ICT 269978

Integrated Project of the 7th Framework Programme

COOPERATION, THEME 3
Information & Communication Technologies
ICT-2009.5.3, Virtual Physiological Human

Work Package: WP2

Data and Compute Cloud Platform

Deliverable: D2.5

Specification and Costs of Bought-in Requirements for Cloud Compute (WP2) and Data (WP3) Services

Version: 1v0

Date: 24-Feb-13

Document Information

IST Project Num	FP7 – ICT - 269978	**Acronym**	VPH-Share
Full title	Virtual Physiological Human: Sharing for Healthcare – A Research Environment		
Project URL	http://www.vph-share.eu		
EU Project officer	Robert Begier		

Work package	**Number** 2		**Title**	Data and Compute Cloud Platform
Deliverable	**Number** 2.5		**Title**	Specification and Costs of Bought-in Requirements for Cloud Compute (WP2) and Data (WP3) Services

Date of delivery	**Contractual**	28-Feb-13	**Actual**	28-Feb-13
Status	Version 1v0		Final ☐	
Nature	Prototype ☒ Report ☐ Dissemination ☐ Other ☐			
Dissemination Level	Public (PU) ☒ Restricted to other Programme Participants (PP) ☐ Consortium (CO) ☐ Restricted to specified group (RE) ☐			

Authors (Partner)	Maciej Malawski, Jan Meizner, Piotr Nowakowski, Marian Bubak (CYFRONET), Susheel Varma (USFD)			
Responsible Author	Maciej Malawski		**Email**	malawski@agh.edu.pl
	Partner	CYFRONET	**Phone**	+4812 328-33-53

Abstract (for dissemination)	This document gives a survey and ranking of public cloud providers and presents results of their performance and compatibility evaluation as well as an assessment of compute and storage requirements of VPH-Share applications. It also overviews the status of federated cloud based on resources provided by project partners are also given. In result, this deliverable is a firm background for specification of resources to be purchased from public cloud providers.

Keywords	Cloud computing, public clouds, cloud performance evaluation, application resource requirements.

Version Log			
Issue Date	**Version**	**Author**	**Change**
24/01/2013	0v1	Maciej Malawski, Jan Meizner, Piotr Nowakowski, Susheel Varma	Initial import from Google doc.
13/02/2013	0v2	Maciej Malawski	Update after comments from WP5 (Susheel Varma)
20/02/2013	0v3	Maciej Malawski	Update after internal review, including resource estimates
21/02/2013	0v4	Maciej Malawski	Add SoftLayer performance results
22/02/2013	0v5	Maciej Malawski, Marian Bubak	References, final corrections
24/02/2013	1v0	PMO	Submission Version

Contents

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The goal of this document is an assessment of public cloud services to be purchased by the project in years 3 and 4 as well as to define mechanisms of federating resources from private clouds operated by project partners. The project attempts to build on top of a hybrid Infrastructure-as-a-Service (IaaS) cloud platforms. Currently, the project operates a private development cloud infrastructure based on OpenStack and hosted by CYFRONET and installations in Sheffield and Vienna are in progress.

We analysed nearly 50 public commercial cloud providers using criteria such as EU location, jclouds API (Application Programming Interface) support, BLOB (large binary object) storage service, public API, published price, hourly billing, VM (Virtual Machine) import feature and relational database support. These criteria were selected to choose only these providers that offer the required elastic and dynamic services needed for the VPH-Share cloud platform. We identified three leading cloud providers, namely Amazon EC2, RackSpace and SoftLayer that fulfilled the three the most important criteria. There are also providers such as CloudSigma, ElasticHosts and Serverlove that fulfil most criteria except BLOB storage.

We have tested jclouds API compatibility of these six top cloud providers and found some minor compatibility issues. We have also analysed the performance of compute instances of Amazon EC2, RackSpace and SoftLayer to evaluate their cost efficiency (price vs. performance) for compute intensive applications. These data will be used to guide the dynamic resource allocation of the Atmosphere cloud platform.

According to the estimates based on cloud survey and price and performance analysis, the budget of EUR 70,000 allocated for public cloud providers within VPH-Share will be sufficient to buy a service of 1,000,000 single-core CPU hours, or operate a 57-core cluster running 24x7 for 2 years, or store 29 TB of data for 2 years.

Current estimates of resource needs from VPH-Share workflows (WP5) give the total of around 180,000 compute hours, 15,000 GB-months of storage and data transfer of 6,000 GBs over 2 years. Based on our cost analysis, these requirements should be satisfied with a large safe margin.

The public cloud resources will be used to supplement the existing VPH-Share private cloud resources, thus delivering a hybrid, scalable and dynamic cloud computing environment for VPH research.

This deliverable provides a technical background for elaboration by the Project management of a specification of resources to be purchased from public cloud providers.

1 INTRODUCTION

The goal of this document is an assessment of public cloud services to be bought by the project in years 3 and 4 as well as to define mechanisms of federating resources from private clouds operated by project partners. The analysis presented in this deliverable is based on our experience gained during research on efficient usage of cloud resources [1] [2] [3] [4]. According to the analysis and design of the Atmosphere cloud platform described in D2.1 and D2.2, the project is building on top of a hybrid Infrastructure-as-a-Service (IaaS) cloud platform (defined according to NIST report [5]). Currently the project operates a private development cloud infrastructure based on OpenStack and hosted by CYFRONET and installations in Sheffield and Vienna are in progress. The project has a budget of EUR 70,000 allocated for public cloud providers (DoW, part B, p. 69), based on the following initial estimate:

- Storage: at a rate of 1TB per month with a 5:1 ratio of download to upload.
- Compute: 3 large instance servers (as defined by Amazon) running 24x7.

A schematic outline of the federated cloud is shown in Figure 1 below.

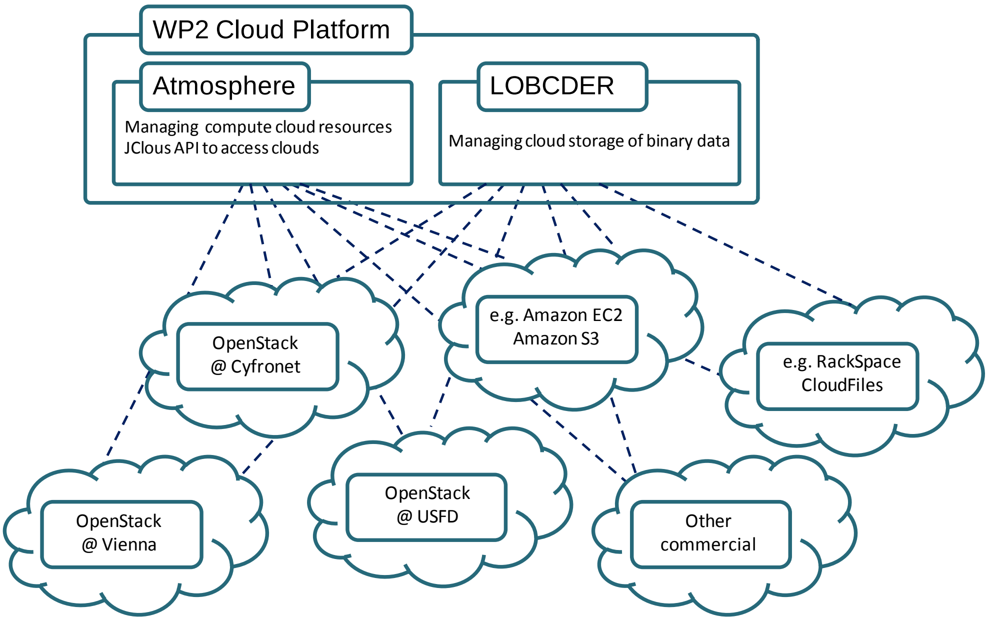

Figure 1 Federated data and compute cloud consisting of private sites (Krakow, Sheffield, Vienna) and selected commercial public cloud providers (to be selected).

The recommendations prepared in this document are based on:

1. Functional requirements from the VPH-Share applications (workflows),
2. Technical requirements from the Atmosphere cloud platform (WP2),
3. Cost estimation based on resource needs from the applications,
4. Results of preliminary performance benchmarks of public clouds.

‘The resources that need to be provided fall into two categories:

1. compute resources, i.e. IaaS cloud services enabling on-demand provisioning of Virtual Machines (VMs) using a public API,
2. cloud storage enabling storage of large binary objects using a public API.

Our plan for the federated cloud is to select 2 or 3 top cloud providers that can be used by the project. Having more than one provider may be necessitated by the fact that not all providers offer complete functionality (e.g. it may be better to use one provider for compute resources and another one for storage). Non-functional requirements such as latency to specific users (e.g. hospitals) may be also important. Moreover, supporting more than one cloud provider will help demonstrate the federation capabilities of the Atmosphere cloud platform and prevent ‘vendor lock-in’ problems, as well as help mitigate potential provider outages.

The document is organised in the following way:

- Section 3 defines the application requirements; we provide the application classification with examples from VPH. The applications are grouped into the categories such as:
  - Web Application/Services (e.g. stateful or stateless services)
  - low latency Applications (e.g. native GUI applications)
  - High CPU Applications (all compute-intensive services)
  - High I/O Applications (e.g. databases)
  - High Memory Applications (e.g. in-memory caching)
  - Cluster Applications (HPC, MPI, CFD)
  - MapReduce (Hadoop) Applications
  - GPGPU Applications
- Section 4 defines evaluation criteria of public cloud providers. These criteria are broken down into two levels:
  - Level 1 - Broad Survey with criteria such as: jclouds API support, European Economic Area (EEA) Zoning, Price, BLOB Storage support, RDS storage
  - Level 2 - Detailed Evaluation of selected clouds from Level 1, using such criteria as Application Benchmarks and results of API tests.
- Section 0 describes the application of evaluation criteria to various commercial and academic IaaS providers
- Section 6 gives conclusions from the broad survey.
- Section 7 describes results of performance and API tests.
- Section 0 provides guidelines for federating private resources from project partners.
- Section 0 gives conclusions and outlines the plans for future.

2 BACKGROUND AND RELATED WORK

2.1 Standards and common APIs

Problem of selecting appropriate cloud providers is not trivial due to a high number of potential providers on the market and the lack of standardised evaluation criteria for cloud services. E.g. there are many examples of providers that claim to offer 'cloud services', while in fact they offer only simple hosting service that does not provide such features as true elasticity and pay-per-use that are essential e.g. in the NIST definition [5]. The EU also observes the problem of the lack of standards preventing from creating truly open cloud service market in Europe and this problem will be addressed by the initiatives such as European Cloud Computing Strategy within the Digital Agenda,[1] which highlights the needs for common standards and practices for procurement of cloud resources by public organisations from commercial providers. Before these standards are established, we have to rely on available research material from industry and research organisations.

There are examples of efforts on the technical level to provide common APIs for accessing public cloud providers. The most relevant include:

- DMTF Cloud Infrastructure Management Interface (CIMI) http://dmtf.org/standards/cloud partially implemented by Delta Cloud;
- Open Cloud Computing Interface (OCCI) by OGF http://occi-wg.org/;
- OpenStack API http://api.openstack.org/ supported by OpenStack consortium.

The standard APIs for managing clouds include:

- jclouds library http://www.jclouds.org/ developed in Java that supports a wide range of public cloud providers;
- Fog library http://fog.io/ developed in Ruby that is used e.g. in Chef infrastructure management framework;
- Apache delta cloud http://deltacloud.apache.org/ that provides a REST interface.

The publicly available reports and benchmarks of IaaS clouds include:

- Gartner Magic Quadrant [6] of IaaS cloud providers;
- Cloud Harmony Benchmarks: http://cloudharmony.com/benchmarks;
- Cloud Sleuth benchmarks: https://cloudsleuth.net/.

[1] http://ec.europa.eu/information_society/activities/cloudcomputing/index_en.htm

3 APPLICATION CLASSIFICATION AND REQUIREMENTS ESTIMATE

VPH-Share and more general the VPH community uses various classes of applications that have different requirements regarding VM types offered by clouds. Here we provide a possibly complete list of application classes with examples from VPH and corresponding requirements.

3.1 Classification of applications

3.1.1 Web Application/Services

This class of applications includes Web portals for human users and SOAP/REST Web services for programmatic API access. These applications typically need at least one instance of VM running to provide required response time and are mainly characterised by variable load of requests. The application can be scaled by using a more powerful instance (CPU, memory) or by adding more instances using a load balancer. Stateless services are easier to maintain and scale since they do not require maintaining a session with a client. Most cloud providers are well suited to support this class of applications. Examples in VPH-Share include ViroLab Drug Ranking System.

3.1.2 Low latency Applications

This class represents applications that typically have a rich graphical user interface (GUI) and require visualisation and interaction. When ported to the cloud their user interface is available to the client using VNC (Virtual Network Computing) protocol, which requires high network bandwidth and low latency for smooth interaction. Choosing a cloud provider that operates a datacentre in geographical proximity of the end user (e.g. hospital) may be crucial for this class of application. An example in VPH-Share is @neurIST workflow that uses GIMIAS visualisation framework.

3.1.3 High CPU Applications

This class includes many scientific applications that are CPU-bound and typical execution times of single jobs range from minutes to hours, which is longer than for typical Web service requests. These applications are mainly sequential or can use multicore machines, but a single job requires only a single node (VM). Such applications often consist of multiple jobs that are independent from each other (e.g. parameter sweep), so they can be processed using high-throughput computing systems consisting of a pool of VMs that can be scaled according to demand. Examples in VPH-Share are the Meshing and Segmentation tools in the euHeart and @neurIST workflow that requires high-CPU image processing.

3.1.4 High I/O Applications

This class includes applications that require disk I/O of high volume and frequency, such as processing of database queries or processing of large data and text files (e.g. genomic databases). Such applications may perform poorly on standard VMs in which I/O overheads are high due to virtualisation and usage of SAN (Storage Access network) for attaching storage. Dedicated high I/O VM instance types may be required for this class of applications.

Examples in VPH-Share are the PatientDB application in the ViroLab workflow and other similar database applications that require at least very high random read speeds.

3.1.5 High Memory Applications

Some applications may require high memory (RAM) for processing large and complex queries, statistical analysis or image processing. VM instances with high memory may be required for databases with in-memory caching, etc. Lookup table applications like a Patient master Index in WP3 and large un-partitionable network graph visualisation applications from ViroLab are example in this category.

3.1.6 Cluster Applications

These are typical High Performance Computing applications using MPI. Examples are CFD or molecular dynamics. These applications in VPH-Share are mainly supported through the AHE and run on dedicated HPC resources, but it is also possible to run them (in smaller scale) on clouds. E.g. Amazon EC2 provides compute cluster instances that can be used to create virtual clusters on demand with VMs co-located in a single placement group for better latency between nodes. Examples in VPH-Share are the parallel CFD applications in euHeart, VPHOP and @neurIST and the bi-domain electrophysiology equation solvers in euHeart.

3.1.7 Workflow applications

Workflow applications consist of multiple interdependent tasks connected by data flow or control flow dependencies. In VPH-Share they are represented as Taverna workflows consisting of multiple applications (atomic services). Computational requirements of individual tasks often vary, so it may be necessary to provide different VM instance types for these tasks. @neurIST, euHeart and VPHOP are all creating a large number of Taverna workflows suitable for such a classification. There may also be applications that are atomic but are tied to each other, for e.g. a front-end web application, a middleware message queue and several worker nodes.

3.1.8 MapReduce (Hadoop) Applications

Data-intensive application operating on large data sets can be efficiently processed using MapReduce model. Some cloud providers offer dedicated solutions for running MapReduce jobs using Apache Hadoop. So far we have not identified such applications in VPH-Share, but they may be of interest for general VPH community, e.g. for processing large-scale genomic data.

3.1.9 GPGPU Applications

Some applications including various molecular dynamics codes can achieve significant speedups when running on GPGPU machines. Some cloud providers offer access to VM instances with dedicated GPGPU for computing. We have not identified such applications in VPH-Share; however, there are groups within our flagship workflows and the larger VPH community that are building such competencies.

3.2 Compute and storage resource requirements

Based on initial experience with application workflows gathered in years 1 and 2 of the project, we summarised the compute and storage requirements of each group of VPH-Share application workflows in Table 1. *On Demand* usage means that the application services are launched only during an interactive workflow session and shut down afterwards. The usage is estimated based on a 2 year estimate. *Always available* services are planned to be running 24x7 e.g. as Web servers and assume 720 hours a month. *Single study run* means a short concentrated run, e.g. sensitivity analysis. *Data egress* is amount of data (in GB) transferred from the cloud to the user, while *data ingress* (transfer from the user to the cloud) is assumed to be free. We include also the requirements of the infrastructure (WP2) that will be used for development, testing and maintenance of the infrastructure services.

The data in Table 1 give total of 177,408 compute hours, 13,640 GB-months of storage and data transfer of 5,232 GBs over 2 years. These estimates do not include external alpha and beta users that are planned to be included to the project during years 3 and 4 or collaboration with p-medicine project.

Comparing this data with the estimates of CPU hours and GB-months of storage that can be provisioned using the budget allocated in VPH-Share (see Section 7.3), we can see that these requirements can be satisfied with a large safe margin, and there should be enough resources to plan more large-scale runs during years 3 and 4 of the project.

Table 1 Summary of compute and storage requirements of VPH-Share application workflows for year 3 and 4 of the project. The estimates are based on preliminary experience gained in first two years of the project.

User (User Application)	Compute (in CPU-Hours / month)	Storage (in GB / month)	Data Egress (in GB / month)	Type of Access	Additional Requirements Comments
Infrastructure					
Provider Evaluation	60		20	On Demand	
VM Deployment Testing	60	50	20	On Demand	
Cloud Optimisation Testing	60			On Demand	
Data Storage Tests	60	50	20	On Demand	
Atomic Service Storage		50		On Demand	
Maintainence & Rolling Tests	60			On Demand	
WP5 Workflow - @neurIST					
Morphological Workflow (20/month)	10	1	1	On Demand	Windows VM
Heamodynamic Workflow (20/month)	500	100	20	On Demand	Windows VM, Cluster Compute
Structural Workflow	100	50	20	On Demand	Windows VM, Cluster Compute
Ancillary Tools	10		5	On Demand	
CRIM Database		10	5		
Medical Images, Meshes		100	10		
WP5 Workflow - euHeart					
Ancillary Tools (Heartgen, VTK2Ex)	10		2	On Demand	
Heart Mechanics Workflow	10	10	5	On Demand	
Parameter Estimation Analysis	5	1	1	Single Study Run (100 hrs)	Cluster Compute
Uncertainty Tools	5	1	1	Single Study Run (100 hrs)	
Cardiac Mesh Data		25	5		
WP5 Workflow - VPHOP					
Femur Workflows (10/month)	5000	10	20	On Demand	Cluster Compute
Spine Workflows (10/month	500	30	20	On Demand	Cluster Compute
Ancillary Tools	100		5	On Demand	
Orthopaedic Datasets		100	10		
WP5 Workflow - Virolab					
WebDRS	720	2	10	Always Available	
Literature Miner	10		10	On Demand	
Abstract Miner	2		1		
Literature Mining Tool - Bootstrap	100	10	1	Single Run (2500 hrs Bootstrap)	
Literature Mining Tool - Update	10		1	Always Available	
Rule Base		10	5		

4 Evaluation criteria of public cloud providers

Since there are many services offered under a common name of 'cloud' we have to define the criteria that will enable to select the most appropriate cloud providers and eliminate the ones that although advertised as 'cloud' in fact do not offer the required functionality. First we define the criteria for a broad survey of commercial cloud providers (Level 1), and then we go into more detailed hands-on evaluation of top provides based on the results from the survey.

4.1 Levels and importance of the criteria

The evaluation criteria are broken down into two levels:

- Level 1 Broad Survey with criteria:
  - API access (jclouds)
  - European Economic Area (EEA) Zoning
  - Published Price
  - On-Demand (Hourly) billing
  - BLOB Storage support
  - Relational Database storage
  - VM Import/Export support,
- Level 2 Detailed Evaluation of selected clouds from Level 1, using such criteria as:
  - Application Benchmark results,
  - Results of functionality tests (e.g. tests of jclouds API).

Below we give the justification of these criteria.

4.2 Justification of the criteria

The main criteria for the evaluation come from the technical assumptions under which the Atmosphere cloud platform is developed. The crucial assumption is that we use truly elastic cloud platforms, which allow dynamic resource provisioning (on demand instance creation with hourly billing). As CYFRONET and other project partners stand to provide the substantial part of resources, the public providers will be used in cloud-burst scenarios to provision additional capacity in peak demands. To prevent vendor lock-in we intend to select 2-3 alternative providers to be able to dynamically switch between them in case of failure or for proximity reasons. The detailed list of criteria is given in the Table 2 below.

We consider two criteria (EEA zoning and jclouds API support) as essential criteria for our choice, so we automatically give score of 0 to all providers that do not meet these criteria.

Table 2 Public cloud provider evaluation criteria with weights and their justification.

Criteria with weight	Justification
EEA Zoning, weight 20	Due to the fact that the VPH-Share is a European project, we prefer the cloud providers that offer their infrastructure through datacentres located in Europe. From technical point of view it means lower network latency and higher throughput, which has significant impact on application performance. From policy perspective this gives better control over billing (invoices) and most importantly it fulfils the requirement that the sensitive data do not cross the EU border [7]. In case of any dispute, the providers are also subject to European jurisdiction. This criterion is a project policy requirement and highly important, and as such it is given the highest weight.
jclouds API support, weight 20	In addition to above, the Atmosphere cloud platform developed in VHP-Share uses Java programming language and is based on Apache Karaf OSGi container for required modularity and extensibility (see D2.1-D2.4 documents for details on the Atmosphere design and implementation [8] [9] [10] [11]). As such, the Atmosphere relies on jclouds open source library to interface cloud providers. jclouds supports most open source cloud stacks, including OpenStack that the project uses for private clouds, and support for commercial cloud providers is systematically added. Choosing a cloud provider that is supported by jclouds out of the box considerably decreases the development effort needed to integrate this cloud. Adding support for a new provider not supported by jclouds can be estimated as of several weeks of developer time, which we consider not justified for the project. As this criterion is a technology requirement and highly important, it is given the highest weight.
BLOB storage support, weight 10	As WP2 develops both data and compute cloud platform, with a special focus on access to large binary objects, we prefer cloud providers that offer object storage in addition to compute services. Although it would be possible to use one provider for compute and another for storage resources, the advantage of choosing one that provides both services comes from the possibility of using data locality for efficient processing. By spawning the processing VMs close to the data (within the same provider) it is also possible to avoid data transfer costs, which can be significant budget item.
API Access, weight 5	VPH-Share is building a cloud platform on top of IaaS clouds with the assumption that the underlying providers offer truly elastic infrastructure. This means that the provider must offer an API for programmatic management of VMs. Some of cloud providers offer only portal-based access, which is not suitable for such automation as is required by VPH-Share so such providers should be rejected.
Per hour instance billing, weight 5	To fulfil the requirement of elasticity, i.e. the possibility of adding and removing resources on demand, we highly prefer such providers that charge the deployed VMs in hourly (or shorter) intervals. In this way we reject such providers that require minimum monthly commitments, as they are not suitable for VPH-Share applications.

Criteria with weight	Justification
Published price, weight 5	As the cloud market is becoming larger, we as customers prefer such providers that announce their pricing openly, in the form of price lists of GB storage or VM instance hours. This is also justified by the assumption that we are interested in building a flexible cloud platform on top of elastic IaaS providers, and not an enterprise solution requiring dedicated business agreements or contracts with a specific provider.
VM image import, weight 3	To provide effective cloud federation and to facilitate migration of VMs between providers, we prefer these ones that support importing of existing VM disk images to their clouds. The VMs (atomic services) developed for VPH-Share on our private cloud can be then easier migrated to the external cloud provider. As an alternative solution it is possible to automate the image building process using such tools as Chef or Puppet, we consider such requirement as not of the highest priority.
Relational DB support, weight 2	Some cloud providers offer relational databases (SQL-based) as a service or via dedicated VM types. As in VPH-Share we assume that most of relational data will reside in the in-premise databases (e.g. in hospitals), this criterion is not of high importance.

5 EVALUATION OF COMMERCIAL AND ACADEMIC CLOUD PROVIDERS

5.1 Commercial cloud providers

Based on the survey of publicly offered information on cloud providers we produced a ranking list of best cloud providers. Table 2 shows the results of evaluation based on criteria and weight.

Table 3 Results of survey of commercial IaaS cloud providers based on the evaluation criteria and weight.

	IaaS Provider	EEA Zoning	jclouds API Support	BLOB storage support	Per-hour instance billing	API Access	Published price	VM Image Import / Export	Relational DB support	Score
	Weight	20	20	10	5	5	5	3	2	
1	Amazon AWS	1	1	1	1	1	1	0	1	27
2	Rackspace	1	1	1	1	1	1	0	1	27
3	SoftLayer	1	1	1	1	1	1	0	0	25
4	CloudSigma	1	1	0	1	1	1	1	0	18
5	ElasticHosts	1	1	0	1	1	1	1	0	18
6	Serverlove	1	1	0	1	1	1	1	0	18
7	GoGrid	1	1	0	1	1	1	0	0	15
8	Terremark ecloud	1	1	0	1	1	0	1	0	13
9	RimuHosting	1	1	0	0	1	1	0	1	12
10	Stratogen	1	1	0	0	1	0	1	0	8
11	Bluelock	1	1	0	0	1	0	0	0	5
12	Fujitsu GCP	1	1	0	0	1	0	0	0	5
13	BitRefinery	0	0	0	0	0	1	0	1	0
14	BrightBox	1	0	0	1	1	1	1	0	0
15	BT Global Services	1	0	0	0	1	0	1	0	0
16	Carpathia Hosting	1	0	0	0	0	0	1	0	0
17	City Cloud	1	0	0	1	1	1	0	0	0
18	Claris Networks	0	0	0	1	0	0	0	0	0
19	Codero	0	0	0	1	1	1	0	0	0
20	CSC	1	0	0	0	0	0	1	0	0
21	Datapipe	1	0	0	1	1	0	0	0	0
22	e24cloud	1	0	0	1	0	1	0	0	0
23	eApps	0	0	0	0	0	1	0	0	0
24	FlexiScale	1	0	0	1	1	1	1	0	0
25	Google GCE	1	0	1	1	1	1	0	1	0
26	Green House Data	0	0	0	0	1	0	1	0	0
27	Hosting.com	0	0	0	0	0	1	1	1	0
28	HP Cloud	0	1	1	1	1	1	1	1	0
29	IBM SmartCloud	0	0	1	1	1	1	0	1	0
30	IIJ GIO	0	0	0	0	0	0	0	0	0
31	iland cloud	1	0	0	1	0	1	1	0	0
32	Internap	0	0	1	1	1	1	0	0	0
33	Joyent	0	0	0	1	1	1	0	0	0
34	LunaCloud	1	0	1	1	1	1	0	0	0
35	Oktawave	1	0	1	1	1	1	0	1	0
36	Openhosting.co.uk	1	0	0	0	0	1	0	0	0
37	Openhosting.com	0	1	0	1	1	1	1	0	0
38	OpSource	1	0	1	1	1	1	1	0	0
39	ProfitBricks	1	0	0	1	1	1	0	0	0
40	Qube	1	0	0	0	0	1	0	0	0
41	ReliaCloud	0	0	0	0	0	0	0	0	0
42	SaavisDirect	0	0	1	1	0	1	0	0	0
43	SkaliCloud	0	1	0	1	1	1	1	0	0
44	Teklinks	0	0	0	0	0	0	0	0	0
45	Terremark vcloud	0	1	0	1	1	1	1	0	0
46	Tier 3	0	0	0	0	1	0	0	0	0
47	Umbee	1	0	0	1	1	1	1	0	0
48	VPS.net	1	0	0	0	1	1	0	0	0
49	Windows Azure	1	0	1	1	1	1	0	1	0

5.2 Academic cloud providers

Academic cloud providers do not easily fit into the criteria for commercial cloud providers. Most of them are experimental test-beds that are under development, and as such, cannot be used for production usage in VPH-Share; however, some of them can be used for experiments with selected applications. Below, we briefly summarise the analysed academic cloud installations.

- EGI Federated Cloud Task Force (https://wiki.egi.eu/wiki/Fedcloud-tf) prepares a test-bed and a blueprint for sharing virtualised resources under the umbrella of European Grid Initiative. The test-bed comprises heterogeneous clusters based mainly on OpenStack and OpenNebula. CYFRONET participates in this initiative and operates a test installation, which is a part of PL-Grid project. This installation can be used as alternative to VPH-Share specific OpenStack installation at CYFRONET, if such demand arises. We plan to continue the collaboration with this initiative of EGI and investigate whether this infrastructure will be of interest to VPH-Share community.
- Eduserv (http://www.eduserv.org.uk) is a non-profit SME providing cloud services for public sector. Education cloud (http://www.slideshare.net/andypowe11/eduserv-education-cloud) is currently more focused on enterprise type of applications and based on VMware vSphere stack, but support for OpenStack is planned. Eduserv also provides storage using WebDAV and SFTP protocols. When OpenStack type of compute service is available, it may be of interest to research community of VPH-Share.
- Open Cloud Consortium (http://opencloudconsortium.org/) is a US-based organisation that operates cloud test-bed for research institutions that contribute their hardware, and Open Science Data Cloud (OSDC) for data intensive applications. OSDC is oriented towards Map-Reduce applications, which currently are not of high priority for VPH-Share project.
- FutureGrid (https://portal.futuregrid.org/) is a US-located cloud-based test-bed for distributed computing experiments and middleware development. The resources are provided using Eucalyptus, Nimbus and OpenStack cloud stacks. Researchers from around the world can apply for access to these resources. Since FutureGrid provides open APIs, it will be possible to use these resources for experiments with the Atmosphere platform and selected applications from VPH, if such need arises.
- Local cloud test-beds. There are numerous cloud test-beds operated by computer centres, such e.g. SARA cloud (https://www.cloud.sara.nl/) or MetaCentrum HPC cloud (http://www.metacentrum.cz/en/devel/cloud/index.html). Access to these resources is usually limited to local or national users, but it is possible to integrate them with the Atmosphere cloud platform of VPH-Share if such demand from the users arises and if there are public APIs provided for interacting with these resources.

Currently, there is no particular public cloud available for research community that would be of particular interest to VPH-Share project. However, we expect that when the aforementioned projects mature and their API interfaces and usage policies stabilise, it will be possible to connect them to the federated cloud managed by Atmosphere in VPH-Share.

6 Conclusions from commercial provider survey

Results gathered in Table 3 show that there indeed a lot of potential cloud providers, totalling 49, however, not all of them meet the required criteria of VPH-Share applications. The breakdown of providers fulfilling the given criteria is following:

- Majority (31) of providers have datacentres in Europe.
- Only 16 of providers are supported by jclouds library, so integrating other providers with the Atmosphere framework would require more effort.
- Only 12 providers offer object storage service, many are only compute service providers.
- Majority of providers offer API support, hourly billing and publish their prices, which are crucial for the requirement of elasticity. Other clouds are either enterprise oriented and are seeking customers for longer commitment via negotiable contracts, or are simply VPS hosting providers, advertising their service as a 'cloud'.
- Image import and relational DB support are not essential features, but they may be useful when choosing a provider that has most complete service.

We identified 3 leading providers that fulfil 3 most important criteria: EU location, jclouds support and BLOB storage service:

- Amazon AWS is the leading cloud service provider, with datacentre in Ireland. EC2 API becomes almost a standard for compute clouds. In addition to our evaluation criteria, there is a wide support for tools and documentation, making the integration with VPH-Share an easier task.
- RackSpace is also a leading cloud provider (datacentre in London, UK), which has the advantage of being involved in development of OpenStack software, used in private clouds of VPH-Share.
- SoftLayer operates in Amsterdam datacentre and its advantage is the BLOB storage service. However, as shown after a more detailed examination in section 7, the jclouds support is limited, so integration of SoftLayer with VPH-Share would require more effort or time.

There are also 3 additional providers that do not offer BLOB storage capabilities, but they may be of interest for compute-only services. They are:

- CloudSigma with a datacentre in Zurich, which offers 5-minute billing cycle. Object storage support is planned in Q2 2013. Image upload feature is also relevant.
- ElasticHosts with a datacentre in UK.
- ServerLove with a datacentre in UK.

These providers should be considered if there is a need for running compute-intensive applications that are not dependent on access to local data, or when the location of specific datacentre provides exceptionally low latency to a particular user.

7 API AND PERFORMANCE TESTS

It would be impossible to make a decision on choosing a cloud provider based only on publicly available information and with no real experience with their service. For that reason, we signed up for the service of these providers and evaluated in practice how jclouds support works with them, and how is the estimated performance of the VMs they offer.

We chose top 6 providers for jclouds runtime tests and compared the performance of 2 top ones: Amazon EC2 and RackSpace.

7.1 jclouds support runtime tests

In order to verify the jclouds support, we conducted simple tests with selected 6 top providers from the ranking list. The purpose of the test was to check if it is possible to create a VM instance from a custom image template (snapshot). The results are following:

- Amazon EC2 is well supported and documented and we have found no problems with using jclouds with Amazon. jclouds allows also using EC2-specific features, e.g. launching spot instances at discounted price, which may be useful for high throughput computing.
- RackSpace support – RackSpace compute jclouds provider uses standard OpenStack Compute (Nova) API. The API is well documented and supported, however we encountered some timeouts when using the new OpenStack API to create servers.
- SoftLayer is supported in jclouds, but we encountered issues with timeouts with certain API calls (e.g. listing images). Moreover, as of version 1.5.3, launching instances from custom image templates in not supported. We can expect this support will be added in the future release of jclouds, but no timeline is available.
- CloudSigma is supported by jclouds, however we encountered issues with timeouts when launching instances. Investigating it in more detail would require more effort.
- ElasticHosts is well supported by jclouds using provider specific API client. We observed no problems with launching an instance from custom template and listing instance parameters.
- ServerLove uses the same API as ElasticHosts (Elastic Stack) and the integration with jclouds library works with no issues.

The conclusion from this hands-on experience is that Amazon EC2 and RackSpace are the only providers that can be integrated with VPH-Share with no major issues. ElasticHosts and ServerLove also seem to be good candidates for compute service, but they do not support BLOB storage. CloudSigma is an interesting provider, but we can expect some issues with jclouds API and BLOB storage is not yet available.

7.2 Performance tests

Cloud providers often offer multiple VM instance types with various performance and price. Choosing the right one for particular application is not straightforward and requires running application specific benchmarks to estimate the performance and costs. Here we present the preliminary results of benchmarks that we run when preparing this evaluation. It should be noted that more tests are planned throughout the project, as we proceed with integrating more applications and get better understanding of their requirements.

From the list of applications described in section 3 we selected high CPU applications, as their performance may be most affected by the VM performance.

7.2.1 High CPU Applications on Amazon EC2, RackSpace and SoftLayer

As a test application we selected the segmentation tool in @neurIST workflow, which was already deployed as Atomic Service on CYFRONET cloud. The goals of the test were: (1) to find the most efficient cloud/instance type (2) to find the most cost-efficient cloud/instance type in terms of price/performance. In the current @neurIST workflow the segmentation tool runs ~130 seconds on a VM at CYFRONET.

Table 4 Tested instance types on EC2, RackSpace and SoftLayer

Instance type	hourly price in $	number of cores	RAM in GB	provider
m1.small	0.065	1	1.7	EC2
m1.medium	0.13	1	3.75	EC2
m1.large	0.26	2	7.5	EC2
m1.xlarge	0.52	4	15	EC2
m2.xlarge	0.46	2	17.1	EC2
m2.2xlarge	0.92	4	34.2	EC2
m2.4xlarge	1.84	8	68.4	EC2
c1.medium	0.165	2	1.7	EC2
c1.xlarge	0.66	8	7	EC2
hi1.4xlarge	3.41	8	60.5	EC2
cc2.8xlarge	2.7	16	60.5	EC2
cg1.4xlarge	2.36	8	22	EC2
m3.xlarge	0.55	4	15	EC2
m3.2xlarge	1.1	8	30	EC2
rs-0.5GB	0.032	1	0.5	RackSpace
rs-1GB	0.064	1	1	RackSpace
rs-2GB	0.128	2	2	RackSpace
rs-4GB	0.256	2	4	RackSpace
rs-8GB	0.512	4	8	RackSpace
rs-15GB	0.96	4	15	RackSpace
rs-30GB	1.6	8	30	RackSpace
sl-1c	0.12	1	1	SoftLayer
sl-2c	0.2	2	2	SoftLayer
sl-4c	0.3	4	4	SoftLayer
sl-8c	0.45	8	8	SoftLayer
sl-16c	1.15	16	16	SoftLayer

We created a VM image on EC2 UE Ireland region using Ubuntu 12.04 64-bit and copied the GIMIAS-1.5-VPHShare installation from image deployed at CYFRONET cloud, together with sample data. Additional image was created for cluster and GPU instances that use HVM virtualisation. We applied the same procedure on RackSpace London UK region and SoftLayer Amsterdam region. We used all instance types available on Amazon EC2 and RackSpace, while in the case of SoftLayer we created instances of 1, 2, 4, 8 and 16 cores with 1GB of RAM per core and the smallest 25GB disk. It should be noted that SoftLayer allows creating custom instances with number of cores ranging from 1 to 16 and 1 to 32 GB of memory, which gives more flexibility and more fine-grained resource control. The summary of instance types is given in Table 4.

The segmentation application is compute-bound and it consumes nearly 100% of single core during the whole run time. For the sample dataset the computing was in the order of 2 minutes on most of the instance types. All instances, except m1.small, are multicore, having from 2 to 16 virtual cores. On each instance type we run single core tests, where only one segmentation process was running; and multi-core tests where we run 2, 4, 8 or 16 processes in parallel, with the maximum number of processes equal to the number of virtual cores of the instance. The single core run represents a scenario where the user needs to run a single computing job and is interested to get the result quickly and with lowest cost. The multicore scenario is useful either if there are multiple users working in parallel or if there is a larger batch of jobs to process: in that case users are interested to find the most cost effective instance type, since there is a possibility to launch multiple instances to further increase the throughput.

We ran the tests by repeatedly launching new instances of each type and running the single and multicore tests on each of them, and after the tests the instances were shut down. We repeated the tests about 10 times at different times of day and week between Dec 2012 and Feb 2013. The figures show the averaged results from these runs. We also observed a variability of performance within the same instance type: the standard deviation of the computing time was less than 10%.

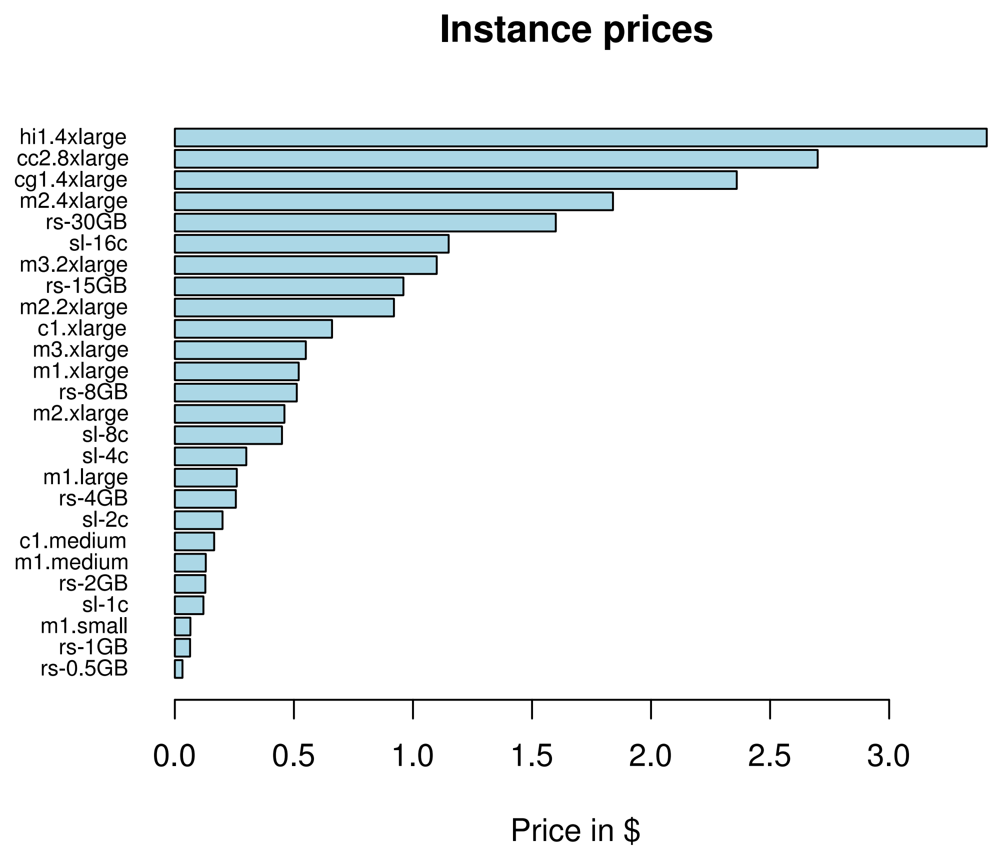

Figure 2 Instance prices in $ per hour on Amazon EC2, RackSpace (rs-*) and SoftLayer (sl-*).

Figure 2 shows the price of on-demand instance per hour of runtime. The prices are of Feb 2013 for EU regions in USD. For individual jobs it is most economical to launch small or medium instances, since additional cores are not used in this case. RackSpace instances are the cheapest, but the memory may be the limit. The smallest SoftLayer instance is 1 core of 2GHz and 1GB RAM, making it more expensive at $0.12 per hour, comparing to $0.03 of the cheapest RackSpace instance.

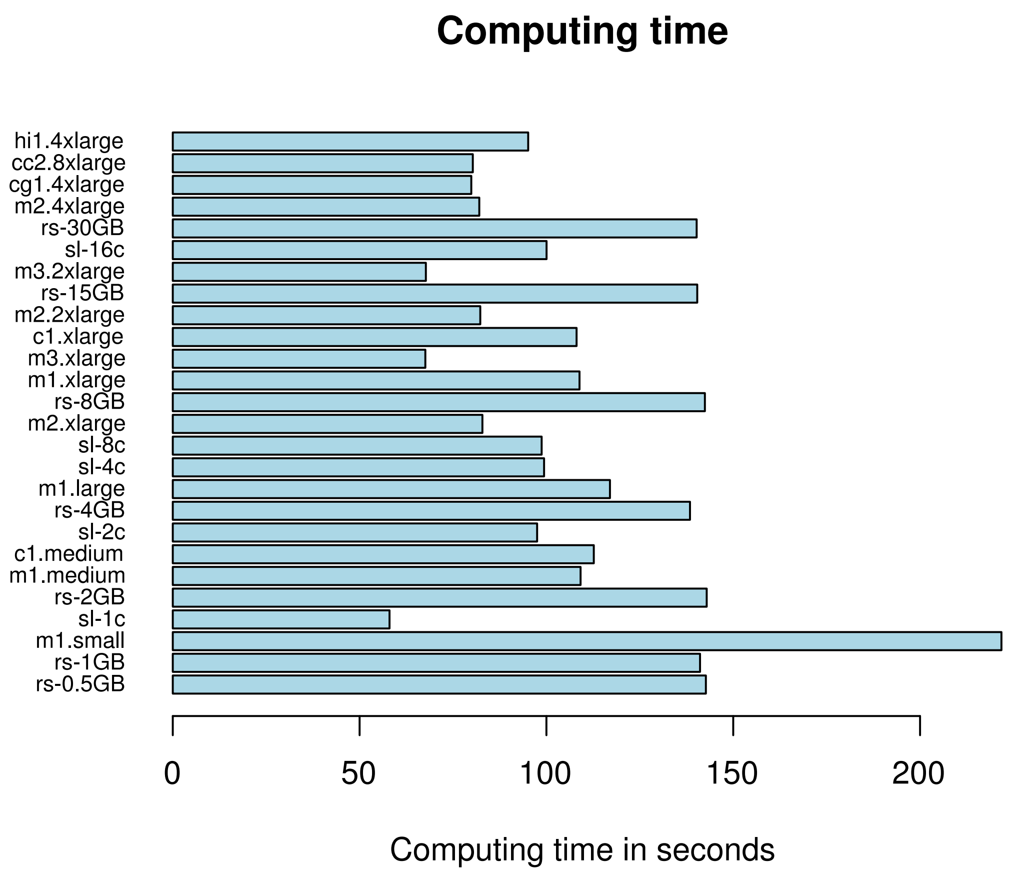

Figure 3 Single core computing time on Amazon EC2, RackSpace and SoftLayer. Plot shows execution time of a single CPU-intensive process.

In Figure 3, we observe that the smallest SoftLayer instance gives the best single-threaded performance. In the case of EC2 two classes of instance type c1 and m1 instances give the computing time over 100 seconds and more powerful m2, cc2, cg1 and hi1 types are about 20% faster. The exception is m1.small instance that is 2 times slower due to CPU performance cap imposed by hypervisor. Second generation instances m3.xlarge and m3.2xlarge are the fastest ones, probably due to new hardware introduced in 2013. RackSpace instances are slower, and, interestingly, their single core performance does not depend on the instance type.

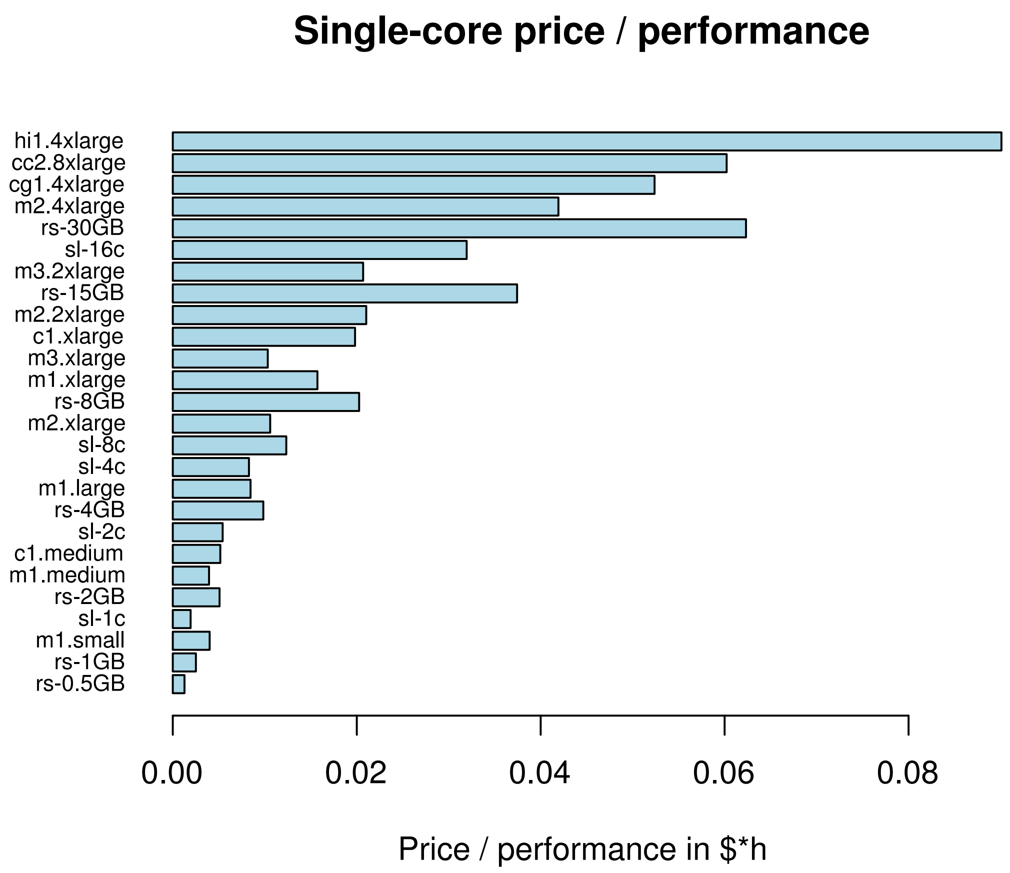

Figure 4 Price to performance ratio for single core usage is computed by dividing the hourly price by performance measured as inverse of computing time.

In Figure 4 the performance measures the throughput in jobs per hour. It can be seen that the cheapest instances of RackSpace and SoftLayer are most cost-efficient. On EC2 when performance is critical, then m3.xlarge instance is the most cost effective option, as it is the cheapest of the instances with the high performance. We can see that it does not make sense to use more powerful instance types, since the sequential application process cannot use all their cores. RackSpace instances appear to be the most cost-efficient ones, if high memory is not essential. The smallest RackSpace instances seem to be most attractive for high throughput computing.

Another perspective on these data is shown in Figure 5, where it is possible to observe the trade-off between computing time and instance price.

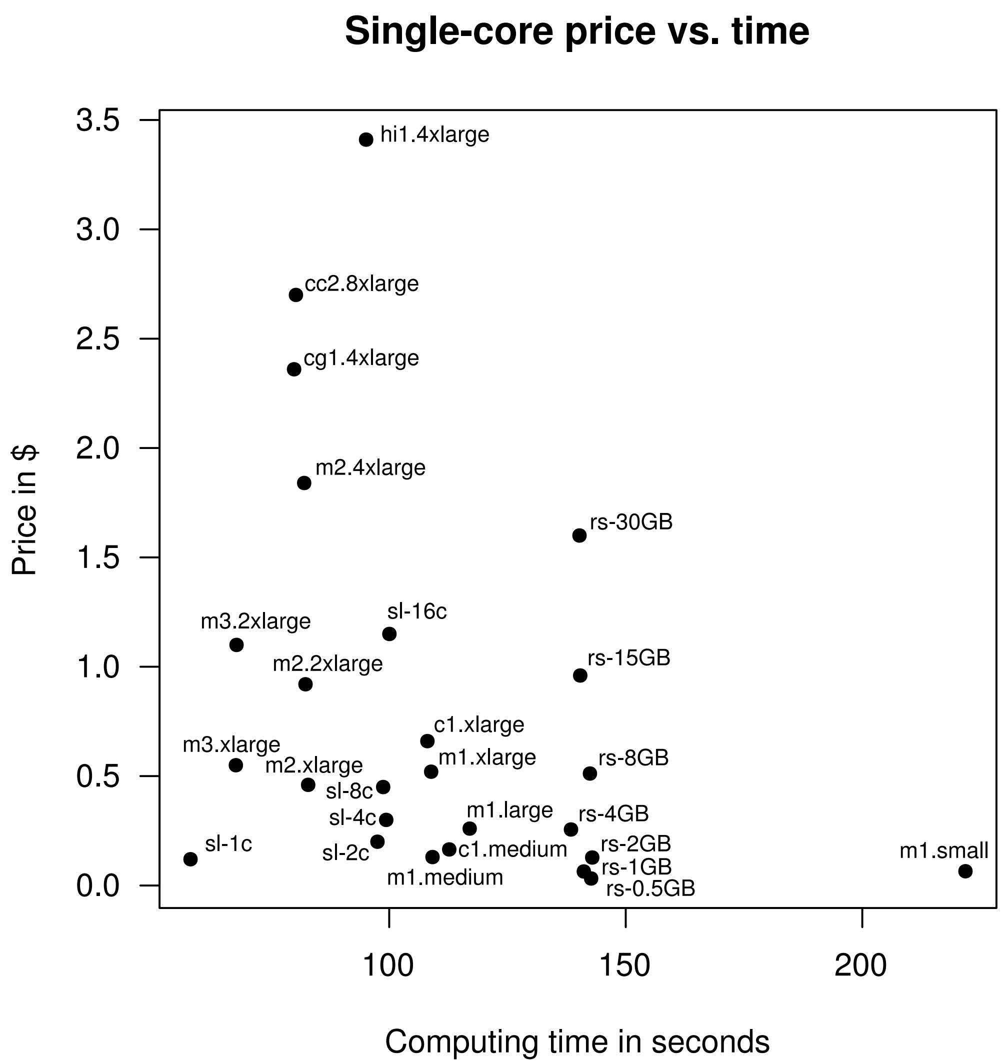

Figure 5 Price vs. computing time of EC2, RackSpace and SoftLayer instances. It can be observed that most RackSpace and SoftLayer instance types give similar single-core performance, while results of EC2 are spread more widely.

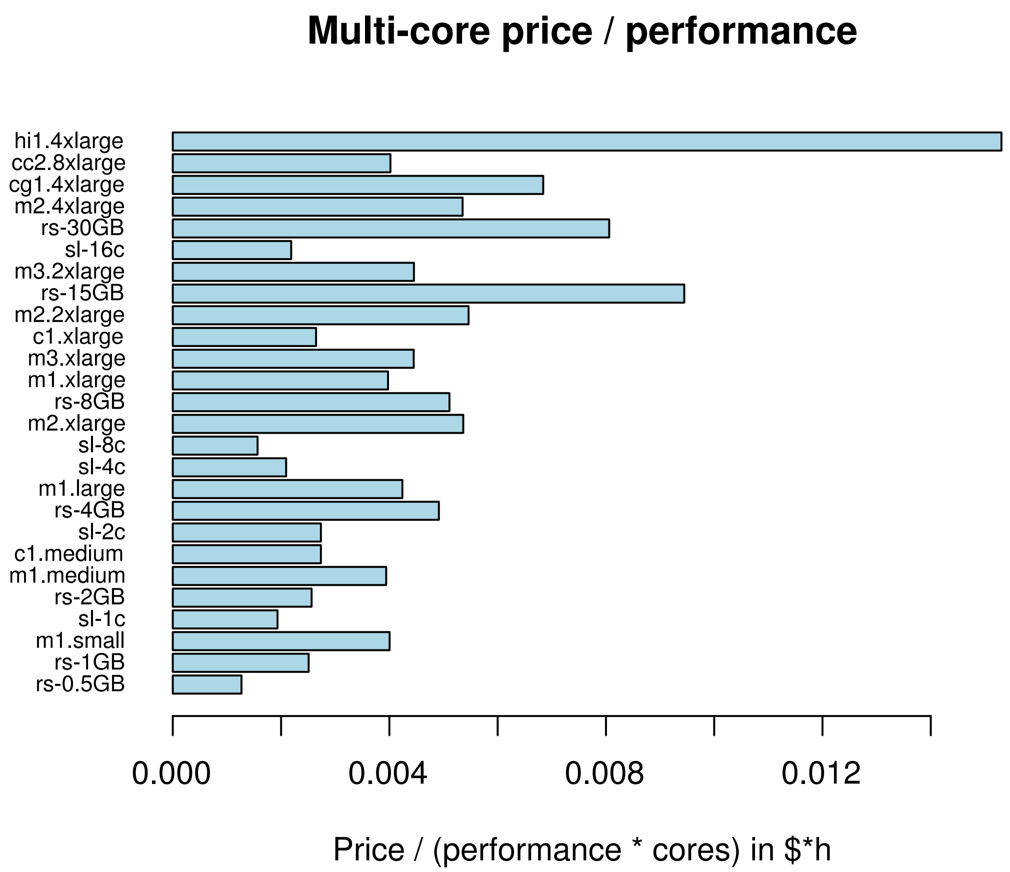

Figure 6 Multi-core price performance ratio shows the instance price divided by the throughput in jobs per hour measured when the number of parallel jobs was equal to the number of cores of the virtual machine.

In Figure 6, when there are multiple CPU-intensive jobs running, it is most economical to use ‘high-CPU’ instance types: sl-1core, c1.medium, c1.xlarge and rs-2GB. SoftLayer sl-8core instance type is the most cost efficient overall thanks to the good performance for 8 cores and reasonable price with 8BG of RAM and 25GB of disk. c1.medium seems to be attractive option, since it has only 2 cores, so it is better suited for auto-scaling. Cluster instances cc2.8xlarge are overall also relatively cost effective, but their 16 cores are better suited for larger workloads. Second-generation EC2 instances m3.large and m3.xlarge are not as efficient for multicore jobs, as we observed that speedup using their virtual cores is not linear.

7.3 Conclusions from performance tests

After evaluating the performance and costs of two leading cloud providers, we can observe that the number of options that these providers offer is substantial and the decisions of selecting the appropriate cost-effective instance type is not trivial. However, the broad decision space can be narrowed by additional application specific criteria. Here we list a set of conclusions and hints.

- For pure compute-intensive applications, not requiring a lot of RAM, the cheapest instances from all the providers are a good choice.
- When single-core performance is critical, SoftLayer 1-core instances and second generation EC2 instances provide the fastest CPU speed.
- For applications requiring more RAM or disk the smallest instances are not sufficient, so the type of instances will be determined by the RAM and disk requirements of a specific application.
- Small instances are interesting for applications that can be scaled horizontally (by adding new instances), since they are more elastic, i.e. they enable finer granularity in scaling the infrastructure by a single core. SoftLayer 1-core instances are the good choice here, but the observed higher provisioning time (not less than 5 minutes) and API issues make using these instances less convenient.
- The performance measurements obtained here will be input to resource allocation and management policies of the Atmosphere platform developed in WP2.

Based on the results of our performance tests and price survey, we can estimate that the 70,000 EUR allocated cloud services in VPH-Share will be sufficient to provide:

- 1,000,000 CPU-hours of single-core virtual machine, or
- 57-core compute cluster running 24x7, or
- 29 TB of storage for 24 months, or
- Any combination of the above.

8 GUIDELINES FOR FEDERATING CLOUD RESOURCES FROM PROJECT PARTNERS

In VPH-Share cloud federation will be managed by the following components of the cloud platform developed in WP2 (see Figure 1):

- Atmosphere will federate compute resources from multiple private and public clouds,
- LOBCDER will federate storage from multiple object stores.

The private cloud of VPH-Share is based on OpenStack, as was decided based on evaluation described in Deliverable D2.1 [8].

8.1 Instructions for private cloud resource providers

Compute resource providers should install OpenStack Compute (Nova). Although other existing open source solutions such as OpenNebula or Eucalyptus may be also used if they provide compatible APIs and are properly managed by local administrators, WP2 will not help with installation and support; Basic requirements:

- Ubuntu http://docs.openstack.org/folsom/openstack-compute/install/apt/content/
- RHEL http://docs.openstack.org/folsom/openstack-compute/install/yum/content/

At CYFRONET we successfully installed **OpenStack Folsom** release and this is the current recommended version.

Compute resource providers → Install OpenStack Compute (Nova) (other existing open source solutions such as OpenNebula or Eucalyptus may be also used if they provide compatible APIs and are properly managed by local administrators - WP2 will not help with installation); Basic requirements:

- One Cloud Controller node (for Keystone, Glance and Nova services except nova-compute) - 64-bit CPU, large HDD (local or via SAN) for images (at Glance) and (optionally) nova-volumes service (Amazon EBS equivalent), single Gigabit Ethernet Card.
- At least one (for production use - multiple) Compute nodes (nova-compute) to run VMs - 64-bit CPU with hardware virtualization support (VT-x / AMD-V), large quantity of RAM (at least 2 GB per concurrent VM), sufficient HDD (local or via SAN) to hold running VMs, Gigabit Ethernet Card (two are recommended).
- (Recommended) 802.1Q (VLAN tagging) capable Ethernet switch - to use most powerful VLAN Network Manager allowing network isolation for each project.

Storage resource providers (large binary data) → Install OpenStack Object Storage (Swift); Basic requirements:

- Proxy node - handling incoming request - node optimised for CPU and high network usage, Gigabit Ethernet connectivity (recommended 10-Gbit), no HDD type/performance requirements.
- Object node (at least 3, recommended 5) - storing actual data - optimised for HDD price and quantity (regular SATA drives, RAID NOT recommended), Gigabit Ethernet.
- (Optional, if not present those services must be installed on Object node) - Container/Account nodes (at least 3, recommended 5) - HDD optimised for IOPS (due to have SQLite usage), Gigabit Ethernet.

8.2 Current status of cloud resources hosted by project partners

The summary of compute and storage resources in VPH-Share is given in Table 5. Currently 3 sites are installing OpenStack (CYFRONET, STH and UNIVIE). CYFRONET installation hosts the development and production cloud service for VPH-Share. Other resources in the table are either data servers for hosting databases or are HPC clusters that will be accessible via dedicated AHE middleware (developed within Task 2.3 High Performance Computing Infrastructure). These resources will not be part of OpenStack cloud, so no installation is required.

Table 5 Status of resources for federated cloud

Resource	Partner	In-kind / funded	Comment	OpenStack installation status	Administrator Contact
2 TFlops (Compute) + 15 Tb (Storage)	Cyfronet	In-Kind	Current location of the development cloud	Folsom release installed	Jan Meizner j.meizner@cyfronet.pl
Cluster of 4 nodes, 96 cores total (2xOpteron 6172 per node)	UNIVIE	In-Kind		Installation in progress	Yuriy Kaniovskyi yk@par.univie.ac.at
Local compute clusters	USFD	In-Kind		Installation in progress	Susheel Varma susheel.varma@sheffield.ac.uk
Upto 4 euHeart machines (after November 2012)	USFD	In-kind	Additional to the DoW commitment		Susheel Varma susheel.varma@sheffield.ac.uk
Local compute clusters	Philips	In-Kind	Status TBD		
Local data clusters	STH	In-Kind	For database services (WP3)		
240K (Dell PE1950/Opteron x2200)	KCL	In-Kind	For HPC applications		
33M (HECTor) CPU hours	KCL	In-Kind	Possibly accessible through the AHE		
Data Server	STH	Funded €22k		Installation in process	Richard Knight richard.knight@sheffield.ac.uk
Data Server	FCRB	Funded	€28k – bought in Period 1		
Mini-Cloud	Cyfronet	Funded	€30k - unspent	In progress	Jan Meizner j.meizner@cyfronet.pl
HPC recources PLX-IBM Cores: 10240 Rpeak= 293.17 Tflops	SCS	In-Kind	More information: http://www.hpc.cineca.it/hardware/ibm-plx	Status TBD	Debora Testi d.testi@scsitaly.com

9 Conclusions

In this report we analysed the requirements of VPH-Share applications with respect to compute and storage resources that need to be procured from public cloud providers.

After analysing the offer of nearly 50 public commercial cloud providers using criteria such as EU location, jclouds API support, BLOB (large binary object) storage service we conclude that there are three leading cloud providers, namely Amazon EC2, RackSpace and SoftLayer that fulfilled three most important criteria. There are also providers such as CloudSigma, ElasticHosts and Serverlove that fulfilled most criteria except BLOB storage.

The tests of jClouds API support of these six top cloud providers reveal that all of them have only minor compatibility problems, except SoftLayer that does not support custom image templates. We measured the performance of compute instances of Amazon EC2, RackSpace and SoftLayer and gathered data will be used to guide the dynamic resource allocation of the Atmosphere cloud platform.

Currently the project operates a private development cloud infrastructure based on OpenStack Folsom release hosted by CYFRONET and installations in Sheffield and Vienna are in progress.

According to the estimates based on cloud survey and price and performance analysis, the budget of EUR 70,000 allocated for public cloud providers within VPH-Share will be sufficient to buy a service of 1,000,000 single-core CPU hours, or operate a 57-core cluster running 24x7 for 2 years, or store 29TB of data for 2 years. The current estimates of application requirements are in the order of 150,000 CPU hours and 20 TB-months of storage. This leads to the conclusion that the application requirements are likely to be satisfied with a safe margin, and that there is enough budget for planning more large-scale experiments using the VPH-Share cloud platform. Since most VPH-Share applications require *on-demand* access to computing resources, we recommend that these resources will be purchased from public cloud providers in a pay-per-use model, based on the demand requested from the application users. To demonstrate the federation capabilities of VPH-Share cloud platform and to prevent the vendor lock-in problem we recommend selecting at least two independent public cloud providers.

We plan to continue the evaluation of cloud providers using such criteria as network latency and to conduct more performance tests of CPU-intensive applications, as more application services are integrated with the Atmosphere cloud platform and their resource demands and usage patterns are better understood throughout the years 3 and 4 of the project. We also plan to follow the dynamic market of cloud providers to provide the required elasticity of resources for VPH-Share workflows. The results described in this deliverable will provide a technical background for elaboration by the project management of a specification of resources to be purchased from public cloud providers.

10 References

1. Malawski M, Meizner J, Bubak M, Gepner P. Component approach to computational applications on clouds. Procedia Computer Science. 2011 May; 4: p. 432–441.

2. Malawski M, Gubała T, Bubak M. Component-based approach for programming and running scientific applications on grids and clouds. International Journal of High Performance Computing Applications. 2012 Aug; 36(3): p. 275–295.

3. Malawski M, Juve G, Deelman E, Nabrzyski J. Cost- and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds. In SC '12 Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis; 2012; Salt Lake City.

4. Malawski M, Kuźniar M, Wójcik P, Bubak M. How to Use Google App Engine for Free Computing. IEEE Internet Computing. 2013 Jan-Feb; 17(1).

5. National Institute of Standards and Technology. The NIST Definition of Cloud Computing.; 2011. Available from: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

6. Haynes L, Leong D, Toombs B, Gill G, Petri T. Magic Quadrant for Cloud Infrastructure as a Service. Gartner; 2012.

7. VPH-Share. Deliverable D7.3 Clinical data governance guidelines. ; 2012.

8. VPH-Share. Deliverable D2.1 Analysis of the state-of-the art, work package definition. CYFRONET; 2011.

9. VPH-Share. Deliverable D2.2 Design of the cloud platform. CYFRONET; 2011.

10. VPH-Share. Deliverable D2.3 First prototype of the cloud platform. CYFRONET; 2012.

11. VPH-Share. Deliverable D2.4 Second prototype of the cloud platform. CYFRONET; 2013.

List of Key Words/Abbreviations

AHE	Application Hosting Environment
API	Application Programming Interface
BLOB	Binary Large Object
CFD	Computational Fluid Dynamics
DB	Data Base
EC2	Amazon Elastic Computing Cloud
EEA	European Economic Area
GPGPU	General Purpose Graphics Processing Unit
HPC	High Performance Computing
IaaS	Infrastructure as a Service
LOBCDER	Large OBject Cloud Data storagE federation
MPI	Message Passing Interface
REST	REpresentational State Transfer
SaaS	Software as a Service
SOAP	Simple Object Access Protocol
VM	Virtual Machine
VNC	Virtual Network Computing
WebDAV	Web-based Distributed Authoring and Versioning
WP	Work Package

This page was intentionally left blank