes on any cloud provider's infrastructure within Helix Nebula federated cloud this will include calls for both single and multiple VMs (US1/US2/US3).

Within Release 1, it will not be possible to provision multi-cloud resources or hybrid cloud solutions.

The provisioning of VMs based upon templates (US6) will be supported and provisioned VMs can be annotated with metadata for easier management.

Contextualisation support within this release will allow during provisioning the specification of a public SSH key that will be available on the machine during boot (US4) and in addition the specification of user data so that it is available within the machine for contained scripts to act upon it (US5).

Placement Criteria will allow users to provision computing resources that are close to the data they wish to process so that large amount of data do not need to be transferred for processing (US15.

5.1.5. Monitoring/Metering

Monitoring and Metering will, in the first instance, be vendor specific i.e. there will be no federated monitoring and metering functionality that will be provided within the Helix Nebula portal or via the API within Release 1. However, from within the portal users will be provided with hyperlinks to the metering and monitoring pages of the relevant suppliers. Where possible, efforts will be made to ensure commonality within the metering and monitoring functionality provided by each supplier.

5.1.6. Data Management

Release 1 will provide data ingest capability and the ability to transfer and export data to other cloud provider's platforms or back to the user's data centre (US14).

HN System Administrators will be able to replicate popular datasets within their cloud so that they do not need to retransfer the data each time they wish to perform computation on it (US16).

5.1.7. Billing/Payment

Billing under the first release will continue to be performed on a per supplier basis. However, efforts will be made to ensure that, for example, the format of the billing will be aligned by the suppliers.

5.1.8. Termination

Account termination will support the de-provisioning of any data and images used and termination reporting will be provided. Users will have the ability to quarantine any associated data.

5.1.9. Security

Release 1 allows full firewall setup and configuration (US28/US29) and basic credential and key rotation functionality will be present (US30/31). Security monitoring and incident response will be performed manually.

5.1.10. Connectivity/Networking

Within the first release, all suppliers who are part of the federated Helix Nebula cloud are expected to have GEANT connectivity (US34).

5.2. Release 2

Key areas of focus within Release 2 are the introduction of the Image Factory component that will provide image conversion between different vendor's infrastructure, Combined Metering and Billing functionality in addition to increased focus on automating the various security aspects within Helix Nebula.

5.2.1. Service Catalogue

Release 2 builds upon the functionality of the catalogue providing reporting functionality to suppliers around exception, use and content e.g. providers will receive details about when they were not chosen as a provider because, for example, they did not support Solid State Disks (SSD) or GPUs.

In addition, the Catalogue will be extended to support Policy Definition i.e. to control the movement of VMs due to legal constraints.

5.2.2. Image Factory & Marketplace

Release 2 will further increase the aspect of federation by providing an Image Conversion tool which will allow users without the support of system experts, convert an image in one format to that of the format supported by another hypervisor (US9/US12)

5.2.3. Provisioning

Provisioning aspects of the system will be further extended to ensure that policy constraints that have been defined within the Catalogue are enforced.

5.2.4. Monitoring/Metering

One of the focuses of Release 2 is the consolidation and unification of the monitoring and metering aspects of Helix Nebula. A common monitoring dashboard within the portal will provide a single point of contact for monitoring and metering functionality, which will be backed through the introduction of an Inbound Supplier API. Integrated Server Level Monitoring will be provided and the enforcement of accounting and quotas.

5.2.5. Billing/Payment

The new Inbound Supplier API will allow the combining of billing data of the various suppliers in the Helix Nebula programme. A single display within the portal will present a combined view of all usage costs within Helix Nebula and users will be presented with a single bill from the Helix Nebula legal entity. Common internal formats such as EDI will be adopted and reporting and lightweight accounting functionality will be included.

5.2.6. Termination

Termination functionality will be extended to support the process of Employee Termination and Key de-provisioning.

5.2.7. Data Management

Functionality will be provided to support the Archiving & Backup of user data and efforts will be made to ensure that a common approach of encryption of data across providers will be agreed.

5.2.8. Security

Security functionality of Helix Nebula is further extended within Release 2 to provide Key Ring Management, Automated Credential and Key Rotation (US30/31), Automated Security Monitoring and Incident Response in addition to Vulnerability and Patch Management.

5.3. Release 3

Release 3 in addition to some added functionality around monitoring and metering mainly focuses upon building value added services on top of the Helix Nebula federated cloud through the introduction of an Image and DataMart marketplace. Additionally, multi-cloud and hybrid cloud provisioning will be introduced.

5.3.1. Image Factory & Marketplace

Along with the introduction of Image Ingest functionality, a marketplace will be introduced to allow vendors to upload images pre-installed and configured with added-value software. In addition, a DevPay-like feature will allow vendors to generate revenue based upon the use of these added-value images paid via the cloud providers (US21).

5.3.2. Provisioning

With Release 3 the ability to provision cloud resources on several clouds at the same time will be introduced allowing users to take advantage of the benefits provided by different clouds e.g. geographical distribution/data proximity (US35).

Additionally, the ability to, out of the box, support a hybrid cloud deployment will be introduced through the creation of a Helix Nebula Connector, installable within a customer's data centre such that cross cloud provisioning can take place (US36/US37).

5.3.3. System Provision

The introduction of System Provisioning builds upon the Provisioning functionality that has been provided to date and provides the ability to provision a set of VMs including configuration and orchestration of the VMs such that together these VMs can deliver a system that can be used for scientific HPC processing (US13)

5.3.4. Monitoring & Metering

The common Monitoring and Metering functionality from Release 2 will be further extended to provide Service Level Agreement (SLA) reporting.

5.3.5. Data Management

Release 3 introduces a number of features related to public/private data, more specifically around functionality to allow the introduction of a DataMart. The DataMart allows to users to upload (US23) and discover high-quality scientific data (US22). Data sets will be identifiable by a unique identifier (US25) and the data providers will have the ability to access basic analytics regarding the data such as popularity of a particular dataset (US27). Users will be able to restrict access to particular datasets shared on the HN federated cloud (US24) and data providers can advertise datasets within a marketplace allowing the generation of revenue for a DevPay type system (US26).

Additionally, further progress within Data Mobility and format will be introduced to ensure that data sets can be moved between providers within the Helix Nebula Federated Cloud ensuring data validity amongst suppliers (US32).

5.3.6. Billing/Payment

The Helix Nebula Billing and Payment framework will be extended to cover DevPay payments related to the Image Marketplace and DataMart functionality introduced within this release.

6. Conclusion

This document brings together the shared technical and architectural knowledge of the supply-side of the Helix Nebula collaboration. This knowledge is based on the participants' background expertise as well as the knowledge acquired from running proof of concept with CERN, ESA and EMBL.

This document is a first attempt to serialize our current body of knowledge and is aimed at creating a technical dialogue between the supply and the demand sides, such that the implementation of the Helix Nebula system follows a realistic path, while delivering the right features in the right order. This will enable the cloud vision all Helix Nebula participants share to become reality in a timely fashion.