

Helix Nebula – The Science Cloud

# Helix Nebula Marketplace Catalogue

**Abstract:** This document describes the context for and the outline content of the catalogue facility within the Helix Nebula Marketplace (HNX). It is intended both to provide orientation to HNX Suppliers as to what they could include and guidance to HNX Users as to what they should expect to find. The catalogue will evolve with use and this document is expected to evolve accordingly.

## Document Information Summary

<b>Deliverable number:</b>	<i>n/a</i>
<b>Deliverable title:</b>	<i>Helix Nebula Marketplace Catalogue</i>
<b>Editor:</b>	<i>Mick Symonds (Atos)</i>
<b>Contributing Authors:</b>	<i>Michel van Adrichem (Atos)</i> <i>HN ServArch group</i>
<b>Reviewer(s):</b>	
<b>Distribution:</b>	
<b>Nature:</b>	<i>Report</i>
<b>Version/Revision:</b>	<i>1.0</i>
<b>Draft/Final:</b>	<i>Final</i>
<b>Keywords:</b>	<i>Marketplace Catalogue HNX</i>

## 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 Marketplace activities and offerings, and it may also contain contributions from distinguished experts who contribute to Helix Nebula. Any reference to content in this document should clearly indicate the authors, source, organisation and publication date. The content of this publication is the sole responsibility of the Helix Nebula Marketplace suppliers, and cannot be considered to reflect the views of the wider Helix Nebula initiative nor associated EC FP7 Project.

## Log Table

Issue	Date	Description	Author/Partner
V0.1	6 Jan 2014	First draft	Mick Symonds (Atos)
V0.2	7 Jan 2014	ServArch updates	MAS
V0.3	7 Jan 2014	... plus those of Michel van Adrichem	MAS
V0.4	16 Jan 2014	Comments from Robert Jenkins (CS), ServArch/FitSM	MAS
V1.0	11 Feb 2014	Final version	Michel van Adrichem (Atos)
	27 May 2014	To Creative Commons licence	MAS

## Document Approval

Issue	Date	Name

## Table of Contents

1. Introduction .....	5
2. The Helix Nebula services in general.....	6
2.1. Generic Infrastructure as a Service.....	6
2.2. How it works .....	7
2.3. Service levels .....	7
2.4. What do you need to use the services? .....	8
2.5. Sandbox facility.....	8
2.6. Additional and future services.....	8
2.7. Costs and payments.....	8
2.8. Sustainability .....	9
2.9. Further future developments.....	9
2.10. Further documentation on Helix Nebula.....	9
3. Nature of the catalogue.....	10
4. Catalogue elements .....	11
4.1. Overall capacity .....	11
4.2. Single VM capacity.....	12
4.3. Availability.....	13
4.4. Price.....	14
4.5. Locations(s).....	15
4.6. Suppliers catalogue.....	15
5. Future developments.....	16

## 1. Introduction

This document is prepared as part of the launch of the Helix Nebula Marketplace (HNX), in which a number of suppliers seek to offer their services, in a competitive but compatible manner, primarily to European research organisations.

Those suppliers must adhere to the Anti-Trust Guidelines (ATGs) as defined within Helix Nebula, thus not acting as a cartel. Therefore this document should not be seen as defining or restricting the services offered by any one supplier. Rather, it serves to ease the discovery and selection process by the potential customers of the services, by describing a mechanism which provides an indication of the range of services which may be available in common between those suppliers, as chosen by each supplier themselves. It is entirely up to each supplier as to how much and what information they include.

This document only reflects the current views and vision of its authors.

The document is structured as follows:

- An introduction to the intentions of the document itself (this section);
- A general outline of the Helix Nebula services;
- An outline of the nature of the catalogue;
- An outline of the initially-expected elements within the catalogue;
- Some indication of possible relevant future developments.

## 2. The Helix Nebula services in general

Helix Nebula is a pioneering development which is intended to support the massive IT requirements of European scientists, at least initially, by establishing a sustainable, federated European cloud computing infrastructure. It provides state-of-the-art infrastructure, deployed or decommissioned easily and rapidly, without the need for capital investments or on-going fixed costs.

Helix Nebula is delivered as a federated, multi-tenant and multi-supplier facility, providing a choice of service providers, using a common standard interface to its Market Place, based on open standards and Open Source software. There are currently four IaaS providers and others providing other added-value services. It was developed under the aegis of three of Europe's most advanced research organisations: CERN, EMBL and ESA.

It delivers a secure European solution, partly to address concerns regarding adherence to European data protection legislation and to avoid unwanted access by foreign agencies.

Helix Nebula allows scientific organisations to concentrate on their science, leaving the provision and administration of IT facilities to the professionals in that field.

### 2.1. Generic Infrastructure as a Service

Infrastructure as a Service (IaaS) consists of a number of matched components:

- Processing: virtual machines (VMs) made up of CPU and memory, in variable or optional standard T-shirt sizes;
- Storage: connection to various technologies and performance characteristics, typically defined in structures of Tiers, permitting use of approaches such as Hierarchical Storage and Information Lifecycle Management (HSM and ILM);
- Network connectivity: secure virtual data centres, using techniques such as VLANs, DMZs, and/or VPNs, accessible on demand via the internet, or via GEANT for researchers.

The services are offered with built-in security and integrity: the supplier's services may be certified to ISO2700x standards, with optional ISAE3402 reporting.

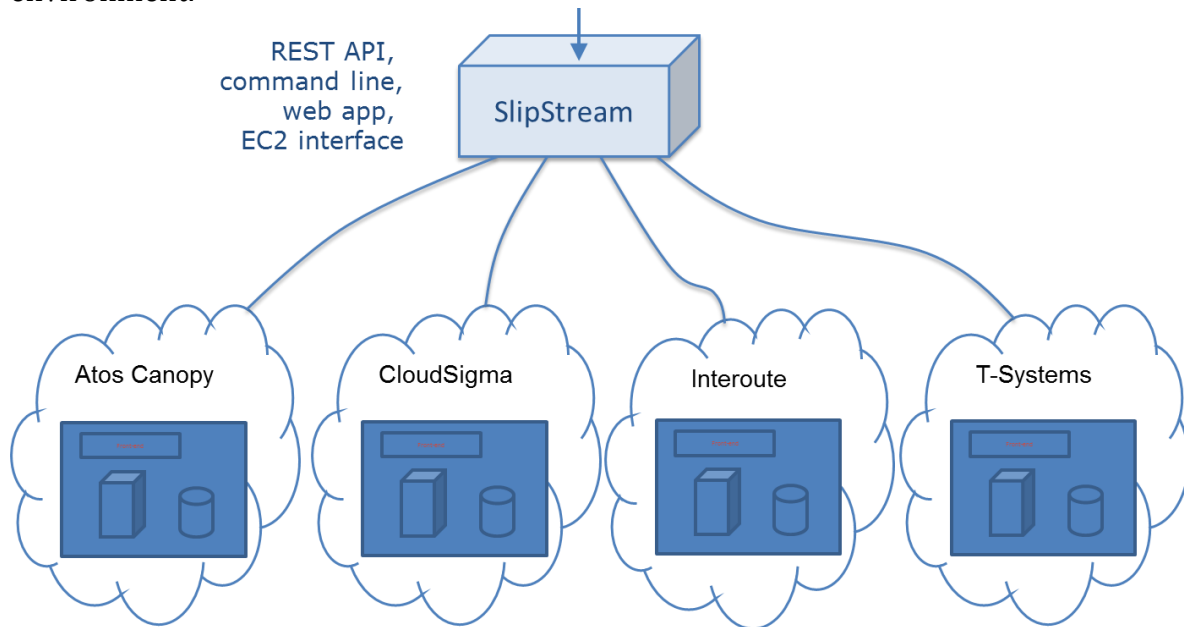
The environments benefit from the economy of scale of major service providers: so they are highly-standardised. They are likely to be kept up-to-date ("evergreen"): with in-built release and patch management, especially for security patches. The environments are highly automated, to ensure responsiveness and predictable behaviour.

## 2.2. How it works

The environments are connected via a common front end (or “Blue Box”, so-called from its original depiction):

- On the user side, providing a self-service portal and/or API which can be accessed from a user’s own workload/scheduling system,
- On the supply side, connected to multiple IaaS suppliers.

Services are invoked via this common Blue Box, which can create, monitor or cease an environment.



The suppliers seek to provide a “joined up” service, by coordinating their service delivery and support structures, with the option of a single point of contact and unified reporting.

Workload is typically deployed as VM images: these may be from the provider or users can bring their own.

## 2.3. Service levels

Quality of Service metrics are in place, and can be monitored and reported on, e.g. availability, which can be defined at pre-specified levels: e.g. bronze, silver, gold or platinum. Undertakings are also given as to speed of changes, e.g. up- and down-scaling.

The providers each define the degrees of warranty they provide and possible penalties if they are not met.

## **2.4. What do you need to use the services?**

User organisations will need to be registered, with accounts to which services can be charged and associated quotas, on both the central Blue Box and each deployed IaaS supplier.

On the IaaS services, users may need to provide their own application stack, possible including licencing where that is applicable, unless they can run using only the (“PaaS”) tools and facilities provided.

## **2.5. Sandbox facility**

There is a “Sandbox” facility, using a segment of the full production environment for a trial period, to allow users to “test and try” before they buy.

## **2.6. Additional and future services**

Various other added-value services are available from the various suppliers, and can be found in the central catalogue or that from each supplier. They include the following:

- Environment selection and configuration: a consultancy and brokerage service, to help users work out what their needs are and how they can best be fulfilled;
- Implementation facilitation and support, including both initial consultancy and on-going support;
- Operating Systems (OS) which can be run, possibly including licensing and support (e.g. for various flavours of Linux or Windows);
- Data backup and archiving;
- High Performance Computing (HPC): the provision of specialist environments which differ from normal IaaS and are configured to allow more extreme “number crunching”, whether for scientific or business use;
- End user and application support, e.g. for scientific or business applications, such as SAP with DB2 or Oracle;
- Business Continuity and Disaster Recovery (BC/DR): including both consultation on needs and adherence to regulations, and technical facilities such as clustering, replicated storage, back-up, etc.

## **2.7. Costs and payments**

Each supplier within Helix Nebula sets their own service definitions and prices, which can be accessed via the central or that supplier’s own portal.

A central indication of a user’s quota allocation and use is provided from the Blue Box, which can be used to give an indication of expected billing, which itself comes directly from the supplier(s).

## 2.8.Sustainability

Suppliers may give indications or warranties as to the Carbon neutral nature of their services, and/ or an indication of their Carbon footprint.

## 2.9.Further future developments

Helix Nebula currently delivers IaaS services, and further service extensions are in preparation, including:

- Data management: the structure for storing and replicating data, and the middleware to manipulate and manage it, such that it is advertised, made available (only) to, and useable by, those who are its targeted users.
- There is an intention that scientific data from a range of sources can be made available via coherent means, such that cross-field synergies may emerge, leading to scientific breakthroughs and benefits to society as a whole.
- Certain European public infrastructure facilities (e.g. EGI, GEANT) will be connected into the Helix Nebula environment, allowing its users to adopt a hybrid cloud approach, making use of both public and commercial services via the same interface.
- Information (or Data Analytics) as a Service, building an environment to extract useful information from the hosted data and make it available as a service, possibly as a paid service, to authorised users.

## 2.10. Further documentation on Helix Nebula

There is a wealth of documentation produced by/on behalf of the HN consortium and published on the internet (at <http://www.helix-nebula.eu/>); only a few documents are kept to NDA signatories.

### 3. Nature of the catalogue

The central catalogue is provided to ensure ease of discovery and access by consumers to the various services available within HNX.

It describes what is available, in general and outline terms, across multiple suppliers. It is intended to supplement the fuller information within individual supplier catalogues, which remain the definitive sources. It is intended simply to provide orientation and some degree of alignment, in order to find what common services might be available across suppliers, and to facilitate comparison between those services.

The catalogue is housed within the portal facility, which forms part of the so-called “Blue Box”, providing a common front-end to the services within HNX. It can be accessed by users via a web interface (at <https://bb1.hnx.helix-nebula.eu/> and/or programmatically by use of a provided API.

## 4. Catalogue elements

The catalogue consists of 6 sections and a variable number of listings per section. The five sections in this description of the catalogue are:

- Overall capacity
- Single VM capacity
- Availability
- Price
- Locations
- Suppliers catalogue

The actual representation within the online catalogue may or may not follow this same pattern.

Each catalogue listing consists of a number of value pairs, which each describe, per supplier:

- Nature: what each element is (e.g. the number of CPU cores available), and
- Its value (e.g. 1,000);
- Explanation: an explanation of the catalogue entry. This is primarily intended for this document; it is yet to be determined how it will be handled within the online catalogue.

Within each section below, examples are given of the initial expected set of those value pairs, with some explanation of their possible use.

Note that this list is likely to be extended, as the services themselves evolve. But it should not become too long and complex, as that would defeat the object of having a common, simple catalogue facility.

### 4.1. Overall capacity

This indicates the overall scale (currently) available within (the relevant part of) the supplier's IaaS environment.

Note that this may vary over time, e.g. increases due to extensions to the installed base and decreases due to assignment elsewhere/to other customers/uses. It is not expected to be maintained on a realtime basis, so is only an indication rather than a warranty of what would be available for a specific request. Some suppliers make offer the ability to "reserve" capacity if firmer commitments are required.

#### cpu/cores

- nature: the number of CPU cores (currently) available within (the relevant part of) the supplier's IaaS environment;
- value: an integer, and possibly approximate, number, e.g. 1,000;

- explanation: to give an indication of the scale of the environment available for use.

#### **ram**

- nature: the amount of random-access memory in total;
- value: expressed in relevant terms, e.g. 10 TB;
- explanation: the amount of memory available across the installation as a whole. See below for what is available on any one system.

#### **storage**

- nature: the amount of persistent storage (e.g. SSD, disk, tape) available within that supplier's environment;
- value: expressed in relevant terms, e.g. 10 PB;
- explanation: possibly multiple values, e.g. per technology type.

### **4.2.Single VM capacity**

The capacity that is available on any one virtual machine within that supplier's environment.

#### **Size of CPU core**

- nature: an indication of the speed of the processor;
- value: a number in GHz;
- explanation: there is no one single measure of processor capacity or speed, but GHz is a reasonable indication.

#### **type of CPU**

- nature: an indication of the processor manufacturer;
- value: manufacturer name (e.g. Intel or AMD generally) and CPU generation (e.g. 63XX, E5YY, etc.);
- explanation: applications may be written to take advantage of facilities incorporated by a particular manufacturer at some stage of their product development.

#### **minimum cpu/cores**

- nature: the minimum number of CPU cores with which this supplier's VMs can be configured;
- value: an integer number, e.g. 1;
- explanation: to give an indication of the minimum configurable environment.

#### **maximum cpu/cores**

- nature: the maximum number of CPU cores with which this supplier's VMs can be configured;

- value: an integer number, e.g. 8;
- explanation: to give an indication of the maximum configurable environment.

#### **minimum ram**

- nature: the minimum amount of random-access memory (currently) available within VMs;
- value: expressed in relevant terms, e.g. 8 GB;
- explanation: the minimum amount of memory available to any one VM within the supplier's IaaS environment.

#### **maximum ram**

- nature: the maximum amount of random-access memory (currently) available within VMs;
- value: expressed in relevant terms, e.g. 128 GB;
- explanation: the amount of memory available to any one VM within the supplier's IaaS environment.

#### **volatile storage**

- nature: the amount of volatile, typically locally-attached disk, storage available locally to that VM;
- value: expressed in relevant terms, e.g. 500 GB;
- explanation: the amount of "scratch" space, which could be used, e.g. to extend the random access memory of a VM. Local disk space is typically slower than ram but faster than persistent storage space.

#### **persistent storage**

- nature: the amount of persistent storage (e.g. SSD, disk, tape) available to that VM, of a given storage access method (e.g. local or network) and storage type (e.g. block device or network mount) and resilience level or equivalent, (e.g. RAID6, RAID5, etc.);  
value: expressed in relevant terms, e.g. 10 TB per drive/block device;
- explanation: possibly multiple values, e.g. per technology type. This presumes that storage is associated with a particular VM, i.e. it is locally attached or via a restricted network. Otherwise, it could be up to the total figure, as above.

### **4.3.Availability**

An indication of the normal, single-system availability offered as a service level within that supplier's IaaS environment. This does not allow for special arrangements, e.g. customer-specific system clustering.

Note that suppliers may differ in their approach to offering levels of availability and granting penalties if they are not met.

#### **availability offerings**

- nature: whether one or more availability offerings are made;
- value: the level designation(s), e.g.: bronze, silver, gold;
- explanation: these could vary per supplier, as there is no standard designation. Work from the EC, the ODCA or Deutsche Bourse could be used to derive such a standard, at least for comparative purposes, in the future.

#### **availability levels**

- nature: the actual values of availability offered;
- value: the level as a percentage, e.g. 99.9%, etc.;
- explanation: to be comparable, service elements must be expressed and measured in a consistent way (e.g. as defined within ITIL).

### **4.4.Price**

An indication of the normal, non-discounted pricing for use of that supplier's IaaS environment. This does not allow for, e.g. customer-specific or volume discounts.

#### **charging unit**

- nature: the unit used for charging;
- value: the pricing unit, e.g. GHz, portion of CPU chip, etc.;
- explanation: this could vary per supplier, as there is no standard unit. Work from the EC, the ODCA or Deutsche Bourse could be used to derive such a standard, at least for comparative purposes, in the future.

#### **charging period**

- nature: the period used for charging;
- value: the pricing period, e.g. hour, month;
- explanation: this could vary per resource, e.g. CPU per hour, storage per month.

#### **cpu/cores/hour**

- nature: the price for use of a unit of processing per period, e.g. hour;
- value: the price in euros, e.g. €0.05;
- explanation: CPU is shown as being priced per hour, but in practice that may differ per supplier and environment.

#### **ram/hour**

- nature: the price for use of a unit (e.g. 1 GB) of memory per hour;
- value: the price in euros, e.g. €0.05;
- explanation: typically, the RAM is associated with the CPU, above.

#### **storage/month**

- nature: the price for use of a unit (e.g. 1 GB) of (e.g. disk) storage per month;
- value: the price in euros, e.g. €0.20 (20 cents);

- explanation: storage is shown as being priced per month, but in practice that may differ per supplier and environment. Note that it is possible that storage is either associated with a particular VM or as a generally-available resource.

#### **network/GB**

- nature: the price for transmitting a unit (e.g. 1 GB) in or out of the environment;
- value: the price in euros, e.g. €0.30;
- explanation: the supplier may charge, at differing rates, for data transmission in and out of the environment. Note that for customers making use of the GEANT connection to their cloud supplier, there is currently no charge from the supplier for that part of the connection.

#### **iops/GB**

- nature: the price for a unit of data (e.g. 1 GB) to or from the storage environment;
- value: provisioned IOPS (e.g. at €0.125 per GB of allocated storage per month);
- explanation: (do we split read and write?).

### **4.5.Locations(s)**

The customer may need assurance as to the geographical location(s) of the supplier and/or in which the environment is housed, e.g. for data protection legislation reasons. This can be by country, and/or whether it is 'within Europe'.

- nature: geographical location of relevant data centre(s) and/or operational company (this can be different from the physical cloud location);
- value: ISO-standard country code and name for cloud location and/or operational company location, e.g. NL The Netherlands;
- explanation: currently, data protection legislation differs per country, and for some legislations the data owner is obliged to ensure the location for some types of data (e.g. personal data).

### **4.6.Suppliers catalogue**

Source of further information, whether of more details of these services or further available services.

- nature: URL of web site with further details.

## 5. Future developments

It is expected that the catalogue will be enhanced and extended in the future, although it is not intended to extend it to encompass all of the information available in all suppliers' catalogues; that would defeat the object of having a simple point of initial reference.