Abstract: This document is scoped to identify and discuss relevant aspects that influence the future sustainability of e-Science cloud infrastructures. Starting with an in-depth market analysis, we then analyse the economic, social, environmental and technological sustainability of the seven VENUS-C scenarios. This cost-benefit assessment is meant to form the basis for the analysis of the future final sustainability in Deliverable 3.10.
**DISCLAIMER**

VENUS-C is a Research Infrastructure Project co-funded by the GÉANT & e-Infrastructure Unit of the Information Society & Media Directorate General of the European Commission.

VENUS-C targets scientific communities, decision makers from SMEs, scientific clusters, incubators and technology parks, as well as European Union policy makers. VENUS-C is supported and evaluated by an Expert International Advisory Committee (EIAC).

This document has been produced with the co-funding of the European Commission. The content of this publication is the sole responsibility of the VENUS-C consortium and cannot be considered to reflect the views of the European Commission.

Grant Agreement Number: 261565

Start Date: 1 June 2010

Duration: 24 Months

---

**Document Information Summary**

<table>
<thead>
<tr>
<th>Deliverable number:</th>
<th>D 3.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable title:</td>
<td>Future sustainability strategies – draft</td>
</tr>
<tr>
<td>Editor</td>
<td>Francesca Spagnoli, Giorgia Fiume/ENG</td>
</tr>
<tr>
<td>Contributing Authors:</td>
<td>Andrea Manieri/ENG, Fotis Karayiannis/MIC-GR</td>
</tr>
<tr>
<td>Reviewer(s):</td>
<td>Jacek Calo/NCL, Goetz Philip Brasche/EMIC</td>
</tr>
<tr>
<td>Work package no.:</td>
<td>WP3</td>
</tr>
<tr>
<td>Work package title:</td>
<td>Communication and Sustainability</td>
</tr>
<tr>
<td>Work package leader:</td>
<td>OGF</td>
</tr>
<tr>
<td>Work package participants:</td>
<td>ENG, EMIC, OGF, BSC, UPVLC, KTH, AEGEAN, COSBI, UNEW, CNR, COLB, MIC-GR</td>
</tr>
<tr>
<td>Distribution:</td>
<td>Public/Private</td>
</tr>
<tr>
<td>Nature:</td>
<td>Report</td>
</tr>
<tr>
<td>Main Target Audiences:</td>
<td>VENUS-C partners, research communities, VENUS-C IT developers, European policy makers, public procurement bodies</td>
</tr>
<tr>
<td>Version/Revision:</td>
<td>2.0</td>
</tr>
<tr>
<td>Draft/Final:</td>
<td>Final</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Sustainable scientific cloud, market analysis, cost models, TCO, outsourcing</td>
</tr>
</tbody>
</table>
Log Table

<table>
<thead>
<tr>
<th>Issue</th>
<th>Date</th>
<th>Description</th>
<th>Author/Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1.0</td>
<td>29/03/2011</td>
<td>Table of content, references, expected</td>
<td>Francesca Spagnoli/ENG, Andrea</td>
</tr>
<tr>
<td></td>
<td></td>
<td>contributions</td>
<td>Manieri/ENG</td>
</tr>
<tr>
<td>v1.1</td>
<td>18/05/2011</td>
<td>Pre-final</td>
<td>Francesca Spagnoli/ENG, Giorgia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fiume/ENG</td>
</tr>
<tr>
<td>V1.2</td>
<td>21/05/2011</td>
<td>Formal check of content</td>
<td>Andrea Manieri/ENG</td>
</tr>
<tr>
<td>v2.0</td>
<td>21/05/2011</td>
<td>Full draft for Internal Reviewers approval</td>
<td>Francesca Spagnoli/ENG, Andrea</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Manieri/ENG, Giorgia Fiume/ENG</td>
</tr>
<tr>
<td>v. final</td>
<td>2/06/2011</td>
<td>Refinement and incorporation of Reviewers</td>
<td>Francesca Spagnoli/ENG, Andrea</td>
</tr>
<tr>
<td></td>
<td></td>
<td>comments</td>
<td>Manieri/ENG, Giorgia Fiume/ENG</td>
</tr>
</tbody>
</table>

Document Approval

<table>
<thead>
<tr>
<th>Approval by</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Management Board</td>
<td>16/06/2011</td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

The interest in cloud computing has been rapidly growing since the last two years. Not only companies, but also institutions and research bodies start to see how useful it can be for their activities to use the different services offered today by cloud computing infrastructures.

Within this context, our study has been conducted to develop a clear representation of the cloud services’ potential for the scientific environment. The main objective of this document is to formulate a sustainable strategy for e-Science cloud-based infrastructures.

An initial analysis of recommendations and sustainability models of European e-Infrastructure cases (e-IRG Reflection Group, EGI-Inspire, PRACE, the KTH case and PDC case) allows to set the scene for an illustration of existing strategies and approaches. Interviews with leaders of these projects, allowed us to compare the different experiences in Europe and to make assumptions for possible approaches for the sustainability of e-Science clouds.

The analysis of these models reveals that in particular the EGI and PRACE initiatives have made relevant contributions to the promotion of the sustainability of the computing-related e-Infrastructures, such as the development of policies, business models and funding schemes for the new required structures. Also the e-IRG recommendations have played a significant role to increase the adoption of cloud technologies within the scientific context. Subsequently, the cloud computing market has been analyzed in order to identify and assess available offers of the different cloud providers. A comparative assessment of the various pricing models, may allow the scientific communities to evaluate who are the best competitors on the market and the services they offer.

A SWOT analysis shows that the benefits outweigh the potential threats and weaknesses (see legal and economic barriers, already depicted in deliverable 3.7 “Potential legal issues”). However, in order to maximize the impact of cloud computing technologies, legal and security issues need to be clearly understood and properly reviewed.

The most significant obstacle for a broad adoption of cloud infrastructures in e-Science context is the immanent loss of data control, due to the nature of the cloud model. A very important challenge for the success of e-Science cloud will be the development of high-level abstractions, which will ensure a good level of security and control for the researchers. According to Craig Lee1, this issue can be easily solved if the scientific communities prioritize the requirements related to the abstraction of cloud infrastructures and increase its development for the benefit of scientific goals.

Then, as a result of the analysis of the different cloud providers’ pricing models, we identify the great value of the so called “short-term contracts”, through which users can benefit from flexible payment models, which meet the customer and business needs more efficiently. Subsequently, we provide an

---

1 Lee A. C., “A Perspective on Scientific Cloud Computing”, Open Grid Forum and Computer Systems Research Department The Aerospace Corporation
evaluation of the different cloud providers’ offers, putting them into relation with the three pricing models defined by Lei Han2: PAYG, Flat rate and Mixture. We analyse the Total Cost of Ownership (TCO) and outsourcing to identify which is the most profitable economic model for the scientific community. Certainly outsourcing will be the most successful solution in the near future; indeed we conclude that “cloudsourcing” will predominate. “Cloudsourcing” is the delivery of outsource services via the cloud and is becoming one of the industry’s pre-eminent trends.

In the last chapter, an analysis of the sustainability of e-Science cloud constitutes a starting point for the reflection of the final sustainability analysis, which will be documented in D3.10. Our assumptions are presented after the evaluation of the seven VENUS-C sustainability scenarios. Our partners participated in a cost-benefit assessment, in order to evaluate the social, economic, environmental and technological sustainability of their scenario. The final evaluation allowed us to then evaluate the potential sustainability of generic e-Science cloud infrastructures.

As a result of the VENUS-C sustainability analysis, we predict that this cloud infrastructure is socially, economically, environmentally and technologically sustainable over time. Indeed, the evaluation of the seven VENUS-C scenarios, suggests a great impact of cloud services in all research fields. However, we see the major influence at social and economic level.

In detail, the general social sustainability shows a great impact on the scientific communities in terms of:

- Improvement of knowledge creation and research networking,
- Increase of quality and quantity of scientific production and data available for the scientists,
- Optimization of the resources consumed and the number of researchers involved in the specific research field.

Furthermore, the social benefits will positively influence the competition of e-Science cloud offers, investment flows and above all the operating costs.

Concerning the environmental sustainability, the assessment indicates that the cloud infrastructures can have a huge impact on the environment, especially in terms of savings in electric power.

Unfortunately, there is not yet a high awareness of the positive impact of cloud infrastructures on the environment. With reference to the technological sustainability of e-Science cloud infrastructures, the assessment highlights a significant impact on:

- Reduction of the activity time needed.
- Exponential increase of number of activities completed and consequently also of total activities.
- Improvement in terms of accessibility and usage of the technological infrastructures.
- Increase of productivity in terms of number of activity's outputs.
- Huge increase of data recorded.

---

2 Lei Han, (2009), “Market acceptance of cloud computing – an analysis of market structure, price models and service requirements, April 2009, Universität Bayreuth, Bayreuth
As a result, it is clear that the adoption of cloud technologies can release benefits to research institutions and other research stakeholders (e.g. the European Commission) improving their working routine and financial resources, and maximizing its positive impact on the scientific environment. Above all the major economic impact is related to reduce operational costs and maximize the market competition and investment flows.
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY**................................................................................................................................. 4

1. **INTRODUCTION**........................................................................................................................................... 11
   1.1 Scope of the document................................................................................................................................. 11
   1.2 Target Audiences......................................................................................................................................... 11
   1.3 Structure of the document ......................................................................................................................... 12
   1.4 Terms and definitions................................................................................................................................. 12

2. **e-INFRASTRUCTURES SUSTAINABILITY MODELS**.................................................................................. 15
   2.1 e-IRG recommendations on sustainability............................................................................................ 15
   2.2 The EGI-Inspire model.............................................................................................................................. 20
   2.3 The PRACE model..................................................................................................................................... 25
   2.4 The KTH/PDC Case.................................................................................................................................... 29
   2.5 The BSC Case............................................................................................................................................. 30
   2.6 e-Infrastructures sustainability’s conclusions.......................................................................................... 33

3. **MARKET ANALYSIS FOR e-SCIENCE CLOUD ADOPTION**................................................................. 35
   3.1 Cloud computing benefits and potential issues for e-Science.................................................................. 35
   3.2 e-Science cloud SWOT analysis.............................................................................................................. 38
   3.3 Cloud computing main actors.................................................................................................................. 39
   3.4 Cloud Computing market research for e-Science..................................................................................... 41
   3.5 Cloud computing pricing models ........................................................................................................... 48
   3.6 Service Level Agreements....................................................................................................................... 54
   3.7 Conclusions............................................................................................................................................... 55

4. **TOTAL COST OF OWNERSHIP OF SCIENTIFIC DATA CENTERS**.................................................... 57
   4.1 TCO calculator analysis approach........................................................................................................... 58
   4.2 Ownership vs. outsourcing........................................................................................................................ 62
   4.3 CAPEX vs. OPEX in e-Science Clouds.................................................................................................... 65

5. **DRAFT SUSTAINABILITY STRATEGY OF e-SCIENCE CLOUDS**......................................................... 69
5.1 Introduction to the seven VENUS-C scenarios questionnaire ................................................. 69
5.2 Approaches to e-Science Cloud Sustainability ........................................................................ 72
6. CONCLUSIONS and future activities ......................................................................................... 77
7. REFERENCES & RESOURCES .................................................................................................. 81
8. ANNEX A Sustainability questionnaire to the 7 VENUS-C scenarios ........................................ 85
   8.1 Building Structural Information .......................................................................................... 85
   8.2 Building Information Management .................................................................................... 94
   8.3 Data For Science .................................................................................................................. 99
   8.4 Civil Protection & Emergencies, Virtual Fire ................................................................. 105
   8.5 Bioinformatics ................................................................................................................... 112
   8.6 Systems Biology ................................................................................................................ 119
   8.7 Drug Discovery ................................................................................................................ 124
9. ANNEX B Short definitions of cloud service models in order of time adoption, as stated by the Gartner’s “Hype Cycle for Cloud Computing” ............................................................... 129

LIST OF TABLES

Table 1 - Terminology ................................................................................................................ 14
Table 2 - Cloud computing e-Science SWOT Analysis ........................................................... 39
Table 3 - Enterprise cloud vendors’ offers (as of 2010) ......................................................... 41
Table 4 - Cloud Services, Worldwide, 2008-2013 (Billions of Dollars) .............................. 46
Table 5 - Main SaaS segment trends ....................................................................................... 46
Table 6 - Comparison of major cloud vendors ........................................................................ 52
Table 7 - The different SLA models ....................................................................................... 55
Table 8 - VENUS-C seven users’ scenarios .......................................................................... 69
Table 9 - Hype Cycle Cloud Computing Adoption Provided by Gartner ........................... 71

LIST OF FIGURES

Figure 1 - Cloud computing families and sub-segments/offerings ........................................ 40
Figure 2 - The Four Worlds of Services, 2001 ..................................................................... 42
Figure 3 - The Four Worlds of Services, 2005 ..................................................................... 43
Figure 4 - The Four Worlds of Services, 2010 ..................................................................... 44
Figure 5 - The future of the software and software based services in Europe .................... 48
Figure 6 - Comparison of market structures of current SPs in cloud market .......... 49
Figure 7 - The relationship between transaction costs and factor specificity .......... 50
Figure 8 - Comparison of price models of SPs in the cloud market .................. 54
Figure 9 - TCO calculator variables .................................................................. 62
Figure 10 - Speed of Cost Reduction, Cost of Change .................................... 66
Figure 11 - Comparison of Microsoft Cloud Carbon Emissions .................. 74
Figure 12 - Drivers for cloud strategy definition process ................................. 78
Figure 13 - VENUS-C benefits for the Building Structural Information ............. 87
Figure 14 - VENUS-C social impact on the Building Structural Information ...... 89
Figure 15 - VENUS-C employment impact on the Building Structural Information ... 90
Figure 16 - VENUS-C potential market share achievable by the Building Structural Information .......................................................... 91
Figure 17 - VENUS-C technological impact Building Structural Information ....... 92
Figure 18 - The value of cloud models for Building Structural Information .......... 93
Figure 19 - VENUS-C benefits for the Building Information Management ....... 95
Figure 20 - VENUS-C social impact on the Building Information Management ...... 96
Figure 21 - VENUS-C employment impact of Building Information Management .... 97
Figure 22 - VENUS-C technological impact of Building Information Management .... 98
Figure 23 - The value of cloud models for Building Information Management ...... 99
Figure 24 - VENUS-C benefits for the Data for Science scenario ...................... 101
Figure 25 - VENUS-C social impact on the Data for Science scenario ............... 102
Figure 26 - VENUS-C employment impact on the Data for Science scenario ...... 103
Figure 27 - VENUS-C technological impact on AquaMaps ................................ 104
Figure 28 - The value of the cloud models for the Data for Science scenario ...... 105
Figure 29 - VENUS-C benefits for the Civil Protection & Emergencies .............. 107
Figure 30 - VENUS-C social impact on the Civil Protection & Emergencies .......... 108
Figure 31 - VENUS-C employment impact on Civil Protection & Emergencies .... 108
Figure 32 - VENUS-C environmental impact on Civil Protection & Emergencies .... 109
Figure 33 - VENUS-C percentage of cost savings on the Civil Protection & Emergencies .................................................................. 110
Figure 34 - VENUS-C technological impact on Civil Protection & Emergencies .... 110
Figure 35 - The value of the cloud models for Civil Protection & Emergencies ..... 111
Figure 36 - VENUS-C benefits for the Bioinformatics community .................... 113
Figure 37 - VENUS-C social impact on the Bioinformatics community ............... 114
Figure 38 - VENUS-C employment impact on the Bioinformatics community ...... 115
Figure 39 - VENUS-C environmental impact on the Bioinformatics community .... 116
Figure 40 - VENUS-C percentage of cost savings on the Bioinformatics community .......................................................... 117
Figure 41 - VENUS-C technologique Bioinformatics scenario .......................................................... 118
Figure 42 - The value of cloud models for Bioinformatics community .......................................................... 119
Figure 43 - VENUS-C benefits on the Systems Biology ........................................................................ 121
Figure 44 - VENUS-C social impact on the Systems Biology ........................................................................ 122
Figure 45 - VENUS-C technological impact Systems Biology scenario .................................................. 123
Figure 46 - The value of the cloud models for the Systems Biology .................................................. 123
Figure 47 - VENUS-C benefits for the Drug Discovery ........................................................................ 125
Figure 48 - VENUS-C social impact on the Drug Discovery .............................................................. 126
Figure 49 - VENUS-C technological impact Drug Discovery scenario .................................................. 127
Figure 50 - The value of the cloud models for the Drug Discovery .................................................. 128
1. INTRODUCTION

This analysis is done following best practices in business modelling and economic analysis: we have taken into account available business models, considering market specificities (segmentation, SWOT analysis, customers and financial flow analysis) and propose a model where overall costs for providing a service are compared with the expected revenues.

1.1 Scope of the document

The scope of the document is to analyse the short and mid-term sustainability3 of e-cloud infrastructures as a new model for doing e-Science, as proposed by the VENUS-C project and enabled by the VENUS-C platform. This document reports the initial findings that will serve as a baseline for the final sustainability analysis to be presented in Deliverable 3.104.

We analyse the current approaches towards sustainability of major e-Infrastructures (i.e. the available business models) and the context of procurement and usage of research resources (i.e. the market analysis). Implications of the previous analysis conducted in Deliverable 3.7, “Potential Legal Issues”, have also been considered and further discussed. This document reports on these findings and the initial thoughts about the suitability of clouds in scientific contexts.

1.2 Target Audiences

This document targets:

- VENUS-C partners to help them identifying opportunities and potential profitability of the VENUS-C business model.
- Research communities, who intend to use the VENUS-C infrastructure to help them evaluating the benefits and the costs of this cloud platform.
- Public procurement bodies dealing with research budgets.
- European policy-makers, who need to identify the potential of cloud computing business models in e-Science.
- Individual researchers, to clarify the profitability of the VENUS-C business model.
- IT developers who have to improve the VENUS-C infrastructure architecture and require a comprehensive understanding of the economic and financial aspects related to this model.

3 The World Commission on Environment and Development, (1987): This is the most commonly quoted definition of sustainability: “Sustainable development is development that meets the needs of the present without compromising the needs of future generations to meet their own needs”. It contains within it two key concepts: the concepts of needs, in particular the essential needs of the world poor, to which overriding priority should be given, and the idea of limitations imposed by the state of technology and social organization on the environments ability to meet present and future needs. For a most recent definition see: United Nations General Assembly (2005), “2005 World Summit Outcome”, Resolution A/60/1, adopted by the General Assembly on 15 September 2005. Retrieved on: 2009-02-17, in which it was noticed that sustainability requires the reconciliation of environmental, social and economic demand, the “three pillars of sustainability”.

4 Deliverable 3.10, “Future Sustainability Strategies Final”
1.3 Structure of the document

The document is structured in five main chapters, namely this introduction and four others, which follow the analysis process introduced above:

- In the second chapter, different existing e-Infrastructures models are introduced to provide a representation of approaches applied over the years (by example of EGI-Inspire, PRACE (Partnership for Advanced Computing in Europe) and e-IRG recommendations) to the sustainability of the e-Infrastructures projects. A similar analysis is provided to identify the sustainability of university projects, such as KTH and BSC ones.
- In the third chapter, we provide an analysis of the research market related to the procurement of the resources to support e-Science and their usage patterns (including a calculation of TCOs). This chapter is the starting point for any further reflection of the e-Infrastructures’ future sustainability based on clouds, taking into consideration the potential legal and economic issues already highlighted in Deliverable 3.7 “Potential Legal Issues”.
- In the fourth chapter, we developed an assessment of potential social, economic, environmental and technological sustainability of e-Science clouds, by means of an analysis of the seven VENUS-C scenarios and their usage contexts. A detailed questionnaire sent to our partners, based on a cost-benefit analysis approach, helped us to evaluate their opinion on the potential sustainability of e-Science cloud communities.
- In the fifth chapter, some interim conclusions are provided and predictions are made for the future sustainability of e-Science clouds.

1.4 Terms and definitions

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AISBL</td>
<td>Association Internationale Sans But Lucratif</td>
</tr>
<tr>
<td>BSC</td>
<td>Barcelona Super Computing</td>
</tr>
<tr>
<td>CAGR</td>
<td>Compound Annual Growth Rate</td>
</tr>
<tr>
<td>CAPEX</td>
<td>Capital Expenditure</td>
</tr>
<tr>
<td>CERN</td>
<td>The European Organization for Nuclear Research</td>
</tr>
<tr>
<td>CDN</td>
<td>Content Delivery Network</td>
</tr>
<tr>
<td>CIC</td>
<td>Cloud Innovation Center</td>
</tr>
<tr>
<td>CRM</td>
<td>Customer Relationship Management</td>
</tr>
<tr>
<td>DCI</td>
<td>Development Co-operation Instrument</td>
</tr>
<tr>
<td>DEISA</td>
<td>Distributed European Infrastructure for Scientific Applications</td>
</tr>
<tr>
<td>EGEE</td>
<td>Enabling Grids for e-Science, developed a production-level middleware (gLite), extensively used in the Venus-C infrastructure</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>EGI-DS</td>
<td>The European Grid Initiative Design Study</td>
</tr>
<tr>
<td>EEF</td>
<td>European e-Infrastructure Forum</td>
</tr>
<tr>
<td>e-Infrastructure</td>
<td>An open system that supports flexible cooperation and optimal use of all electronically available resources</td>
</tr>
<tr>
<td>e-IRG</td>
<td>e-Infrastructure Reflection Group</td>
</tr>
<tr>
<td>EIRO</td>
<td>European Industrial Relations Observatory</td>
</tr>
<tr>
<td>ERA</td>
<td>European Research Area</td>
</tr>
<tr>
<td>ERC</td>
<td>European Research Council</td>
</tr>
<tr>
<td>ERP</td>
<td>Enterprise Resource Planning</td>
</tr>
<tr>
<td>e-Science</td>
<td>Computationally intensive science deploys highly distributed network environments, or uses immense data sets. The term was created by John Taylor, the Director General of the United Kingdom’s Office of Science and Technology in 1999.</td>
</tr>
<tr>
<td>ESFRI</td>
<td>ESFRI, the European Strategy Forum on Research Infrastructures, is a strategic instrument to develop the scientific integration of Europe and to strengthen its international outreach.</td>
</tr>
<tr>
<td>ESF</td>
<td>European Science Foundation</td>
</tr>
<tr>
<td>FP8</td>
<td>8th Framework Programme</td>
</tr>
<tr>
<td>GEANT</td>
<td>It is the world’s largest multi-gigabit computer network dedicated to research and education</td>
</tr>
<tr>
<td>HEP</td>
<td>High Energy Physics</td>
</tr>
<tr>
<td>HPC</td>
<td>High Performance Computing</td>
</tr>
<tr>
<td>HTC</td>
<td>High Throughput Computing</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicators</td>
</tr>
<tr>
<td>KTH</td>
<td>Kungliga Tekniska Högskolan</td>
</tr>
<tr>
<td>NREN</td>
<td>National Research and Education Network</td>
</tr>
<tr>
<td>NSC</td>
<td>National Supercomputer Centre in Sweden</td>
</tr>
<tr>
<td>OECD</td>
<td>The Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>OPEX</td>
<td>Operational Expenditure</td>
</tr>
<tr>
<td>OSS</td>
<td>Open Source Software</td>
</tr>
<tr>
<td>PDC</td>
<td>Center for High Performance Computing</td>
</tr>
<tr>
<td>PPP</td>
<td>Point-to-Point Protocol</td>
</tr>
</tbody>
</table>
### ACKNOWLEDGMENTS

The Engineering Ingegneria Informatica SpA team wants to especially thank the experts who participated in the interviews. In particular, we highly appreciate the unfunded effort and time from Matti Heikkurinen (ex-deputy director of the e-IRG support program), Sergio Andreozzi (Policy Development Manager of EGI), Ludek Matyska (Chairmain of the Project Management Board of EGI-Inspire), Thomas Eickermann (Project Manager of PRACE).

We are pleased to express our gratefulness to our partners who contributed to the interviews and questionnaire: Erwin Laure and Ake Edlund (KTH), Sergi Girona and Daniele Lezzi (BSC), José M. Alonso and Ignacio Blanquer (Universidad Politécnica de Valencia), Simon Woodman (University of Newcastle), Corrado Priami (COSBI), Furio Barzon (Collaboratorio), Christos Vasilakos (University of the Aegean), Pasquale Pagano and Marko Mikulicic (Consiglio Nazionale delle Ricerche).

Special thanks to Fotis Karayiannis (Microsoft Greece) and Stephanie Parker (OGF.eeig) for their continuous support to assure the best results for the deliverable.

<table>
<thead>
<tr>
<th>PRACE</th>
<th>Partnership for Advanced Computing in Europe initiative aims at developing a European top Supercomputer</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D</td>
<td>Research &amp; Development</td>
</tr>
<tr>
<td>RES</td>
<td>Red Española de Supercomputación</td>
</tr>
<tr>
<td>SERC</td>
<td>Swedish e-Science Research Center</td>
</tr>
<tr>
<td>SICS</td>
<td>Swedish Institute for Computer Science</td>
</tr>
<tr>
<td>SOA</td>
<td>Service Oriented Architecture</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>SWOT</td>
<td>SWOT analysis is a strategic planning method used to evaluate the Strengths, Weaknesses, Opportunities, and Threats involved in a project or in a business venture.</td>
</tr>
<tr>
<td>TCO</td>
<td>Total Cost of Ownership</td>
</tr>
<tr>
<td>USD</td>
<td>United States Dollar</td>
</tr>
<tr>
<td>VO</td>
<td>Virtual Organization</td>
</tr>
<tr>
<td>VRC</td>
<td>Virtual Research Community</td>
</tr>
<tr>
<td>WeNMR</td>
<td>A World Wide Infrastructure for Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>WISDOM</td>
<td>Integrated Watershed Management</td>
</tr>
</tbody>
</table>

*Table 1 - Terminology*
2. e-INFRASTRUCTURES SUSTAINABILITY MODELS

In this chapter, we review major e-infrastructures sustainability models and approaches with the objective to provide a concise representation of roles, services and benefits, which the e-infrastructures bring into the European Research Area (ERA). These approaches are analysed by identifying sources of funding, decision-making processes, potential beneficiaries and enablers or barriers that need to be considered to verify the suitability of cloud computing for the scientific context.

The analysis has been performed by studying official documents made available from different flagship projects as well as interviewing key people of the various initiatives (e-IRG, EGI and PRACE). A similar analysis has been done with the institutes responsible for the resource provision in VENUS-C at university level, namely PDC at KTH, Stockholm and BSC, Barcelona representing the Spanish e-research context.

2.1 e-IRG recommendations on sustainability

The main objective of the e-IRG initiative is to support the creation of a political, technological and administrative framework for the easy and cost-effective shared use of distributed electronic resources across Europe. The e-IRG mission is to pave the way towards a general-purpose European e-Infrastructure. The vision of e-IRG is an open e-Infrastructure enabling flexible cooperation and optimal use of all electronically available resources, which should have a positive high-societal impact.

Governance

The e-Infrastructure Reflection Group (e-IRG), founded to define and recommend best practices for the (pan-)European distributed e-Infrastructures, is a policy committee of representatives appointed by Ministries in EU Member States, Associated States to the EU Research Framework Programme and the European Commission.

Beneficiaries

The e-IRG target audience are policy makers at the (inter-)governmental level (mainly dealing with strategy and funding issues), e-Infrastructure service providers, existing and new user communities, and technology developers. Interaction with all the above mentioned bodies is crucial to foster a broader uptake of the e-Infrastructure–enabled technologies and processes.

Sustainability

e-IRG has worked out a series of recommendations on the sustainability of e-Infrastructures, the first of which is developed by a dedicated e-IRG Task Force on “Sustainable e-Infrastructures”.

The main recommendations are the following:

---

5 www.e-irg.eu
• **Investments based on coherent policies and interoperability**: governments and the Commission should develop policies and mechanisms to encourage increased investment in a more coherent and interoperable way across Europe.

• **Sustainable structures and initiatives**: the existing e-Infrastructure projects must be superseded by integrated sustainable services at national and European levels.

• **General purpose e-Infrastructures**: e-Infrastructures must be application-neutral and open to all user communities and resource providers. National funding agencies should be encouraged to fund multi-disciplinary and inclusive infrastructures rather than disciplinary-specific alternatives.

• **Interoperability and standards**: e-Infrastructures must inter-operate and adopt international standard services and protocols in order to qualify for funding.

• **European integration**: the Commission should, within the Seventh Framework Programme, develop a pan-European e-Infrastructure which explicitly encourages the further integration of national e-Infrastructure initiatives.

Moreover, the following actions are recommended:

• **Innovation aspect**: infrastructures need to remain state of the art; therefore new technologies should be evaluated and introduced continuously. In order to make investments in R&D more efficient, the take up of new technology in production infrastructures should be improved by appointing e-Infrastructure providers as stakeholders in relevant R&D efforts.

• **Ease of use**: to accelerate and expand the adoption of e-Infrastructures attention must be paid to their ease of use. Investing in improving the usability (e.g. by hiding complexity and increasing interoperability) will broaden their user base, adding significant value to the science community and increasing European competitiveness. Also cloud computing technologies should acknowledge these recommendations, in order to increase the adoption of this kind of infrastructures.

• **Training**: the new opportunities presented by distributed infrastructures require increased training and an improved skills base for the research community, which also needs to form part of any national or European strategy for e-Infrastructure. This will require further advisory and guidance services that collaborate across Europe.

• **Integration of Industry**: to strengthen the integration of industrial efforts and SMEs into sustainable infrastructures in this context, industry has to be seen as both a potential user and a major partner for service provision. Clear policies have to be established for access from industrial research projects in pre-competitive domains, industrial production projects accessing innovative technologies or deploying innovative strategies and industrial production projects with occasional exceptional requirements (critical computing on demand), this recommendation is very relevant also regarding cloud computing technologies and the Venus-C project.
The e-IRG notes the importance of the steps undertaken by the EGI and PRACE initiatives to promote sustainability of cloud computing, such as the development of policies, business models and funding schemes for the new required structures.

To achieve this goal, e-IRG states that adequate levels of funding should be granted by the EC and Member States for the development of new governing structures both on the national and European levels.

The e-IRG sustains that current fragmented efforts (projects, facilities, initiatives) should necessarily evolve into a rich service-oriented e-infrastructure economy, both attractive and sustainable able to support a wide variety of applications, services and differentiated user communities. The e-IRG recommends that Member States and the EC promote and actively facilitate sharing of information and best practices related to allocation, accounting and economic models for e-infrastructure resources and services.

The last e-IRG paper on e-infrastructure service sustainability was produced in 2011 and states: \(^7\)

1. The need, benefits and risks of the change to e-Infra structure services should be underlined and communicated at all policy levels.
2. Standardization, fast development and flexible exploitation of e-Infrastructure services should engage e-Infrastructure users in the widest sense.
3. Facilitate the participation of the user communities to the definition and exploitation of e-Infrastructure services.
4. Virtualization and SOA should be widely exploited in developing and introducing new e-Infrastructure services in order to improve availability, accessibility, efficiency, and cost-effectiveness of IaaS operation.
5. Simplified access, transparent service offerings, customized support, improved governance models and sustainable business models should be key elements in the joint deployment of e-Infrastructure services.
6. Facilitate the creation of community infrastructures (or community clouds) where organizational or scientific requirements suggest/necessitate shared IaaS services.
7. Co-operation in the e-Infrastructure area with other public sectors, like government and health care, should be promoted, in order to exploit economies of scale and intensify the contribution of research and education e-Infrastructures in facing societal challenges at large, while safeguarding the necessary legal position as a non-public service provider in the sense of the European directives.
8. Innovation by PPP activities should be strengthened through jointly creating a market for e-Infrastructure resources and services by co-operating Public and Private Partnerships (PPP).

According to the opinion of Matti Heikkurinen, ex-deputy director of the e-IRG support programme: “Long-term sustainability is possible only through somehow linking the e-Infrastructure directly or indirectly - to either increasing or securing socioeconomic benefits for European citizens. A bit like

making a PR-case of being either like trans-national high-speed rail network or the Interpol or enabling technology to one or both of them.\(^8\)

Moreover, regarding the involvement of industrial players for the improvement of the sustainability of e-Infrastructure projects, Heikkurinen states:

“If industrial players can convince that they will provide cheap services in a sustainable manner, they will take away users from e-Infrastructure projects/initiatives. However, I’d imagine any initiative needing resources long-term (order of years) will not trust commercial providers directly. So I’d see their role most likely complementary keeping the e-Infra projects honest by providing cost/pricing baseline and in some specific cases providing resources for certain high-visibility research activities. Real threat to traditional e-Infrastructure projects might come from a joint venture of industrial players that would offer long-term public-private partnership committing to providing resources for science clearly at less than 50% of the market price”.

Furthermore, reflecting the possibility that the integration of the European e-infrastructures initiatives could increase the sustainability of different projects (GEANT, EGI, PRACE), Heikkurinen supports that it would certainly solve certain problems, like mapping user problems to appropriate services; however actually this integration has not yet been possible because NREN/GEANT has a sort of “monopoly” in relation to the other projects EGI and PRACE. This lack of competition can allow the providers to be less responsive to the users’ discontent.

The solution to this urgent issue, according to Heikkurinen, could be the development of a model based on a common funding and user support structure, which also holds the budget for all the e-Infrastructure services and directs the "clients" to NREN/GEANT, EGI and PRACE services. Regarding to the common user support for all projects, Heikkurinen states that they could perhaps emerge from voluntary collaboration between the different initiatives. While this kind of support could reduce the project’s costs, it could probably make this model less sustainable, since the funds do not come from a unique entity.

Moreover, not to create differences between the various subjects, the e-Infrastructure projects require a very good understanding on how the contributions to common parts of the organisation will be divided between GEANT, EGI and PRACE. In order to improve the integration process between the e-Infrastructure projects, the European Commission, the Governments and the National e-Infrastructure providers should have a common vision of how and where the funds are needed and accounted for enabling research tools.

Furthermore, Heikkurinen states that:

“The current uncertainty of whether research groups get extra funds to use for e-Infrastructure or if they get their funding (effectively) cut and are given access to common e-Infrastructure for "free" is the worst

---

\(^8\) VENUS-C Interview with Matti Heikkurinen, Ex-deputy Director of the e-IRG support programme, on the e-IRG sustainability,(April 2011).
possible situation. This problem is probably going to become more acute with the long-term requirements related to data services and their relationship with open access and open data goals”.

One of the possible solutions, from Heikkurinen’s point of view, is to mandate the European Commission to develop a recommendation for a usage model or at least communicate to researchers the model chosen at the national level. Then, Heikkurinen recommends a gradual tweaking of funding models in FP8, ERC, ESF and national funding agencies.

Finally, for a comprehensive settlement of the problem Heikkurinen states:

“Common policy-level principles for accounting for the use and exchange of resources would also be needed, since the technical solutions for cross-domain and cross-service accounting will not emerge if the policy-level requirements are not clear. For instance, if you can’t account for the use of resources, you will not have comprehensive incentives for sharing”.

Within this context, the “e-IRG e-Infrastructures roadmap” provides a vision for the future to motivate continuing efforts in order to create links between stakeholders and maximize the socioeconomic value of a common research e-infrastructure. The latest version of the roadmap, published in 2010, outlines the role e-IRG intends to play in the on-going development of a common e-infrastructure, and the organization’s plan to expand the scope of its mission. Key e-IRG recommendations from the e-IRG roadmap relevant to VENUS-C and the DCI environment are related to the following topics:

- **Commodity computing**
  e-IRG recommends that organizational structures and incentives are put in place that ensure that all the actors in the commodity computing domain will have an interest in –when technically appropriate – bringing new users and user communities into contact with other components of the e-infrastructure.

  Commodity computing encompasses also cloud computing, so again this is relevant to VENUS-C and the cooperation among the DCI projects.

- **Standards and interoperability**
  e-IRG recommends the continuation of interoperability benchmarking through global standardisation. In addition to making it easy to benchmark in terms of suitability, dependability and cost-effectiveness in different application domains, this will ensure long-term interoperability of different implementation technologies used for providing e-Infrastructure services and create a marketplace for commercial offerings.

- **Commercial uptake**

---

e-IRG recommends the gathering of information about successful commercial uptake of e-Infrastructure–related innovations to identify policy, funding and other mechanisms that would support seamless transition of proven e-Infrastructure collaboration models into broader use in society.

- **New user communities**

  e-IRG recommends that the Member States, European Commission and the various sustainable e-Infrastructures initiatives propose and provide resources for mechanisms that will:

  - Accelerate the adoption of sustainable e-Infrastructure services by new user communities.
  - Complement the competitive “call for projects” approach with a faster mechanism to target resources to popular e-Infrastructure services.
  - Identify partners and collaborative processes to support the organizational development of new user communities.

Finally, e-IRG recently published the so-called e-IRG Blue Paper\(^\text{10}\), a document specifically prepared for ESFRI\(^\text{11}\). The document was endorsed by both e-IRG and ESFRI in September 2010.

The e-IRG Blue Paper reports on current trends and issues and sets out policy recommendations for several key areas including Grids, Cloud computing and virtualization. It promotes collaboration among Grid, cloud infrastructure providers and users to raise awareness of the range of available technologies and the best way to use them.

### 2.2 The EGI-Inspire model

The EGI-InSPIRE project supports the transition from a project-based system to a sustainable pan-European e-Infrastructure, by supporting ‘grids’ of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI-InSPIRE collects user requirements and provides support for the current and potential new user communities, such as the ESFRI projects. The main goals of the project are to:

1. Continue operation and expansion of today’s production infrastructure by transitioning to a governance model and operational infrastructure that can be increasingly sustained outside of specific project funding.
2. Continue support of researchers within Europe and their international collaborators that are using the current production infrastructure.
3. Support current heavy users of the infrastructure in earth science, astronomy and astrophysics, fusion, computational chemistry and materials science technology, life sciences and high energy physics (HEP) as they move to sustainable support models for their own communities.

---

\(^{10}\) E-Infrastructure Reflection Group, (2010), “e-IRG Blue Paper”

4. Provide mechanisms to integrate existing infrastructure providers in Europe and around the world into the production infrastructure, so as to provide transparent access to all authorised users.

5. Establish processes and procedures to allow the integration of new DCI technologies (e.g. clouds, volunteer desktop grids) and heterogeneous resources (e.g. HTC and HPC) into a seamless production infrastructure as they mature and demonstrate value to the EGI community.

The distributed computing grid was originally conceived in 1999 to analyse the experimental data produced by the Large Hadron Collider at CERN – the European particle physics laboratory located on the Swiss/French border. The European DataGrid Project, which started in January 2001, led the research and development of Grid technologies.

It established the organisational structure, gathered and analysed requirements, developed middleware (the software that links hardware resources), and provided training to its users. The project proved the successful application of Grids in various research fields – high energy physics, earth observation and bioinformatics. Upon completion in March 2004, a new project called EGEE (Enabling Grid for E-science) took over the further Grid development in what would result in three successive two-year phases.

EGEE provided access to computing resources on demand for researchers, from anywhere in the world and at any time of the day. Ease of access and the ability to analyse a larger amount of data within a shorter timescale than before attracted participation from a wider range of scientific disciplines. By April 2010 when the last EGEE project phase was completed, there were about 13 million jobs per month running on the Grid, hosted by a network of 300 computer centres worldwide.

In order to support science and innovation, a longstanding operational model for e-Science is needed, both for coordinating the infrastructure and for delivering integrated services that cross national borders. A dedicated design study (EGI-DS, European Grid Initiative Design Study) took place from September 2007 to December 2009. It established the conceptual and logistical framework for a permanent organisation to oversee the operation and development of the Grid on a Europe-wide level.

EGI.eu, a foundation established under the Dutch law, was created in February 2010 to coordinate and maintain a sustainable pan-European infrastructure to support European research communities and their international collaborators. Its work builds upon the previous EU-funded projects in order to sustainably operate a scalable, federated, distributed computing system. EGI-InSPIRE, a four-year project, coordinated by EGI.eu supports and co-funds operations within the EGI Community with a common goal to create a seamless system ready to serve the demands of present and future scientific endeavours.

The mission of EGI-Inspire is to establish an EGI (European Grid Initiative) that strives to create, maintain and extend large scale distributed computing (and to some extent related data) infrastructure (Grid) that spans the whole Europe and is fully connected with similar activities worldwide (some of those are also initiated and supported by EGI). To achieve this primary goal EGI develops operational and financial

---

13 http://web.eu-egi.eu/
models for sustainability. EGI also coordinates activities of National Grid initiatives (NGI) to draw on synergies and complementarities of experience available at national level.

EGI is funded through a combination of European Commission (EC) funding provided by the EGI-InSPIRE project and national funding sources that co-fund national activities that in turn contribute in part to the EGI-InSPIRE project and directly to the operating costs of EGI.eu. Even though the maturity of NGIs varies widely across Europe, over 30 NGIs and EIROs have committed to paying the annual EGI.eu participation fees as specified in the EGI.eu statutes for services provided to the community as a whole, and to continue to invest significant additional funds in the development of their own national e-Infrastructures and supporting staff.

The associated effort and costs of providing the production infrastructure in the context of the EGI-InSPIRE project is conservatively estimated at more than €330M with the €25M contribution from the EC specifically for the project costs of €72M. The EC contribution therefore, represents just 7.5% of the overall costs of this community activity. This provides a clear demonstration of the commitment by the partners and their funding organisations to EGI.

**Governance**

EGI.eu is governed by a Council, which has representatives from all of its participants and is responsible for providing the long-term direction of the organisation through a voting scheme where the votes are proportional to the participant’s financial contribution. An Executive Board provides frequent guidance to the Director, who leads the organisation on a day-to-day basis. The EGI-InSPIRE project is managed on a daily basis by the work packages leaders and the Project Director through the Activity Management Board. The Collaboration Board is the senior governance body for the project where each lead beneficiary within the project is represented, however the detailed management of the project is undertaken by the Project Management Board where the partners are represented. An External Advisory Committee provides additional input to the project on its strategic evolution.

The EGI ecosystem comprises EGI.eu as the central coordination body, the participants within EGI such as NGIs and EIROs as resource infrastructure providers, technology providers, user communities, and funding bodies.

**Beneficiaries**

EGI’s user community has two important structures: Virtual Organisations (VOs) and Virtual Research Communities (VRCs). A VO is a group of people (e.g. scientists, researchers) with common interests and requirements, which need to work collaboratively with other members of their collaboration and/or share resources (e.g. data, software, expertise, CPU, storage space) regardless of geographical location.

---

14 VENUS-C Interview with Sergio Andreozzi, Policy Development Manager of EGI, on the EGI sustainability, (April 2011)
They join a VO in order to gain access to resources by agreeing to a set of rules and policies that govern the access and security rights for the users, resources and data in question. EGI currently hosts more than 200 VOs for communities with interests as diverse as earth sciences, computer sciences and mathematics, fusion, life sciences or high-energy physics. A VRC is a group of large-scale scientific research collaborations, either covering multiple VOs that are part of a larger domain area. The VRC model allows a community to have bi-directional interactions through defined points of contact with EGI across broader domain areas.

Expanding the breadth of EGI’s user community by evolving its services, increasing its flexibility, and lowering the complexity of its use are key objectives supported through the EGI-InSPIRE project. The VRC model allows EGI to identify the communication points within each community to collect feedback and new requirements on the offered services in order to evolve them through a virtuous cycle of Discover, Design and Deliver. This cycle is supported and championed through strong dissemination and outreach activities frequently in partnership with other projects and initiative as well as at various themed workshops and events.

The EGI-InSPIRE “inherited” the High Energy Physics community as its major user community. Nowadays, this community is extremely active due to the experiments going on in CERN (LHC) that requires tremendous amounts of data storage and computing power for their processing. Also some other physics communities are heavy users and beneficiaries of EGI (Astrophysics around the Auger\textsuperscript{15} virtual organization being a prominent example). Other large and most important communities come from the Life Sciences and Computational Chemistry.

Virtualization

Sergio Andreozzi, Policy Development Manager of EGI, concerning the virtualization of EGI states in:

“Virtualization is already being frequently used ‘behind the scenes’ in many of EGI’s data centres. Indeed there are sites currently in production use where the services being directly consumed by end-users that are completely hosted on virtualised resources. The challenge facing EGI is how these new technologies can enable dynamic execution environments or on-demand elastic service deployment that can be accessed by end-users, or by experts acting on their behalf. Such functionality is being frequently requested from the current EGI user base, and is seen by many new communities as being an essential pre-requisite for them to use EGI as a resource provider”.

The impact of adopting virtualization locally and across the whole production infrastructure would allow the rapid and flexible deployment of different software environments, customised to the needs of individual user communities, with minimal overhead to the individual resource centre and on demand from the end-user.

\textsuperscript{15} The Pierre Auger Cosmic Ray Observatory, http://www.auger.org/
EGI three major success stories for enhanced scientific research as a direct result of using the grid infrastructure

According to Sergio Andreozzi\textsuperscript{16}, the EGI’ three major success stories of EGI are:

1. **The Applications Database**: stores information about applications that are ready to be used directly by scientists. Advantages: scientists save time and energy to focus on their research; it lowers the ‘entry barrier’ to the Grid world – scientists are able to use the grid without having to develop software; it represents a forum for dialogue between scientists and developers.

2. **Statistics**: number of users increased 10% since March 2010; the number of jobs submitted to the Grid by scientists working in fields other than HEP has doubled in the last year (from about 554,000 to 909,000). This means that a) Grid computing is becoming more popular as a tool amongst scientists; b) Grid computing has the potential to make useful contributions in many fields of science, not just HEP.

3. **Examples of Grid computing research projects making a difference in people’s lives**: a) an application that helps doctors to diagnose Alzheimer’s Disease at an early stage (SPM application, part of the DECIDE infrastructure); b) the WeNMR\textsuperscript{17} infrastructure uses Grid computing to screen millions of chemical compounds in the search for cures against serious diseases – doing the same thing in the lab would take tens of years; c) the WISDOM\textsuperscript{18} project is an initiative for grid-enabled drug discovery against neglected and emergent diseases, for example malaria and avian flu) Researchers in Lithuania are developing grid-based algorithms to help doctors to interpret electroencephalograms from children with epilepsy or to assess the severity of arthritis or the necessity of hip replacement surgeries, for example; e) the Dutch BiG Grid infrastructure is implementing a platform to implement popular DNA re-sequencing experiments on the Grid.

**Sustainability**

Due to the interdependencies of the ecosystem and in order for EGI to function, it is paramount that the NGIs and, in turn, EGI.eu are sustainable. As coordinating and maintaining a high-quality infrastructure costs money, the focus has been to ensure sustainability by transferring recurring operational costs to the communities that benefit from using the services.

It is expected that NGIs, institutions and/or users will eventually need to pay in some form or another for the resources and services they currently may use for free. As the community shifts to a more service oriented model, the pressure on the providers to ensure that the services offer ‘value for money’ will certainly increase, as will the community’s ability to compare costs to commercial cloud providers and evaluate the value that e-Infrastructures provide.

\textsuperscript{16} Interview with Sergio Andreozzi, Policy Development Manager of EGI, about EGI sustainability

\textsuperscript{17} http://www.wenmr.eu/

\textsuperscript{18} http://www.wisdom.caf.dlr.de/
The recently published EGI Sustainability Plan, reference has provided a comprehensive list of the wide range of services that the EGI provides and has taken a proactive approach in defining potential business models in order to sustain these services. A preliminary assessment was made which of the single services could be supported by a particular business model, providing a means to structure and direct future discussions and dedicated activities within the project.

With reference to the EGI social sustainability, the relevance of collaboration and sharing is a very important variable which influences the EGI sustainability model. The EGI affirms that there is a benefit from connecting own resources to a larger infrastructure; this benefit comes both from more efficient use of the resources (e.g., it is not necessary to buy resources at peak demand as in a large infrastructure there will be some spare capacity at any moment to cover such peaks) and from providing a platform for large scale collaboration (when a team at one particular institute starts to collaborate with a different team elsewhere, they will not need to create and setup a dedicated IT infrastructure to give each other access to data and computing resources but being both connected to a Grid and all this functionality will be available from day zero).

According to Ludek Matyska, Chairman of the Project Management Board of EGI-Inspire, reference here: “This means that the basic sustainability model uses national and regional funding for resources and through synergies to increase efficiency. At the social level, if there will be no interest in collaboration between research groups (coming from different countries), there will be no interest in grid and its technologies. Also, if nobody will be buying and operating local computing and storage resources, there will be no resources to connect in a grid”.

This model will be influenced by the success of ESFRI and EGIprojects. A positive solution should combine the projects’ best practices in order to provide a comprehensive solution.

About the priorities of EGI-Inspire, Ludek Matyska states:

“Acceptance of the EGI model by the research communities is the fundamental condition for the long term EGI sustainability. This defines the policy priorities, namely work with research communities and funding bodies, the former to convince to use the grid infrastructure and actively contribute to its further development, to the later to show that actually the efficiency (and quality and extent of results) of research communities is extended through access to the grid (distributed computing and storage) infrastructure”.

2.3 The PRACE model

PRACE, the Partnership for Advanced Computing, is a permanent European Research Infrastructure to “enable world-class science through world-class systems and services”19. The implementation of PRACE is supported by projects funded in part by the European Commission.

19 http://www2.fz-juelich.de/jsc/grid/prace1ip
The main goals are:

- Development and provision of an Infrastructure at European level which allows the scientific communities, including those within industry, to access European High-end Computing (HeC) systems (Tier-0).
- Management of the coordination between the Infrastructure and existing national computation centres (Tier-1) and also, if agreed, regional computation centres (Tier-2), to allow for the establishment of relationships with the HeC user communities.
- Provision and rationalization of access to the Infrastructure by qualified European and international scientific communities, either academic or industrial, whose projects may be evaluated for such purpose.

The PRACE Research Infrastructure has many commonalities with more traditional European RIs: driven and funded by the member states, independent scientific governance, peer review based access. Main differences are the:

- Short investment cycles compared to physical RIs like telescopes or accelerators: a Tier-0 HPC systems has a typical life-time between 3 and 5 years.
- Distributed nature: the physical Tier-0 resources and the user support structures are distributed over different countries; users usually use the resources via Internet and do not physically access them.

PRACE main focus is on providing a high-end HPC service on European level. However, it also strives to manage the coordination between this infrastructure and existing national HPC centres (Tier-1) or of regional initiatives in Europe.

This is currently actively pursued through the integration of DEISA-like services by the PRACE implementation phase project. Concerning the e-Infrastructure landscape, PRACE positions itself as a partner in a cooperation of independent e-Infrastructures that also operate on a European scale, such as GEANT, EGI, and an envisioned future data infrastructure.

**Governance**

PRACE has the legal form of an Association Internationale Sans But Lucratif (AISBL), an international non-profit association under Belgian law with its seat in Brussels. Members are the national HPC coordinators of 20 European countries, each mandated for that purpose from its national government. The governing body is the European Council, composed of all members.

The user access to the RI is governed by a peer review process that is driven by the Scientific Steering Committee (SSC), an independent body composed of distinguished scientists from all major scientific user communities. Access is provided purely on the basis of scientific excellence.

---

20 Tier 0 deploy systems of the highest performance level from 3 to 6 European Centres, instead, Tier-1 provide coordination with national centres or HPC organizations and different access rules and procedures on national level limit integration. Tier-2 coordinates regional or university centres.

21 The aim of data infrastructure is to minimize programmer’s effort by providing a set of built-in services and to enforce good coding practices and a homogeneous interface.
Four hosting member countries (France, Germany, Italy, and Spain) provide currently € 100 Million each over a 5-year period for the Tier-0 systems and services offered by PRACE. All 20 members of PRACE AISBL, the legal entity representing the PRACE Research Infrastructure (RI) mentioned above, pay an annual membership fee to fund the headquarter operation and related activities. Furthermore, the European Commission supports the implementation of the RI through a series of FP7 projects (with € 48 Million so far and additional € 20 Million envisaged for a future project). The project partners complement the EC contribution by a similar amount.

**Beneficiaries:**

Computer simulation has evolved into a third research paradigm (in addition to theory and experiment) of scientific discovery. All user communities that require high-end HPC resources for simulations benefit directly from PRACE. These are essentially all natural sciences (physics, chemistry, life-science) and engineering. Through HPC resources at the highest performance level that PRACE offers to its users, PRACE contributes to the world-wide competitiveness of European Science and Industry.

**Services**

The main services provided by PRACE are: access to Tier-0 resources; expert support for porting, enabling and scaling of applications to make the most effective use of the unique resources; training in advanced, novel programming languages and techniques. Currently, virtualisation plays little to no role for HPC, especially on the high-end. The main reasons for this are: the focus on highest performance in combination with the differences in HPC-system characteristics. With MPI and OpenMP as widely used standards, the main effort of porting an application to a specific system is not making it run on the system, but adapting it to system performance characteristics related to the CPUs, node layout, inter-node network, memory size and bandwidth, I/O, etc. These characteristics are intrinsic and hiding them through virtualisation is still a research topic.

**PRACE major success stories for enhanced scientific research as a direct result of using the HPC services**

HPC is a tool and enabling technology in many areas of science and engineering. Examples are:

- The prediction of the climate change is unthinkable without simulations on HPC systems. The IPCC modelling and simulation approach has been quite successful.
- The calculation of the hadron masses in a computer simulation (performed on the Tier-0 system Jugene) from first principles has been rated as among the top 10 scientific breakthroughs in
2008, and is considered as a proof of the Quantum Chromo Dynamics, the fundamental theory of strong interaction (quarks and gluons).

Applied research and engineering relies to a large extent on HPC simulations. Different aircraft manufacturers and oil companies now use HPC almost routinely resp. for aircraft design and oil exploration. This is becoming true also for new materials and drugs design.

**Sustainability**

The European research and engineering communities have articulated their demand for a sustainable high-end computing service as now provided by PRACE through a scientific case, prepared by the HPC in Europe Taskforce. This scientific case has been instrumental to bringing PRACE on the ESFRI roadmap and to trigger the financial commitments of the hosting members for the first 5 years.

For the long-term economic sustainability, it will be essential that PRACE meets the demand and provides unique and comprehensive, high-quality services, and continues to adapt to the changing demands of its user communities. The representation of the users in the governance model of PRACE is ensuring this.

By providing a service that is available to all European researchers, and by strengthening the collaboration and coordination of HPC centres in Europe, PRACE is providing an important contribution to the European Research Area.

In the light of this, the EC considers a significant increase of its financial contribution to PRACE in the successor of FP7, which will further strengthen the financial sustainability of PRACE. By investigating and investing into energy-efficient HPC technology in cooperation with vendors and other technology providers, PRACE is also working towards a reduced energy-footprint of its service and thus environmental sustainability.

With reference to the potential integration of the European e-Infrastructure initiatives to increase the sustainability of the different projects, Thomas Eickermann, Project Manager of PRACE, states:

“Several user communities use more than one of the three horizontal e-RIs GEANT, EGI, and PRACE. A cooperation of these e-RIs towards interoperability of certain services (e.g. AAA) can be of benefit for these users”.

The European e-Infrastructure Forum (EEF) has been created as a forum to achieve such cooperation. Concerning a closer integration, it needs to be considered that the three above mentioned e-RIs have overlapping, yet different user groups, largely different business models and interaction patterns with their users: e.g. basically every scientist uses GEANT, but without any direct interaction; Grids currently rely mainly on a sharing of compute resources within and to some extent between communities and granting mutual access. PRACE provides peer reviewed access to nationally funded high-end HPC services.
resources. Due to these differences, it is still unclear what the benefit of an integration of the different e-RIIs could be.

PRACE considers itself mainly as a user service, providing resources and services to individual user groups and projects. A current route of extension is the establishment of collaborations with scientific communities to match their long term demands and development cycles (programme access). On a policy level, PRACE is mainly working with its funding bodies, the national governments (through the members), ESFRI, and the EC to achieve this goal.

2.4 The KTH/PDC Case

PDC operates leading-edge, high-performance computers on a national level. PDC offers easily accessible computational resources that primarily cater to the needs of Swedish academic research and education.

PDC also takes part in major international projects to develop HPC for the future and stay a leading national resource in parallel computing. PDC-HPC is part of the Swedish National Infrastructure for Computing (SNIC).

At a national level, KTH founded together with SICS, the “Cloud Innovation Centre” (CIC), an open consortium promoting cloud technologies in academia and industry, which specifically will:

- Provide a forum to disseminate and discuss cloud initiatives in Sweden.
- Foster research collaborations in national and international projects.
- Coordinate education activities in cloud technologies.
- Interact with industry through agile business driven projects.
- Contribute to Swedish seed accelerator programs.

PDC is also involved in various international e-Infrastructures and cloud computing projects, such as:

- Distributed European Infrastructures for Supercomputing Applications (DEISA);
- Enabling Clouds for e-Science (ECEE);
- European Grid Infrastructure (EGI);
- Partnership for Advanced Computing in Europe (PRACE);
- Scalable Software Services for Life Science (Scalalife);
- Northern Europe Cloud Computing (NEON);
- Virtualized Storage Services Foundation for the Future Internet (VISION).

---

25 www.pdc.kth.se
26 www.snic.vr.se
27 www.sics.se
28 http://www.pdc.kth.se/research/cloud-computing
29 http://www.pdc.kth.se/projects/international-projects
30 www.deisa.eu
31 http://www.scientific-cloud.org/
32 www.egi.eu
33 www.prace-project.eu
34 www.scalalife.eu
35 www.necloud.org
The main goal of PDC is to provide HPC services (e.g. Supercomputers, clusters, storage, archival storage) to Swedish and international researchers. Indeed, PDC is involved in various national and international e-Infrastructure and cloud computing projects.

**Governance**

KTH is a member of the Swedish e-Science Research Centre (SeRC\(^37\)), which brings together a core of nationally leading IT research teams (tool makers) and leading scientists in selected strategic application areas focusing on:

- Formation of e-Science communities that connect application groups with relevant core e-Science groups and computer experts at PDC and NSC\(^38\).
- Research in core e-Science methods.
- Much closer collaboration between PDC/NSC, and a substantial increase in advanced support staff, which will turn the centers into comprehensive e-Science enablers.

**Sustainability**

The sustainability model of PDC is based on long-term funding through the Swedish Research Council (via SNIC) and KTH. SNIC is a meta centre dedicated to Swedish researchers and their requirements. SNIC’s mandate is to provide unified access to the leading supercomputers in Sweden.

The Swedish Resource Council provide long-term funding for high performance computing resources in Sweden and coordinate competence profiles between Sweden’s existing supercomputer centre, so the resources are developed and used optimally. The SNIC funds are mainly devolved to institution and researchers working on computers, grid, storage, network, visualization.

Within this context, Erwin Laure, director of PDC/HPC, states that the sustainability model of KTH is very stable, indeed:

*“PDC has existed for over 20 years now and there are no signs that things will change”*\(^39\).

### 2.5 The BSC Case

BSC-CNS\(^40\) (Barcelona Supercomputing Center – Centro Nacional de Supercomputación) is the National Supercomputing Facility in Spain. Officially constituted in April 2005, BSC-CNS manages the

---

\(^{36}\) [http://www.sics.se/project/vision](http://www.sics.se/project/vision)


\(^{38}\) [www.nsc.liu.se](http://www.nsc.liu.se)

\(^{39}\) Interview with Erwin Laure, Director of PDC/HPC, on PDC sustainability (March 2011)

\(^{40}\) [www.bsc.es](http://www.bsc.es)
supercomputer MareNostrum\textsuperscript{41}. The mission of BSC-CNS is to investigate, develop and manage information technology in order to facilitate scientific progress. With this aim, special dedication has been taken to areas such as computational sciences, life sciences and earth sciences. All these activities are complementary to each other and very tightly related. In this way, a multidisciplinary loop is set up: the exposure to industrial and non-computer science academic practices improves BSC's understanding of the needs and helps focusing its basic research towards improving those practices.

BSC manages the MareNostrum supercomputer, which is composed by 10240 IBM PowerPC processors contained in 2560 JS21 Blades from IBM. It has 20TB of main memory and 480 TB of disk storage. In addition, it manages an Altix 4700 SMP supercomputer and a cluster of ~60 JS20 Blades from IBM. The RES consists of a virtual distributed infrastructure of supercomputers located in different sites, each of which contributes to the total processing power available to users of different R&D groups in Spain. Generally, each node reserves 20% of capacity for use by researchers at the host institute, and the other 80% is made available to the general scientific community via the Access Committee.

- The total processing capacity of the RES is 138.5 TFLOP/s (138.5 trillion floating point operations per second). The RES is coordinated by the Operations Department of the BSC-CNS, which includes support for global maintenance and upgrades, training of users and technicians, facilitation of access and all aspects of user support.
- The powerful resources of the MareNostrum supercomputer and the RES nodes are accessed by a broad spectrum of Spanish and international scientists. Computing time is allocated by the Access Committee, composed of a Core Team and four Expert Panels of prestigious Spanish scientists external to the BSC-CNS. Additionally, a percentage of computing time is reserved for commercial projects to enable Spanish companies to maintain international competitiveness.

\textcolor{red}{Governance}

The BSC-CNS is a legally autonomous, public consortium, with three founding partners, the Spanish Ministry of Science & Innovation (MICINN), the Departament d’Innovació, Universitats i Empresa (DIUE) of the Catalan government and the Universitat Politècnica de Catalunya (UPC). The voting representation is divided between MICINN (51%), DIUE (37%), and UPC (12%). Overall governance of the BSC-CNS is provided by the Board of Trustees, formed by members of the three institutions that are partners of the BSC-CNS Consortium, and will be further supported by the Scientific and Business Advisory Boards (still in formation). Strategic direction is provided by the Executive Commission and this devolves to day-to-day management via the Management Board. Reporting to the Management Board are the various scientific and support departments.

The main BSC-CNS Governance Boards are the following ones: Board of Trustees formed by members of the three institutions that are partners of the BSC-CNS Consortium (Spanish Ministry of Education and Science, Department of Education and Universities of the Generalitat of Catalonia, and the Technical

\textsuperscript{41} http://www.bsc.es/plantillaA.php?cat_id=5
University of Catalonia). With the aim to follow up and execute all activities of BSC-CNS, the Executive Commission was constituted.

Furthermore, the BSC-CNS is headed by a director appointed by the Board of Trustees. The decision about the scientific use of MareNostrum is taken over by the Access Committee, formed by prestigious Spanish scientists external to the BSC-CNS.

20% of MareNostrum is reserved for internal research in BSC, and the rest is available for science research teams, companies and for contributing to several supercomputing infrastructures: at a national level the Spanish Supercomputing Network (RES, Red Española de Supercomputación) tries to provide the computing resources needed by researchers in Spain and Europe and the Spanish e-science Network coordinates the scientific activity in Spain by allowing the collaborative usage of geographically distributed resources. At an international level, BSC is involved in the European projects PRACE, DEISA and VENUS-C, in which BSC will contribute with two types of resources: virtualized commodity resources and supercomputing resources.

**Beneficiaries**

- Researchers from the departments of BSC and Spanish and international researchers through an evaluation procedure. The Access Committee allocates more than 25 million computational work hours every four months. Many scientific projects often request numerous periods of access to the RES42 in order to perform different work activities (new activities or continuation activities). Each request is treated separately and must pass the evaluation procedure of the Access Committee.
- The BSC-CNS has played a leading role in the development of PRACE, and as one of the founding host sites for a Tier-0 node will reap significant advantages, including adding value to other Spanish infrastructures, strengthening the interface between science and industry through interdisciplinary research contacts, attracting young and enquiring minds, creating technology clusters of associated industries and attracting high-tech firms to install R&D facilities nearby, spawning new spin-off products and start-up companies, increasing Spain's international reputation and visibility in scientific and high-tech fields, and improving the trans-European and public-private mobility of researchers and new technologies.

**Services**

The Operations Department of BSC ensures the continued availability and accessibility of RES systems 24 hours a day, 7 days a week and to provide support to all the users of the RES. Further core objectives are to manage upgrades to the MareNostrum and other RES nodes; facilitate access to RES facilities, including online electronic applications, remote access, and porting of code; manage the environmental aspect of the BSC-CNS installations; manage the technical aspects of integration of the MareNostrum in the DEISA and European HPC network grids; and ensure that RES staff receive appropriate training and

---

skills development in order to be able to professionally carry out their duties in an environment of constant technological change and advancement.

The Operations Department organises technical seminars where engineers share knowledge and experience, for example with manufacturing companies in order to inform attendees about their products and technical processes.

**Sustainability**

With reference to the BSC financial revenues, the consortium income of €20,126,170 recognised for fiscal year 2009 derived from public administration contributions from the Ministry of Science and Innovation (MICINN), Generalitat de Catalunya (GdC) and European Commission (EC), as well as from agreements, contracts or other collaborative agreements with private organisations. Furthermore, the consolidated budget of the BSC-CNS included the assignment of internal resources carried over from the previous year’s provisions.

Concerning the predictions of the BSC sustainability model, the supercomputer at BSC will be upgraded at the end of the year according the usual 3 years upgrading plan. Also, the acquisition of a new supercomputer is foreseen in 4 years.

### 2.6 e-Infrastructures sustainability’s conclusions

The e-IRG recommendations on sustainability, clearly identify the need to develop policies and mechanisms to encourage increased investment in a more coherent and interoperable way across Europe. The e-Infrastructure environment, in which we include also the cloud, must inter-operate and adopt international standard services and protocols in order to qualify for funding and increase the sustainability of the projects, also in the scientific context.

In order to make investments in e-Science more efficiently, cloud technology should be improved and the European Commission has to expand the adoption of e-Infrastructures, increasing European competitiveness.

The e-IRG states the relevance of a rich service-oriented e-Infrastructure economy which can be sustainable and support a wide variety of applications, services and user communities. To achieve this objective, e-IRG recommends continuing interoperability benchmarking through global standardisation.

An e-Science cloud infrastructure, in order to be sustainable, has to:

- Expand the participation of user research communities.
- Increase virtualization, in order to improve availability, accessibility, efficiency, and cost-effectiveness of e-Science cloud platforms.
- Simplify access, transparent service offerings, customized support, improved governance models.
- Facilitate the co-operation of different e-Science cloud communities.
• Develop a partnership between private and public sectors in order to create a market for e-Science cloud infrastructure, attracting new investors.

The EGI basic sustainability model is based on national and regional funding for resources and develops synergies to increase e-Infrastructures efficiency. The EGI social sustainability model is primarily based on the collaboration and sharing relevance, which is also the same objective of the VENUS-C project. The interest of e-Science in interoperability and cooperation increase the cloud adoption. The wide acceptance of the cloud model is fundamental for the development of the related sustainability model. Concerning the economic sustainability, to maintain a high-quality of EGI’s e-Infrastructure, recurring operational costs has been transferred to the communities which benefit from the services. On PRACE opinion, to develop a long-term sustainability, e-Infrastructure projects have to first analyse the demand and provide unique and comprehensive high-quality services. Furthermore, the most important point to consider is to continuously adapt the services in accordance with the evolving needs of e-Science communities. PRACE also recognize the relevance of the environmental sustainability. In order to reduce carbon footprint, PRACE is investing into energy-efficient HPC technology, in cooperation with vendors and other technology providers. All the e-Science communities should consider the environmental impact as an important factor to manage.
3. MARKET ANALYSIS FOR e-SCIENCE CLOUD ADOPTION

In order to better describe the target market applicable for the e-Science, it is useful to analyse also the different actors involved, the procurement process and best-practices, the usage schemas for local, national or European resource providers (when available) and the potential benefits of e-Science clouds, as enabled by the VENUS-C platform. The market analysis includes a SWOT analysis of e-Science cloud, also including references to the results reported in Deliverable D 7.

3.1 Cloud computing benefits and potential issues for e-Science

Cloud computing technologies improve existing IT technologies and bring a lot of social, economic, environmental and technological benefits. It is useful to understand how Europe can benefit from these technologies and implement them in all the society sectors, especially in the growing e-Science Cloud market. The pure technological benefits of cloud technologies, such as flexibility, scalability, easy use, greater utilization rates and efficiency of provision, have higher impact on e-Science clouds because they allow to eliminate CAPEX, increase OPEX and improve ROI, which are attractive propositions for new large scientific communities. In the current economic scenario, these capabilities allows companies to construct new systems without the need of huge capital investments in new IT infrastructure and further to grow the new systems to meet ever-changing business requirements at a rapid pace, which undoubtedly is a drift for innovation.

Another relevant benefit of e-Science cloud is the development of virtual ownership resources and user-friendly interfaces, which improve Identity and Federation management mechanisms for large e-Science cloud communities. According to OECD, cloud computing technologies contributed to handle the economic crisis with high-speed internet, green ICTs and smart applications, increasing ICT specialists employment. Moreover, clouds develop a “Green ICT” technology which may really reduce carbon footprint, creating a new high value for the specialist scientific market. Thanks to the virtualization feature, cloud technologies also allow to develop a new way of working and doing research improving collaborations among users. This is particularly relevant in the sector of health, education, research and university.

The cloud computing profitability is based on economies of scale which, as a result, leads to better optimized rates. The huge business challenge for the future is the development of pay-per-use licensing, which will certainly increase the ICT demand and the global cloud ecosystem, made secure from interoperability standards, especially concerning large research communities. However, the most relevant and interesting benefit of cloud adoption for scientific communities is the opportunity to use similar configurations set up to reproduce and share research results.

---

Indeed, research groups need to share data with other collaborators and using a cloud infrastructure, many large-scale datasets, which are hard to manage by a single group, can be efficiently used by scientific research groups. Obviously, like any other technology, also cloud computing it is not immune from defects and, a relevant concern relate to legal and ethical issues such as confidentiality, privacy, identity & access management policies and data security. As we stated in D3.7, the formulation of a common international regulatory seems to be an appropriate measure to address these concerns.

Custom-made contracts with specific target of individual or community users, should be deployed in order to satisfy the specific needs of the e-Science cloud customers and possibly eliminate these legal issues.

Moreover, we see an urgent need to plan global audit and compliance processes applicable to all kind of cloud services and applications. From the business perspective, the most relevant weakness of cloud infrastructures is the lack of a validated and widely accepted business model concerning pricing, ROI, cost structure, TCO vs. outsourcing comparison and predictions about eventual costs of service interruption or disruption. Furthermore, another considerable issue is related to marginal costs: indeed, depending on the volume of data and compute, operating on a cloud provider’s infrastructure may become more expensive than providing the necessary IT infrastructure in-house. A mixed-use strategy, in which some of the applications and services are delivered in-house and others continue to be hosted in the cloud, should be considered in order to allow a slow but safe migration of the users and scientific-Science communities to the cloud. The cloud weaknesses can have a negative impact on e-Science because they can affect the scientists’ reputation, trust and confidence due to data loss, account phishing and fraud, and cause even more damages for the researchers at social and personal level. To date, there are no developed model to quantify the economic and social cost played up by these losses. Furthermore, at the economic level, an exact forecast method to predict economic performances of cloud technologies has not yet been developed. This issue makes the economic evaluation process of e-Science cloud products even more cumbersome. However, the most outstanding problem for a fast adoption of cloud computing infrastructures in the scientific environment is the loss of data control due to the cloud abstractions. The success of performance management issues for scientific applications will depend on what minimal abstractions can be exposed to users, in order to enable the needed performance behaviours and be adequately controlled. Unfortunately, scientific communities need high-level abstractions to achieve their research goals.

A very important challenge for the success of e-Science cloud communities will be the development of high-level abstractions, which will ensure a good level of security and control for the researchers. Craig Lee, in his study “A Perspective on Scientific Cloud Computing”⁴⁴, to reduce the issues related to e-Science clouds, identifies the following necessary requirements:

---

• “Applications be transferable out and back into the same cloud and between different clouds, and still retain desired properties, such as performance, numerical stability, monitoring for fault detection/resolution, etc.

• A registry or brokerage be available for the discovery of available resources and service level agreements for using them.

• Support for virtual organizations across different science clouds and national clouds.

Support for virtual missions and virtual data centres allocated out of different science clouds and national clouds”.

Lee states that the most relevant issue to be solved is how to prioritize these requirements and structure their development for the benefit of scientific goals.

The success of cloud infrastructures stems from the huge benefits that the platform generates, such as:

1. Easier access, faster cycle times, ease of deployment, optimised utilisation, greater utilisation rates.
2. Enhanced business continuity and refocusing IT on business value.
3. Best-in challenge technology adoption, reuse and share of resources.
4. Greater efficiency of provision due mainly to the scalability of services.
5. Flexibility, allowing to create new computing resource to experiment with.
6. More rapid and increased ROI or time to value with lower upfront investments.
7. Reduced development, delivery and operation costs, support and maintenance.
8. Agility, reducing time to market. Lower support, develop, deploy, operate or change costs.
9. Reduced and/or removal of capital expenditure allowing operational expenditure savings.

According to Pierre Audoin Consultants, who analysed the “Economic and Social impact of software & software-based services” 45 in Europe in 2010, to clearly understand the cloud computing market, it is useful to first identify drivers and barriers of cloud adoption.

Drivers:

• Commoditization of software products: cloud computing is one of the major factors which will increase the commoditization of the software product together with Open Source Software.

• Crises-related topics: cloud computing models offer interesting value propositions such as cost reduction, transition from CAPEX46 to OPEX47, flexibility, simplicity...

• Cloud computing provides similar value proposition as infrastructure or application outsourcing, just in a more efficient way.

---

46 Capital Expenditure is the amount a company spends on buying fixed assets, other than as part of acquisitions.
47 Operational Expenditure is an ongoing cost for running a product, business, or system.
• Industrialization of the IT market, which is becoming, due to the relevant infrastructure needs and similarly to the telecom or utility industries, a more capital intensive industry in similar fashion as the telecom or utility industries.
• Increase of Services Oriented Architecture adoption: SOAs are key enablers of the SaaS and PaaS models.

Barriers:
• Magnitude of the data centres investments.
• Open standards.
• Data security and privacy.
• US cloud computing software companies.
• Scarcity of bandwidth and computing
• Change management: the changing enterprise investment habits is not immediate and takes time.

3.2 e-Science cloud SWOT analysis

<table>
<thead>
<tr>
<th>STRENGTHS</th>
<th>WEAKENESSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>• flexibility</td>
<td>• data security breach</td>
</tr>
<tr>
<td>• scalability</td>
<td>• lack of a single digital market</td>
</tr>
<tr>
<td>• easy access</td>
<td>• no single certification model</td>
</tr>
<tr>
<td>• faster cycle times</td>
<td>• no disaster recovery model</td>
</tr>
<tr>
<td>• optimised utilisation</td>
<td>• loss of control on data exchanging</td>
</tr>
<tr>
<td>• greater utilisation rates</td>
<td>• lack of cloud customer oriented SLA</td>
</tr>
<tr>
<td>• virtual ownership of resources</td>
<td>• scarcity of bandwidth and computing</td>
</tr>
<tr>
<td>• no CAPEX, less OPEX</td>
<td>• data privacy and confidentiality issues</td>
</tr>
<tr>
<td>• more rapid and increased ROI or time to value</td>
<td>• multiple access and authentication problems</td>
</tr>
<tr>
<td>• enhanced continuity</td>
<td>• Intellectual Property Rights issue</td>
</tr>
<tr>
<td>• greater efficiency of provision</td>
<td>• no single audit and compliance processes</td>
</tr>
<tr>
<td>• no licensing of subscription</td>
<td>• low level of due diligence process</td>
</tr>
<tr>
<td>• agility, reducing time to market</td>
<td>• no standard pricing models, cost structure and ROI model yet available</td>
</tr>
<tr>
<td>• software products commoditization</td>
<td>• lack of a single regulatory framework on cloud aspect, also especially in the scientific field</td>
</tr>
<tr>
<td>• reduced application integration costs</td>
<td>• no monitored system environment for e-Science cloud customers</td>
</tr>
<tr>
<td>• pricing based on consumptions or use (pay-per-use)</td>
<td>• lack of custom made contracts</td>
</tr>
<tr>
<td>• reuse and share of resources</td>
<td>• higher marginal costs</td>
</tr>
<tr>
<td>• multiple service options (the three layers)</td>
<td>• economic and legal implications of service interruption or disruption</td>
</tr>
<tr>
<td>• further adoption of SOAs</td>
<td>• abstraction vs. control</td>
</tr>
<tr>
<td>• reduced development, delivery and operational costs</td>
<td>• cloud ethical issues</td>
</tr>
<tr>
<td>• lower barriers to scientific applications development</td>
<td></td>
</tr>
<tr>
<td>• Reduced carbon footprint</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OPPORTUNITIES</th>
<th>THREATS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• allowing create innovation in all sectors</td>
<td>• data audit-ability</td>
</tr>
</tbody>
</table>
• developing new ICT jobs and increase employment
• building data observatories
• supporting reproducible research results
• increasing IT services demand
• developing new SME and start-up business, due to improved ROI and less up-front investments
• implementing and disseminate green ICT
• pay-per-use licensing development
• improving Identity and Federation management mechanisms
• improving data sharing
• developing a new way of doing research
• creating a global cloud ecosystem for interoperability and standardization of scientific services
• spreading virtualization and collaboration

• jurisdictional issues
• account phishing and fraud
• loss of reputation and gains
• performance unpredictability
• data lock-in and data vulnerability
• significant cost of moving to the cloud
• lack of trust confidence on cloud infrastructures

Table 2 - Cloud computing e-Science SWOT Analysis

3.3 Cloud computing main actors

Before to starting the analysis of the cloud computing market, it is important to identify the cloud market actors, who play primary and significant roles within this context. In Deloitte opinion\(^\text{48}\), there are four main actors:

- **Consumers:** including enterprises and individuals. Consumers are relevant to set standards and drive the aggregation of players in mature context, that actually the cloud not yet develop.
- **Regulators:** should identify and reduce cloud risks, in order to increase the benefits and empower the spread of cloud computing infrastructures.
- **Integrators:** act to support the users to internalize new services and products. They allow an easier and faster adoption of new technologies by customers.
- **Providers:** deliver cloud services to consumers. Within the current cloud market, we notice small and large providers, which driven technology innovations to customers.

Cloud computing markets also need *intermediates* (brokers) in order to reduce the service transaction costs\(^\text{49}\) and simplify the transactions for both service buyers and providers. According to Deloitte, the figure below illustrates the main cloud actors, divided into three different market segments and highlights the possible competitors. The three different groups, are broken down into further sub-segments:

---


\(^{49}\) Transaction cost: cost incurred when buying or selling securities. These include brokers’ commissions and spreads (the difference between the price the dealer paid for a security and the price at which it can be sold).
To date, it is not so simple and intuitive to distinguish between the different cloud providers and actors; indeed it is increasingly common to profit from various cloud vendors’ offers and use a mixed strategy in SaaS, IaaS and PaaS cloud platform services. In the following table, starting from the previous table of Capgemini\textsuperscript{51}, we summarise the different cloud offers that were available in 2010.

![Cloud computing families and sub-segments/offerings](image)

**Figure 1 - Cloud computing families and sub-segments/offers**\textsuperscript{50}

<table>
<thead>
<tr>
<th></th>
<th>SaaS</th>
<th>PaaS</th>
<th>IaaS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Servers</td>
<td>Storage</td>
<td></td>
</tr>
<tr>
<td>Oracle on Demand</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Cisco Webex</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>IBM Louts Live</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Salesforce Force.com</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Google App Engine</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Microsoft Azure</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
</tbody>
</table>

\textsuperscript{50} Ibidem
Table 3 - Enterprise cloud vendors’ offers (as of 2010)

<table>
<thead>
<tr>
<th></th>
<th>Amazon EC2</th>
<th>Rackspace</th>
<th>Eucalyptus</th>
<th>OpenNebula</th>
<th>RedHat OpenShift</th>
<th>Engineering CLOE</th>
<th>Telecom Italia Nuvola</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>x</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>√</td>
</tr>
</tbody>
</table>

Once the cloud computing reference scenario and related actors are analysed, it is necessary to evaluate the estimated market share of these cloud providers by analysing the worldwide cloud market and also the European one.

3.4 Cloud Computing market research for e-Science

The revolutionary changes in how users buy and providers deliver IT services were predicted by many authors in the last decade. The IT market is still in this transition and Gartner states that it will be entirely fulfilled by 2015, when industrialized services will represent more than 30% of the IT service market. In 2001, Gartner formulated the “Two axes and the four worlds of services” scheme, to classify the IT market offers. In this graph we see the delivery and the number of people that use the service on the horizontal axis and on the vertical axis the value, from IT efficiency to business outcome.

---

52 Gartner, (2010), “The 2001-2010 reshaping of the IT services market: was Gartner right 10 years ago?” 29 July 2010, p.4
The two axes (delivery and value), create four quadrants, which are named with the four terms *optimization, creation, management* and *access*. The changes over the years illustrated in figures 2 and 3 create a visual representation of how the market and the services have evolved.

In ten years, the offers across the four quadrants have changed frequently. To date, we are in the “fourth wave” of industrialized services. Within this context, cloud computing is positioned as the “third wave” of IT industrialization. During the “second wave”, in 2005, the most relevant innovation was the appearance of the data center in the *Management* quadrant. ASP, together with SaaS Application Utility, has migrated from the *Access* quadrant to the *Creation* one.
However, in 2009, the “third wave” of IT services industrialization represents the biggest change by showing a really complex and differentiated market.

The next figure illustrates the main forces driving these revolutionary changes:

- Commodity and standardization.
- Business-oriented and business-led IT spending.
- Differentiation to commodity.

These three forces allowed the users’ migration from:

- Traditional to industrialized/lower-cost services.
- From non-differentiating IT, to value-added IT spending.
- Differentiation back to parity and commoditization.
Currently, the market is experiencing the “fourth wave”, which power is going to increase with high acceleration, having achieved the point of no-return and more than 50% of the expected impact. The market Creation quadrant will be the major competing battleground for the IT industries, due to the high value of its services and growing business revenues.

In 2010, according to the Gartner “Hype Cycle for Cloud Computing”, this kind of infrastructure is in the “Peak of Inflated Expectations” phase and it starts to get into the “Through of disillusionment” one. Cloud infrastructure has not yet completely reached the maturity phase, however its adopters have already taken advantage of its benefits. The cloud market is growing quickly, the “big players”, such as Microsoft, IBM, Amazon, are providing new cloud computing services. Many companies are also trying to expand their business to more than one market, to increase the interest for cloud computing adoption also in the educational, health, institutional and scientific environments.
Furthermore, the cloud market successfully starts to embrace open source projects, benefitting from lower costs and initial investments. Gartner predicts that the market for cloud products and services will move from US$46.4 billion in 2009 to US 150.1 billion in 2013.

The following table shows Gartner’s sizing of the market for cloud services in 2008 and a forecast through 2013 at a worldwide level.

<table>
<thead>
<tr>
<th>Business Process Services</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>CAGR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud-Based Advertising</td>
<td>28</td>
<td>33</td>
<td>38.9</td>
<td>47.4</td>
<td>59.2</td>
<td>76.9</td>
<td>22.1</td>
</tr>
<tr>
<td>E-Commerce</td>
<td>1.3</td>
<td>1.8</td>
<td>2.5</td>
<td>4</td>
<td>7.2</td>
<td>12</td>
<td>56</td>
</tr>
<tr>
<td>Human Resources</td>
<td>7.5</td>
<td>8.9</td>
<td>11.3</td>
<td>14.1</td>
<td>16.2</td>
<td>18</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>CAGR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payment Processing</td>
<td>0.3</td>
<td>0.5</td>
<td>0.75</td>
<td>1</td>
<td>1.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Others</td>
<td>1.8</td>
<td>2.4</td>
<td>3.5</td>
<td>4.9</td>
<td>7</td>
<td>9.8</td>
</tr>
</tbody>
</table>

| Business Process Services Total | 38.9 | 46.6 | 57.0 | 71.4 | 91.5 | 119.3 | 25.1     |

| Applications Total          | 5.04 | 6.52 | 9.6  | 11.4 | 14.6 | 20.2 | 32       |

<table>
<thead>
<tr>
<th>Application Infrastructure</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>CAGR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform Infrastructure</td>
<td>0.05</td>
<td>0.07</td>
<td>0.09</td>
<td>0.13</td>
<td>0.2</td>
<td>0.4</td>
<td>51.6</td>
</tr>
<tr>
<td>Integration Services</td>
<td>1.47</td>
<td>1.54</td>
<td>1.62</td>
<td>1.7</td>
<td>1.78</td>
<td>1.86</td>
<td>5</td>
</tr>
</tbody>
</table>

| Application Infrastructure Total | 1.52 | 1.61 | 1.71 | 1.83 | 1.98 | 2.26 | 8.3      |

<table>
<thead>
<tr>
<th>System Infrastructure</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>CAGR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Services</td>
<td>0.66</td>
<td>1.17</td>
<td>2.00</td>
<td>3.40</td>
<td>4.90</td>
<td>6.80</td>
<td>59.5</td>
</tr>
<tr>
<td>Storage Services</td>
<td>0.009</td>
<td>0.025</td>
<td>0.084</td>
<td>0.241</td>
<td>0.525</td>
<td>0.750</td>
<td>52.8</td>
</tr>
<tr>
<td>Backup Services</td>
<td>0.30</td>
<td>0.37</td>
<td>0.45</td>
<td>0.55</td>
<td>0.67</td>
<td>0.82</td>
<td>22.3</td>
</tr>
</tbody>
</table>

| System Infrastructure Total | 0.96 | 1.56 | 2.53 | 4.19 | 6.10 | 8.37 | 53.8     |

| Infrastructure Total        | 2.6  | 3.4  | 5.0  | 6.0  | 8.1  | 10.6 | 33.5     |

| Cloud Services Total        | 46.4 | 56.3 | 70.8 | 88.8 | 114.2 | 150.1 | 26.5     |

Note: Technically, the cloud services market is a composite or “meta” market; that is, it is an aggregate of other categories that have the true characteristics of market.
According to these predictions, the most relevant feature of the cloud service market is the cloud-based advertising which, in 2008, represented 60% of the cloud services market and by 2013 will still represent 50%.

While the actual economic downturn has certainly led to a reduction of the IT investment, cloud computing services are still attracting the attention of a growing number of investors because of its benefits. Considering the cloud market, it is useful to evaluate the different segments, which show differentiated adoption levels.

According to Deloitte, SaaS will be the higher contributor for the overall revenues. Indeed, the SaaS market accounts in 2009 to $8.1 billion worldwide, which represents 7.7% of total enterprise application revenues and 89% of the cloud market.

Due to this favourable trend, the SaaS market will reach $17.8 billion in 2013, which represents a CAGR of 17.2%. We need to specify the actual different software segments in SaaS, which show different revenue rates. Within this context, it is useful to take a look at the following table of Gartner, which illustrates the main trends in the SaaS segment.

<table>
<thead>
<tr>
<th>Main software segments</th>
<th>SaaS revenue in billions of dollars (USD)</th>
<th>CAGR</th>
<th>Share of SaaS of the total enterprise application market</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2008</td>
<td>2013</td>
<td>2008</td>
</tr>
<tr>
<td>CCC</td>
<td>2.16</td>
<td>5.07</td>
<td>18.7%</td>
</tr>
<tr>
<td>CRM</td>
<td>1.84</td>
<td>4.02</td>
<td>17.0%</td>
</tr>
<tr>
<td>Int.aaS</td>
<td>1.47</td>
<td>1.86</td>
<td>4.8%</td>
</tr>
<tr>
<td>ERP</td>
<td>1.26</td>
<td>1.96</td>
<td>9.3%</td>
</tr>
<tr>
<td>SCM</td>
<td>0.75</td>
<td>1.65</td>
<td>17.1%</td>
</tr>
<tr>
<td>DCC</td>
<td>0.07</td>
<td>0.37</td>
<td>39.1%</td>
</tr>
</tbody>
</table>

Table 5 - Main SaaS segment trends

---

53 Gartner, (2009), “Forecast: sizing the cloud; understanding the opportunities in cloud services”
55 Gartner, (2009), “Forecast: sizing the cloud; understanding the opportunities in cloud services”
According to Gartner’s analysis, this table shows the greater speed growth SaaS segments. Between these segments, the fastest growing are Digital Content Creation (DCC) and CCC (Content, Communication and Collaboration).
In 2013, the Customer Relationship Management market, with the subarea of sales, marketing and services, is expected to contribute 30% to the total CRM market. Sales will still remain the largest contributors, 70% of the SaaS revenues. The much profitable segment in Content, Communication and Collaboration (CCC) market in the SaaS applications are represented by web-conferencing and e-learning. The Supply Chain Management segment is expected to grow and reach 16% of the SaaS market in 2013, and this success is primary related to the on demand services (6%) versus on premise solutions. However, SaaS will decline, leaving the scene for IaaS and PaaS solutions.
Many SaaS vendors, as a matter of fact, have already begun to switch to a PaaS solution. Gartner states that, in the next five years, applications in the cloud will be a mix of existing and new solutions. By 2013, only the PaaS, IaaS and Infrastructure Utility Market is supposed to generate more than $130 billion in annual spending. According to Gartner (2009), IaaS revenue was 969 million USD in 2008 and is expected to increase to more than 8 billion USD in 2013, which represent a CAGR of 53.6% and a share of 31% of the total cloud market.
Instead, PaaS, is in the early-stage of evolution in the market with revenue of 50 million USD, which represents 1.5% of the total application development market. Anyway, Gartner’s predictions highlight that the PaaS market, due to Independent Software Vendors leveraging PaaS offerings to reduce their costs, is expected to show a high growth and to reach 400 million USD by 2013. The figure below represents the cloud computing market segmentation in Europe in 2007 and the predicted market up to 2012.

---

The figure confirms the assumptions mentioned previously. The past situation for the years 2007, 2008, 2009 and 2010 shows a slow revenues increase in the IaaS and PaaS segment, while SaaS total revenues have reached a fast and exponential development intended to increment again over time. Infrastructure-as-a-Service has reached great interest also from the investment perspective. The most relevant innovation is the success of Infrastructure Utility within the outsourcing model, which is becoming always more used as a IaaS cloud offer, due to the flexibility and scalability offered by the cloud platforms. Platform-as-a-Service is still in the early stage of maturity within the cloud environment. However, some authors argue that PaaS has the potential to overwhelm in the future even the success and popularity of Software as a Service solutions (SaaS). This representation concerns only the European cloud market, but to identify the best approach to apply to the e-Science cloud communities, it is necessary to think in global terms. According to Pierre Audoin consultants opinion by 2013, cloud computing will represent 4.3% of the total software & software-based services share. Starting from the success of a “private cloud” market up to 2013, cloud providers will try to solve security and legal issues by developing more sophisticated SLAs. European public cloud computing will emerge slowly and will benefit from the collaboration with telecom operators and U.S. based IT companies. At this stage of the market expansion, European ASP and U.S. Software-as-a-Service players will still prevail. The predictions for the market penetration in 2020, suggest that cloud computing will represents 12.1% of the SSBS market and IT architecture will be mainly provided by European enterprises.

3.5 Cloud computing pricing models

Within a business analysis, pricing is a very relevant element to be considered and supposed to, influence the business model. All the pricing models described subsequently are related to three different contract providers’ models:

- Short-term contract: when users can buy the service at anytime.
- In-house transaction: buyers own the infrastructure.
- Long-term contract: a mix between short-term contract and in-house transactions.

---

Lei Han\textsuperscript{59} has analysed the different cloud computing market structures and has represented in the figure above the detailed percentage of Service Providers’ contract models in the current cloud computing market. With regard to transaction costs, Williamson’s theory states that they are influenced mainly by three parameters:

- **Asset specificity:** “the degree to which durable investments that are undertaken in support of particular transaction, the transaction-specific skills and assets that are utilized in the production processes and provision of services for particular customers. There are four different types of asset specificity: human, physical, site specificity and dedicated assets, produced for particular clients.
- **Uncertainty:** the cost associated with explaining and understanding products.
- **Frequency of transaction:** whether the transactions are occasional or recurrent”.

The relationship between these parameters determines the transaction costs, which are specific and differentiated for each kind of market structure. As represented in the figure below, the relationship between asset specificity, transaction costs and different market structure, shows the distribution of the different the different factor specificity and transaction costs for the public, hybrid and private cloud models. The factor specificity is defined by Williamson as “the degree to which durable investments that are undertaken in support of particular transaction (...) are utilized in the production processes and provision of services for particular customers”\textsuperscript{60}. Williamson classified the factor specificity into four types:

- Human asset specificity.
- Physical asset specificity.
- Site specificity, by investments with great setup and relocation costs;

\textsuperscript{59} Ibidem
• dedicated assets, which are usually purchased or produced on special requirements of certain clients.

The three lines highlighted in the next figure, represent the transaction costs function for each of the three market structures. The transaction costs can be categorized into two types:

• **Ex ante transaction costs**: according to Williamson, they are “the costs of drafting, negotiating and safeguarding an agreement”\(^\text{61}\). For cloud computing applications which provide standardized services, there are less ex ante transaction costs because the service can be defined with few parameters and applications can benefit from economies of scale. However, in the case that the user needs custom-made services, these costs can be high.

• **Ex post transaction costs**: these costs are often caused by contract misalignments. For cloud computing services typical ex post transaction costs occur because of system outages that cause business losses. Moreover, for particular services (not standardized) the switching costs can be quite significant. Hence, according to Williamson’s assertion, users should prefer public cloud for services with low factor specificity and private cloud for services with high factor specificity.

The following figure illustrates the relationship between transaction costs and factor specificity of various cloud computing market structures.

![Figure 7 - The relationship between transaction costs and factor specificity](image.png)

According to Han’s opinion⁶², concerning the purchasing costs of cloud technologies, there are three pricing models applicable to cloud computing services:

- **PAYG model**: named “pay-as-you-go” or “usage-based price model”, users are charged according to their real and actual usage of resources. This model is becoming even more commonly implemented; however there are still some issues to manage, which could impact the PAYG success. For instance technical short comings of current accounting and billing implementations, the matching between price and costs; indeed is not very clear if it is an information good, for which the marginal cost pricing is zero, or if cloud technologies are not information goods.

- **Flat rate model**: users pay a fixed amount per time unit, regardless of real usage of cloud services and applications. This model benefits from inaccurate measurement of billing and accounting; however it provides no incentive for optimizing the resource allocation.

- **Mixture model**: it is characterized by the union of PAYG and Flat Rate models. Users pay a fee for resource usage within a pre-defined period. This model is based on an almost fixed payment and it does not vary in response to changes in use.

According to Shenker⁶³, there is no clear boundary between Flat Rate and PAYG price model. Considering a user purchasing a software package, the possible payment structure include a:

- One-time selling price, which he pays before using the software and covers all the purchasing costs for the user.
- Fixed price for a certain period of use.
- Unit price for the usage of the software calculated by the actual user numbers, which can be also combined with an annually, quarterly or monthly payment structure.

For hardware resources, the situation is similar. The payment structure includes a:

- One-time purchase price, which directly transfers the hardware to the user.
- Fixed price for a certain period of use.
- Unit price for the actual usage of resources calculated by CPU hours and used storage space.

Hence, the flat rate model is based on periodical payment and one-time purchase, while the PAYG model differs between hardware and software payment: the first is usage-based and the second is subscription based.

---

⁶² Lei Han, (2009), “Market acceptance of cloud computing – an analysis of market structure, price models and service requirements, April 2009, Universität Bayreuth, Bayreuth

The mixture model is often related only to hardware resources and is based on periodical fee with payment for extra use. The table below represents a brief overview and comparison of the major cloud vendors and their adopted pricing model.

There are various offers which differ from bare metal servers to entire virtual infrastructures customized for web applications. Some vendors own the physical infrastructure and others provide additional value such as system design and application packaging. Taking this representation into account, the most commonly used pricing model by these vendors is pay-as-you-go (PAYG); however, the monthly pricing model is also widespread.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Service type</th>
<th>Product</th>
<th>Pricing model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle on demand</td>
<td>Platform, Infrastructure,</td>
<td>Oracle on demand (Cnt), (S), (C)</td>
<td>PAYG</td>
</tr>
<tr>
<td></td>
<td>Service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cisco WebEX</td>
<td>Platform</td>
<td>Web Ex (Cnt)</td>
<td>Flat Rate</td>
</tr>
<tr>
<td>IBM Lotus Live</td>
<td>Service</td>
<td>Lotus Live (Cnt)</td>
<td>Flat Rate</td>
</tr>
<tr>
<td>Salesforce Force.com</td>
<td>Platform</td>
<td>Force.com (Cnt)</td>
<td>Flat Rate</td>
</tr>
<tr>
<td>Google</td>
<td>Platform</td>
<td>AppEngine (C)</td>
<td>PAYG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Service type</th>
<th>Product</th>
<th>Pricing model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft</td>
<td>Platform, Infrastructure</td>
<td>Azure (C,S), Dynamic Data Center</td>
<td>PAYG</td>
</tr>
<tr>
<td></td>
<td>(M)</td>
<td>Toolkit (C,S)</td>
<td></td>
</tr>
<tr>
<td>Amazon EC2</td>
<td>Infrastructure</td>
<td>EC2 (C), S3 (S), CloudFront (Cnt)</td>
<td>PAYG</td>
</tr>
<tr>
<td>RackSpace</td>
<td>Infrastructure</td>
<td>Cloud Server (C), Cloud File (S)</td>
<td>PAYG</td>
</tr>
<tr>
<td>Eucalyptus</td>
<td>Infrastructure (M)</td>
<td>Eucalyptus (C)</td>
<td>Consulting</td>
</tr>
<tr>
<td>C12G</td>
<td>Infrastructure (M)</td>
<td>OpenNebula</td>
<td>Consulting</td>
</tr>
<tr>
<td>RedHat OpenShift</td>
<td>Platform</td>
<td>OpenShift (SD)</td>
<td>N/A</td>
</tr>
<tr>
<td>Engineering CLOE</td>
<td>Infrastructure</td>
<td>Cloe (C), (S)</td>
<td>Flat Rate</td>
</tr>
<tr>
<td>Telecom Italia Nuvola</td>
<td>Platform, Infrastructure</td>
<td>Nuvola (C), (S)</td>
<td>Flat Rate or</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PAYG</td>
</tr>
</tbody>
</table>

Table 6 - Comparison of major cloud vendors

Legend

C: Compute
S: Storage
SD: System Design and Application Packaging
M: the vendor does not provide any infrastructure support but only the management software
Denne\textsuperscript{64} discusses different ways to implement the most popular pricing model, the PAYG one:

- **Time-based pricing (subscription pricing):** pricing is based on consumed time units. The difference with respect to the actual subscription pricing mechanism is that users do not sign a fixed contract.
- **Park-level pricing:** pricing is based on peak consumption within a defined window.
- **User-based pricing:** pricing is based on the number of distinct users presenting to the system.
- **Ticked-based pricing:** pricing is based on fixed price electronic tickets that services provider issues for use of the service (for a specific period of time).
- **Integral pricing ("under the curve"):** pricing is based on peak utilization of defined capacity unit divided by average utilization.
- **Overage charges:** pricing changes in the case the customer exceeds the average consumption of the service.
- **Consumption commitments:** pricing is based on estimated average consumption and exceeding or undercutting the consumption commitment affects the price.

In an effort to compare the different pricing models, Lei Han made a good analysis which highlights the superiority of the PAYG model, especially in case of low uncertainty, see figure 8 below. Indeed, at present, this model is chosen by 50% of users. With pay-per-use mechanisms, capacity units such as transactions, gigabytes of storage or memory or units per time or per hour are associated with resources and assigned with fixed price values. Hence, the user pays according to his metered usage of resources. The flat-rate model is less used by customers, due to the lack of flexibility and also because the payment is based on the predicted and not real usage of services and resources. Actually, customers prefer to pay for the effective usage, rather than for a fixed fee.

3.6 Service Level Agreements

The management of SLAs (Service Level Agreements) play an important role. A SLA describes the expected level of service between the consumer and the provider. The SLA should be clearly defined, otherwise it will be useless.

Many items have to be monitored as part of an SLA, but the scheme should be kept as simple as possible, to avoid confusion, excessive cost and allowing to choose the right metric.

Depending on the service, the types of metric to monitor may include:

- Service availability.
- Defect rates.
- Technical quality.
- Security.

One of the main contention points in negotiating an SLA is around outage credits and how they are applied. The most frequent questions are:

- Does the customer get a refund for the lost services or is the SLA applied to a future credit?
- How does a customer prove an outage to get credit?
- How does the credit get applied?

SLA management is a business challenge; it represents a critical part of any supplier agreement, and it will pay-off in the long-term if it is properly codified at the beginning.

---

Lei Han, (2009), "Market acceptance of cloud computing – an analysis of market structure, price models and service requirements, p. 46, April 2009, Universität Bayreuth, Bayreuth

---
The most diffused SLA model is named “99.999%” or “Five Nines”, which is a gold standard in the enterprise, equating to about 5 minutes of outage per year. However, sometimes, a strategic business cannot admit this kind of outage, hence service providers sell protection plans and addictive custom SLA plans, which cover 100% of service and sell extra support time.

In the following table we list the different SLA models adopted by the major cloud vendors:

<table>
<thead>
<tr>
<th>Vendor</th>
<th>SLA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle on demand</td>
<td>99.95</td>
</tr>
<tr>
<td>Cisco WebEX</td>
<td>99.99%</td>
</tr>
<tr>
<td>IBM Lotus Live</td>
<td>99.9%</td>
</tr>
<tr>
<td>Salesforce Force.com</td>
<td>N/A</td>
</tr>
<tr>
<td>Google App Engine</td>
<td>99.95%</td>
</tr>
<tr>
<td>Microsoft Azure</td>
<td>99.95% connectivity, 99.99% service</td>
</tr>
<tr>
<td>Amazon EC2</td>
<td>99.95%</td>
</tr>
<tr>
<td>RackSpace</td>
<td>99.99%</td>
</tr>
<tr>
<td>Eucalyptus</td>
<td>N/A</td>
</tr>
<tr>
<td>C12G</td>
<td>N/A</td>
</tr>
<tr>
<td>RedHat OpenShift</td>
<td>N/A</td>
</tr>
<tr>
<td>Engineering CLOE</td>
<td>N/A</td>
</tr>
<tr>
<td>Telecom Italia Nuvola</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 7 - The different SLA models

3.7 Conclusions

The market analysis described in this document has identified, through a specific SWOT analysis, the main benefits of e-Science cloud infrastructures, the actors involved, the most profitable pricing models, and a brief analysis of the current cloud market share. This assessment has been made in order to provide starting points for the final sustainability strategy, which will be developed in Deliverable 3.10.

According to the Gartner’s evaluation, the research is related to the estimated market share of different cloud providers. Our analysis can be useful for scientists who want to move to cloud infrastructures, giving a general overview of the different cloud pricing models. Finally a great interest within the cloud context is represented by Service Level Agreements’ management.
According to Gartner’s considerations, today the market is experiencing the “fourth wave” of IT services industrialization, which represents a relevant change by showing a really complex and differentiated market. The main forces driving these revolutionary changes: are commodity and standardization, business-oriented and business-led IT spending, differentiation to commodity. These three forces are pushing the users to migrate from: traditional technologies to cloud infrastructures, which results in a constantly growing market for cloud solutions.

Gartner predicts also that the market for cloud products and services will move from US$46.4 billion in 2009 to US$150.1 billion in 2013. The most increasing cloud sector is the SaaS, but also the PaaS one is evolving, positioning itself as the most promising segment for the future. These considerations can help researchers and scientists to understand the importance to move towards cloud offering PaaS services, as the VENUS-C cloud platform.

Our analysis also highlighted that the most profitable market structure currently available for scientists is based on flexible short-term contracts allowing users to buy services at any time. With regards to the different pricing models, the PAYG is the most profitable and suitable for researchers, because users are charged according to the actual usage of resources. As a result, this pricing model is the most commonly used.

Researchers have to consider all the following requirements, in order to identify the best gainful cloud provider to choose:

- Analyse the different pricing model and possibly adopt the pay-per-use one, based on a short-term contract;
- Move to PaaS offers, which develop custom-made applications for researchers;
- Evaluate if the cloud provider develop tailor-made SLA, also for groups and communities;

If researchers will take into account all these considerations within the scientific context, then it will be easier to choose the best-effective solution.
4. TOTAL COST OF OWNERSHIP OF SCIENTIFIC DATA CENTERS

The cost structure of data centers of service providers plays a key role in the analysis of a cost model for cloud computing. In the IT industry, the Total Cost of Ownership (TCO) is a fundamental instrument for acquisition and planning decisions and provides a cost basis for determining the economic value of an investment, by comparing lifetime costs associated to the different IT infrastructure scenarios.

TCO provides a realistic measure of the long-term costs required to acquire and operate technology solutions. Return on investment (ROI) is another method to evaluate and prioritize technology investments in a company. This measure is typically used to compare investments that uncover new revenue and growth opportunities. However, ROI tends to be more subjective in nature than TCO, because it evaluates the business benefits, which often cannot be measured as objectively as costs.

The traditional ROI formula does not consider time. To evaluate costs and benefits of a project, we must take into account that neither costs nor benefits materialize instantaneously, they could play out over time. Instead, the Net Present Value, can make this kind of assessment which is very useful for projects evaluations. Indeed, costs and benefits of a project will be evaluated over multiple periods, which form a time series of cash flows. According to James Downey opinion66, to estimate the Net Present Value of a project, we have to calculate the net of the cash inflows and outflows of each period and then discount each period’s net:

\[
Net \text{ Present Value} = \sum_{t=0}^{n} \frac{CF_t}{(1 + i)^t}
\]

In the chart above, time period zero represents the present, which falls at the very beginning of a project when the first costs are encountered. All other costs and benefits are discounted to their present value so as to be made comparable with those of period zero.

A limitation of NPV is that it does not in itself enable comparisons across projects with different levels of investment.

A TCO comparison provides a more tangible assessment of the total costs involved in deploying cloud business solutions. According to Hurwitz & Associates67, the TCO calculation model compares clouds and on premise business applications over a period of four years. The TCO model should consider the following components:

1. **Evaluation and electron**
   - Solution evaluation and functionality analysis of competitive products.
   - Vendor review and SLA analysis (if applicable), evaluate also security requirements.

---

Security requirements evaluation.

2. **IT Infrastructure hardware, software and support:**
   - Server and storage hardware and maintenance.
   - Operating system, database, security, backup software and maintenance.
   - Administrative IT costs for systems and database.

3. **Application subscription costs or application license costs:**
   - ERP and CRM application subscription or license cost: which are calculated considering an equivalent number of users. Within cloud solutions, these costs represent operational costs which remain constant for the life of the subscription service. For on-premise solutions, these are initial deployment capital expenses;
   - application maintenance: this represents operational costs for bug-fixes and upgrades to new versions of the software.

4. **Application solution deployment costs:**
   - Detailed design: this activity define project objectives and scope, existing processes and gaps, data sources, document business requirements and, develop final project plan.
   - configuration and deployment: this task requires the application configuration to specifications, integration between front and back office functions, custom integration if required, data migration, system testing, cutover to new solution.

5. **Initial and on-going training costs:**
   - User training: activity related to the provision of software training for end users, aimed to minimize any productivity losses associated with the software transition.
   - Administrative training: costs associated to transition daily system administration to the customer’s internal staff.

The study provided by Hurwitz and associates, which compare the TCO in the cloud vs. on-premise business solution for 100 user scenarios, clearly shows that the first year costs associated to the on-premise ones are much higher with respect to the cloud platform, since the first solution requires more upfront capital expenditure costs.

It is assumed that, in average, a typical data center is only used to 30% of its capacity\(^{68}\). However, the utilization of a data center varies during its lifetime. Fraction of utilization and time variation are important inputs to consider for the development of a TCO model. For this reason, it is useful to provide a long-term assessment, which includes at least a five years life cycle.

### 4.1 TCO calculator analysis approach

TCO refers to the complete set of costs associated with acquiring, deploying, managing and eventually phasing out any form of IT investment and data center is one prominent example.

---

\(^{68}\) APC White Paper 37, (2003), “Avoiding Costs from Oversizing Data Center and Network Room Infrastructure”.
The prediction and measurement of TCO is generally required for ROI analyses and other business decision processes. Though there are many models for owning and maintaining a data center, to date there are no recognized standards for measuring the related TCO.

In this paragraph, we describe the method we developed for determining the TCO of data centers. This assessment is useful to identify, in the next paragraph, the best model to choose between ownership and outsourcing. These starting assumptions will be exploited in the next Deliverable 3.10, aimed to describe the cloud model final sustainability analysis.

The primary objective of this study is to develop a valid and efficient TCO calculator model to analyze the cost of owning and operating a data center. In this context, both the equipment needed to provide power, cooling, physical protection of IT equipment and the IT equipment itself will be considered. In order to make the study as impartial and balanced as possible, a structured research approach has been developed including a step-by-step validation:

- Develop a TCO calculator containing general categories, which include more specific costs and percentage information.
- Collect data, synthesize and compile findings.
- Make assumptions and compare the TCO of different data centers in Deliverable D3.10, over a period of five years. A long-term assessment is needed in order to compare start-up costs with the related subsequent costs. The results of this analysis provide a valid support for any researcher, in order to assess future ownership costs.

The starting point is a definition of TCO, which for this study refers to “the cost for the owner to build the data center and the cost overtime to maintain and operate it”. Based on this definition, we established an evaluation method which considers start-up costs and a cycle of four years, analyzing costs generated in the entire implementation lifecycle, which will typically include acquisition, implementation, day-to-day operations management and eventual retirements.

The TCO metric used in our research approach is based on Cost/Server ratio (divided in traditional servers and blade servers). Within the power analysis of the data center, a Cost/kWh metric has been chosen to estimate the power availability of servers, blades, monitors and Cost/m² per rack. In order to develop an accurate cost analysis, the derived evaluation, expressed in a per-rack basis, implies the TCO allocation across the utilized racks of the IT infrastructure of a data center.

The first phase of the analysis, dedicated to the preparation of the TCO calculator, requires an assessment of the following five general cost categories:

1. **Infrastructure’s deployment costs**: this category includes all the data derived from the initial Capital Expenditures and those related to the infrastructure’s renewal within the next five years.

---

69 “The term “rack” refers to an open frame rack or rack enclosure and it also refers to proprietary cabinets such as mainframe computers and large disk systems”. Rasmussen N., “Determining Total Cost of Ownership for data center and network room infrastructure”, White Paper 6, APC
2. **Power costs**: this category estimates information based on the real kWh power consumed and identifies the total cost of power by crossing the percentage and the average of the total power costs (including green power).

3. **Human capital and training costs**: these elements are related to the internal and external staff costs, client and application porting and replacement costs.

4. **Management costs**: this category includes all the costs concerning the data centers’ administration, comprising also the security and disaster recovery costs.

5. **Risks costs**: are fundamental for the TCO analysis because their evaluation allows to calculate the total productivity lost costs. Even if, many TCO calculators do not include this indicator, we believe it is fundamental to take into consideration this cost category and comprehensively evaluate the true Total Cost of Ownership.

Within these general costs categories, we defined the following specific sub-categories considered relevant to the TCO estimation:

**Infrastructure/deployment costs**

- **Data center costs**: number of racks, data center space, cost per m², cost of wiring per rack, m² per rack, port costs per rack, construction costs, fire prevention and surveillance costs.

- **Server costs**: number of servers and racks, cost of single server and rack.

- **Blade pc costs**: number of blades, number of enclosures and racks, cost of blades, enclosures and racks.

**Power costs**

- **Server power costs**: power consumed per year while running by all servers, power consumed while stand by all servers, power consumed while turned off by all servers, power loss factor, total cooling power for servers (kWh).

- **Blade power costs**: power consuming per year while running all blades, power consumed per year while stand by all blades and turned off all blades, power loss factor, total cooling power per blades (kWh).

- **Monitor power costs**: power consumed per year while running by all monitors, power consumed while stand by all monitors, power consumed while turned off by all monitors, power loss factor, total cooling power for monitors (kWh).

- **Green power costs**: percentage and average of green power costs.

- **Carbon power costs**: percentage and average of carbon power costs.

**Human capital and training costs**

- **Implementation costs**: internal staff, outside consultant, special contractors.
• **Training costs**: number of IT staff, use training base cost, user training total cost, IT training total cost.

• **Application porting and replacements costs**: client application porting and replacement purchases.

**Management costs**

• **Inventory**: number of clients inventoried remotely and manually; cost per inventory for remote and manual inventory.

• **Patch management**: total annual patch distributions, total cost for a single distribution.

• **Support**: total number of calls/tickets, total time spent in minute for single ticket, cost of single hour.

• **Configure/add/delete**: number of services, cost of configure a single service, cost to add and delete a single service, cost to add and delete all services.

• **Security**: cost of unscheduled IT time per service.

• **Compliance**: cost of compliance per service.

• **Other manageability costs**: cost of managing all servers, cost of other service management functions.

**“Risks” costs**

• **Productive minutes lost per day**: average of productive minutes lost due to rebooting, to failures and other IT issues, to serve congestion, minutes lost due to restrictions of the environment, productive minutes lost during compute-intensive tasks.

• **Productivity**: percentage of productive time lost, cost of non productive time per employee.

The following figure summarizes the hierarchical subdivision and correlation between all the primary and secondary TCO components used in our assessment.
4.2 Ownership vs. outsourcing

According to Gartner\textsuperscript{70}, the growth of cloud computing technologies and related services, will have a profound impact on the outsourcing industry, during the next five to ten years. Thus, it seems relevant to examine the history of outsourcing, to better understand the role of the emerging cloud environment. At the beginning of the outsourcing phenomenon the focus was on the decision whether to build a solution, to acquire a product or outsource the development solutions.

The reasons that have pushed the companies to outsource some of its services ground mainly in economic benefits, such as cost flexibility and saving, technological advantages, innovation, strategic aims, business-oriented benefits and an increasing business quality and flexibility.

\textsuperscript{70} Gartner, (2010), “Cloud computing: the next generation of outsourcing”, 1 November 2010, ID number G00207255
According to Lee et al. (2003), the evolution of the outsourcing concept has gone through several phases:

- **Make or buy**: the choice between internally develop technology and its external acquisition.
- **Motivation**: the stage of the benefits and risks of the outsourcing study. Between these two phases, the “Kodak effect” in 1989 was instrumental for the development of the concept (outsourcing decision).
- **Scope**: development stage regarding to the degree, period of outsourcing, the number of vendors and the types of outsourcing.
- **Performance**: benefits phase concerning user and business satisfaction, service quality and cost reduction.
- **Insource or outsource**: trade-off between contingent factors in outsourcing.
- **Contract (formal)**: well-designed contract to reduce unexpected contingencies.
- **Partnership (informal)**: key factors and effective way for building outsourcing partnerships.

During the past fifteen years, the commoditization, virtualization, standardization of technologies and the success of the Internet and service-oriented software architectures, have modified the relationship between the IT vendors and the IT users.

The IT outsourcing markets expand more and more. Customers require cost-effective efficient and flexible IT services from the provider, at flexible costs models (like pay-as-you-go models). Cloud computing infrastructures satisfy these business demands, entering the traditional outsourcing value chain and competing with established outsourcing providers. The innovation of cloud computing outsourcing models resides in the pay-per-use payment and satisfies the effective needs of customers through producing cost savings. According to Böhm et al., in traditional IT service outsourcing, the value chain can be divided into the areas of infrastructure hardware network, applications data, business processes and strategy consulting models. Through these four parts, the value chain synthesizes a whole cycle of IT services (“plan, build, run”). Each step may be outsourced.

Cloud computing infrastructures link the service-oriented hardware outsourcing and the “as-a-service” software concept. New innovative cloud platforms allow integrating hardware and software as-a-service offerings, through a single interface. The cloud infrastructures are merging on-premise and off-premise services together with cloud applications, in order to provide the desired services to customers. As a result, enterprises are investing in more modular and energy-efficient data center designs.

---

72 Böhm M., Leimeister S., Riedl C., Krimar H., “Cloud computing – Outsourcing 2.0 or a new business model for IT provisioning?”
According to Gartner\(^7\), the benefits of cloud outsourcing are:

- Pay for what you use.
- Costs derive from an operational budget.
- More basic functionality that you really need not bells and whistles that you might need.
- No operational management worries.
- No infrastructure overhead/management.
- Midterm lower Total Cost of Ownership.
- Faster implementations.
- Easier integration.

Certainly, some of these points are debatable, in fact these benefits do not always appear using cloud infrastructures and we have to consider also some disadvantages, such as security concerns, long term TCO uncertainties, less central management of IT. Surely it is fact that cloud platforms can benefit from the value of the network, which is a relevant element to consider when choosing between traditional IT outsourcing and cloud outsourcing offers.

Within this context three actors play fundamental roles:

- The service provider, who develops applications deployed on the cloud platform.
- The platform provider, who is responsible for the provisioning of hardware and storage capacities to users.
- The infrastructure provider, who provides cloud applications.

Finally, the aggregator is another relevant player who combines pre-existing services with cloud solutions and provides a customized solution to customers. In the cloud world, the traditional IT outsourcing service is replaced by a network of different service providers, who offer a greater quantity of services and products in a very cost-effective way.

Any company may act in more than one role. The innovative characteristic of outsourced cloud offers is the flexible deployment of virtual and asset-free resources and services, which will enable IT service providers to deliver end-to-end services, regardless of the various platforms, applications and technologies involved. Hence, providers can break up the outsourcing value chain, position themselves and enter the market with a new service offering. Cloud-based resources will radically change the outsourcing business on the service providers’ perspective.

The cloud infrastructure does not focus more on the technological aspects, but on the business needs in terms of flexibility, availability and reliability. Due to this huge impact on the traditional outsourcing model, many outsourcing vendors are contrasting the cloud trends, but they should keep in mind, apart

\(^7\) Gartner, (2010), “Cloud computing: the next generation of outsourcing”, 1 November 2010, ID number G00207255
from the initial new obstacles and complexity, the inherent huge potential in terms of capabilities, cost savings and efficiency benefits.

Ram Prasad Kan, Chief Technologist at Wipro Technologies, states that cloud technologies are a potential game-changer: “They encompass infrastructure, platforms, applications and BPO services – and this “IT-as-a-service” model may create a whole new wave of outsourcing.”

4.3 CAPEX vs. OPEX in e-Science Clouds

Within an analysis of cloud computing benefits, the necessary transition from CAPEX to OPEX plays an important role. Indeed, we consider incorrect the current debate which argument a clear contrast and challenges between these two elements.

Cloud computing infrastructures are the key drivers for reducing the TCO as well as cost of operation and maintenance (OPEX) of IT services. Within the cloud environment, the payment may occur through the following three different procedures: pay-as-you-go, pay-per-user and pay-per-transaction. A customer pays only for services and for the time effectively used and he does not have to pay for upfront infrastructure investments. Using a cloud offer, which benefits from relevant economies of scale, for instance through a pay-as-you-go manner, the service falls into on-going operating expenditure and does not requires capital expenditures: thus is the “CAPEX vs. OPEX” contrast. Furthermore the OPEX can be controlled by the cloud customers, unlike CAPEX.

According to Greg Baker, there are essentially four strategies to move from CAPEX to OPEX; most of them are related to the cloud infrastructures:

- **Software-as-a-Service (SaaS):** applications eliminate front-loaded software and implementation fees. The user does not own the software and pay only for time or services that he really uses. This model affects benefits switching from CAPEX to OPEX.
- **Cloud computing:** through scalability, flexibility and capital expenditure avoidance, the cloud is the most important infrastructure that really could reduce economic risks for the customer. Indeed, without the need to own the capital assets, the user may choose and use different kind of services whenever he wants.
- **Outsourcing:** using this model, the business risks and volatility are under the responsibility of the provider. The customers, companies or institutions, shift the equipment, the infrastructure and the liability costs to other subjects.

---


- **Leasing**: the main advantage is that, by providing an alternative to ownership, it allows benefiting from not having resources invested in equipment and, as a result, leads to the migration of the asset costs into monthly operating costs.

In order to maximize benefits, IT operators should take into consideration both CAPEX and OPEX elements: it is primarily necessary to evaluate not only the technical implications, but also the financial ones. Making a long-term assessment of the OPEX costs, could reduce the necessity of large investments on behalf of annual cheaper OPEX budgets. Besides, by triggering this activity, the organization could develop more frequently and easily new technology opportunities further to the business flexibility.

The adoption of the cloud computing paradigm transforms cost of ownership and changes the dynamics of the provisioning cycle.

As reported by the Open Group\(^76\): “The speed and rate of change of cost reduction can be much faster using Cloud Computing than traditional investment and divestment of IT assets. In Cloud Computing the buyer can move from a CAPEX to an OPEX model through purchasing the use of the service rather than having to own and manage the assets of that service. This responsibility is transferred to the service provider”.

---


\(^77\) Ibidem
Figure 10 illustrates that the maximization of benefits for the users is related to his ability to adopt and remove the service either at the point of use (to scale up and down) or to make choices to use new services or change service provider.

The cloud billing model, mainly based on the PAYG pricing model, imply different rates and contractual obligations, compared to traditional IT ownership. This billing model gives the opportunity to scale the resources on the users’ needs and to enhance the level of services provided.

Within the e-Science cloud communities, there are some challenges migrating from CAPEX to OPEX. It is common that researchers, at the beginning of their experiments, do not have all the information to exactly quantify the compute and storage requirements needed. This situation could cause some risks for the scientists. Furthermore, a change of the platform or provision model will affect efficiency and costs. This inability to make a long-term prediction is a strong limitation to the different e-Science cloud communities, which should be assessed at the beginning of the research project. The only reasonable solution, seems to be the conclusion of highly flexible contracts between the cloud provider and scientists, who allows to amend the amount of storage and compute resources, according to the needs identified by the researchers during the experiments.

4.4 Conclusions

To date, considering all the different available approaches, the TCO represents the best model to assess the cost structure of a (scientific) data center. Indeed, with respect to the Return on Investment and the Net Present Value, the Total Cost of Ownership allows to provide a realistic measure of the long-term costs required to acquire and operate technology solutions. The traditional ROI formula does not consider time, but in order to evaluate costs and benefits of a project, we must consider that these may materialize in a long time. On the other hand, the Net Present Value approach evaluates multiple periods, but cannot be used to compare the project with different investments.

Within these evaluations, the TCO can provide more tangible crossed assessments. Considering these reasons we have decided to develop a TCO calculator in order to assess the total cost of ownership of scientific data centers in Deliverable 3.10.

The analysis highlighted that for scientists it is more profitable to use cloud computing infrastructures, which can better satisfy their business demand, entering the traditional outsourcing value chain and competing with established outsourcing providers. Through the pay-per-use payment, cloud providers satisfy the effective needs of customers, often producing high cost-savings.

Furthermore, we have also explained how cloud computing infrastructures can really eliminate Capital Expenditure and reduce cost of operation and maintenance (OPEX) of IT services. Through cloud outsourcing, which develops a long-term assessment of the OPEX costs, it is possible to reduce the necessity of large investments on behalf of annual cheaper OPEX budgets. Thus, organizations and research institutions are able to couple business flexibility with more frequent new technology opportunities.
Starting from these general considerations coming from an analysis of the literature and comparisons of ownership vs. outsourcing models, we would like to investigate, in Deliverable 3.10, how outsourcing generated by cloud platforms could be more profitable for scientists than hosting the infrastructure in house.
5. DRAFT SUSTAINABILITY STRATEGY OF e-SCIENCE CLOUDS

5.1 Introduction to the seven VENUS-C scenarios questionnaire

This section introduces an analysis which aims to evaluate the social, economic, environmental and technological sustainability of the seven VENUS-C scenarios. A cost-benefit analysis based on a multi-criteria methodology has been used to estimate benefits and costs of the e-Science cloud paradigm and the activities developed using the cloud platform.

<table>
<thead>
<tr>
<th>VENUS-C scenario</th>
<th>Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building Structural Information</td>
<td>Universidad Politécnica de Valencia</td>
</tr>
<tr>
<td>Building Information Management</td>
<td>Collaboratorio</td>
</tr>
<tr>
<td>Data fro Science</td>
<td>Consiglio Nazionale delle Ricerche</td>
</tr>
<tr>
<td>Virtual Fire</td>
<td>University of the Aegean</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>Universidad Politécnica de Valencia</td>
</tr>
<tr>
<td>Systems Biology</td>
<td>The Microsoft Research University of Trento</td>
</tr>
<tr>
<td>Drug Discovery</td>
<td>Newcastle University</td>
</tr>
</tbody>
</table>

Table 8 - VENUS-C seven users’ scenarios

The questionnaire is approximately twenty pages long and contains five sections:

1. **Introduction**: the first part of the questionnaire is organized to gather general information about the specific scenarios, in order to learn more about our partners’ activities. For example, we ask them to provide a description of the main problem that their scenario addresses. The so-called “baseline scenario” represents the ex-ante situation. Through the baseline scenario, it is possible to have a clear view of the value added by the cloud infrastructure to each scientific scenario. Researchers introduced and described one use case enabled by the project, in which substantial differences emerge adopting the cloud computing capabilities. This study also highlighted direct and indirect beneficiaries and the related expected impact of the VENUS-C project on scientists.

2. **Social impact**: this section is scoped to identify the impact of each VENUS-C scenario on the society. In particular, we estimated how much the use of cloud computing technologies increase the number of research activities and results and the quality of research in each community.

3. Furthermore, we identified, for each institution, the impact of the cloud platform on the
employment and knowledge creation. It is also useful to analyse the improvement on quantity and quality of scientific production and data available for the communities using the cloud platform.

4. **Environmental impact**: the section evaluates the environmental impact of each scenario on the society, such as savings in kWh, savings in consuming and selling off paper, films, CD, DVD, savings in storage items, reduction of technological waste and travels.

5. **Economic impact**: the main objective of this section is to make an economic assessment of the potential impact of the cloud on market and competition, investment flow, operating costs, regional development and improvement on the quality of research processes. Then, we estimate the cost reduction percentage of the following categories: hardware costs, maintenance, connectivity, development, software re-usability and “test-deploy-rework”-circle management. Finally, we investigate the cost reduction associated to the compliance with the regulatory framework and lower process system failures.

This analysis allows to identify the current market trend related to the scenarios’ outcome in terms of global market value, potential percentage of market share achievable, number of competitors. The competition assessment identifies, for each scenario, at least three potential major competitors. Furthermore, we evaluate the potential financial revenues of the seven scenarios, in order to estimate service sales, fees and pay per use approaches, royalties and other sources of founding.

6. **Technological impact**: in this final section we assess the technological impact of the cloud in terms of users involved in each scenario, the percentage of cloud utilisation and their goals. The metric is designed to evaluate the overall effectiveness with a view of a real use and is based on ISO 9126-278 and ISO 9126-479. The indicators used to evaluate the technological impact are: time frame analysis, activity time, number of activities completed, total activities, accessibility, error frequency, productivity, total cost of the activity, potential incidents to user and damages, correctness, data recorded, data loss, crash functionality and understand ability. Finally, we estimate the real adoption and relevance of “cloud service models” for the seven VENUS-C scenarios, using the Gartner “Hype Cycle for Cloud Computing” definitions and rated the cloud service models from 1 to 5 (where 1 is the worst and 5 is the best).

The cloud service models analysed are reported below in time adoption order:

---

<table>
<thead>
<tr>
<th>HYPE CYCLE CLOUD COMPUTING SERVICE MODELS ADOPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cloud parallel processing</td>
</tr>
<tr>
<td>2. Cloud services brokerage</td>
</tr>
<tr>
<td>3. Cloudbursting/overdraft</td>
</tr>
<tr>
<td>4. Cloud management platforms</td>
</tr>
<tr>
<td>5. Business and IT services</td>
</tr>
<tr>
<td>6. Community cloud</td>
</tr>
<tr>
<td>7. Virtual private cloud computing</td>
</tr>
<tr>
<td>8. Browser client-OS</td>
</tr>
<tr>
<td>9. Cloud application development tools</td>
</tr>
<tr>
<td>10. Cloud testing tools and service</td>
</tr>
<tr>
<td>11. Hybrid cloud computing</td>
</tr>
<tr>
<td>12. cloud-enabled BPM platforms</td>
</tr>
<tr>
<td>13. Cloud e-mail</td>
</tr>
<tr>
<td>14. DBMS as a cloud service</td>
</tr>
<tr>
<td>15. Enterprise portals as a Service</td>
</tr>
<tr>
<td>16. Cloud APaaS</td>
</tr>
<tr>
<td>17. Cloud computing for the enterprise</td>
</tr>
<tr>
<td>18. Compute infrastructure services</td>
</tr>
<tr>
<td>19. Private cloud computing</td>
</tr>
<tr>
<td>20. Cloud computing security concerns</td>
</tr>
<tr>
<td>21. Cloud service integration</td>
</tr>
<tr>
<td>22. Cloud storage</td>
</tr>
<tr>
<td>23. Elasticity</td>
</tr>
<tr>
<td>24. PaaS</td>
</tr>
<tr>
<td>25. Cloud computing</td>
</tr>
<tr>
<td>26. Cloud web platform</td>
</tr>
<tr>
<td>27. Public cloud computing/the cloud</td>
</tr>
<tr>
<td>28. “In the cloud” security services</td>
</tr>
<tr>
<td>29. Real time infrastructure</td>
</tr>
<tr>
<td>30. Dedicate e-mail services</td>
</tr>
<tr>
<td>31. Enhanced network delivery</td>
</tr>
<tr>
<td>32. IT infrastructure utility</td>
</tr>
<tr>
<td>33. SaaS</td>
</tr>
<tr>
<td>34. SaaS sales force automation</td>
</tr>
<tr>
<td>35. Virtualization</td>
</tr>
<tr>
<td>36. Cloud advertising</td>
</tr>
<tr>
<td>37. Integration as a Service</td>
</tr>
<tr>
<td>38. Security as a Service</td>
</tr>
</tbody>
</table>

Table 9 - Hype Cycle Cloud Computing Adoption Provided by Gartner

Annex A contains the exact definition of each cloud service model. This technological assessment may be used by the VENUS-C developers to modify the requirements and properties of the cloud platform and adapt them to the VENUS-C users’ actual needs.
5.2 Approaches to e-Science Cloud Sustainability

In this section we provide a synthesis of the results of the questionnaire and summarize the specific social, economic, environmental and technological sustainability of each scenario. Then, we make assumptions about the sustainability of e-Science cloud infrastructures.

Starting from the social, economic, environmental and technological analysis of the seven VENUS-C scenarios, it is possible to develop assumptions on potential approaches for the sustainability of e-Science clouds.

First of all, the assessment highlighted the positive impact of cloud infrastructures in all the sustainability contexts. Indeed, the general social sustainability shows a great impact on the scientific communities, in terms of:

- Improve knowledge creation and research networking.
- Increase the quality and quantity of scientific production, and data available for the scientists.
- Increase the optimization of resources and the number of researchers involved in the specific research field.

Obviously, these social benefits provided by the cloud platforms also increment the economic sustainability of the scientific community within the cloud environment, in terms of creating new collaboration agreements among research institutions or industry partners, with the consequent increase of researchers’ employment.

Furthermore, the social benefits will positively influence the competition of e-Science cloud offers, investment flows and most of all the operating costs efficiency. Regarding these costs, the most relevant savings are related to:

- Hardware costs.
- Maintenance costs.
- Software development costs.

The scientific communities can economically profit from the following benefits, which the cloud infrastructure can provide to researchers:

- Reach more users or beneficiaries.
- Offer services that do not exist at present.
- Improve the quality of pre-existing services.
- Reduce costs.
- Positively modify the working routine.
- Optimize the resources.
- Reduce the time needed to deliver the service or to reach the research goal.
• Expand the range and typologies of research activities and service provided.

Concerning the environmental sustainability, this assessment shows that the cloud infrastructures can have a huge impact on the environment, especially in terms of savings of electric power consumption.

Unfortunately, there is still not a good awareness about the positive impact of cloud infrastructures on the environment. Institutions and companies should develop strategic long-term assessments, in order to reduce the negative influence of IT infrastructures on the environment. Certainly, a participated and wide adoption of cloud services can help to reduce the carbon footprint, as Scott Charney, corporate Vice President of Trustworthy Computing at Microsoft, pointed out:

“As it turns out, the cloud has many green properties. By aggregating data in data centers that are well run for efficiency, you’re able to drive the cost down of energy and actually get huge green benefits. So there are a lot of reasons to move everybody into the cloud”\(^81\).

To assess the environmental impact of cloud computing, Microsoft recently partnered with Accenture and WSP Environment & Energy to compare the energy use and carbon footprint of Microsoft cloud offerings for businesses, with more traditional on-premise deployments\(^82\). The analysis focused on three of Microsoft’s core business applications: Microsoft Exchange, Microsoft SharePoint and Microsoft Dynamics CRM. Each application is available both as a traditional on-premise offering and as a cloud one.


\(^{82}\) Accenture, (2010), “Cloud computing and sustainability: the environmental benefits of moving to the cloud”, Whitepaper November 2010
Concerning the technological sustainability of e-Science cloud infrastructures, the first part of our analysis suggests that, on average across the different applications, typical carbon emission reductions by deployment size are:

- more than 90 percent for small deployments of about 100 users;
- from 60 to 90 percent for medium-sized deployments of about 1,000 users;
- from 30 to 60 percent for large deployments of about 10,000 users.

The figure and the data illustrated above clearly shows that cloud computing infrastructures can really help institutions and enterprises to achieve great environmental sustainability results.

Concerning the technological sustainability of e-Science cloud infrastructures, the first part of our technological assessment highlighted a significant impact on:

---

• Reduction of the activity time needed.
• An exponential increase of number of activities completed and consequently also of total activities.
• An improvement in terms of accessibility and usage of the technological infrastructures.
• An increase of productivity in terms of number of activity’s outputs.
• A huge increase of data recorded.

This data shows the positive impact of cloud infrastructures also within the technological field, which can affect also the economic and social sustainability. Indeed, the technological benefits listed above, can improve the general working routine of researchers, breeding further economic profits for the research communities.

The second part of the technological impact assessment has determined, through a rate from 1 to 5, which is the relevance of the different cloud service models defined by Gartner in the “Hype Cycle for Cloud Computing”. The results of this analysis can really be high value information for the institutions which need to implement cloud services for researchers and scientists.

It is really important to evaluate the potential best cloud service models to provide to the e-Science cloud communities, in order to offer tailor-made solutions. For instance, in the case of the seven VENUS-C scenarios, researchers argue that the cloud platform should offer them the following cloud services:

• Cloud parallel processing.
• Cloud services brokerage.
• Cloudbursting/overdraft.
• Community cloud.
• Cloud application development tools.
• Cloud testing tools and service.
• Compute infrastructure services.
• Cloud computing security concerns.
• Cloud storage.
• Elasticity.
• PaaS.
• SaaS.
• Virtualization.
• Security as a Service.

Hence, the selected VENUS-C researchers are looking for parallel processing techniques which enable the parallelization of program functions, a service brokerage, which can make cheaper, easier and safer the navigation and consume of cloud services. Indeed the security in the cloud is a very important topic for the VENUS-C researchers, who need custom-made security controls depending on the different kind of data used and shared in the cloud. The cloud provider should also implement cloudbursting or
overdraft for the VENUS-C scientists, to increase the capacity by getting it from another cloud provider, in case the cloud infrastructure is not sufficient.

This process should be transparent and based on a unique flow. Within this context also the cloud storage offer should be flexible and based on a pay-per-use model. The most relevant cloud service model for researchers is based on community cloud computing, a shared and virtualized environment where the scientists can work together and can be protected by the same security, privacy and compliance requirements.

Also the Platform as a Service model is required by the VENUS-C researchers, which is a suite that provides technologies of application servers, database management systems, portals, application and data integration, business process management suites, messaging and other different application infrastructures. Instead, the SaaS model is considered less useful by the VENUS-C scientists, but it is also required in order to install software on premise.

Obviously all these cloud service models should possess two characteristics: elasticity and virtualization. The first cloud feature is related to the ability to increase/decrease the amount of the system capacity available for a cloud service on demand, depending on the real needs of researchers. The second characteristic concerns the abstraction of IT resources, to mask the nature and boundaries of physical resources. The virtualization process develops an easy to use and intuitive service to the VENUS-C researchers.

The scientific communities should take into account the services provided by the cloud infrastructure, which can really develop tailor-made solutions to the scientists, improving the social, economic, environmental and technological sustainability of the research communities.
6. CONCLUSIONS and future activities

The fundamental starting point for an organization, an institution or a research body looking at cloud computing platforms, is to examine essential factors such as: its IT architecture, the services to offer, the expected cost reduction, the social impact of their activities developed on a cloud and the technological benefits or disadvantages that the cloud trigger to researchers, also in terms of environmental impact. Last but not least, CIO managers should clearly investigate for which services the cloud environment represents an effective option or for which services it does not make sense. With that knowledge, institutions, researchers and individual users can make informed decisions about the right infrastructure to adopt.

The following figure illustrates that the best cloud offering mix to scientists, in terms of efficacy and efficiency, is related to the integration of the:

- **Suitability**: related to the capability of the cloud provider to offer to scientists scalable and flexible services. Furthermore, elasticity and just-in time services allow the researchers to use high customized and pay-per-use solutions. This kind of infrastructures requires no up-front commitment and scientist can focus only on their main objective: the research.
- **Reliability**: an important factor to consider when the researcher chooses between different cloud providers’ offers. Indeed, a high level of protection is required in case of scientists which treat sensitive and personal data. The cloud provider has to provide scientists tailored and comprehensive SLAs, in order to ensure the security of data. Moreover, a long-term technological user support is required, especially in terms of disaster recovery, performance monitoring and service availability.
- **Profitability drivers**: the economic benefits of cloud platforms are huge for scientific organizations that can profit from cost reductions, revenues efficiency, economies of scale, CAPEX to OPEX transition and green cost savings. Furthermore, the cloud can reduce the cost of ownership, fostering also the transition to outsourcing models.

The cloud is not without limitations even if it promises to deliver a wide and powerful range of capabilities. In fact, the potential usage of cloud technologies is raising and is difficult to manage and foretell. CIO managers, also from research institutions, should make a thorough assessment to find the right trade-off between revenue optimization and cloud computing complexity, using different tools such as TCO calculator, ROI and KPI analysis.
Within this evolving context, IT strategists have to investigate how cloud technologies can impact on their business and they have to continuously track the evolution of the technology, in order to reap the benefits of cloud computing, and take advantage of its potentiality for incremental improvement and disruptive transformation of business processes.

Due to this complex and evolving environment, we consider that it will take time to complete the transition to cloud computing. The capabilities and potential savings are so important to be ignored, thus it is important to consider cloud technologies not only an option, but an efficient model to adopt, also for the scientific context. This implies that, IT innovation, in the next future could be likely based on clouds, rather than conventional computing.

The next big challenge for the IT department will be to provide disruptive hardware and software innovations, in order to enable new capabilities, reducing the cloud implementation and operating costs.
The main issues to solve are related to the actual infrastructure complexity, and evolving targets, who everyday change opinions and needs. Complexity arises from the deep interdependence of design choices in infrastructure, hardware, service software and applications. Changes in any one of these elements can affect the others.

With regards to the sustainability of e-Science cloud communities, our study highlights that the cloud computing infrastructure will be an efficient and cost effective solution to accelerate and solve research challenges. In the opinion of VENUS-C researchers, the cloud platform can provide good levels of sustainability in all the fields of our study, such as the social, economic, environmental and technological ones. Unfortunately, there is still little awareness about how to increase the cloud potentiality in the scientific context. Indeed, research bodies and cloud providers, still need to work, in order to make cloud computing infrastructures more sustainable for scientists. For instance, with reference to the environmental impact of cloud computing, this technology has surely the potential to be greener if it is based on low or net zero carbon data centers, but each cloud provider and institution, respectively, must ensure and verify in time the efforts provided to ensure these energy savings.

Concerning the social impact, cloud providers must evaluate, for example through surveys, which kind of services should be implemented for each category of scientists and offer tailor-made solutions, also distinguishing the offer depending on the needs of the individual researcher.

Even the economic impact can be increased for the e-Science cloud communities when cloud providers will start up standardized and secure procedures for the removal and the prevention of potential damages.

Indeed the most relevant drivers for the maximization of the scientists’ productivity are the reduction of data loss, lower occurrences of economic or software damages or deficiencies in the data exchanged with other subjects and applications.

Surely valuable Service Level Agreements and Disaster Recovery Systems will extend the positive economic impact of cloud technologies on the scientific research communities, as well as the technological one, and performance improvements, bringing benefits to all the research domains.

From the researchers’ perspective, it is necessary to assess the economic advantages of cloud computing infrastructure, evaluating the potential financial revenues and the impact on competitiveness and productivity. Furthermore, the scientists should develop an exploitation strategy of their outcome, identifying the current market trend, the number and offers of competitors, in order to develop a long-term economic analysis which allows them to increase the financial revenues and modify the wrong strategies.
In the future, the cloud computing architecture could move to a hybrid model allowing the provision of a homogeneous set of applications anywhere, anytime and from any device. These changes will constitute relevant drivers for scientific communities which will benefit from the advantages of cloud technologies.

New IT architectures, such as the “cloudsourcing” model, will play an important role in this context. This emerging trend is considered the next logical step after cloud computing and is designed to build complete business solutions through a combination of multiple cloud applications, cloud platforms and cloud infrastructure. The “cloudsourcer” provides cloud applications, platforms and infrastructures together. Cloudsourcing spans applications, utilities, and services, by allowing the delivery of business applications, as well as the sourcing of complete managed processes via cloud applications plus associated services. Using clouds to outsource aspects of a project, researchers can cooperate with experts in other locations, thanks to flexibility and scalability of services, this means quicker completion, reduced capital expenditures and a high capacity of storage and CPUs time. More collaboration and less competition between different scientists will spread the full potential of innovations worldwide. The “cloudsourcing” has the potentiality to enable the freedom to choose who will be your software provider every time you run it, even if there are still some issues to consider about the need to access remote data or transfer it closer to the application. If data is large this may have a major impact on the application.

The growing use of cloud technologies creates new challenges, questions and opportunities in areas like TCO, security, outsourcing, support, and performance management. IT as a service can bring many benefits including increased agility and the ability to achieve far more with much less. Processes, workflows, behaviors and expectations of all the involved actors have to be adapted. This is often a huge and long lasting effort. Therefore it is not so easy to clearly forecast the net value coming from the cloud.

In synthesis, the critical issue is not whether cloud computing will become a fundamental technology in the next years, but how organizations, institutions, research bodies and users will embrace this evolving and innovative process.

---

A composition of at least one private cloud and at least one public cloud.
7. REFERENCES & RESOURCES


WEB REFERENCES


[16] www.bsc.es

[17] www.deisa.eu

[18] www.egi.eu

[19] www.e-irg.eu


[21] www.nsc.liu.se

[22] www.pdc.kth.se

[23] www.prace-project.eu


[25] www.sics.se

[26] www.snic.vr.se

[27] www.taskforceeurope.com
8. ANNEX A Sustainability questionnaire to the 7 VENUS-C scenarios

8.1 Building Structural Information

When designing a building or a civil engineering structure, security must be assured for persons and goods. Thus, accuracy in structural simulations is a key factor, where 3D realistic structural models together with accurate and numerically efficient methods of analysis must be applied. However, the realistic 3D dynamic analysis of large scale structures requires such a computational power that traditionally has been solved by introducing a variety of simplifications (unsuitable for complex structures) in order to reduce the problem size and obtain the results in reasonable simulation times. Our aim is to expose the functionality of an HPC-based structural simulator, in the shape of a service-oriented system, to perform on demand realistic 3D analysis over large cloud deployments.

The University of Valencia, in collaboration with the Department of Mechanics of the Continuous and Structure Theory (DMMCTE) of the UPV, has developed Architrave, an advanced software environment for the design, 3D linear static and dynamic analysis and visualization of buildings and civil engineering structures.

Architrave is composed of three different and independent components, although interacting among them:

- The Design Component: An interactive GUI AUTOCAD-based application where the user can draw the model and define the structural properties.
- The Analysis Component: An interactive GUI application where the user can modify the structural properties, analyse the structure and see the results.
- The Structural Simulator: A batch HPC application used by the Analysis Component to simulate the response of the structure by means of the Finite Element Method.

The user installs and runs all these parts in his computer, and the time spent on the calculations depends on the performance of this resource. When designing a building, or a civil engineering structure, the most appropriate structural solution must be found, according to distinct criteria of safety, economic limitations or construction constraints. For that, a large amount of different configurations have to be simulated, following a trial-error process. Each of these alternatives is defined by the structural engineer varying the size of the structural elements, the material that composes them (concrete, steel, etc.), or the external loads applied, together with any combination of them.

For instance, the Spanish Earthquake-Resistant Construction Standards (NCSE-02) demands a building to be analysed with at least five different and representative earthquakes. Then, all these structural alternatives must be analysed and results are interpreted, maybe giving place to a new iteration in this trial-error scheme. Obviously, this trial-error phase largely increases the computational cost of the
problem, where the structural analysis becomes one of the most time-consuming phases when designing a structure.

For the use case, a large dimension structure will be designed. The user will interact with the Architrave Analysis Component and it will submit to the cloud the remote analysis of a set of different structural solutions, each of them simulated with at least five representative earthquakes. As a result, this situation will give place to different and independent simulations that will be computed on the cloud simultaneously. For any of them, the user will be able to download the results and visualize them.

From the user point of view, the traditional approach to work implies:

- To simplify the problem for solving it on a PC. As a result, the accuracy and the reliability of the results will be reduced.
- To analyse the structure with high simulation times.
- To limit the size of the structures to be analysed.
- To reduce the desirable number of different structural solutions to be analysed.
- To increase the quantity of employed constructive materials (reinforced concrete, steel, wood, and so on), mainly by security reasons, and the final cost of the building under construction, as a consequence.

Before the VENUS-C project, the calculations of the Structural Simulator Component were just executed locally, on the user machine. The cloud infrastructure can develop these potential benefits to our community:

- To remove the need of acquiring software licenses in property (just pay per use).
- To tackle in a realistic way large-scale structural problems, providing highly reliable results.
- To analyze complex structural systems under static and dynamic loads.
- To reduce the time involved in the structural analysis.
- To simulate concurrently a large number of different structural solutions reducing the time and cost for designing building and civil engineering structures and increasing the result reliability.

Concerning the direct beneficiaries, with reference to the high impact on availability of a cloud service for executing on-demand time-consuming tasks related to the structural analysis, these are the subjects involved:

- Architects, structural and civil engineers.
- Architectural studios, construction and engineering enterprises.
- Universities and research centers.

Instead, regarding the indirect beneficiaries, a high impact is expected on public authorities and citizens in order to increase the security of the building and civil engineering structures under external actions, such as earthquakes. Medium impact will also be related to cloud infrastructure providers, concerning the increasing of building and civil engineering structures under external actions security.

Furthermore, the cloud infrastructure can reduce the final cost of the structure under construction.
We evaluated the potential benefits that the cloud platform could develop to the Building Structural Information community with a rating from 1 to 5.

The most valuable benefits detected are: offering services that do not exist at present, reduce the time needed to deliver the service and increase the quality/quantity of data available.

The worthless benefits, but still relevant are related to: increase the optimisation of resources, the quality of pre-existing services, lower costs, reach more users and expand the range of research activities.

![Benefits](image)

**Figure 13 - VENUS-C benefits for the Building Structural Information**

**SOCIAL IMPACT**

Concerning the social impact of the cloud services on the Building Structural Information community, researchers support that the project can:

- Facilitates new research creation.
- Improve the research quality on the community.
- Improve knowledge creation.
- Increase researchers networking.

Regarding the cloud impact on research quality, the community affirms that thanks to the project, researchers will be able to solve large-scale structural problems, increase the complexity of the structure to be analysed and carry out a larger number of realistic dynamic simulations without simplifications. In this way, the reliability and safety of the results obtained will be improved. New structural problems will be tackled.
The cloud infrastructure is also able to increase the quantity of scientific production of the Building Structural Information community. Indeed, the size and the complexity of the structure to be analysed, the type of structural analysis employed and the total number of the different structural solutions or even earthquakes evaluated were limited by the performance of the computational resources used by researchers. With the cloud infrastructure, researchers will have a huge number of resources available to be on-demand employed and lots of simulations will be launched simultaneously. Thus, more structural experiments will be able to be analysed per time unit, increasing the number of structures simulated and speeding up the research process. If researchers get more results in less time, they can get more and better conclusions in their research and reducing the time-to-publication.

Instead, with reference to the quality of scientific production, the community support that, since dynamic simulations of earthquakes are computational expensive and they generate an important volume of results, experts often use approximate methods. Notwithstanding, they can achieve more realistic and accurate results using cloud infrastructures, where neither the structural complexity of the problem addressed nor the type of the applied structural analysis will be limited by the computational cost of simulations.

The cloud can also increase the quality of data available for the community. Using the cloud platform, researchers will be able to easily increase the number of realistic simulations to be evaluated when designing a building or civil engineering structure, for example increasing the number of different and representative earthquakes that can acts over the structure in its life-time. Thus, the accuracy and security of the data obtained for a structure will be better.

Finally, the project will also impact on the development of researcher communities in terms of networking, generating a new collaboration agreement with research institution and two new collaboration agreements with industry partners.
The Building Structural Information community expects also that the cloud will have an impact on their researchers’ employment. In the figure below is represented this impact:

**Social Impact**

- Improve researchers networking
- Increase the quantity of scientific production
- Improve the research quality on your community
- Increase the quality of scientific production
- Improve knowledge creation

**Employment impact**

- At the end of the project
- Within the next 5 years

Number of PhD scholarships sponsored: 1
Number of post-doctoral scholarships sponsored: 1
Number of new job position generated through collaboration agreements: 1
Number of new job position generated through spin-offs: 2

At the end of the project: 1
Within the next 5 years: 2

Figure 14 - VENUS-C social impact on the Building Structural Information
ENVIRONMENTAL IMPACT

Concerning the environmental impact, the Building Structural Information community supports that the cloud infrastructure will positively impact on the environment, in terms of energy savings. However, actually, the scenario is not able to predict an estimation of the saving percentage.

ECONOMIC IMPACT

The Building Structural Information community will expect to observe the most relevant economic impact of the VENUS-C project on:

- Market and competition.
- Investment flow.
- Operating costs.
- Process quality improvement.

Concerning operating costs, the community believes that the cloud infrastructure will not need to buy the software licenses (just pay per use) and expensive hardware in order to solve large-scale structural problems. Moreover, the scientist will not have to worry about new software updates. The community states also that a cloud infrastructure will reduce maintenance and hardware costs, but they are not able to estimate the predicted percentage of cost savings. Moreover, other costs reduction is expected, due to increased software re-usability.

Finally, thanks to the mentioned benefits, the engineering companies and the architectural studios will increase easily their productivity and volume of business. The market exploitation strategy foreseen by the Building Structural Information community is for-profit model based on a fee scheme.

In the figure below is represented the potential market share achievable by the Building Structural Information community:
The three major competitors detected by the community are all using a traditional approach, based on computing sequential structural analysis on the users’ local machine. These are the potential competitors:

- **SAP2000**: with a large market share in Europe and America.
- **CYPECAD**: market share in Spain, Portugal and Latin America.
- **TRICALC**: important market share in Spain.

The Building Structural Information community predicts that competitors at the end of the project will be up to 5 and within the next three and five years up to 10.

### TECHNOLOGICAL IMPACT

This section aims to provide a clear description of users involved in each scenario and the percentage of the cloud utilization. The table below shows the Building Structural Information users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building and civil engineering researchers</td>
<td></td>
<td>10%</td>
<td>Structural analysis for research</td>
<td>Calculate the results using a local machine or our own clusters</td>
</tr>
</tbody>
</table>

Figure 16 - VENUS-C potential market share achievable by the Building Structural Information
Building and civil construction companies, engineering companies, architectural studios | 90% | Structural analysis for commercial projects | Calculate the results using a local machine or our own clusters

Figure 17 - VENUS-C technological impact Building Structural Information

According to the Gartner “Hype Cycle for Cloud Computing” definition, we identify the relevance of these models within the Building Structural Information community.

The most relevant cloud service models for the Building Structural Information community are brokerage, cloud application and development tools, elasticity, public cloud computing, real time infrastructure and SaaS.

Instead the less relevant cloud service models are related to: Virtual private cloud computing, Browser client-OS, cloud enabled BPM platforms, cloud e-mail. Cloud computing security concerns, as they do not use sensitive and personal data, dedicated e-mail services and cloud service integration are the less interesting and useful cloud service models for this community.
### Cloud service models

![Cloud service models diagram]

Figure 18 - The value of cloud models for Building Structural Information
8.2 Building Information Management

The Building Information Management Scenario focuses on integrating the Green Prefab Building Information Management for Green construction\(^{85}\) with cloud applications. Green Prefab platform integrates different software that performs specific tasks in codified phases of construction, such as: digitalization, design, engineering, executive output, production, construction, and maintenance.

The main relevant objective of the Building Information Management scenario, within the VENUS-C project context, is to provide a rendering tool to the community of Green Prefab consisting of Real Estate, industrial supplier, building contractor and designers such as architects and engineers. Currently, and so far, stakeholders use rendering tools working on desktop applications and producing images and videos. Those outputs are useful within the design process, in client negotiations, in engineering controls and in the construction site to produce details and assembling instructions. Developing a single high quality image may nowadays require hours of computation on a local single machine and days for video.

During the VENUS-C project, the Building Information Management scenario will explore a pilot building case, in order to evaluate all the potentialities of cloud applications for rendering. The use case is based in Rovereto (Italy) and provided by the technopole Manifattura Domani, where Green prefab is settled for business development. This use case will allow the involvement of targeted stakeholders and communities within the new building production process. The specific scenario is related to a green prefab show case building, based on prefabrication production’s method, which will involve architects, engineers in close synergy with the Structural Analysis for Civil Engineering, public investors and industrial suppliers.

The cloud application for rendering will be used in a design competition where architects will use the application to allow jury and clients to visualise the architectural projects. The main improvement of the cloud infrastructure is to allow players in the building industry to generate graphical representation of their building projects in a simpler and faster way, compared to traditional systems, saving costs and time to finalize a design project.

Using cloud computing, users have rapid results instead of waiting hours or days to obtain a rendering image or video, giving a significant advantage for speeding up the whole building cycle process.

Furthermore, the cloud infrastructure will reduce costs in terms of investing in expensive workstations. Concerning the direct beneficiaries of the Building Information Management scenario in SMEs, architects, engineers, suppliers, construction firms and Real Estates are the main subjects identified.

However, this is not the only one direct beneficiary and also service providers in terms of rendering for 3D digital models could be interested and involved within this context.

A medium high level of expected impact is also associated to the decision makers of the public sectors. Instead, with reference to the indirect beneficiaries, a high impact is expected on research institutes, university and research centres, especially like scientific parks. We evaluated also the potential benefits

\(^{85}\) www.greenprefab.com
that the cloud infrastructure could develop to the Building Information Management community with a rating from 1 to 5.

The most valuable benefits detected are:

- Offering services that do not exist at present.
- Increase the quality of pre-existing services.
- Reduce the time needed to deliver the service or to reach the research goal.
- Positively modify the internal working routine of the researchers and improve internal processes in user’s institutions.

Even if the following benefits are not the most valuable, also lower costs and the ability to meet users needs are relevant advantages that the cloud can develop to the Building Information Management scenario. In the table below are illustrated the value of the potential benefits stated previously.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positively modify the internal working routine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduce the time needed to deliver the service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offering services that do not exist at present</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase quality of pre-existing services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improve internal process in user’s institutions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability to better meet users/beneficiaries’ needs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase the quality/quantity of data available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 19 - VENUS-C benefits for the Building Information Management**

**SOCIAL IMPACT**

Concerning the social impact of the cloud services on the Building Information Management community, it has been found that this project could really facilitates new research creation and improve the research quality on this community.

The Building Information Management scenario predicts that at the end of the project, a new job position will be generated through the creation/enlargement of spin-offs and within the next five years, the job positions will increase, until reach a value of four.
At the end of the project no new job positions will be generated through collaboration agreements, but within the next five years, other two units could be employed.

With reference to the impact of cloud services on the research quality of the community, the Building Information Management scenario states that visualization of rendered images in architecture is a key decision tool for users and clients. Having high definition rendered pictures and videos in a faster way, will make the process of decision faster and more efficient. Furthermore, the cloud infrastructure could also increase the quantity of the community scientific production, indeed, raising the number of pictures for a design project in a time unit, will increase the options to the decision making processes.

![Social Impact Diagram](image)

**Figure 20 - VENUS-C social impact on the Building Information Management**

Also the quantity and quality of data available for the Building Information Management researchers will be improved by using the cloud platform. Concerning the quantity of data, all stakeholders will have more 3D renderings at same price and time for the same design project (obviously, number will depend on project kind and size). On the other hand, the cloud infrastructure allows more detailed image visualizations (high definition of the picture = more pixels), improving the quality of data.

In the following figures are represented the social and the employment impact of the cloud services on the Building Information Management scenario.
ECONOMIC IMPACT

The Building Information Management community expects to observe the most relevant economic impact of the VENUS-C project on the following categories:

- Market and competition.
- Operating cost.
- Process quality improvement.

Regarding operating costs, the community affirms that in the traditional model users produce rendering visualisations in their own desktop application running on their pc/workstations. Each time they launch a rendering, they wait for the output and they stop the chain of the production until they obtain it.

With cloud services users will have faster results and outputs, so operating costs may decrease or may be used in a more efficient way. Hence, the cost saving is related only to the hardware infrastructure and the percentage of this cost saving is 50%.

With reference to the exploitation strategy, the Building Information Management scenario estimates free or potential indirect revenues. Within the next three and five years, the exploitation strategy will remain the same, but while in the next three years is estimated about 10.000$/year of pure consumption, the situation will change within the next five years and the value will rise to 50.000$/year.

We analysed the potential competitors range, which in terms of concrete numbers will be up to five at the end of the project, up to 50 within the next 3 years and more than 50 within the next five years. The three major competitors of this offer in the market are:
• Green Button\(^86\): provides web applications for desktop software at 0.30\$ per core hour.
• Felix Render by Stack! Studios\(^87\): based on credits.
• Vswarm, Ilexius\(^88\): which provides community based rendering for free.

**TECHNOLOGICAL IMPACT**

This section aims to provide a clear description of the users involved in each scenario and the percentage of the cloud utilization. The table below illustrates the Building Information Management users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architects and Engineers</td>
<td>About 18.000 users in the community</td>
<td>80</td>
<td>Visualize architectural</td>
<td>Desktop software</td>
</tr>
<tr>
<td></td>
<td>expected tests on 100 users</td>
<td></td>
<td>design projects</td>
<td></td>
</tr>
<tr>
<td>Suppliers, Construction firms</td>
<td>About 100 users in the community</td>
<td>20</td>
<td>Visualize architectural</td>
<td>Desktop software</td>
</tr>
<tr>
<td></td>
<td>expected tests on 5 users</td>
<td></td>
<td>design projects and details</td>
<td></td>
</tr>
</tbody>
</table>

According to the Gartner “Hype Cycle for cloud computing” definition, we identify the relevance of these model within the Building Information Management community. In this scenario, the most relevant cloud services and characteristics detected are related to business and IT services, compute infrastructure services, cloud service integration, elasticity, public cloud computing and “in the cloud” security services. Compared to the previous Venus-C scenarios, the Building Information Management community considers more valuable cloud service models related to the enterprise needs. Very evident is the focus on security concerns and standard controls that the cloud community has to respect and provide to the scenario. Less important, but also useful for the community are cloud storage and parallel processing services. A very contrasting information, compared to the other scenarios, is the little value attributed to the virtualization within the cloud environment, which is much more appreciated and relevant for the other communities.

\(^86\) www.greenbutton.net
\(^87\) www.felixrender.com
\(^88\) www.vswarm.com
Figure 23 - The value of cloud models for Building Information Management

8.3 Data For Science

AquaMaps is an approach for the generation, standardized dissemination and mapped visualization of model-based, large-scale predictions of currently known natural occurrence of marine species. It is based on an holistic approach to species distribution modelling, which resorts to expert knowledge and habitat usage data to compensate for the potential unreliability and/or biases, which are associated with sampling and collation of species occurrence data.

AquaMaps is widely used within the biodiversity community as a basis for species distribution inference. AquaMaps complete potential is presently hampered by the difficulty of access to necessary existing data sources and by the computationally intensive nature of the calculations.

The Data for Science scenario provides an implementation of the AquaMaps approach. It delivers to fisheries and aquaculture scientists collections of species distribution maps generated using data...
sources managed by the D4Science e-Infrastructure. The Data for Science scenario relies on a SOA-based implementation exploiting the gCube technology. It is currently deployed in the D4Science infrastructure and it is used by the creator of the AquaMaps approach and by biologists and fisheries management researchers. Species distribution maps are images showing the likelihood that a species or a combination of species lives in a specific area. Species distribution images are an important tool for species assessment and for the investigation of key issues. Species distribution is used by marine ecologists and biologists, biodiversity scientists, fisheries and aquaculture resources for management scientists and decision makers.

With reference to a use case that the VENUS-C project enabled in substantial differences, gCube empowered by the cloud technology, offers a complete innovative suite for the generation of maps characterised by the following new capabilities with respect to the ex-ante scenario:

- On-demand bulk generation of maps for an arbitrary selection set of species.
- 14 complementary spatial projections for each species distribution map.
- GIS-enabled view for each species distribution map.
- Spatial overlay of different species distribution maps.
- Reach array of metadata (including provenance metadata) for each species distribution map.
- Transect visualization of the data plotted as a map.

Moreover, the following enhancements have been introduced:

- **Precision**: scientists would like the algorithm to compute distribution at lower scales along the coasts (at least 10’ side, and 5’ will be tested). The algorithm used for creating the model currently calculates occurrence in a 30’ x 30’ square. For small-scale fisheries management at these latitudes, the scale is too big. The realisation of this feature has not yet been attempted since it would heavily increase the amount of data and quantity of computation needed.
- **Different resolution levels**: decision-makers designing marine protected areas would be very interested in a variety of focussed maps. In particular, they would like to be able to produce biodiversity maps at local, national, regional and global levels including up to 200,000 species. Due to the computational cost AquaMaps implementation is currently limited to single resolution type maps.

The main VENUS-C differences to ex-ante scenario implementation are: increased precision, different resolution support levels, increased throughput (number of records processed per minute), increased performance (number of species distribution maps elaborated per hour).

The direct beneficiaries of the AquaMaps products are: researchers working on biodiversity, fisheries and aquaculture, niche modelling developers and service providers for species distribution predictions; non-EU institutions such as FAO and the WorldFish Center.

The indirect beneficiaries of the AquaMaps products are: research institutes on biodiversity and Aquaculture, Infra or e-Infrastructures providers and non-EU institutions, such as OBIS, GBIF, FishBase and SeaLife Base.
Specifically, concerning the direct beneficiaries, an high impact is expected on researchers and service providers. Indeed, the cloud implementation of the algorithm, underlying the species distribution map production, is able to compute distribution at lower scales. In particular the new approach is able to increase the precision of 25 times.

Instead, with reference to the indirect beneficiaries, the cloud may have a high impact on non-EU researchers, institutions and companies in terms of maps generation use case implementation supports for the production of biodiversity maps at local, national, regional and global levels including up to 200,000 species.

In the figure below are represented the expected benefits of the VENUS-C project for the scientists accessing the D4Science infrastructure powered by Venus-C Cloud infrastructure. The most relevant benefits detected are related to the quantity and quality of data available to researchers, reducing the time needed to deliver the service and increasing the quality of pre-existing services. Positive impacts are also expected on resources optimisation, lower costs and keeping pace with competitors, positively modify the internal working routine and expand the range of research fields.

**Figure 24 - VENUS-C benefits for the Data for Science scenario**
SOCIAL IMPACT

Concerning the social impact of the cloud services on the AquaMaps community, researchers believe that the project can:

- Facilitate new research creation.
- Improve the research quality on the community.
- Improve knowledge creation and researchers networking.

Regarding the impact of the cloud platform on the research quality, the Data for Science researchers affirm that is expected a simplification of the working routine.

The quantity of scientific production is incremented because it is possible to offer different resolution support levels. Finally, also the quality of scientific production is increased because it is possible, for the first time, to work with data with a considerable increased precision. This becomes possible due to the increased throughput and the performance offered by the exploitation of the cloud technology on top of the cloud technology.

Also the quality and quantity of data available for the community and beneficiaries will be increased, in terms of entries per repository and more species distribution maps available at an increased precision.
The following figure illustrates the social impact of the cloud infrastructure on the community exploiting the AquaMaps approach.

ENVIRONMENTAL IMPACT

Unfortunately, no information about the environmental impact has been provided by the Data for Science scenario since this information cannot be calculated yet, because the cloud technology is not fully available.

ECONOMIC IMPACT

Regarding the economic impact of the cloud infrastructure on the Data for Science scenario, researchers believe that the most relevant positive advantages are expected in terms of: operating costs, investment flow.

Concerning operating costs, scientists exploiting the AquaMaps approach expect to observe a reduction on the costs for the management of hardware resources. Since the cloud empowered a version of the application with an increased throughput, it is possible to generate more maps in a time unit, considerably reducing the cost per map generation. The exploitation strategy foreseen by researchers is non-profit scheme with indirect revenues. Other information about the economic impact of the cloud on
the Data for Science scenario is not available at the moment for the same reason of the environmental impact.

TECHNOLOGICAL IMPACT

This section aims to provide a clear description of the users involved in each scenario and the percentage of the cloud utilization. The table below shows the Data for Science users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilization</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biodiversity scientists</td>
<td>Up to 50</td>
<td>Up to 75%</td>
<td>Generate and validate species distribution maps</td>
<td>Standalone applications capable at lower resolution and requiring long desktop session to obtain less results</td>
</tr>
<tr>
<td>Aquaculture resources management scientists</td>
<td>Up to 100</td>
<td>15%</td>
<td>To compare data, to analyse distribution, to validate data</td>
<td>Products already generated and available on the Web. The quality and the provenance is not usually reported and the data are not customisable</td>
</tr>
<tr>
<td>Niche Modelling developers</td>
<td>Up to 50</td>
<td>10%</td>
<td>To develop new niche modelling approaches.</td>
<td>Standalone frameworks allow to generate and test new models but requires ad-hoc hardware architecture costly to install and maintain.</td>
</tr>
</tbody>
</table>

Figure 27 - VENUS-C technological impact on AquaMaps

The community related to the Data for Science scenario provided an assessment of the technological effectiveness of the cloud infrastructure, with a view of its real use. This metric is based on the ISO 9136-2 and ISO 9126-4.

The VENUS-C cloud infrastructure decreases the activity time needed of the execution of a task, from 7 hours to 30 minutes, the number of activities completed correctly from 18 to 19 and the number of data recorded from 5 GB to 12 GB.

According to the Gartner “Hype Cycle for Cloud Computing” definition, we identify the relevance of these models within the Data for Science community. The most relevant cloud service models detected by the Aquamaps community are related to cloud parallel processing, cloud bursting, cloud testing tools and services, elasticity and virtualization.
8.4 Civil Protection & Emergencies, Virtual Fire

The main problems that the Civil Protection & Emergencies scenario addresses are the wildfire risk estimation and wildfire propagation simulation. Fire risk is the estimation of the probability that a fire will be ignited at a specific place and time based on various factors. This process is executed automatically. Fire propagation simulation is the estimation of the propagation of the fire front during a fire event. This kind of simulation is executed on demand by the end-user. All the interactions and the presentations of the outputs are handled within a web-based spatial/geographical interface. Before the
The involvement of this community in the VENUS-C project, fire risk estimation map was calculated once every day based on the forecasted weather data of the following day.

The output was provided to the end-users through a web map server. Fire propagation simulation was provided to end-users only as a “per request” service running locally by the developers’ team. Using the cloud infrastructure, many relevant improvements have been made to the Civil Protection & Emergencies community. Indeed, actually the fire risk is based on an hourly calculation rather than daily.

Furthermore, the cloud platform allows the researcher to obtain faster processing at lower costs, mainly due to the optimization of the resources. Simultaneous fire propagation modelling is now available to the Civil Protection & Emergencies scenario. During VENUS-C, three use cases are developed:

- Batch processing (once every day) for the calculation of the forecasted fire risk. Fire risk processor process 120 weather input files in order to produce the forecasted fire risk for the next 120 hours.
- Calculation of current fire risk based on real time weather data. It is almost identical with case 1 except that is based on real weather data and is calculated once every hour.
- Calculation of fire propagation. End-user will initialize fire ignition at a specific geographic place within map server. Simulation process downloads real time weather data and calculates the expected propagation of the fire front for a given time period.

With reference to the direct beneficiaries of the Civil Protection & Emergencies scenario, the VENUS-C cloud infrastructure could have a very high impact on fire department and civil protection agencies in terms of using better decision support tools for wildfire confrontation at lower cost. A relevant impact is also expected on the community by studying environmental and health protection on the society as a whole.

The indirect beneficiaries detected are fire researchers and academia, also from non EU countries; a high impact is here expected in terms of resources optimization.

We evaluated the potential benefits that the cloud infrastructure could develop to the Civil Protection & Emergencies community with a rating from 1 to 5. The most valuable benefits detected are: Increase the quality and quantity of data available to users, increment the optimisation of resources, reduce time needed to deliver the service, lower costs, reach more users, offering services that do not exist at present.

Despite the lower mark still remain the capability to positively modify the internal working routine, reduce knowledge gaps among users or foster a better distribution of knowledge and information, improve internal processes in users’ institutions. In the opinion of the Civil Protection & Emergencies community, the cloud infrastructure can really provide economic and social benefits at a roughly similar level.
SOCIAL IMPACT

Concerning the social impact of the cloud services on the Civil Protection & Emergencies community, researchers support that the project can:

- Facilitate new research creation.
- Improve the research quality on the community.
- Improve knowledge creation and researchers networking.

Regarding the impact of cloud services on the research quality, the researchers state that Fire risk application developers will optimize their resources’ usage and will have a better balanced time spent for research. The indirect expected economic benefits include the investments in research rather than in infrastructures, hence the quality of the research output will be considerably improved.

Faster processing allows availability of time and resources for more an increasing number of tests on the research models that are under development. This will increase the quantity and the quality of scientific production, thanks to continuous evaluations and validations made available by the VENUS-C project.

![Figure 29 - VENUS-C benefits for the Civil Protection & Emergencies](image-url)
Moreover, also the quality and quantity of data available for Civil Protection & Emergencies community will increase, especially regarding the entries per repository which will rise, reaching a final output of 120 files. The following figure illustrates the social impact of the cloud infrastructure on the Civil Protection & Emergencies community.

Figure 30 - VENUS-C social impact on the Civil Protection & Emergencies

Figure 31 - VENUS-C employment impact on Civil Protection & Emergencies
The VENUS-C project will also have a valuable impact on the University of the Aegean’s employment. Indeed, within the project life-time, researchers expect three new job positions generated through collaboration agreements and a new PhD scholarships sponsored.

With reference to the impact of the cloud on the development of research communities, the Civil Protection & Emergencies scenario detected an increasing of new networking events, but also a decrement of new virtual communities and new collaboration agreements with research institutions.

**ENVIRONMENTAL IMPACT**

A highly positive impact of the VENUS-C project on the environment was detected by the Civil Protection & Emergencies community, which predicts a 50% of savings in kWh, that will be constant by the end of the project up to the following five years.

![Environmental impact graph](image)

**Figure 32 - VENUS-C environmental impact on Civil Protection & Emergencies**

**ECONOMIC IMPACT**

The community expect a relevant economic impact of the cloud on the next categories:

- Investment flow.
- Operating costs.
- Process quality improvement.
Also in this case, the exploitation strategy foreseen for the outcome of the VENUS-C community will be based on public funding or for free. No competitors are detected by the Civil Protection & Emergencies community. Concerning operating costs, this scenario supports that the most relevant economic impact will depend on reduced maintenance and hardware costs, cost reduction due to less process breaks or system failures even if this last variable is difficult to estimate.

**Economic Impact**

Figure 33 - VENUS-C percentage of cost savings on the Civil Protection & Emergencies

**TECHNOLOGICAL IMPACT**

This section aims to provide a clear description of the users involved in each scenario and the percentage of the cloud utilization. The table below shows the Civil Protection & Emergencies users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fire Department and Central Civil Protection Agency</td>
<td>&gt;50 for Greece</td>
<td>100%</td>
<td>Provide the ability for hourly fire risk and on demand fire propagation simulation at low cost</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Figure 34 - VENUS-C technological impact on Civil Protection & Emergencies
The Civil Protection & Emergencies community provided an assessment of the technological effectiveness of the VENUS-C cloud infrastructure, with a view of its real use. This metric is based on the ISO 9136-2 and ISO 9126-4.

The cloud infrastructure reduces the time required to complete an activity from more than one hour to less than seven minutes.

Furthermore, the project will increase the number of tasks completed correctly in the time frame and also the total tasks executed in the time frame from 1 to 120 tasks. Finally, also the productivity will be highly improved in terms of activity’ output numbers and information recorded (again from 1 to 120).

According to the Gartner “Hype Cycle for Cloud Computing” definition, we identify the relevance of these models within the Civil Protection & Emergencies community. The community rate only the most relevant cloud service models, which are cloud parallel processing, elasticity, cloud storage and finally, less relevant, but also considerable is the real time infrastructure.

![Figure 35 - The value of the cloud models for Civil Protection & Emergencies](image-url)
8.5 Bioinformatics

This scenario aims at the migration of a reasonable set of bioinformatics applications widely used by the community from on premise clusters. As Ignacio Blanquer, professor at the University of Valencia states: “Scientists are overwhelmed by the exponential increase of data to process in every field, which is referred to as the data deluge”.

Massive Bioinformatics experiments, which imply the alignment of thousand of sequences against a database (in the Bioinformatics context, a file with a series of reference sequences) demand huge amounts of computational (CPU years) and storage resources.

A good example is the application domain of DNA sequence analyses, where DNA fragments are produced at massive rates by state-of-the-art DNA Sequencer machines. One of the most widely used methods for analysing these DNA fragments is the ‘alignment’ between the newfound sequences and reference sequences, in order to identify functional relations. Nevertheless, aligning thousands of new sequences against a huge set of reference sequences, using popular algorithms such as BLAST\(^9\), is a very computationally intensive process that cannot be used efficiently following traditional approaches using the researchers’ own resources. Consequently, conventional research groups are forced to search for alternative parallel approaches.

The first alternative consists of acquiring a cluster of certain number of compute nodes and using a parallel version of the alignment program with the intention of maximizing the performance obtained. If the research group cannot afford a cluster, a good alternative could be requesting the use of a Supercomputer or a Grid infrastructure. Within this context, the cloud infrastructure – as a reference example of applying cloud computing to e-Science applications represents also a good alternative.

In order to assess the potential benefit of VENUS-C cloud services for e-Science, we asked our partners to describe one use case, which shows significant improvements and substantial differences in the way of doing research before and after their involvement in the VENUS-C project. The Bioinformatics community identified metagenomics as use case. Metagenomics is the study of genetic material recovered directly from environmental samples and as its name suggests, it comprises a genome of genomes (i.e. a very huge amount of sequences) that must be analysed (or ‘aligned’ in this context).

This is a good example of a use case that becomes feasible by means of VENUS-C. Indeed, a biologist could use the VENUS-C framework to perform massive experiments, such as an ‘all versus all’ comparison between the sequences from different organisms groups. Something that was not possible to make without using the cloud infrastructure.

Furthermore, taking as a reference a sequential strategy, benefits obtained with the VENUS-C framework are obvious. The response time of the experiments is reduced from various years to only a couple of weeks. Instead of acquiring a cluster, the cloud approach allows a research group to temporary rent the necessary resources for the study, avoiding the investment on new hardware and

---

\(^9\) \url{http://blast.ncbi.nlm.nih.gov/Blast.cgi}
maintaining costs. The improvements obtained from the cloud, instead of Supercomputer or Grid infrastructures are the exclusivity in the use of the resources and a higher quality of the service.

Subsequently, we determine the main direct and indirect beneficiaries of the Bioinformatics scenario, in order to identify the expected general positive impact of the cloud infrastructure:

- **Direct beneficiaries**: Bioinformatics researchers, either SMEs or research centres from EU, with reference to the availability of a computing service on the cloud for executing long-runs of sequence alignments. Also application developers can benefit from the availability of service points for processing pipelines.

- **Indirect beneficiaries**: a medium impact is expected on the availability of cloud providers services, especially PaaS or IaaS; on the increasing of DCI e-Infrastructure projects usages and on public authorities and society as whole, in terms of health delivery improvement.

This cost-benefit analysis has determined the most relevant benefits that the cloud platform will bring to the Bioinformatics community on a scale from 1 to 5. The highest value in terms of relevance was assigned to the following categories: expand typologies of research activities with lower costs and increase the resource optimisation. Furthermore, VENUS-C will also reduce the time needed to deliver the service or to achieve the research scope, increase the service’s quality and reach more beneficiaries. The figure highlights also that cloud infrastructure can offer services that do not exist at present and positively modify the internal working routine of scientists.

![Figure 36 - VENUS-C benefits for the Bioinformatics community](image-url)

**SOCIAL IMPACT**

The cloud infrastructure facilitates new research and knowledge creation, improves the research quality of the Bioinformatics community. However, the most relevant benefit that the cloud could develop is to improve researchers’ networking between the Bioinformatics community.
With reference to the expected cloud impact on the University of Valencia, the Bioinformatics predicts that, at the end of the project, a new job position will be generated through collaboration agreements and within the next five years this value will be doubled. Moreover, in this long term prediction, also two PhD scholarships and a post-doctoral scholarship will be sponsored. The Bioinformatics scenario also states that the cloud infrastructure will impact their research quality, providing different effects on two categories of stakeholders: final Bioinformatics users and bioinformatic application developers.

- **Final bioinformatic users**: Bioinformatics rely on intensive computational resources with a learning curve that could be reduced by the use of the VENUS-C enabled bioinformatics tools. Moreover, they will have an enhanced flexibility for planning the experiments which could end in a reduced time-to-publication.

- **Bioinformatics application developers**: will have reduced development time by using the cloud services. The expertise gained in the Bioinformatics context, will increase processing pipelines commonly used in Bioinformatics bringing new capabilities.

The cloud platform can increase the quantity of scientific production, by enhancing the size of the experiments, amplifying the appeal of an article. Regarding the cloud impact on the quality of scientific production, a representative of the Bioinformatics community states that: “Experiment-size and parameter variability are bound to the availability of computing resources accessible for the researchers. If more resources can be brought on demand, experiments could be larger and more accurate. This will lead to more post-processed data that could lead to more research publications”.

![Social Impact](image)

**Figure 37 - VENUS-C social impact on the Bioinformatics community**

The Bioinformatics community affirms also that the cloud impacts the growth of their research communities through three new collaboration agreements among multidisciplinary teams.
ENVIRONMENTAL IMPACT

The VENUS-C Bioinformatics community recognizes the ability of the project to develop a positive environmental impact. Indeed, the cloud infrastructure produces consistent savings in the magnitude of power electricity. The assessment of the environmental impact is based on three specific time periods in order to evaluate the percentage of reduction: at the end of the project, within the next 3 years and within the next 5 years.

The Bioinformatics community evaluate the environmental impact making assumptions on two potential cases: the best case and the most adjusted case. The resulting outcome is significant enough at the end of the project; it will slowly increase within the next 3 years and almost double within the next 5 years.

Unfortunately, no information about the related environmental savings are provided, such as savings in consuming and selling of papers, films, CD, DVD, in storage-items, travels and reduction of technological waste. In the figure below, the two environmental impact cases are represented.
ECONOMIC IMPACT

The Bioinformatics expects to observe the most relevant economic impact of the project on:

- Market and competition.
- Investment flow.
- Operating cost.

Concerning the operating cost, the most relevant economic impact is expected on reduced maintenance and hardware costs. Also software operating costs could be reduced, but only slightly. Reduction in licenses is not foreseen since most of the community work is based on free access and on open-source models. Cost reduction may come due to incremental increase in software re-usability, because of an improvement of test-deploy-rework circle management and lower software development costs. However, the Bioinformatics community affirms that is difficult to estimate the exact percentage of this cost saving. The cloud infrastructure could also have a high impact on competitiveness and productivity, increasing flexibility, reducing time to market and costs, due to the cloud characteristics.
With reference to the exploitation strategy of the Bioinformatics scenario results, the community global market value support is based on Open Source tools and the potential market share achievable is based on the coverage of the two key tools in alignment by the next three years. Furthermore, a largest market share could be reached increasing the development of other new tools. Below we analyse the main potential three competitors of the current Bioinformatics VENUS-C community, only concerning BLAST:

- NCI BLAST on Windows Azure\(^90\): which enables researchers to take advantage from the scalability of the Windows Azure platform, in order to perform analysis of vast proteomics and genomic data in the cloud.
- Windows HPC BLAST with Windows azure burst\(^91\).
- Amazon EC2.

Unfortunately, no information about the potential financial revenues (at the end of the project, within the next 3 years and within the next 5 years) are provided by the Bioinformatics community for the cloud economic impact analysis.

**TECHNOLOGICAL IMPACT**


This section aims to provide a description of the users involved in each scenario and the percentage of the VENUS-C cloud utilization. The table below shows the Bioinformatics users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioinformatics scientists</td>
<td>10</td>
<td>15% of the global pipeline</td>
<td>Basic component for many processing pipelines</td>
<td>Use of own clusters</td>
</tr>
<tr>
<td>Bioinformatic Data Center managers</td>
<td>1</td>
<td>15% of the global pipeline</td>
<td>Basic component for many processing pipelines</td>
<td>Use of own clusters</td>
</tr>
</tbody>
</table>

*Figure 41 - VENUS-C technologique Bioinformatics scenario*

In our metric we provide also a set of indicators, which helps the scientists to assess the technological advantages of the cloud platform and to estimate costs and potential damages caused for the Bioinformatics community. At the moment, this scenario is not able to provide this information; consequently this kind of analysis was not performed.

Following the Gartner “Hype Cycle for Cloud Computing”\(^92\) definition, we used Gartner’s list of the different cloud service models, in order to identify the relevance of these models within the Bioinformatics community and, rating from 1 to 5 (where one is not relevant and 5 is the most important). For a better comprehension of the different cloud service models, please read a summary of Gartner’s definitions in the Annex.

The most relevant cloud services and characteristics detected are related to the cloud parallel processing and service brokerage, elasticity, cloud web platforms, public cloud computing and virtualization. The bioinformatic community considers the PaaS and SaaS cloud service models very useful for their research processes. In addition to the obvious relevance of this cloud services, the bioinformatic scenario support the need also of using a browser client-OS, cloud testing tools and service, Security as a Service controls and Enterprise portals as a Service model. In the following figure is represented the rating of the relevance of each Gartner “Hype Cycle” category for the Venus-C bioinformatic researchers. The lowest value was associated to the service models typically related to the enterprise environment such as: business and IT services, cloud e-mail, cloud computing for the enterprise, dedicated e-mail services and enhanced network delivery. In our metric we provided also a set of indicators which help the scientists to assess in detail the technological advantages of the Venus-C bioinformatic scenario.

cloud platform and to estimate costs and damages caused for the bioinformatic community. Unfortunately, this scenario does not have yet this information and it was not possible to achieve this kind of analysis. In the table below is illustrated in detail the estimated value of type of cloud service models, according to the Gartner “Hype Cycle for cloud computing” definition.

8.6 Systems Biology

CoSBi provides tools for modelling and simulating biological systems, which can be used also to model and simulate biochemical pathways (genetic regulatory networks, metabolic networks, signalling pathways as well as ecological networks (food webs)). To be accurate and close to the simulated reality,
CoSBi uses stochastic, Monte Carlo simulation algorithms, which are both computationally intensive and require many runs in order to collect statistically significant results.

Biological systems are characterized by incomplete knowledge of mechanistic behavior. Therefore, when building models of these systems, we have to cope with lack knowledge. A possible approach is to run many simulations with different parameters and configurations in order to fill the gap and/or explore a wide space solution. Before the involvement in the VENUS-C project, this solution was viable only with an in-house HPC cluster.

The aim of the task is to provide to users tools for Systems Biology modeling and simulation which can be run directly on a cloud infrastructure. In particular, users from a community of academic labs and companies doing research in either medicine, or biology, or pharmacy, will be able to simulate and analyses the dynamics of in-silico models of complex biological systems leveraging the computational power of the cloud, without the need for an in-house HPC system. This task will be accomplished by porting COSBI simulation tools to VENUS-C service oriented computing facilities.

The main improvements of the cloud infrastructure on the Systems Biology research are related to:

- Scalability.
- Ease of access.
- Ease of sharing data and resources.
- Security.

Regarding the direct beneficiaries, higher impact of the cloud services is expected on systems and computational biology, biochemistry researchers/ institutions/ companies, in the context of simulation and analysis of complex biological networks and interactions. Instead, a medium impact on the indirect beneficiaries is expected on drug targets discoveries and the development of new drugs within health and medicine sectors.

We evaluated then the potential benefits that the cloud platform could develop to the Building Structural Information community, with a rating from 1 to 5.

The most valuable benefits detected are: offering services that do not exist at present, reach more users or beneficiaries and increase the optimisation of resources.

The worthless benefits, but still relevant are related to: increase the number of researchers involved in research activities and lower costs.
SOCIAL IMPACT

Concerning the social impact of the cloud services on the Systems Biology community, the researchers state that the project can:

- Improve your research quality on the community.
- Improve knowledge creation.
- Improve research networking.

No impact on the employment of the community is detected.
Figure 44 - VENUS-C social impact on the Systems Biology

Instead, with reference to the positive impact of the VENUS-C project on the research quality, the Systems Biology community states that the cloud platform get these good results providing an ease of access to services thanks to the ability to perform heavy experiments without the need to buy and administrate expensive hardware resources. The ease of access to services will also increase the quantity of scientific production, making viable to perform heavy experiments, expensive hardware resources and enabling small research institutes/individual researchers to produce scientific data. The cloud infrastructure can also improve the development of researcher communities in terms of networking, producing some new collaboration agreements with research institutions and industry partners.

ENVIRONMENTAL IMPACT

Concerning the environmental impact, the Systems Biology community affirms that the cloud infrastructure will positively impact on the environment, in terms of energy savings. However, actually, they are not able to predict an estimation of exact percentage of savings.

ECONOMIC IMPACT

The Systems Biology community expects to observe the most relevant economic impact of cloud services on:

- Market and competition.
- Operating costs.
- Accessibility.

Concerning the operating costs, savings in kWh investment and on power usage, will be the main drivers of the economic impact on the Systems Biology scenario. The community states also that the cloud will reduce maintenance, hardware and connectivity costs, unfortunately, they are not able to estimate the predicted percentage of cost savings. The exploitation strategy that the Systems Biology community has foreseen for their selected outcome is based on fee, not-for-profit.

TECHNOLOGICAL IMPACT

This section aims to provide a clear description of the users involved in each scenario and the percentage of the cloud utilization.
The table below shows the Systems Biology users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systems biology researcher</td>
<td>50</td>
<td>50</td>
<td>HPC cluster</td>
<td></td>
</tr>
<tr>
<td>Students (bio and bioinformatics)</td>
<td>50</td>
<td>50</td>
<td>HPC cluster</td>
<td></td>
</tr>
</tbody>
</table>

Figure 45 - VENUS-C technological impact Systems Biology scenario

According to the Gartner “Hype Cycle for cloud computing” definition, we identify the relevance of these model within the Systems Biology community. The most relevant cloud service models for the Systems Biology scenario are related to cloud parallel processing, cloudbursting, compute infrastructure services, cloud storage, elasticity and PaaS. Instead, the less important cloud service models detected are: cloud services brokerage, DBMS as a cloud service, Enterprise portal as a Service, private cloud computing, cloud service integration, SaaS, integration as a Service, Security as a Service.

Figure 46 - The value of the cloud models for the Systems Biology
8.7 Drug Discovery

This scenario is aimed to building QSAR (quantitative structure-activity relationship) models that can be used during the drug development process to determine the likelihood of a compound reacting in a certain way. The scenario builds models using different techniques such as Neural Nets, Linear Regression and Recursive Partitioning.

The creation of these models involves a novel branching workflow, which is embarrassingly parallel in places. The input data for the models comes from the open source ChEMBL, a database of chemical activities. Prior to VENUS-C, the large scale creation of QSAR models was only achievable using a legacy system called the Discovery Bus. Although this system was powerful, it was not designed to run in a cloud scale and was limited to a very small number of machines due to scalability issues. In order to identify the benefit of the cloud services within the scientific context, we asked to our partners to describe one use case, which could show significant improvements and substantial differences in the way of doing research before and after their involvement of researchers in the VENUS-C project. Indeed, before the starting of the project, utilising cloud resources for large scale QSAR model building was very hard and did not scale beyond a few 10s of machines.

The model built on a single machine was estimated to take five years, and on a few machines this would come down to a few months. Using the cloud platform, the Drug Discovery community, hopes to be limited only by the number of resources (within reason) and so the process should last a small number of days.

Researchers are also able to show the viability of this solution to other model building ones apart from QSAR. Hence, the main improvement that the cloud provides to the drug discovery community is the re-implementation of the workflow coordination logic, in order to allow scalability across many nodes and use scalable queuing and storage services offered by cloud computing providers.

With reference to the direct beneficiaries of this scenario, the highest cloud impact is expected on researchers, large companies, SMEs and non-European researchers, institution and companies within the medical and chemical environment. Instead, regarding the indirect beneficiaries, the community did not identify any. We asked to the drug discovery community to evaluate the relevance of potential benefits that the cloud can produce to this scenario rating from 1 to 5.

The highest value was accorded to increase the quality and quantity of data available for the drug discovery scientist by using the cloud infrastructure and the reduction of time needed to deliver the service or to reach the research goals.

Less valuable, but still relevant are the benefit indicators related to the development of offer services not existing before the project, increasing quality of pre-existing services, optimised and increased resources, better distribution of information and reduced knowledge gaps among users. The Drug Discovery community consider less important the cloud potential benefits in terms of keeping pace with competitors or with research advances in the Drug Discovery field and the improve of internal processes in user’s institutions.
The following figure illustrates the drug discovery community perception of cloud benefits rating from 1 to 5.

**Figure 47 - VENUS-C benefits for the Drug Discovery**

### Benefits

- Reduce the time needed to deliver the service
- Increase the quality/quantity of data available
- Reduce knowledge gaps among users
- Increase the optimisation of resources
- Offering services that do not exist at present
- Increase quality of pre-existing services
- Keeping pace with competitors or with the...
- Improve internal process in user’s institutions

### SOCIAL IMPACT

With reference to the social impact of the cloud infrastructure, the Drug Discovery community states that the project could really improve the research quality and the knowledge creation, providing more high quality model of publicly available data reducing also time and costs of developing drugs. Quality of data could have a huge impact on the automation of the decision points in QSAR model building, for instance by removing the human bias that is sometimes present. Furthermore, the cloud platform, allows the Drug Discovery scenario to increase the quantity of scientific production, because it becomes easier to produce QSAR models in a very large scale. Concerning also to the quantity of data available for the Drug Discovery community, the cloud could have a significant impact on the entries per repository, which are increased by one million.

While, with reference to the number of repositories, the cloud infrastructure does not have a relevant impact on them (only one) nor a valuable influence on Meta data. Researchers affirm that the cloud platform is not able to have a significant impact on the development of researcher communities and networking.

In the figure below are illustrated the indicators concerning the social impact of the VENUS-C project on the Drug Discovery scenario.
ENVIRONMENTAL AND ECONOMIC IMPACT

Unfortunately, no information about the environmental and the economic impact are provided by the Drug Discovery community.

With reference to the economic impact, this scenario affirms that actually they are only interested in the process research quality improvement and, acting as an Academic Institution (university), they do not have competitors and do not have drafted an economic exploitation strategy of their research outcome. All the revenues of the University of Newcastle scenarios are public funding and indirect revenues. No information about indirect revenues has been provided at the moment.

TECHNOLOGICAL IMPACT

This section aims to provide a clear description of the users involved in each scenario and the percentage of the cloud utilization. The table below shows the Drug Discovery users’ scenario information:

<table>
<thead>
<tr>
<th>Users</th>
<th>Number</th>
<th>% of utilisation</th>
<th>Goal</th>
<th>Describe the major alternative to reach the same goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological or chemical researchers</td>
<td>100%</td>
<td>Build and get predictions from the models</td>
<td>Building their own system</td>
<td></td>
</tr>
</tbody>
</table>
According to the Gartner “Hype Cycle for cloud computing” definition, we provided a list of the different cloud service models available. In this scenario, the most relevant cloud services and characteristics detected are related to cloud parallel processing, cloud bursting/overdraft, cloud storage and elasticity. The most valuable cloud service model identified by the community is the Platform as a Service (PaaS). Less valuable, but also important is the DBMS as a cloud service, to run scalable and elastic service on the Venus-C cloud infrastructure, the community cloud in terms of a shared environment and the potential of virtualization which abstracts the physical IT resources, developing easy-to-use services. Instead, this technological value is not recognized to cloud service models and characteristics such as business and IT services, cloud e-mail, cloud-enabled BPM platforms, enterprise portals as a Service, real time infrastructures, IT infrastructure utility and SaaS sales force automation, i.e. to the services that are most related to the business and enterprise world. An unexpected data concerns the low relevance assigned by “Security as a Service” controls, that apparently are not necessary to protect the Intellectual Property Rights or the confidentiality of scientist’s information.
### Cloud service models

<table>
<thead>
<tr>
<th>Service Model</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud parallel processing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud bursting/overdraft</td>
<td>4</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud multi-service</td>
<td>5</td>
</tr>
<tr>
<td>Cloud computing</td>
<td>5</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>3</td>
</tr>
</tbody>
</table>
9. ANNEX B Short definitions of cloud service models in order of time adoption, as stated by the Gartner's “Hype Cycle for Cloud Computing”

1. **Parallel processing** techniques are algorithmic and code-structuring methods that enable the parallelization of program functions. The adoption of cloud architectures creates an opportunity to apply these techniques to application system design.

2. A **Cloud Service Brokerage** typically leverages some combination of the following technologies:
   - Cloud infrastructure — some combination of infrastructure as a service and platform as a service.
   - Integration as a service — to integrate with cloud services consumers and cloud services providers.
   - Governance technology — for security and policy compliance of cloud services consumption.
   - Community management — to manage the provisioning of consumers and providers.
   - SaaS — where appropriate, some combination of SaaS functionality as part of service enrichment.

A viable CSB provider can make it cheaper, easier, safer and more productive for companies to navigate, integrate, consume and extend cloud services, particularly when they span multiple, diverse cloud services providers.

1. **Cloudbursting/overdraft** is the ability to automatically get more capacity from a different cloud infrastructure when the primary cloud infrastructure is overloaded.

2. **Cloud management platforms** are integrated products that provide management capability for external (public) and internal (private) cloud environments to enterprise consumers and IT operations teams. Included in this category are products that provision system images, monitor performance and availability, enable metering and billing, monitor and manage applications, and integrate with enterprise management systems.

3. There are two components of **cloud-driven business and IT services**: the first area includes all types of consulting, advisory, business analysis, IT architecture analysis, application portfolio and cloud readiness, system integration, deployment and testing services delivered by service providers to enterprises. The main objective is focused on assisting enterprises navigate and implement various areas of cloud computing technologies and determining the impact on business and IT within the enterprise. The second area includes all types of solutions that are developed, bundled and packaged as outsourcing offerings, where the business or IT service

---

provider leverages one or more cloud computing technologies within the solution's overall architecture.

4. **Community cloud computing** refers to a shared cloud-computing service environment that is targeted to a limited set of organizations or individuals which share similar security, privacy and compliance requirements. Community cloud computing offers a blend of public and private cloud benefits and challenges, and is likely to serve as an intermediary stage between private and public cloud computing, which provide not only a common set of shared infrastructure services, but also specialized application, information and business process services of unique value and interest.

5. A **VPC** refers to the partitioning of a portion of a public cloud-computing service provider's environment into an isolated environment that is dedicated for use by a single entity or group of related entities (such as multiple departments within a company). In addition, a VPC may be isolated from the Internet, utilizing a private network (virtual private network [VPN] or private connectivity) and/or a virtual LAN for access to the services to add an additional level of performance, security and control.

6. In **browser client OSs**, all applications are Web-based, and many traditional PC OS functions are missing. However, they can provide a much-simpler and lower-cost browsing experience.

7. **Cloud application development (AD) tools** are used to create custom software applications deployed on an application platform as a service (APaaS), a cloud-enabled application platform (CEAP), or a cloud system infrastructure. These applications can range from simple, situational business process management (BPM) to complex, mission-critical, line-of-business systems. The distinguishing features of these tools include awareness of, integration with and control of the target cloud runtime environment where the finished application executes. Cloud AD tools can be categorized along two axes: target audience and runtime environment.

8. **Cloud testing tools and service** are deployed to public or private clouds. These solutions require large scalability, strong technology coverage and the ability to work across applications using a mix of technologies. To improve testing with public cloud solutions, SOAS architectures are used. Moving test labs to use a virtualized infrastructure in private and public clouds can reduce the cost of management, hardware, software and power and should be a key element of any center of excellence effort for software quality.

9. **Hybrid cloud computing** refers to the combination of external public cloud-computing services and dedicated IT resources in a coordinated fashion to assemble a particular solution. It implies some degree of integration or coordination between the dedicated and public cloud environments at the data, process, management or security layers.

10. **Cloud enabled BPM platforms** is software that uses BPM technologies (BPMTs) to construct and optimize process-centric solutions in a software-as-a-service (SaaS) or cloud service delivery model. BPMTs include high-level process modeling tools, business process analysis software, workflow, automated business process discovery (ABPD) tools, BPM suites (BPMSs), business activity monitoring (BAM) and business rule management systems.

11. **Cloud e-mail** describes a vendor-offered, multi-tenant, Internet-delivered e-mail service that is scalable and flexible.

12. **DBMS as a cloud service** are DBMSs engineered to run as a scalable, elastic service available on a cloud infrastructure. These DBMSs are available only as a cloud offering and are not necessarily relational. Cloud-based DBMS services are provided in a multi-tenancy environment.
with elastic resource allocation, for use in simple to complex transactions. DBMSs as a cloud service excludes those DBMSs that will run on cloud infrastructure, but are not purpose-built as a cloud service.

13. **Enterprise portals as a Service** provide organizations with a unified Web environment on which to manage user access, and integrate and deploy Web applications. From a user perspective, an enterprise portal is a single, personalized point of access to enterprise information, processes and people. Enterprise portals as a service employ cloud services, whether at a hardware, middleware and/or software layer, to suit enterprise portal needs.

14. **Cloud application platform as a service (APaaS)** is a general-purpose platform for building SaaS applications, which intermediates such concerns as multitenancy and scalability. Cloud APaaS is, in principle, a specialized application server (and dedicated development toolset) that is deployed "in the cloud" and offered as a service to software developers. The technology internal to cloud APaaS that facilitates the service is referred to as a cloud-enabled application platform (CEAP). Cloud APaaS is part of a larger-scope application infrastructure platform as a service (PaaS), which encompasses other middleware-style service offerings, such as database as a service and message-oriented middleware as a service.

15. **Cloud computing for the enterprise**: cloud technologies can have significant business impact on an enterprise, including reduced total cost of ownership, increased agility and significantly reduced cycle time to provision new IT resources.

16. **Compute infrastructure services** are a type of infrastructure as a service (IaaS) offering. They offer on-demand computing capacity from a service provider. Rather than buying servers and running them within its own data center, a business simply obtains the necessary infrastructure from a service provider in a shared, scalable, "elastic" way and accesses it via the Internet or a private network.

17. **Private cloud computing** has access to the service completely open, and the service implementation is completely hidden from the customer. Private cloud computing will reduce the cost of operations and enable faster service delivery.

18. **Cloud computing security concerns** about the relevance of evaluating the security of any service provider.

19. **Cloud service integration** is a relatively new application of integration software or integration as a service to IT project scenarios that involve the direct integration of cloud service with each other (cloud-to-cloud service integration), or with on-premises applications and data (cloud service to on-premises integration). Cloud service integration can be solved using integration technology and service that are similar to integration solutions used to implement on-premises (internal or integration) and traditional e-commerce (B2B integration) projects.

20. **Cloud storage** is a storage utility offering that is defined by the following characteristics: pay-per-use model, software-agnostic, reservation-less provisioning and provider-owned; it is also frequently geographically separated from the servers that are using it.

21. In the service provider view, **cloud service elasticity** is the ability to increase or decrease the amount of system capacity (for example, CPU, storage, memory and input/output [I/O] bandwidth) that is available for a given cloud service on demand in an automated fashion. The degree of automation of elasticity is determined by the service provider.

22. **A platform as a service (PaaS)** is a generally accepted reference to the middle layer of the cloud technology stack that Gartner refers to as application infrastructure services. A comprehensive
A PaaS suite, usually depicted in cloud diagrams, is a broad collection of application infrastructure services offered by a cloud service provider. Such comprehensive PaaS suites would include technologies of application servers, database management systems (DBMSs), portals, application and data integration, business process management suites (BPMSs), messaging and many other forms of application infrastructure — all formatted to be offered as a service.

23. Gartner defines "cloud computing" as a style of computing where scalable and elastic IT-enabled capabilities are delivered as a service using Internet technologies.

24. Cloud/Web platforms use Web technologies to provide programmatic access to functionality on the Web, including capabilities enabled not only by technology, but also by community and business aspects. This includes, but is not limited to, storage and computing power.

25. Gartner's definition describes public cloud computing as a style of computing where scalable and elastic IT-enabled capabilities are provided "as a service" to external customers using Internet technologies. Therefore, public cloud computing is the use of cloud-computing technologies to support customers that are external to the provider's organization. It is through public consumption of cloud services that the types of economies of scale and the sharing of resources will be generated to reduce cost and to increase choices available to consumers.

26. "In the cloud" security services can be delivered as a part of Internet bandwidth services or as independent cloud-style offerings. Threat-facing services, such as firewalls, intrusion detection systems, intrusion prevention systems, antivirus services, distributed denial-of-service protection services, messaging security and Web gateway security services have seen the earliest adoption.

27. Real-time infrastructure (RTI) represents a shared IT infrastructure (across customers, business units or applications) in which business policies and service-level agreements drive dynamic and automatic allocation and optimization of IT resources (that is, services are elastic), so that service levels are predictable and consistent, despite the unpredictable demand for IT services. Where resources are constrained, business policies determine how resources are allocated to meet business goals. It provides the elasticity functionality as well as dynamic optimization and tuning of the runtime environment based on policies and priorities.

28. Dedicated e-mail services (DES) are one of three off-premises ways of provisioning e-mail. The other two are outsourced e-mail services and true multitenant mail (MTM) services (the only type that meets Gartner's definition of cloud-based services).

29. Enhanced network delivery comprises a combination of network-based WAN optimization services and WAN optimization controller (WOC)-equipment-based deployments. Enhanced network delivery uses a combination of protocol spoofing, route control, compression and caching to provide enterprises with improved application performance across the WAN for cloud-based services.

30. An IT infrastructure utility (IU) is a shared IT infrastructure architecture provided through on-demand services. Pricing is based on service use and proven, ongoing reductions in the fixed baseline (or subscription fees) and unit costs. The IU is open, flexible, predesigned and standardized, as well as virtualized, highly automated, secure and reliable.

31. Software as a service (SaaS) is software that is owned, delivered and managed remotely by one or more providers. If the vendor requires user organizations to install software on-premises using their infrastructures, then the application isn't SaaS. SaaS delivery requires a vendor to provide remote, outsourced access to the application, as well as maintenance and upgrade.
services for it. The infrastructure and IT operations supporting the applications must also be outsourced to the vendor or another provider.

32. Although there are many components of a SaaS sales force automation (SFA) solution (such as proposal generation, sales configuration and pricing management), the primary focus of software-as-a-service (SaaS) SFA is opportunity management. This is the practice of systemizing how sales channels pursue sales opportunities in the context of preferred philosophies, methodologies and strategies, which are tracked and updated.

33. IT virtualization is the abstraction of IT resources in a way that masks the physical nature and boundaries of those resources from resource users. An IT resource can be a server, a client, storage, networks, applications or operating systems. Essentially, any IT building block can potentially be abstracted from resource users. Abstraction enables better flexibility in how different parts of an IT stack are delivered, thus enabling better efficiency (through consolidation or variable usage) and mobility (shifting which resources are used behind the abstraction interface), and even alternative sourcing (shifting the service provider behind the abstraction interface, such as in cloud computing).

34. Cloud advertising is a business process cloud service defined as the capability to deliver advertising where the content and the fee charged are determined at the time of end-user access, usually by an auction mechanism that matches bidders with "spots" as they become available. Search engine marketing (SEM) and various forms of online display advertising (for example, banners) are the most developed formats, but the concept is also evolving to other channels and platforms, such as online video, mobile devices, addressable television and out-of-home digital signage.

35. Integration as a service (IaaS) is integration functionality — i.e., secure B2B communications, data and message translation, and adapters for applications, data and cloud APIs — delivered as a service. IaaS is always scalable, sometimes elastic, but is almost always deployed with enough multi-tenancy capabilities such that one instance of a provider’s IaaS functionality can support multiple B2B integration projects across multiple B2B communities. There are two categories of IaaS: IaaS for traditional e-commerce projects and IaaS for cloud service integration.

36. "Security as a service" means security controls that are owned, delivered and managed remotely by one or more providers. The provider delivers the security function based on a shared set of security technology and data definitions that are consumed in a one-to-many model by all contracted customers anytime on a pay-for-use basis, or as a subscription based on use metrics. There can be some customer premises equipment involved, but only where dictated by the service provider, not by the enterprise.