Ref. Ares(2013)2703248 - 18/07/2013

European Commission

Cloud for science and public authorities

FINAL REPORT

A study prepared for the European Commission
DG Communications Networks, Content & Technology

Digital Agenda for Europe

This study was carried out for the European Commission by

Authors: Gabriella Cattaneo, Massimiliano Claps, Steve Conway, Marcello Bardellini (IDC), Silvana Muscella, Stephanie Parker, Nicholas Ferguson (Trust IT)

Internal identification
Contract number: 30-CE-0456927/00-41
SMART number: 2011/0055

DISCLAIMER

By the European Commission, Directorate-General of Communications Networks, Content & Technology.

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

ISBN number 978-92-79-31142-0
DOI: number 10.2759/25446

Table of Contents

EXECUTIVE SUMMARY 7
1. Introduction and Background 14
1.1. Background 14
1.2. Definition of Cloud Computing 14
1.3. European e-infrastructures: definitions 16
1.4. Structure of the report 16
2. THE DEVELOPMENT OF CLOUDS FOR GOVERNMENT AND SCIENCE 17
2.1. Introduction 17
2.2. Cloud Computing in the EU: key findings 17
2.2.1. Cloud policies for government 17
2.2.2. Cloud policies for sciences and research 18
2.2.3. Conclusions on cloud development models 19
2.3. Towards cloud cross-border e-infrastructures for research 21
2.3.1. NREN cloud strategies 21
2.3.2. Emerging challenges for NREN and Cloud services 22
2.4. European cross-border e-infrastructures initiatives 23
2.4.1. The SIENA Roadmap 25
2.4.2. Helix Nebula: developing the European Science Cloud 25
2.4.3. The EGI InSPIRE project 26
2.4.4. e-Fiscal: analyzing the costs of e-infrastructures 27
2.5. Case Studies 27
2.5.1. Overview 28
2.5.2. Type of services and sharing of infrastructures 29
2.5.3. Comparative analysis of governance, funding models and potential synergies 30
2.5.4. Key Findings 32
2.5.5. VENUS-C: pilot case studies with a user-centric approach 33
2.5.6. Improved business opportunities for SMEs: Green Prefab 34
Benefits identified & future prospects 35
2.5.7. Improved business opportunities for SMEs: Molplex Ltd 35
Benefits identified 36
2.5.8. Technical University of Valencia, Spain – bioinformatics 36
Benefits identified & future prospects 36
2.5.9. A new project to watch: PRISMA 37
3. ANALISYS OF DEMAND FOR E-SCIENCE CLOUDS 39
3.1. Introduction 39
3.2. The main user communities 39
3.2.1. The types of research activities 39
3.2.2. The research domains 39

3.2.3. The main Stakeholders involved in research 40
3.2.4. The main communities 41
3.2.5. Distributed Computing in Europe and Clouds 42
3.2.6. The level of openness of research 43
3.3. Cloud Computing and Science 2.0 44
3.4. Drivers and Barriers of Cloud Deployment in e-Science 45
3.5. Drivers and Barriers of pan-European cloud e-infrastructures in Europe 47
3.6. Potential synergies between e-Science and e-Government clouds 49
3.6.1. Perspectives of change 51
4. FORECASTING DEMAND FOR THE EU SCIENCE CLOUD 52
4.1. The European Scientific Cloud Forecast Methodology 52
4.1.1. Taxonomy and scope 52
4.1.2. Macro-Assumptions 52
4.1.3. Growth-Assumptions 53
4.1.4. Demand of Big Data in e-Science 54
4.1.5. Forecast 55
4.1.6. Summary Forecast Scenarios and System Value 59
4.1.7. Concluding remarks 61
5. ANALISYS OF GOVERNANCE AND FUNDING MODELS 63
5.1. Governance, Funding and Organizational Attributes 63
5.2. Actors 64
5.3. E-Science Grids for Clouds 65
5.3.1. Definition 65
5.3.2. Governance 65
5.3.3. Funding 66
5.3.4. Organization 66
5.4. Private or Community Clouds 67
5.4.1. Definition 67
5.4.2. Governance 68
5.4.3. Funding 68
5.4.4. Organization 69
5.5. Public e-Science Cloud 69
5.5.1. Governance 70
5.5.2. Funding 70
5.5.3. Organization 70
5.6. Concluding considerations 72
6. CONCLUSIONS AND POLICY RECOMMENDATIONS 75
6.1. Conclusions 75
The need for pan-European cloud e-infrastructures 75
Potential synergies between e-Government and e-Science Clouds 76
Emerging Initiatives and Funding Models 76

6.1.1. Technical challenges 77
6.1.2. Market and Business challenges 78
6.2. Recommendations 79
6.2.1. Promote the integration and federation of clouds 80
6.2.2. Promote the use of clouds across multiple scientific domains 81
6.2.3. Require full costs assessments in Research projects 82
6.2.4. Promote new business models 83
6.2.5. Create the next generation of cloud enthusiasts 83
6.2.6. Promote the development of innovative SMEs in cloud services 84
6.2.7. PRACE to start offering cloud services with a pay-as-you-go model 84
MAIN REFERENCES 86
GLOSSARY 88
FINAL WORKSHOP REPORT 91
Scenario & Key Findings 91
Summary of results 91
Case Studies Overview & Main Findings 92
Workshop – EC Policy 94
Potential demand and provisioning scenarios 95
eFiscal initial findings 95
Steven Newhouse presented the EGI perspectives 96
Bob Jones presented the Helix Nebula perspectives 96
Interactive Discussion – Main points 97
EC Conclusions 98
Agenda and Participants 100

Table of Figures

Figure 1 Stakeholders in scientific research 40
Figure 2 Distributed Computing Ecosystem in Europe 42
Figure 3 Forecast of Cloud-Friendly portion of demand in Europe (Number of Cores, 000). 57
Figure 4 Forecast of Cloud-friendly portion of pent-up demand in Europe (Number of cores, 000) 59
Figure 5 - Average Weight of Cost Components in HPC Systems (based on 2012 spending data) 61
Figure 6 Summary view of E-Science Cloud Provisioning Scenarios 74

Table of Tables

Table 1 Classification of MS by level of activity in Cloud policies for Government 17
Table 2 Classification of MS by level of activity in Cloud policies for Science and Research. 19
Table 3 - Snapshot of NREN Cloud computing initiatives & pilots 21
Table 4 Comparative analysis of case studies/ Type of services 29
Table 5 Comparative analysis of case studies/ Synergies and scalability potential 31
Table 6 Increasing Openness of the Scientific Research Process 45
Table 7 Main Drivers and Barriers of Cloud adoption in e-Science 46
Table 8 Main Drivers of pan-European cloud e-infrastructures for science and research 48
Table 9 Main Barriers to pan-European/ cross-border cloud e-infrastructures for science and research 49
Table 10 IDC Worldwide Data Intensive (Big Data) Focused HPC Server Revenues ($ Millions) 54
Table 11 Forecasts of Europe computing demand from the "core" scientific community 55
Table 12 Forecast of Europe computing demand from the "core" scientific community: cloud friendly portion of existing demand 56
Table 13 Forecast of Europe computing demand from the "core" scientific community: cloud friendly portion of pent-up demand 58
Table 14 Total Forecast Scenarios - Impact of Additional Government Investments on potential Cloud demand 60
Table 15 Total Forecast Scenarios - Total System Value to meet demand (Mill €)............. 60
Table 16 - Main Attributes of Clouds Provisioning Scenario 1....................................... 67
Table 17 - Main Attributes of Clouds Provisioning Scenario 2....................................... 69
Table 18 - Main Attributes of Clouds Provisioning Scenario 3....................................... 71
Table 19 - Summary Assessment of Attributes of E-Science Cloud Provisioning Scenarios 73

EXECUTIVE SUMMARY

Emerging demand for European clouds for e-Science

The European policy strategy for the implementation of the ERA and the pursuit of excellence in science by Horizon 2020 reflect the increasing relevance of scientific and research activities for growth and innovation. The development of e-infrastructures across Europe is becoming an enabling condition for the "fifth freedom" of knowledge and data mobility in the EU, responding to emerging demand for open, flexible and scalable computing capacity that national resources cannot satisfy. As the demand for e-Science grows beyond the traditional boundaries of national research networks and big science projects, there is a clear potential for cloud computing infrastructures and services to fill the gap between traditional offerings and emerging demand, as documented by this report.

European investments in e-infrastructures and a wide range of pilot projects have already provided early demonstrations of the potential of cloud to transform science, address big data challenges and enable collaboration across a much wider research community, as documented by this report. This broader community numbers in the thousands for each traditional supercomputer user and is important to ensure sustainable e-infrastructure. There are also clear examples of researchers pursuing new lines of research, creating start-ups and growing businesses that we need to drive the economy. Clouds are well suited to respond to peak usage or oscillatory demand for computing power, as well as to a range of emerging applications combining research and governmental risk management activities for example in the healthcare and environmental protection fields. Therefore the existence of a strong potential demand for e-Science clouds, including not only physical resources, but also data sources, services leading right through to computation, is clearly proven. However, demand cannot be satisfied only or mainly by public cloud commercial offerings for science, which exhibit several limitations in terms of actual capability, lack of transparency, compliance with regulation, and sometimes even higher pricing than traditional DCI.

The need for pan-European cloud e-infrastructures

There is a natural pan-European dimension for cloud e-infrastructures, given the transnational nature of science and research, and the strong demand by users for freedom of choice beyond national boundaries. The two forecast scenarios highlight how sensitive potential demand is not only to the capacity, but also to the technical capability of e-infrastructures. This requires the investment of considerable resources into the extension and upgrade of e-infrastructures for cloud extended to the entire EU, since the pent-up demand is distributed across the EU27, as is the emerging demand from small research centers and researchers engaged in smaller projects (sometimes called SMS, Small and Medium Science, in contrast to the Big Science projects typical of physics and other natural sciences requiring very high investments for research). A pan-European infrastructure for cloud will be able to avoid any potential "digital divide" between the large/small countries and between Big Science/Small and Medium Science (SMS) projects insuring equal access to computing resources across the EU.

While market drivers pushing the development of e-Science clouds exist, there are considerable barriers and challenges to be overcome in order to achieve EU-wide cloud infrastructures, avoiding potential digital divides between MS, duplication of investments at the national level, and fighting the proliferation of standards and lack of interoperability, which may reduce the potential benefits of cloud services. Given the current fragmented scenario, there is a need for a coherent set of policies, standards and services supporting the development of pan-European cloud infrastructures to achieve economies of scale and scope. Without such an approach, much of the potential demand may risk remaining unsatisfied. There is, however, still considerable uncertainty about the best business models for supporting these infrastructures.

The EGI.eu Federated Task Force and the Helix Nebula project are both high profile initiatives working to solve these problems. Helix Nebula has established a growing public-private partnership of 30 commercial cloud providers (suppliers) and publicly funded research organisations (users). Three high-profile flagships sponsored by CERN (high energy physics), EMBL (life sciences) and ESA (earth science) have been deployed and extensively tested across a series of cloud service suppliers. These commitments behind these initial flagships have created a critical mass that attracts suppliers to the initiative, to work together and make investments. Links have been established with DANTE and a number of NRENs so that the commercial data centres around Europe have been accessed by the user organisations via the GÉANT network. These deployments and tests have revealed a series of gaps in the current set of offerings on the cloud market and the appreciation that the best means of promoting Europe's leadership is to create an open standards based multi-vendor federated market which will allow the diversity of Europe's suppliers to compete with global leaders such as Amazon, Rackspace and Google.

Based on the experience gathered from these "proof of concept" deployments, Helix Nebula's architecture group, led by a series of cloud-savvy SMEs, have defined standards based federated cloud architecture to enable an open platform for science innovation. EGI.eu is contributing to the development of the architecture so that the EGI publicly funded e-infrastructure can be interfaced with Helix Nebula. Flagship applications from more research disciplines that will stretch the functionality and impact of Helix Nebula have been identified for deployment during 2013.

Estimate of potential demand of cloud in e-Science

The universe of science and research is in rapid evolution, no less because of the trend towards greater transparency and communication with non professional scientists, a phenomenon that is also in the early stages of research (the emergence of science 2.0). But there is also increasing demand for computing power by social science and humanities and an explosion of demand for high-volume data integration and analysis (big data). To assess potential demand the report defines a taxonomy of research domains, types of research, and types of stakeholders (research communities). The demand model is focused on open, non-proprietary research (excluding classified and protected research, or industry research not peer-reviewed).

The main user communities identified are the following:

- A "core" scientific community, spanning the tens of thousands of scientists active in the European research projects and using European e-Infrastructures. This will involve the scientists using the distributed computing platforms (EGI - European Grid Initiative or e-Science grids), and the HPC initiatives (PRACE, DEISA).
- An extended research and higher education community, including also the university professors and students that use shared infrastructure to collaborate on research (for example the GÉANT2 high-bandwidth, academic Internet serving Europe's research and education community, which serves over 30 million researchers with a multi-domain topology spanning 34 European countries and links to a number of other world regions).
- An open and wider research community, including the non- professional scientists.

The forecast model assumes that the demand for clouds in e-Science will expand within the "core" scientific community and also from this community to the extended research community. At this stage, we are not able to estimate the demand of the wider research community, including non-professional scientists, which may involve the interaction with individual PCs and so on.

The forecast model estimates the demand for government-supported community and public clouds of e-Science infrastructures, from 2011 to 2016. The starting point is the estimate of pent-up demand, which is approximately four times the current level of HPC demand (which in IDC's taxonomy extends to all of the technical and scientific computing market).

According to IDC estimates, based on our ongoing surveys of the market, the total number of computing cores installed in technical computing systems worldwide in 2011 was 51 million; approximately 16 million are estimated to be in Europe. Of these, IDC estimates that there were approximately 2.5 million computing cores in use in 2011 for open science, which could run in the cloud given today's technology attributes (such as performance, interoperability, security). The percentage of demand which is expected to migrate to the cloud by the year 2016 varies from 45% of existing cores, corresponding to 4.2 million cores (in the no additional investments, low satisfaction of pent-up demand scenario) to 80% of existing cores, corresponding to 63 million cores (in the additional investment, high satisfaction of pent-up demand scenario).

The two forecast scenarios highlight how sensitive potential demand is not only to the capacity, but also to the technical capability of e-infrastructures. This requires the investment of considerable resources into the extension and upgrade of e-infrastructures for cloud extended to the entire EU, since the pent-up demand is distributed across the EU27, as is the emerging demand from small research centers and researchers engaged in smaller projects (sometimes called SMS, Small and Medium Science, in contrast to the Big Science projects typical of physics and other natural sciences requiring very high investments for research). Cloud Provisioning Scenarios

The report carries out a comparative assessment of the governance, funding models and technical attributes of the 3 main typologies of cloud provisioning scenarios, e-Science grids, private clouds and public clouds. In reality, "hybrid" provisioning scenarios are more likely to be dominating the EU future developments. For example, the experience of the Helix Nebula project is moving towards the development of a standards based federated cloud architecture to enable an open platform for science innovation, based on the federation of publicly funded, community grids (such as the EGI) and commercial providers. While the HN project started out as a "community" provisioning scenario open to the research centres of the communities involved, it is moving towards open access based also on commercial services (that is, open to anybody who can pay, even if with some differences compared to the completely commercial platforms such as Amazon's).

Main challenges for the development of pan-European cloud e-infrastructures

Within this context, the main challenges identified by this study for the evolution of flexible, scalable and quality cloud services in Europe are the following.

Technical challenges

- **Need to develop open standards** and technical requirements enabling the development of open, interoperable and federated cloud services;
- **Need to design and develop innovative and integrated services** for e-infrastructure, aimed at diversifying service offers for different communities as much as possible.
- **Need to advance the capabilities of these clouds**, especially faster communications to support scientific work that is less embarrassingly parallel and more tightly coupled.
- **Need to fill the technical gaps in the offering**. There is yet a lack of maturity of SLAs on performance of development, testing and production environment, business continuity, data access, migration and helpdesk for cloud services and how they apply to different usage scenarios.
- **Lack of commitment by commercial cloud providers** for the development of applications and services needed by e-Science, due to uncertainty in the assessment of the potential demand and the risk-benefits balance.

Market and Business challenges

- **Need to develop new business and funding models**, where users start paying for the e-infrastructures services they consume and providers compete for innovation money and to generate revenue from their customer base. New funding models should be easy to manage, because otherwise the cost of fee for service mechanisms outweighs the benefits

- **Need to move beyond the CAPex funding model in research funding**. Cloud computing does not fit well with the typical science grant budgets. There are two main challenges: fostering a change in funding policies and raising awareness of the many opportunities of using cloud computing for research.
- **Need for user-centric approaches, possibly creating a market for e-research**. The main challenge lies in re-directing the e-infrastructure provision strategy focusing it on user needs. A possible approach to do so could be the creation of a "data and e-research market" in Europe both for scientific and social science applications. This approach sees the technology (grid, cloud, combined approaches) mainly as a tool that glues together data and user communities across diverse fields. The aim is to develop an environment where research data can be generated through the cloud, widely shared and create valuable knowledge. One possible model to do so is the creation of application store-like services charging a small fee, payable to the data/software provider with a high flux of transactions, plus a consultation fee.
- **Need for more agile provisioning model for science**. Traditional peer review is an effective model for determining which scientific projects are granted access to HPC resources, But the peer review process can take months to complete. This does not exploit the elastic ability of clouds to respond to more immediate needs of scientists and other researchers for additional capacity or special capabilities. Peer review bodies should consider defining circumstances in which researchers can gain more rapid access to cloud-based resources.
- **Solve legal and compliance issues around contracts and SLAs**. These issues concern service providers' accountability, liability, compliance with data and privacy protection regulation, both national and cross-border. They are well known and are being addressed by the European Cloud strategy. However, the e-Science environment has specific requirements which need to be addressed (for example, because of the different balance of requirements in terms of open access to data and knowledge but also intellectual property issues).
- **Promote Open Access to data. Open e-Infrastructures are a strategic resource that need open policies and open access to ensure that they remain the platforms for Innovation**. Not only is open government data opening up new possibilities in areas such as environmental data but there are many opportunities to develop new applications as part of the drive to address societal challenges and accept the risks involved.

Policy Recommendations

Given the multiple market and technical challenges already discussed, the European Commission has an important role to play to promote the development of pan-European cloud e-infrastructures, insuring the availability of cloud-based quality services to research and science, as well as the public sector.

The EC should promote and sustain the spontaneous movement towards the integration and federation of clouds at the EU level, avoiding the risk to develop top-down infrastructures totally dependent on public funding and unable to adapt to the multidimensional characteristics of demand. This will allow supporting the provision of cloud services for science and research across Europe, filling the gap between the actual offering of computing resources and the emerging, pent-up demand.

The EC should however make sure that the provision of cloud services for science and research covers the whole of the EU27, avoiding any potential "digital divide" between the large/small countries and between Big Science/Small and Medium Science (SMS) projects, insuring equal access to clouds for all researchers.

The development of pan-European cloud e-infrastructures and services will support the achievement of the main challenges of the Horizon 2020 Programme as follows:

- **Achieving excellence in science**: the availability of cloud services for science and research across the EU27 will fill the gap between the actual offering of computing resources and the emerging and pent-up demand by researchers across the EU.
- **Meeting social challenges**: the development of an ecosystem of value-added cloud services addressed to research and the public sector, driven by demand and based on user-centric approaches, will help the emergence of e-Science-as-a-service business models, provide an open, scalable and flexible environment for large-scale collaboration between scientists and citizens in an open science/ science 2.0 perspective and for increasing collaboration between science and government.
- **Promoting Industrial Leadership and Competitive Industries**: this will be achieved by the promotion of interoperability through open standards for the provision of cloud services across the EU, responding to emerging research and industry demand needs; on the other hand by supporting the transition of European e-infrastructures and cloud service providers towards more sustainable business models. This should include the development of the business case for EU cloud e-infrastructures through the achievement of economies of scale and scope.

In order to implement these recommendations, the key elements of an EC strategy in this area should be the following:

1. ***Use EC funding and initiatives to promote the integration and federation of clouds and enable the migration from e-infrastructures towards a European marketplace of connectivity and cloud services for e-Research***
2. ***Promote and extend the use of clouds across multiple scientific domains and the development of a cloud services ecosystem, in order to narrow the gap between the supply and user communities and overcome cultural and resistance barriers to cloud.***
3. ***Support the consistent, comprehensive and business-case oriented analysis of cloud computing costs compared to other computing resources, requiring full costs assessment in all public funded projects***
4. ***Promote the transformation of the business models and organizational structure of e-infrastructure providers***
5. ***Create the next-generation of cloud enthusiasts, supporting the change of mindsets the development of the new skill sets needed for new clouds services and e-infrastructures***
6. ***Promote the development of innovative SMEs developing cloud-based services, also leveraging spin-offs and start-ups***
7. ***Continue strong support for PRACE and promote the addition of cloud capabilities and at least partially "pay-as-you-go" access models to HPC centers, to extend their utilization and best exploit their resources***

Background

This is the Final report of the study “Clouds for Science and Public Authorities” entrusted by the European Commission, DG CONNECT to IDC and TRUST-IT. The main goal of this study is to analyze the current and prospective development of cloud computing infrastructures for e-Science and e-Government in Europe, contributing to the development of a EU strategy for cloud computing in this field, as foreseen by the Digital Agenda for Europe (DAE). The ultimate goal of the study is to promote the development of cloud-based quality services to researchers, public sector employees and the public at large.

This report builds on desk research, a wide range of interviews with key stakeholders and case studies in Europe and the world, collected in national profiles of cloud policies and initiatives for e-Science and e-Government for the 27 EU and 14 other world countries. IDC's worldwide databases on cloud computing, HPC and technical computing were used to develop a forecast model of potential cloud demand in e-Science.

The final conclusions and recommendations were discussed and validated in a workshop with main stakeholders at the EC premises, in Brussels, on November 26, 2012. We wish to thank all the workshop participants for their extremely useful feedback and comments. Special thanks are due to:

- **Bob Jones**, CERN & Helix-Nebula The Science Cloud Coordinator
- **Fotis Karagiannis**, eFiscal coordinator & e-IRG Task Force on cloud Computing
- **Nikos Athanasis**, University of the Aegean & Wildfire
- **Steven Newhouse**, EGI-InSPIRE Project Director and EGI.eu Director
- **Radu Prodan**, Associate Professor, Institute for Computer Science University of Innsbruck
- **Linda Strick**, Fraunhofer FOKUS, Germany
- **Nicola Westmore**, Deputy Programme Director - G-Cloud and Common Infrastructure
- **Christopher Cox**, Programme Manager, Cheltenham Borough Council
- **Giuliano Franceschi**, DG Lepida
- **Vangelis Floris**, GRNET in-house cloud solution OKEANOS
- **Ignacio Blanquer**, Professor, Technical University of Valencia and VENUS-C Community Manager

1. INTRODUCTION AND BACKGROUND

1.1. Background

This is the revision of the Final Study Report of the study "Clouds for Science and Public Authorities" entrusted by the European Commission, DG CONNECT to IDC and TRUST-IT. The main goal of this study is to analyze the current and perspective development of cloud computing infrastructures for e-Science and e-Government in Europe, contributing to the development of a EU strategy for cloud computing in this field, as foreseen by the Digital Agenda for Europe (DAE).

Cloud computing represents a fundamental change in the way computing power is generated and distributed, transforming the delivery of IT tools and products into elastic, on-demand services characterised by flexible 'pay-as-you-go' payment models.

This study focuses on the potential use of cloud computing to support both e-Science and e-Government benefiting from previous research and activities in these areas. It builds on the assessment of the European environment of clouds for science and government, opportunities and barriers to their development, and possible synergies between the two environments. The study takes a broad approach to cloud computing in order to cater for a diverse and fast-evolving landscape, where best practices and new ecosystems are beginning to emerge.

The ultimate goal of the study is to promote the development of cloud-based quality services to researchers, public sector employees and the public at large.

This report builds on the research carried out in the first phase of the study and is mainly focused on the feasibility of a potential e-Science cloud computing infrastructure in Europe (including size, characteristics, governance and funding models), in order to provide recommendations on the best way to leverage public demand for its development.

1.2. Definition of Cloud Computing

Within the context of this study, IDC has adopted a definition of cloud computing, based on 10 years of market research:

"**Cloud computing services are consumer and business IT products, services, and solutions delivered and consumed in real time over the Internet**".

IDC's definition builds on a granular taxonomy of the functions and services offered, useful to analyse the supply offering. This definition allows leveraging for this study the accumulated research by IDC on public and private clouds.

The study has also taken into account the definition of cloud computing most diffused in the scientific environment proposed by the US National Institute for Standards and Technology (NIST)[1]:

"**Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction**."

More specifically, within the context of this study we have used the following definitions of the main delivery models of cloud services:

- In private clouds, services are provided exclusively to trusted users via a single-tenant operating environment. Essentially, an organisation's data centre delivers cloud computing services to clients who may or may not be on the premises.

---

[1] Mell P & Grance T 2009. The NIST definition of cloud computing: http:// csrc.nist.gov/ groups/SNS /cloud-computing/cloud-def-v15.doc

- Public clouds are the opposite: services are offered to individuals and organisations who want to retain elasticity and accountability without absorbing the full costs of in-house infrastructures. Public cloud users are by default treated as untrustworthy.
- Hybrid clouds combine both private and public cloud service offerings. . A typical scenario in hybrid cloud services is an infrastructure service provided from a private cloud that will "burst out" into the public cloud when the capacity of the private cloud vendor is insufficient to meet demand.
- A “community cloud” is one where the cloud infrastructure is shared by several organizations and supports a specific community that has common concerns (e.g., mission, security requirements, policy, and compliance considerations). The cloud may be managed by the organisations or a third-party and may exist on or off-premise.

List of Abbreviations

CapEx	Capital Expenditure
CCUCDG	Cloud Computing Use Cases Group
DAE	Digital Agenda for Europe
DANTE	Delivery of Advanced Network Technology to Europe
DMTF	Distributed Management Task Force
FCC	Federal Communication Commission, US
EGI	European Grid Initiative
e-IRG	e-Infrastructures Reflection Group
EIRO	European Intergovernmental Research Organisation
ERA	European Research Area
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategy Forum on Research Infrastructures
HN	Helix Nebula
HPC	High Performance Computing
HTC	High Throughput Computing
IaaS	Infrastructure as a Service
NGI	National Grid Infrastructure
NREN	National Research and Education Network
OpEx	Operation Expenditure
PaaS	Platform as a Service
PRACE	Partnership for Advanced computing in Europe
SaaS	Software as a service
SPoF	Single Point of Failure

SLA	Service Level Agreement
TCO	Total Cost of Ownership
VRC	Virtual Research Community

1.3. European e-infrastructures: definitions

The Digital Agenda for Europe identified a new, emerging demand of joint ICT infrastructures and innovation clusters, where the provision of cloud-based quality services may satisfy emerging needs for open, flexible and scalable computing services.

The analysis of a potential pan-European e-Science cloud computing infrastructure must take as a starting point the past evolution and current landscape of e-infrastructures.

Within the EU context, **(ICT-based) e-infrastructures for science and research** are defined as electronic services which integrate physical computing, storage, networking and other hardware to connect researchers from all disciplines with the reliable and innovative ICT services for uniform access to commodity computing, archiving, and management of distributed data that they need to undertake their collaborative world-class research.

By 2020, these capabilities need to be deployable on demand in order to provide a foundation for the digital European Research Area (ERA). They foster the emergence of new working methods, based on the shared use of resources across different disciplines and technology domains, enabling sustainable collaborations and partnerships between researchers in "virtual research communities" in all e-Science fields. This will create a single European space for "online research".

1.4. Structure of the report

The report is articulated as follows:

- The first chapter presents the background and the main definitions used in the study;
- The second chapter reviews the level of development of national and cross-border policies and initiatives of clouds for science in Europe, including some case studies;
- The third chapter presents the analysis of the current and potential demand for e-Science cloud infrastructures in the EU, including the main drivers and barriers;
- The fourth chapter presents the forecast of demand for the EU science cloud;
- The fifth chapter analyses governance and funding models issues under the main possible provisioning scenarios;
- The 6th chapter presents the final conclusions and recommendations, revised after the discussion with the main stakeholders at the final workshop of the study, in Brussels, on November 26, 2012.

The annexes to the report include:

- The report on the workshop held on November 26, 2012
- The main references
- A glossary of the main terms used in the report

2. THE DEVELOPMENT OF CLOUDS FOR GOVERNMENT AND SCIENCE

2.1. Introduction

This chapter presents an overview of the main national policies and strategies of cloud computing, already implemented or planned for the near future in the EU and the main world countries. This is based on the Cartography developed for this study through desk research and interviews[2]. The goal is to provide a review of the concrete initiatives of development and deployment of cloud computing services and infrastructures currently ongoing for government and science, to provide a baseline scenario to investigate potential commonalities and differences.

2.2. Cloud Computing in the EU: key findings

2.2.1. Cloud policies for government

The current landscape in Europe is very diverse both in terms of strategies and on-going initiatives and with regard to both government and science policy approaches.

The following table presents a snapshot of the level of development of cloud policies and initiatives for government in the EU27 (based on the data from the cartography). It includes the following information:

- Existence of national strategies promoting the adoption of cloud computing in government organizations (Yes, Planned, No);
- Existence of national government initiatives (including competitive calls) to develop-deploy cloud infrastructures and tools for public authorities (Yes, Planned, No); we tend to call these for short top-down initiatives.
- Whether public organizations (e.g. public authorities or academic institutions) are starting to use cloud services on their own, as bottom-up initiatives.

Table 1 Classification of MS by level of activity in Cloud policies for Government

Cloud Policies and Initiatives for Government	MS	N
Cloud Policy (actual or planned), Initiatives (actual or planned) and bottom-up adoption	DE, DK, ES, FI, FR, IE, IT, UK	8
Cloud Policy (actual or planned), Initiatives (actual or planned) , no bottom-up adoption	AT, NL, PL	3
No Policy, No Government initiatives, Yes bottom-up adoption	BE, CZ, SE, SI, SK	5
Policy planned, no Initiatives and no adoption	CY, EE, EL, HU, MT, PT	6
No Policy, No initiatives, No bottom-up adoption	BG, LT, LU, LV, RO	5
Total		27

Source: IDC, Trust IT: Clouds for science and public authorities 2012

[2] Published as an annex to the Interim report of this study and validated through a survey of the EU CIO Network

As shown above, we have identified 5 main groups of MS, from the most active to the least active in cloud computing (top to bottom in the table 1). They are:

- **Germany, Denmark, Spain, Finland, France, Ireland, Italy** and the **UK** are the most active, with cloud policies either existing or planned, government-led initiatives existing or planned, and also bottom-up adoption of cloud by public authorities;
- **Austria, the Netherlands** and **Poland** have cloud policies existing or planned, government initiatives existing or planned, but we found little evidence of bottom-up adoption of clouds.
- Another group of countries, including **Belgium, Czech Republic, Sweden, Slovenia** and **Slovakia** have no current or planned cloud policies or government initiatives, but spontaneous bottom-up adoption by administrations. Belgium and Sweden have a decentralized administrative structure, which may be the reason for the lack of a central cloud policy.
- In **Cyprus**, **Estonia**, **Greece, Hungary, Malta** and **Portugal** governments are planning cloud policies, but there are no implementation initiatives yet and little evidence of bottom-up adoption.
- Finally there is a group of 5 MS (**Bulgaria, Lithuania, Luxemburg, Latvia** and **Romania**) who do not seem to place cloud high in their policy agenda, without policy strategies, government initiatives or evidence of bottom-up adoption.

Bearing in mind that we may well have missed some initiatives and some bottom-up activity, this seems to be a quite positive picture, with the majority of the MS actively engaged in promoting the development of cloud computing in the government sector. On the other hand, as will be discussed below, national plans are still very much on paper, initiatives are at the early development stage, and most often concern private or hybrid clouds (for groups of administrations) than public clouds.

Moreover, there seems to be a positive correlation between the country size and the level of activity, with the largest MS being most actively engaged, probably trying to bring some order to the diffusion of cloud services and chasing potential economies of scale.

There is also a positive correlation with the level of intensity and diffusion of IT investments, with the Scandinavian countries and the UK (the MS with the highest levels of IT spending) leading the way also in terms of propensity to adopt clouds. The laggards tend to be the Eastern European MS, particularly the smallest ones.

2.2.2. Cloud policies for sciences and research

The research and education landscape is undergoing radical changes, moving towards on-line learning and digital content, global virtual communities, large-scale collaborations. In this context, cloud-computing promises new answers to emerging needs for more computing capacity but also collaborative services and a wider researchers community.

The snapshot emerging from the classification of the MS in terms of level of activity in cloud policies for science and research is similar to the one presented above. The main groups are the following:

- 7 MS (Germany, Denmark, France, Ireland, Italy, Netherlands and the UK) have both a cloud policy for science, and at least one of the two types of cloud initiatives presented in the table;
- 6 MS (Austria, Greece, Spain, Finland, Hungary and Sweden) do not have a policy (or are planning it), but have cloud initiatives for science or research;
- 3 MS (Estonia, Malta and Poland) are planning a policy, but have no current initiatives;
- Finally in 11 MS (Belgium, Bulgaria, Cyprus, Czech Republic, Latvia, Luxemburg, Lithuania, Portugal, Romania, Slovenia and Slovakia) we found no evidence of national cloud policies, or of nationally funded science and research cloud initiatives.

Table 2 Classification of MS by level of activity in Cloud policies for Science and Research

Development of Cloud Policies and initiatives for science and research	MS	N
Cloud Policy and initiatives for research	DE, DK, FR, IE, IT, NL, UK	7
No Policy (or planned), yes initiatives	AT, EL, ES, FI, HU, SE	6
Policy planned, no initiatives	EE, MT, PL	3
No Policy No initiatives	BE, BG, CY, CZ, LT, LU, LV, PT, RO, SI, SK	11
Total		27

Source: IDC, Trust IT: Clouds for science and public authorities 2012

We see again a correlation between the size and level of IT investment in a country, and the propensity to engage in centralized cloud policies. There is a group of MS, including the UK, Germany, Denmark, France, Italy, Ireland, where national governments are active in top-down policies in both domains of government and science. In the case of science and research, however, decentralized initiatives seem to be more frequent than is the case for government. In both domains, a sizable minority of MS are not actively engaged in cloud policies and maybe lagging behind in the adoption of clouds in the public sector. This group includes many of the Eastern European countries.

2.2.3. Conclusions on cloud development models

A key consideration emerging from the Cartography analysis is that cloud computing is still in its nascent stage, particularly with regard to government and public authorities and with strategies still at early planning or investigation stage. This is particularly true for Europe, where the "time to market" from policies to actual implementation is particularly long, certainly much longer than in the US or China.

Another important finding is that patterns of adoption of government and science are moving along two different trajectories, driven by different demand drivers (explored more in depth in the next chapter). In the e-Government field, the main driver so far has been IT cost optimization/ affordability of alternative IT service delivery models; in academic research, the main driver is the sustainability of collaborative efforts and access to sufficient computing power and better performance. In some cases (as shown by the Magellan project analysis in the US and by the Cost analysis report in the UK) cloud computing may be effective but not necessarily at cheaper costs.

While funding agencies are concerned about the cost of research computing, evidence shows that this is not always a top priority, especially for researchers looking for new functionalities. Funding agencies are slowly beginning to respond to these needs. For example, the UK's National Research and Education Network (NREN) Janet now plays a broker role to facilitate smart adoption. Janet's funding agency, JISC, is driving a policy for adopters to use mechanisms to assess costs such as an "impact calculator". Sustainability is, without doubt, one of the major challenges facing the research and scientific community today and where more concerted efforts are needed.

In terms of government-science synergies, there is little evidence to date of potential collaboration at the higher, macro level, mainly due to very different requirements and solutions sought. This finding is aligned with the findings of a Fraunhofer FOKUS 2011 report and the German government's approach to funding in general. The report, "Cloud concepts for public sector in Germany" (August 2011), summarizes the findings of a comparison between the requirements identified for the public sector with the requirements identified by the EU project, VENUS-C (www.venus-c.eu), a PaaS for research and SMEs. The scenarios identified by VENUS-C combine computing needs in high-throughput and high-performance paradigms, workflow management, intensive data, and integration with external sources and different types of users. Most of the scenarios use some degree of parallel computation approach; require accessing external sources of data and network connections, and require authenticated access mainly through login and password. Anonymous access is not demanded.

"*The eScience related requirements for an eScience-cloud-platform can hardly be compared with the public sector related requirements discussed in this White Paper, because it addresses different issues influenced by eScience Grid computing environments. Nevertheless a deeper look into future VENUS-C deliverables may be worthwhile*"[3].

There is, however, evidence of current and potential synergies at the micro level and wider deployment including examples of solutions aimed at addressing challenges such as climate modeling and weather forecasting, as well as public healthcare solutions. A number of these initiatives have been defined in synergy with local and regional government.

There is an interesting trend towards initiatives focused on the **procurement** of cloud services, where national authorities act as "brokers" for public authorities for cloud applications and services, at the same time providing accreditation of the vendors for the data protection and regulation compliance aspects. This is the case for example of the GoBerlin platform for cloud computing promoted by the German Trusted Cloud Technology Programme and of the SURFconext collaborative research environment and brokerage service promoted by the Dutch NREN.

The U.K. CloudStore is the most innovative example, because the framework procurement approach has enabled leading edge authorities to purchase services quickly for low risk use-cases, such as document collaboration environments for projects with non-sensitive information. However, the complexities of pan-government information assurance accreditation have limited the scalability to a wider audience of more risk-averse public sector entities, or use-cases where security concerns are higher. Again, this is a case to watch since it is only in its initial stage.

The comparative analysis of EU and other world policies and initiatives highlights some approaches and solutions, which should be considered closely for a EU cloud strategy in this field. Some of them are already implemented by some MS. They are the following:

- **Cloud-First and government-wide approaches** as spearheaded by the US, placing cloud at the heart of the ICT policy. European leaders include Denmark, Ireland, Netherlands, UK, Finland, France, and Germany.
- **Shared services approaches** pursued by Canada, New Zealand and the US, with the aim of avoiding duplication and reducing costs wherever common requirements and concerns can be satisfied with the same approach. The UK's shared government services strategy is focused on business services such as payroll processing, order management, invoice/ account payable management. However, it could serve as a model to extend the cloud approach to other e-Government and e-Science services.
- **Actions to boost competitiveness and economic growth as a common priority**. The **UK App Store** (CloudStore) is a leading example of how public sector procurement can significantly increase SMEs role as suppliers. Another interesting example is **the Finnish Cloud Software Program** animated by TiViT, the largest ICT-related industrial research initiative funded in Finland in recent years with a total budget of €60 million over four-years (2010-2013). The initiative brings together 22 companies and 8 research institutions with the aim of creating new business activity in Finland in the value chains of internet services and in the areas of sustainable development, user experience and information security. The Cloud Software Program has focused on creating a new ecosystem that prioritizes the most profitable cloud services.

---

[3] http://www.interoperability-center.com/c/document_library/get_file?uuid=9176f0fa-1ea2-4771-b8e0-a3f9c685199f&groupId=12725, August 2011, report page 33.

- **Collaborative work between government and research** funding agencies, sharing experiences on current cloud service providers to pinpoint the benefits and issues. This has proved highly valuable for the Irish strategy and to define its implementation roadmap.
- **Open data initiatives** with the release of open APIs to drive new service creation, including by small businesses with countries like the US leading the way.

2.3. Towards cloud cross-border e-infrastructures for research

2.3.1. NREN cloud strategies

E-infrastructures for science were built to support scientific collaboration and enable digital research to tackle modern challenges. However, the research and education landscape is undergoing radical changes, moving towards on-line courses and digital content, global virtual communities, large-scale collaborations. The commoditization of virtualized technologies and the emergence of cloud computing from commercial providers are becoming attractive to several scientific domains or groups.

These new trends are impacting directly the role of NREN. According to this study research, the most relevant cross-border initiatives of cloud services for science and research (besides public cloud commercial services, of course) are generated by NREN. It is worthwhile to analyse more in depth the evolution of NREN and their main challenges.

European NREN is showing increasing activity in the field of cloud computing, within the context of the evolution of their strategic role and positioning. NREN are under pressure to respond to increasing demand of connectivity and computing, as well as to globalisation trends making the national user constituency model obsolete. NRENs are experimenting with new business models and services portfolios, and cloud services seem to be a promising direction to reinforce and renew their positioning.

The following table shows our findings about the development and implementation of cloud policy and pilot initiatives by European NRENs.

Table 3 - Snapshot of NREN Cloud computing initiatives & pilots

Level of development of Cloud Policies	MS*	N
Cloud strategy implementation & service deployment	EL, ES, HU, IE, NL, UK	6
Cloud pilot and/case studies	IT (GARR)	1
Planning of strategy & service deployment	DE, DK*, FI*, FR, SE*	5
No strategy or planning	AT, BE, BG, CY, CZ, EE, LT, LU, LV, MT, PL, PT, RO, SI, SK	15
Total		27

Source: IDC, Trust IT for Clouds for science and public authorities study, 2012

***Note**: NORDUnet comprises 3 MS – Denmark, Finland Sweden and 2 Associated countries, Norway and Iceland.

EU Member States that have already deployed some type of cloud service include a range of different approaches, cost and business models as follows: **GRNET**[4], **Greece** (cloud services to the Greek research and education community); **RedIRIS**[5] , **Spain** (commodity services); **SURFnet**[6] and **SURFconext**, **Netherlands** (provides email, conferencing and document sharing services for academic and research institutions and a pilot implementation of MS Office 365 for education, see also the case study in chapter 6); **NIIF/Hungarnet**[7], **Hungary** (private IaaS cloud solution); **HEAnet**[8], **Ireland** provides (EduStorage to academic and research institutions based on a pay per TB/p.a. business model and negotiates preferential rates with providers for Internet services); the **Danish e-Infrastructure Cooperation – DeIC**[9], **Denmark** (developing plans for a national cloud infrastructure serving the wider research community); **Janet**[10], **UK** (provides cloud brokerage advice and services, guidance on legal and regulatory issues and is developing Janet6 as the next-generation e-infrastructure backbone). Finally, **NORDunet** is planning to develop new cloud-based services.

GARR, **Italy** is conducting a pilot initiative (web hosting) and is evaluating the broader use of the cloud along with an assessment of benefits and challenges. A complete strategy including new cloud services is planned for early 2013.

2.3.2. Emerging challenges for NREN and Cloud services

The shift away from a technical focus towards a commodity sphere means that NRENs will need to ensure added value as technology enablers and digital knowledge transfer assistants. Traditionally, NREN customers have been higher education and research institutions (both public and private). The user base is now becoming more diversified (e.g. US and UK – US UCAN[11]), spanning all of education, health and culture institutions and the public sector at large. The FCC[12] (Federal Communication Commission) in the US has given a clear mandate to extend the user base widely. The UK is one of the first countries to accept this mandate. However, challenges around the EU legal framework for NREN, which is rooted on national user constituencies, remain.

Main challenges for example are:

- Who will provide cross-border common services, e.g. AAI services (Authentication, Authorisation and Identification)? Networks or Grids? According to which regulation, since national requirements are still different, notwithstanding the EU efforts to develop common environments?
- Heterogeneous user communities expect more integrated services (IaaS) that are tailored to the community, not institutions. This means for NREN to grow beyond their original, captive market and become more user-centered.
- Data storage issues, including environmental aspects, conservation and long-term preservation.
- Funding models, which are changing rapidly. n the last years, governments are asking NREN to move from a complete dependence on public funding, to collect at least some payment for their services. For example, in 2005 NORDUnet received 100% funding contribution, but in 2012 public funding represented only 68% of their budget, with 32% coming from service revenues. The provision of cloud services, with their pay-as-you go business model, is particularly attractive in this context.

---

4 http://www.grnet.gr/default.asp?pid=1&la=2

5 http://www.rediris.es/

6 http://www.surfnet.nl/en/Pages/default.aspx

7 http://ipv6.niif.hu/

8 http://www.heanet.ie

9 http://www.forskningsnettet.dk/drupal/node/105?language=en

10 https://www.ja.net/

11 http://www.usucan.org/.

12 http://www.fcc.gov/

This evolution is forcing NREN to rethink drastically their governance, management models but especially to change their mindsets.

Cloud Services is one of the four areas the members of the NREN Global CEO Forum have agreed to work collaboratively on. The first meeting took place in September 2012, bringing together thirteen leaders of National Research and Education Network (NREN) organizations from around the world as a Global CEO Forum to discuss global NREN strategy. The Forum was represented by CEOs of AARNet (Australia), CANARIE (Canada), CERNET (China), CUDI (Mexico), DFN (Germany), Internet2 (USA), Janet (UK), NORDUnet (European Nordics), REANNZ (New Zealand), RedCLARA (Latin America), RENATER (France), RNP (Brazil), and SURFnet (The Netherlands) to deliberate common strategic challenges faced in delivering advanced ICT services to the Research and Education communities.

The TERENA Trusted Cloud Drive is a pilot to develop a personal data storage service for national research and education network organisations to offer their academic and research communities. The pilot is built upon an existing and continually developing federated software platform called the 'cloud broker platform'. This allows the connection of both private and public cloud storage back-ends and stores users' data in the cloud in a secure and privacy preserving way.

The CEOs recognized an urgent need to prepare a seamless global service delivery for users in the Research and Education community. Four major challenges in delivering a high-performance global cyber-infrastructure (e-infrastructure) identified are:

1. **Global Network Architecture** – A well-defined, inclusive, global architecture for, and a roadmap towards, interconnecting the Research and Education Networks on a global scale, taking into account input from the large science & education projects.
2. **Global Federated Identity Management** – A global, interworking architecture for, and a roadmap towards, the delivery of federated identity management for the R&E community to fully interoperate, using open standards and enabling global service delivery.
3. **Global Real-time Communications Exchange** – An interworking system for multi-domain video/audio conferencing systems, with directory systems that interwork based on open standards, using their identity federation & directories, capable of supporting virtual organizations.
4. **Global service delivery** – A model for global above-the-net service delivery to the NREN's constituencies, leveraging aggregation of supply and demand through scale.

For each of these challenges, the NREN leaders decided that a project owner be made available as soon as possible to drive the creation of an action plan and to deliver results as planned. Future activities will also focus on strengthening global NREN collaboration to achieve a high-performance global cyber-infrastructure. The next meeting will be in spring 2013.

Standards for globalizing cloud services

A new initiative, Charter Member - Global Service Partner Program has been established bringing together Australia, Canada, Mexico, the Netherlands, UK and US. Starting in January 2013, NREN CEOs and their representatives will be working on a set of standards for 'globalizing' cloud services with the goal of deploying these services at scale and breaking the model of institution by institution procurement. There are plenty of challenges to overcome, but the benefits are potentially quite substantial.

2.4. European cross-border e-infrastructures initiatives

Over the last ten years the European countries and the EC have made significant investments in e-Infrastructures for scientific computing, notably High Throughput Computing (HTC) and High Performance Computing (HPC) services. Their contribution in pushing the European research towards a leading position and addressing global challenges has been validated by a growing number of research initiatives, most notably the ESFRI (European Strategic Forum for Research Infrastructures) projects.

Currently, the broad definition of e-infrastructures in Europe includes as a minimum the following initiatives and organizations, which in turn federate similar initiatives and research centres at national level:

- GÉANT, the pan-European research network formed by the 32 European NREN (National Research and Education Networks), the Trans-European Research and Education Networking Association (TERENA), plus an additional four Associate NRENs. GÉANT interconnects over 40 million users across Europe.
- DANTE (Delivery of Advanced Network Technology to Europe) is a not for profit organization based in Cambridge set up, and is owned, by a group of National Research and Education Networks (NRENs). Its purpose is to plan, build and operate pan-European research networks, playing a pivotal role in the development of GÉANT.
- PRACE, Partnership for Advanced Computing in Europe: a scientific network providing access to distributed pan-European world class high performance computing and data management resources and services located in Germany, France, Spain and Italy.
- EGI, the European Grid Initiative: a federation of shared computing, storage and data resources from national and intergovernmental resource providers that delivers sustainable, integrated and secure distributed computing services to European researchers and their international partners.

These infrastructures are connected through high-speed networking to support the activities of the European research communities of up to 1.8 million publicly funded and 1.0 million privately funded researchers spread across Europe.

The European Commission hosts the ESFRI Secretariat and a Commission representative is a member of the Forum. The Commission manages EU budget contributions of up to €6 million per project to support their preparatory phase. The EU also supports clusters of projects in the same field to respond to common implementation needs, notably regarding data management and access issues.

The total EU contribution to ESFRI projects so far amounts to about €700 million, out of which €282 million has come from EU research programmes. For example, the EU granted about €48 million for the implementation of PRACE - Partnership for Advanced Computing in Europe. In addition, one project, ELI – Extreme Light Infrastructure, has been granted €416 million of structural funds for the construction of two facilities to be built in the Czech Republic and Romania. A third ELI facility is to be located in Hungary, which will also benefit from structural funds.

Further support is planned under the next research programme Horizon 2020 which will allocate a dedicated budget to support the implementation and operation of pan-European world class research infrastructures.

The emerging commercial offerings (Cloud-based HTC and HPC solutions) pose additional challenges and opportunities for the sustainability of European e-infrastructures. They represent new ways of explicit monetisation of computing and open the way to new business models (e.g. service fees or of lease industrial resources). The current economic and financial crisis is adding further pressure on public budgets. Even if policy makers recognise the need to maintain research and innovation funding as an investment in the future growth and development, the competition for scarce resources keeps growing. It is important both to demonstrate the cost-benefits ratio of the investments in research infrastructures, and to select the most efficient and effective governance and funding models for them.

There are several high profile projects and initiatives in Europe dealing with these issues. The most important are the following ones.

2.4.1. The SIENA Roadmap

The SIENA initiative (Standards and Interoperability for eInfrastructure implemeNtation initiAtive) was funded by the European Commission (EC) under the 7th Framework Programme (FP7) (June 2010 – June 2012). The project worked very closely with all the relevant standards groups to support the analysis of open-standards based interoperable grid and cloud computing infrastructures with the main result being the publication of the far-reaching Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe. The recommendations of the SIENA Roadmap constitute a "Call to Action" in the short and medium term. Within this overall plan there is a role for governments where the trend towards procurement of commercial cloud services by the public sector, including research will generate interest in standards. There is also an important role for industry to play in the international standards dialogue, implementation and certification processes, as well as in continued investments aimed at boosting European innovation. The SIENA Roadmap "Calls to Action" raised awareness about the requirements and demands of cloud for science in Europe and paved the way for several of the initiatives described in this report.

2.4.2. Helix Nebula: developing the European Science Cloud

Helix Nebula (HN) Helix Nebula has established a growing public private partnership of 30 commercial cloud providers (suppliers) and publicly funded research organisations (users). Three high-profile flagships sponsored by CERN (high energy physics), EMBL (life sciences) and ESA (earth science) have been deployed and extensively tested across a series of cloud service suppliers. According to their proponents, these commitments behind these initial flagships have created a critical mass that attracts suppliers to the initiative, to work together and make investments. These deployments and tests have revealed a series of gaps in the current set of offerings on the cloud market and the appreciation that the best means of promoting Europe's leadership is to create an open standards based multi-vendor federated market which will allow the diversity of Europe's suppliers to compete with global leaders such as Amazon, Rackspace and Google.

The Science Cloud Strategic Plan was adopted by representatives of all stakeholder groups at a workshop hosted by ESA/ESRIN in June 2011. GÉANT, represented by DANTE, joined the partnership in November 2012. The strategic goals of HN are the following:

1. Goal #1 Establish a cloud computing Infrastructure for the European Research Area serving as a platform for innovation and evolution of the overall infrastructure.
2. Goal #2 Identify and adopt suitable policies for trust, security and privacy on a European-level can be provided by the European cloud computing framework and infrastructure.
3. Goal #3 Create a light-weight governance structure for the future European Scientific Cloud Computing Infrastructure that involves all the stakeholders and can evolve over time as the infrastructure, services and user-base grows.
4. Goal #4 Define a funding scheme involving all the stake-holder groups (service suppliers, users, EC and national funding agencies) into a Public-Private-Partnership model to implement a Cloud Computing Infrastructure that delivers a sustainable and profitable business environment adhering to European-level policies.

Robert Jones, coordinator of the Science Cloud project summarised as follows the main challenges in the final workshop of the study:

- **About the business model**: the PPP model allows strong demand-supply engagement, but there is tension on who should assume the main risks of development costs:
  - Supply-side want to see demand-side purchasing commitment. Currently the larger IaaS companies are paying the SMEs to develop the services that run on their resources

   - Demand-side wants to see proven functionality/quality and competitive prices. The users are prepared to pay for the services consumed but will not pay for the development of general services.
- HN expects to use more than a single business model
- The basic premise of Helix Nebula is to be open by expanding the set of suppliers and users
- **About the provisioning scenario**:
   - The work of Helix Nebula with 30 commercial providers shows that a hybrid (public/private) cloud is attractive to the research communities. "Private" includes not only "in-house" but also GÉANT/EGI/PRACE (EGI is a member of the consortium). It would be important to move beyond the distinction between private and public cloud to focus on a newer distinction between publicly owned cloud and commercially offered cloud.
   - Use big-science to create a critical mass of demand leading the supply-side to put in place sustainable services which can be exploited by the long-tail of science.
- **The role of GÉANT/NRENs**: they should provide networking connectivity to the commercial cloud providers offering services to research and academia. They should also provide/facilitate federated identity management services. DANTE and a number of engaged NRENs in GÉANT have recently joined the Helix Nebula consortium and have agreed to offer network connectivity to Helix Nebula commercial suppliers for research traffic during the two year pilot phase.
- There is a need for federated e-Identity solutions enabled by third parties.

2.4.3. The EGI InSPIRE project

The EGI Strategic Plan – Seeing New Horizons: EGI's Role in 2020 – describes how EGI will evolve into a universal federated platform for supporting compute and data intensive Research and Education communities. EGI will evolve to provide a framework that will host a range of high-throughput solutions, including both grid and cloud approaches. The EGI-InSPIRE project will support 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 should also be well positioned to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit user communities within the European Research Area.

The EGI.eu Federated Cloud Task Force has been mandated to investigate the methodologies, services and standards necessary to support a research community that are uncomfortable relying on a single provider, are of varying scales from the individual user right through to international collaborations. The Task Force mandate lasts eighteen months, from Sept 2011 to March 2013. The Task Force activities are organised in three, six-months long phases. During each phase, the Task Force evaluates a set of scenarios that an EGI federation of clouds should support. The scenarios are chosen by collecting use cases and requirements among user communities, resource providers and technology providers that have already adopted cloud computing or are planning to do so in a near future.

Being eventually integrated with the European Grid Infrastructure (EGI), "the largest, most powerful, and most comprehensive distributed computing e-infrastructure supporting research in the world"[13], the EGI.eu Federated Clouds Task Force will be able to base its work on the extensive experience gathered within EGI.eu and its member NGIs to coordinate many different stakeholders (resource centres offering cloud services, technology providers tendering solutions, and user communities consuming cloud services) at scale.

---

[13] http://www.sienainitiative.eu/Pages/Static.aspx "The SIENA Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe – May 2012"

2.4.4. e-Fiscal: analyzing the costs of e-infrastructures

The **e-Fiscal** project aims at contributing to a more accurate understanding of the costs of HTC and HPC services, both at a national and European level. The project analyses the costs of dedicated computing e-Infrastructures, as expressed by EGI and PRACE initiatives, facilitated by national entities (NGIs and HPC centres); compare these costs with equivalent commercial leased or on-demand offerings and provide an evaluation report.

- Initial findings show that European HPC/HTC e-Infrastructures are cost-effective where they have relatively high utilization rates and depreciation rates. Operational expenditures are dominating the costs (70%), while personnel costs are roughly half of the total costs.
- It is important to note that comparisons are not at all easy and use case views need to complement the overall picture (i.e. specific applications examined in both environments).
    - A key recommendation is to evaluate public cloud challenges in-field during H2020, including central procurement of commercial resources, allocation and monitoring of resources (metering, effective use etc.), legal compliance, interoperability, prevention of vendor lock-in, governance and user satisfaction.
- Expected trend is a combined approach to distributed computing with both grid and cloud, evolving the current EGI environment into a virtualised service-oriented computing e-infrastructure with a public cloud mainly for smaller computing requirements ("the long tail") with a centralised pool of resources procured centrally from commercial public clouds.

According to e-Fiscal research, therefore, the way forward towards the development of cloud services for the scientific and research community in Europe is the development of a centralised pool of cloud resources, which offers the following advantages:

- Better economies of scale/prices at EU level, co-funded by EC;
- Better stimulation of a cloud market for research at the EU level;
- Better stimulation of interoperable, standard and recoverable (no SPoF, single point of failure) solutions from multiple vendors. Such an approach could help Europe advance towards interoperability and ultimately move towards standard cloud stacks and interfaces;
- Better central control of legal, financial and policy issues.

2.5. Case Studies

A key objective of this study was to identify and analyse case studies of implementation of cloud services in the e-Government and e-Science environments, to investigate actual or potential synergies and hybrid models emerging in industry, science and government. However, the majority of cloud initiatives identified are in the initial phase of development, and very few present any kind of synergy between e-Science and e-Government. Nevertheless, we have selected 4 case studies which represent well the main typologies of business and governance models emerging in the deployment of cloud services.

The case studies are the following:

For e-Science:

- **RainCloud** is an Austrian e-Science pilot project, where hybrid cloud computing infrastructures provided by Eucalyptus are used by the University of Innsbruck's Institute of Meteorology and Geophysics for weather forecasting. The results are being used by two spin off services, the daily avalanche bulletin of the Avalanche Service Tyrol and the Tyrolean Hydrographical Service, run by the local government. This is a good example of a range of innovative applications in the field of environmental control and management, based on the collaboration between research and scientific institutes and local governments. The cloud computing model is very well suited to satisfy this emerging demand.

- **Surfconext** is a component of the Dutch National Research and Education Network (NREN), SURFnet. Its main objective is to provide a Collaborative Research Networking environment and access to cloud services to 200 higher education and research institutions with 1 million researchers. This is an example of how the collaborative research environments are evolving, exploiting the cloud delivery model. SURFconext also develops tailor-made solutions for the public sector, such as for university hospitals in the Netherlands and connect all the services to the SURFconext network.

For e-Government:

- **goBerlin**, is a project of the German Trusted Cloud Technology Programme. The main objective of the goBerlin project is to create a platform for the development of innovative applications for citizens, industry and administration with cloud-based services revolving around different use case scenarios. Trustworthiness, security and legal compliance of service providers are deemed particularly important. goBerlin offers IaaS, PaaS and SaaS services.
- The **Cheltenham Borough Council**, a small town in the Cotswold Hills in Britain, has implemented a cloud-based shared document management service provided by Huddle, a private English company, in cooperation with 6 other local administrations of the area. This is a good example of the bottom-up adoption of cloud services by governments most frequently found across Europe.

In addition we present here the key results **15 pilot cases promoted by VENUS-C**, which explored the implementation of cloud computing in various research and scientific environments, following an open call for proposals which had a huge success.

The lessons learnt from these case studies will help to provide insights on the emerging demand of cloud services in the public sector and how it can be satisfied by e-infrastructures in the public sector.

2.5.1. Overview

The case studies presented are very different but they reflect the most frequent typologies of the current use of clouds, respectively in the e-Science and e-Government domains, that is:

- **RainCloud** is one of the multiple research and pilot projects testing clouds to respond to increased demand of high throughput computational power for research;
- **SURFconext** shows the potential application of clouds as complementary services to research networking in the NREN environment;
- **goBerlin** is an example of the type of projects promoted by national governments with the cooperation of regional/local authorities to develop cloud solutions for government and promote the adoption of clouds, solving security and feasibility problems at national level; most of them are in the pilot or early implementation phase.
- The **Cheltenham Borough case** is an example of the actual implementation of cloud services by public administrations, but on a small scale, focused on non-sensitive data, with a bottom-up approach with little scalability potential (even if imitation and duplication is possible). This is the only case of use of the public cloud (the other 3 are all hybrid clouds). The main drivers of adoption were the need to save money and to facilitate collaboration for shared services.

Since these experiences are all in the pilot phase, we were not able to collect proper cost-benefits data (Cheltenham did not want to provide any). In the case of RainCloud, there was an analysis of cost scenarios contrasting the use of traditional in house HPC data centre vs the cloud computing service, but was essentially a simulation. Overall, the cost and flexibility advantages of cloud computing seems to be accepted and taken for granted. The range of cost savings is coherent with the estimates found in literature, and fits within the range of savings documented by IDC's survey of 1056 European organizations[14] using clouds (78% of users saved money, with a majority clustering around a range of 10 to 20% of cost savings and a sizable minority with more than 30% savings).

2.5.2. Type of services and sharing of infrastructures

Comparing our case studies, we notice that three out of four employ a hybrid cloud, while the only one using the public cloud (Cheltenham) confines the application to non-sensitive data. There is a definite trend by government users to prefer private cloud infrastructures to solve data protection and regulation compliance problems. In the science sector the preference for private clouds is based on the need to insure capacity and traditional reliance on internal or trusted datacentres.

Looking at whether the cloud infrastructures and services are shared, we notice that fragmentation, rather than integration is dominant. More specifically:

- SURFConext is a collaborative research networking initiative, so it leverages sharing at all levels (infrastructure, platform and services);
- The other 3 case studies do not employ shared infrastructures or shared platforms, with the exception of the GoBerlin project which plans to let the platform be shared between public administrations (so in a community mode);
- At the service level, only RainCloud does not share, while SURFConext and goBerlin do. In the case of Cheltenham the SaaS are not so much shared, but they are interoperable (maintaining a degree of separation between the information systems of the administrations collaborating in the cloud service).

This analysis in any case confirms that the main trend is towards hybrid models with interoperable rather than shared infrastructures. The pursuit of economies of scale thanks to cloud does not seem at the forefront of the cloud users strategies, if these cases are any guide to the prevalent mind frame in e-Government and e-Science.

Table 4 Comparative analysis of case studies/ Type of services

Case study Name	**RainCloud, Austria**	**SURFconext, Netherlands**	**goBerlin, Germany**	**Cheltenham Borough council, UK**
Main focus	Metereological weather forecasting	Collaborative Research Networking, Providing cloud applications to researchers	Innovative, cloud-based applications for enterprises (SMEs), citizens and public authorities in Berlin	Cloud-based shared document management
Duration	January 2010-December 2013	Ongoing	January 2012-March 2014	ongoing
Stage of development	Pilot Project	Pilot phase started in 2011	Design and development. Pilot implementatio	fully implemented

[14] Final report of the study "Quantitative Estimates of the Demand for Cloud Computing in Europe and the Likely Barriers to Up-take" - SMART 2011/0045, by IDC EMEA, http://ec.europa.eu/information_society/activities/cloudcomputing/cloud_strategy/index_en.htm

			n planned from 2013	
Users	University of Innsbruck and Tyrol administrations, Austria	200 Higher education and research institutions with 1 million researchers, in the Netherlands	Open to citizens, enterprises and public administrations in Berlin, Germany	300 users across 7 local administrations
Type of cloud	Hybrid	Hybrid	Hybrid	Public Cloud
Type of Service	IaaS	SaaS and PaaS	Iaas, PaaS, Saas	SaaS
Service Providers	Amazon & Eucalyptus	PaaS: Greenqloud, Iceland; SaaS: Google, Microsoft, CISCO, SAP, IBM + services shared by partner institutions	Berlin IT Service Centre (coordinator); Atos GmbH(Germany); HSH Kommunal Software	Huddle, private cloud services provider http://www.huddle.com/
Are the infrastructures shared?	NO	YES	NO	No
Is the platform shared?	NO	YES	Only between public administrations	No
Are the cloud services and applications shared?	NO	YES	Yes	Not shared, but interoperable

Source: IDC and Trust IT 2012

2.5.3. Comparative analysis of governance, funding models and potential synergies

Concerning governance and funding models, we find little evidence of potential convergence or synergies between science and government. Three out of four case studies concern projects funded by public research funding and/or led by the main user entity. In none of these three cases there is real concern about business model sustainability, which is certainly premature in the pilot phase, but should at least be taken into account. The RainCloud project is testing so-called "spin-off" services, which may become sustainable if the public administrations using the weather forecasting data pay for the services. SURFConext acts as a cloud broker service and depends on the public funding for NRENs. GoBerlin at the moment is an innovative platform, may become self-sustainable if the targeted users (SMEs) will pay for the services, but this is all to be demonstrated.

In the case of the Cheltenham Borough Council, the business model is typical of the private market with the public organization buying the services of a public cloud provider.

We have analysed the potential synergies between e-Science and e-Government for each of the investigated case studies, as well as their scalability potential (table 2 below). According to the summary table, the RainCloud case is the only case where we see a possible convergence between science and government, with possible synergies in the future between the university weather forecasting services and the government agencies managing environmental protection and safety. This is highlighted in the summary table as medium level operational and technical synergies for services, low level for IaaS and PaaS. However there is low integration between infrastructures (basically the university sends data to the local administrations IT systems).

The SURFConext case will connect with a public social network in the Netherlands (PLEIO) acting as an access point to it. This can be considered a synergy between science and government, or at least a step towards interoperability of different infrastructures. However the two infrastructures are not integrated, nor will they be.

goBerlin is focused on the e-Government cloud application areas and presents no current or potential synergies with the science and research environment, nor any integration with e-Science infrastructures.

The scalability potential is high for the RainCloud service, which could be extended at national level and duplicated at international level. In the case of SURFConext, the service already has a national scope and is building links at the EU level. The goBerlin project should be scaled up in the future at the national level. On the contrary, in the case of the Cheltenham Borough services there is a high potential of diffusion (e.g. other administrations implementing similar services) but not so much scalability, if it is defined as the extension of the same service at the national level.

Table 5 Comparative analysis of case studies/ Synergies and scalability potential

	Are there synergies between e-Science and e-Government?			
Type of synergy	**RainCloud, Austria**	**Surfconext, Netherlands**	**goBerlin, Germany**	**Cheltenham Borough council, UK**
Operational	Low for Iaas and PaaS, Medium for services	High (with PLEIO public social network)	No	No
Technical	Low for Iaas and PaaS, Medium for services	High (with PLEIO public social network)	No	No
Governance/ financial	Low for Iaas and PaaS, Medium for services	Low	No	No
Level of Integration	Low	Non- existent	non-existent	non-existent
Scalability potential	High	National scope, building links at EU level	Planned to be extended across Germany	Medium

Source: IDC and Trust IT 2012

2.5.4. Key Findings

The 4 case studies analysed reflect the most frequent typologies of adoption of cloud services in the e-Government and e-Science environment:

- **RainCloud** is one of the multiple research and pilot projects testing clouds to respond to increased demand of high computational power for research;
- **SURFconext** shows the potential application of clouds as complementary services to research networking in the NREN environment;
- **GoBerlin** is an example of the type of projects promoted by national governments with the cooperation of regional/local authorities to develop cloud solutions for government and promote the adoption of clouds, solving security and feasibility problems at national level; most of them are in the pilot or early implementation phase.
- The **Cheltenham Borough case** is an example of the actual implementation of cloud services by public administrations, but on a small scale, focused on non-sensitive data, with a bottom-up approach with little scalability potential (even if imitation and duplication is possible).

Comparing our case studies, we find the following similarities and differences:

- Three out of four employ a hybrid cloud, while the only one using the public cloud (Cheltenham) confines the application to non-sensitive data. There is a definite trend by government users to prefer private cloud infrastructures to solve data protection and regulation compliance problems. In the science sector the preference for private (or community) clouds is based on the traditional reliance on internal or trusted datacenters.
- The case studies are self-standing initiatives, with no sharing of infrastructures with other initiatives. It is more likely to find interoperable rather than shared infrastructures. Only at the services level (SaaS) there is some sharing.
- Concerning governance and funding models, we find little evidence of potential convergence or synergies between science and government.
- Only in the RainCloud case, the weather forecasting data provided by the University of Innsbruck is of interest to local government agencies managing environmental protection and safety, and this may represent spin-off services.
- SURFConext, which acts as a cloud services broker, plans to provide access to a public social network in the Netherlands (PLEIO). However the two infrastructures are not integrated, nor will they be.
- Three out of four case studies concern projects funded by public research funding and/or led by the main user entity. In none of these three cases there is real concern about business model sustainability (RainCloud, SURFConext and goBerlin). In the case of the Cheltenham Borough Council, the business model is typical of the private market with the public organization buying the services of a private cloud provider.
- The scalability potential has a different meaning for all the case studies. For RainCloud, it means both scaling up the technical infrastructures and service developed extending it to other users in the research and government environments. For SURFConext, which has already a national scope, it means building the OpenConext platform to share services at the EU level. For goBerlin it means to extend/duplicate services at the national level. The Cheltenham case could be seen as a good practice, or a forerunner of an effective way to use clouds for shared document and coordination services between local government administrations.

2.5.5. VENUS-C: pilot case studies with a user-centric approach

The VENUS-C[15] (project co-funded by the EC June 2010-May 2012 with an extra year of free resources and technical support until May 2013 from Microsoft) has demonstrated the value of adopting a **user-centric approach**. The project had the objective of evaluating the benefits of Platform-as-a-Service-style cloud computing for the European e-Science community and its applicability to compute-intensive problems that European researchers are facing.

Thierry Priol, senior scientist at INRIA, declared that: "*Participation to the VENUS-C EIAC brings me new insight on how to use Cloud technologies in several scientific domains. The results that have been achieved during the second year of the project are impressive and they contributed to position Cloud as an effective paradigm to provide computing power, not only to the research community, but also to small companies for which HPC systems are not economically affordable*".

The new software platform for scientific applications is built around two frameworks: "Generic Worker" and COMPS superscalar (COMPSs). The Generic Worker runs exclusively on top of Microsoft's public cloud (Azure) while COMPSs (Barcelona Supercomputing Center), in the context of the project, is used on top of Linux and OpenNebula (Engineering, KTH).

Reference User Communities come from 7 scenarios led by 6 partners; a second set of applications comes from 15 pilots recruited through an Open Call process. A total of €400,000 in seed funds was equally distributed among the 15 successful candidates (sub-contracting)[16].

VENUS-C has demonstrated that public cloud infrastructures (i.e. Windows Azure) are practical for scientific research, and that the use of VENUS-C subsystems improves user experience. Open-source, private infrastructures have also been tested, with similar conclusions regarding user experience. The three usage patterns are: **sporadic peak usage**; **oscillatory demand**; **plateau of resources**.

Benefits demonstrated

VENUS-C end-users have demonstrated benefits of the cloud 1) for applications that have previously run in grids and clusters; 2) new research disciplines moving straight to the cloud; 3) usability by domain specialists and SMEs, not only users with advanced IT skills.

Specific benefits include but are not limited to:

Reduced response time; increased problem size:

- Speed-ups of up to 94x with 100 cores.
- Data increase of up to 25x.
- Users are interested in two different offering models
  - o "free", accessing a reduced pool of local resources.
  - o "subscribers only" accessing a larger amount of resources from public Clouds.

Overall consumption for the 7 scenarios and 15 pilots (May 2012 data) usage was over 1.5 million of CPU hours in total (more than 1.3 Billion SpecInt2k hours in EGI terms), 30 TB of cumulative data stored and 80 TB of data transfers.

Feasible alternatives to grid and clusters:

- Bioinformatics (BLAST) – integration of a legacy application with interoperability across target platforms (Partner – Valencia Technical University).
- Earthquake propagation simulation portal with automatic data capture from seismic registers with near real-time information. For this type of computing, it would be

[15] http://www.venus-c.eu

[16] The analysis of selected use cases is based on partner user scenario deliverables and presentations of results at the end of the two-year project funding on reports after 1-year usage, submitted by pilot representatives as part of the sub-contract stipulations (May 2012)

ineffective to have a cluster of 100 nodes for processing rare events (Pilot – Aristotle University).

- Radiotherapy planning based on Monte-Carlo methods is well-known problem in research & already adapted to grids and clusters. However, research resources cannot be exploited in a hospital environment & usage ratios make the infrastructure costs unaffordable. Usage of VENUS-C has therefore led to the identification of new exploitation models (Pilot - CESGA leading hospital pilot).

Possibility to combine approaches:

- Biodiversity (Biogeography modelling) – integrating e-infrastructures with elastic processing; removing the technical difficulties for end-user community (domain specialists, very large international communities), increasing productivity by transparently wrapping existing software components designed for standalone platforms and running them in the cloud (Partner – CNR).
- Benefitting from the elastic provisioning of the cloud to guarantee throughput for computationally intensive molecular docking in a combined environment with volunteer computing (Pilot – Westminster University & EDGI partner).

New Lines of research:

- Systems Biology – users can define own script to control data flows and focus on core research as domain specialists (Partner – COSBI).
- Drug Discovery – integration of fine-grain workflows for domain specialists. This interactive model is far different from the traditional batch approach used in supercomputing facilities (Partner – Newcastle University, eScience Central).
- Social trends analysis through cloud computing (Pilot – Department of Computer Science, Aristotle University).
- A repository of ICU vital signs for studying early predictors (Pilot – University of Cyprus).

2.5.6. Improved business opportunities for SMEs: Green Prefab

Green Prefab is a new spin-off driven by Collaboratorio thanks to participation in VENUS-C. "Green Prefab is one of the top VENUS-C success stories as it has an entrepreneurial spirit coupled with small-scale funding at regional government level and private investment to help kick-start this new company"[17] **Digital Agenda for Europe**, 29 June 2012.

The main focus of the user scenario for building information management in VENUS-C was to explore a prototype rendering visualization. In summer 2012, Green Prefab started incorporating new tools into the system through a new joint venture.

The user scenario was aimed at developing a 3D rendering visualization service, which many architectural firms outsource (Collaboratorio survey sample of 428 professionals: 60% outsource services for rendering and video) in order to accelerate project completion (55%) and due to lack of skills within the company (34%). This service for immersive rendering provides a complete view of architectural concepts, enriches the computer-aided design project that published online, and helps the jury to define the best project in a call for bids system.

The various steps range from logging in, browsing the product library and selecting the project to downloading 3D files, completing the project and using the export plugin to render. Users then log into the Green Prefab cloud application to start a new job, set render quality, wait for render completion and view the log of progress. The example shown in the demo completes a job in 1 hour and 10 minutes, where 1 instance uses 8 cores[18].

[17] http://goo.gl/scxgF
[18] http://www.venus-c.eu/Content/MediaRoom.aspx?id=505f4aa8-2550-4299-b001-01110ba19396

Benefits identified & future prospects

Increased computation power: Green Prefab with cloud computing ensures intensive data computing and readily available secure information anytime anywhere, thus enabling massive communities to cloud computing resources.

Increased problem size: Higher throughput resources.

Increased cost efficiency: the quantitative validation sample for high-quality images estimated a cost of €0.42 (computing only) for a total time of 14 minutes, 17 seconds. This calculation is based on input of (lxs + lxm + lxo): 10 MB + (zip) 0.3 MB; an output volume (png) of 1.6 MB, a render size of 1024 x 768 pixels and computational size of 8 cores, 4 VMs.

Improved business opportunities: outsourcing the service to a target market of 20,000 professionals and new business model defined. Green Prefab is negotiating seed funding for $1 million (US and Italian VCs) with expected revenue from the Trento Cluster (Progetto Manifattura) aimed at target markets such as high-performance buildings (datacenters, healthcare facilities, emergency shelters, residential) and will continue to seek more foreign funding in 2013.

Business model: Two services are currently available: high-quality images using Windows Azure on a 'pay-per-use' basis (resolution: 800x600 pixels; quality: 300 haltspp) and low quality images that can be used for free until September 2013 running on BSC (resolution: 1,024x768 pixels; quality: 1,500 haltspp).

The new joint venture is initially with the VENUS-C pilot led by the School of Architecture at the Royal Danish Academy of Fine Arts for eco-efficiency analysis (see below: EnergyPlus, Royal Danish Academy), by aggregating users and stakeholder communities through HUB-E[19]. This portal is used to integrate the SaaS solutions. HUB-E development plans (summer 2012-2013) include securing seed funding for approximately $200,000 from VCs in the US to port more software to the portal with a focus on target markets such as energy analysis, immersive rendering and building automation. This joint venture has also opened up new business and job opportunities for E3Lab[20], a consultancy founded in Italy, which now employs a former researcher from the Academy and is currently expanding its business in the UK, including its SaaS for energy efficiency after more testing and usage in late 2012.

2.5.7. Improved business opportunities for SMEs: Molplex Ltd

Founded in 2010, Molplex (UK) is a small privately-held company, backed by a mix of private capital VCs and government grants. The Molplex VENUS-C pilot, CAD – Cloud Against Diseases, is aimed at providing a framework to calculate molecular virtual profiles that include shape/docking characteristics and QSAR biological activity predictions. The shape/docking calculation offers an embarrassingly parallel execution model, and has been parallelised with the use of OpenMP threads.

"The launch of the Molplex 'Clouds Against Disease' pilot study shows how the innovations at the heart of the European e-Science community can be brought to bear on commercial premises, providing quick access to large scale compute resource via a well-defined, standards based interface. This is very much the future of high performance computing for science and business. Capitalising on the research infrastructure investments made to enable such breakthroughs in science and discovery", Ian Osborne, Director Cloud and Government IT, ICT KTN.

The virtual profiles are calculated using two techniques: shape/docking profile and QSAR profile. The deployment of former is supported by the Barcelona Supercomputing Center via

---

[19] www.hub-e.com

[20] http://www.e3lab.org/

the COMPSS interface, while part of the QSAR application is deployed on Azure using a legacy system from Newcastle University.

Molplex requires regular access to computer resources to calculate the virtual profiles of molecules. This becomes a computational-intensive process when the number of molecules to profile grows above 100,000. The involvement in VENUS-C is driven by the need to boost the performance of the company's systems and reduce costs by allocating computing resources more efficiently. Molplex is now able to solve a higher number of scientific problems (virtual profiling). According to Molplex, the CDMI Proxy provided by VENUS-C shows very good performance, meets all related requirements and ensures interoperability through easy operation on Linux.

Estimated usage for the period June 2012-May 2013:

- CPU: 600,000
- Storage: 500 GB
- I/O: 500 GB

Benefits identified

Increased computation power: VENUS-C and the platforms available have given Molplex the opportunity to access computing resources 'on-demand'. This is a feature of our scientific needs as computation requirements come in batches, so a burst of computing activity is needed as generally the data used by our system is released in batches.

Higher throughput Resources: provide the means to increase computation throughput and calculate many more problems than before.

Increased cost-efficiency: the costs per computation problem solved are reduced since services are launched when needed. This is a feature of the scientific needs of this type of companies, since computation requirements come in batches and require a burst of computing activity.

Improved Business Opportunities: higher number of scientific problems solved (virtual profiling) giving better market exposure to the company.

2.5.8. Technical University of Valencia, Spain – bioinformatics

The Technical University of Valencia has implemented a cloud-based service for BLAST in Azure[21]. The aim of this partner user scenario was to validate the VENUS-C components and infrastructure by adapting tools used to map sequences over a reference database[22]. The application targets bioinformatics researchers working in the annotation of a large number of sequences for phylogenetics and characterisation. These researchers need a high-performance tool with minimum client-side requirements and minimum usage complexity.

Benefits identified & future prospects

Reduced response time: enable the execution of BLAST more quickly than using local resources or free-access portals while also allowing more predictable behavior. Sequence mapping for the alignment of Sargassos Sea Metagenome using 800 m bases and 7 m seq. ref. DB had a response time of 7.5 days and a speed-up of 63x, using Azure, compared with an estimated sequential time of 1.3 CPU years.

Increased cost efficiency: The cost of the validation case described above is €547 (537 in CPU and 7 for I/O.

---

[21] The demo is available at http://www.venus-c.eu/Content/MediaRoom.aspx?id=17675f01-391d-4799-b1b8-9eb19986bfd9.

[22] This use case has been submitted to the ETSI Cloud Standards Coordination Group.

Reduce adoption risks: ability to use legacy tools with minimal changes in the interface of existing services; easing the integration in existing scripts already using conventional versions.

Enabling developers: the application leverages the Generic Worker, COMPSs, CDMI and accounting. CDMI and BES enable the quick migration of the intermediate service between different platforms and java client is the same. Reduced effort on application porting (task-oriented model, staging of files, synchronization and scaling provided by the API; quick prototyping). Reduced effort on deployment (pre-configured packages for different virtual HW, compatible with different versions of BLAST and BLAT; pre-existing available end-points eased the deployment and migration of services.

New business opportunities: The tool has been interfaced with the **Blast2Go** [23] application, which provides free users with the access to a limited set of features and resources, and provides subscribers with additional features and dedicated resources, dynamically on the cloud. A Spanish company (BIOBAM) is interested in exploiting the BLAST porting with the aim of increasing the Quality of Service while minimizing investment risks.

2.5.9. A new project to watch: PRISMA

A national initiative of interest is the Italian **PRISMA** project (Platforms Cloud Interoperable for SMArt Government). The project is financed within the context of a Call issued by the Italian Ministry of Education, Universities and Research (MIUR) with a focus on Smart Cities and Communities and Social Innovation, in particular for the area "cloud computing technologies for smart government". It has been chosen by the Italian government as a reference initiative at the IaaS level for all national Smart City initiatives and has signed a Memorandum of Understanding with Smart Health, an initiative funded through the same Call.

The main aim of PRISMA is to develop and make available a set of cloud-enabled applications in cloud datacenters, particularly in the south of Italy. This will include, for example, eHealth applications in the region of Puglia, eGovernment applications in several city councils in Sicily and Campania, as well as earthquake mitigation in Sicily. However, the project's role as a potential model should take into account the tangible benefits it will ultimately demonstrate, the ability to avoid re-inventing the wheel (e.g. at SaaS level, where existing solutions already developed in Europe could be scaled out), the speed at which it can move beyond the prototype phase.

PRISMA brings together 7 enterprises, 2 research organizations, 3 universities, and 1 computing center. It draws on expertise gained through significant investments in grid infrastructures, primarily the EGEE project series, EGI and the activities of its Cloud Federation Task Force, as well as funding at national level through the involvement of the Institute for Nuclear Physics (INFN) in its role as the coordinator of the Italian Grid Infrastructure (IGI).

The main infrastructural goal of PRISMA is to develop and make widely available and easily installable an entirely open source IaaS (Infrastructure as a Service) stack, together with some reference PaaS Cloud platforms, to enable the creation of advanced modern Cloud Data Centers federated through standard interfaces into distributed interoperable and shared cloud infrastructures. These infrastructures will target both various e-Science domains and different sectors within Public Administration with a prototype Application Store, which will be used to search, download and potentially purchase, manage and reuse different types of cloud enabled applications.

INFN is responsible for coordinating the development of the open source IaaS platform, whereas the coordinated development of the PaaS platform is led by an industrial partner. PRISMA will focus on R&D with the goal of providing a **prototype of an open cloud stack**

[23] http://www.blast2go.com

that integrates and aims to improve existing open source solutions in order to achieve the necessary scalability, robustness and flexibility. The development process will aim to foster collaborations with initiatives sharing similar goals at both the European and international level.

PRISMA proposes intermediate steps towards the adoption of cloud by public authorities, whereby current PA datacenters are transformed into IaaS providers for internal customers but federated at national level with the aim of more effectively exploiting all resources and data available. It will only subsequently be expanded to incorporate external commercial offers. The approach is therefore considered to be similar to the development of e-infrastructures for e-Science.

With regard to standards interfaces, implementation will depend on the level of maturity of relevant specifications and their practical level of adoption in production centers, drawing also on efforts to establish federated cloud infrastructures coordinated by EGI.

From an open-source perspective, PRISMA is driving efforts towards ensuring users and developers are able to leverage an open source repository that provides:

- Tools to build federated, shared cloud-based e-infrastructure.
- Tools to support access to large amounts of distributed data, e.g. new natively distributed databases, services with standard interfaces to access and manage distributed storage, etc.
- Tools to establish single sign-on, hierarchical security levels, data privacy with encryption etc.
- Tools to support collaborative efforts.

The vision behind PRISMA has several connections with the SIENA Roadmap and its recommendations. A key conclusion of the Roadmap analysis is that there are important areas in cloud computing where Europe can establish leadership without going into head-to-head competition with leading commercial providers. One could be the development of a large scale private or hybrid cloud, available to researchers and optimized for research. Another is to continue the efforts undertaken by EGI-InSPIRE to help develop and implement standards and software to support clouds by federating multiple smaller cloud providers, with the possible extension to commercial public clouds.

3. ANALISYS OF DEMAND FOR E-SCIENCE CLOUDS

3.1. Introduction

This chapter presents IDC's analysis of the main drivers and characteristics of the potential demand for the cloud(s) needed by the scientific community in Europe. This starts from identifying the main categories of potential users in the scientific community and estimating their number, taking into account the evolution of e-Science. This is based on a suggested taxonomy of the scientific community. The second part of the chapter analyses the main drivers and barriers for e-Science clouds, with a specific focus on cross-border e-infrastructures.

3.2. The main user communities

There are no precisely defined boundaries to the scientific community, particularly if we wish to include also amateurs, non-professional scientists in the vision of science 2.0 developments. However, five variables should be considered when defining the taxonomy for e-Science cloud computing demand:

- The types of scientific research activities that are conducted
- The research domains
- The groups' stakeholders that actively participate in research
- The layers of communities of stakeholders
- The "openness" of research

3.2.1. The types of research activities

According to the statistical classification of economic activities in the European Community (NACE Rev. 2 – division 72), there are three types of research:

- Basic research: experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without particular application or use in view.
- Applied research: original investigation undertaken in order to acquire new knowledge, directed primarily towards a specific practical aim or objective.
- Experimental development: systematic work, drawing on existing knowledge gained from research and/or practical experience, directed to producing new materials, products and devices, to installing new processes, systems and services, and to improving substantially those already produced or installed.

3.2.2. The research domains

According to the statistical classification of economic activities in the European Community (NACE Rev. 2 – division 72), there are two macro fields of research:

- Natural sciences and engineering, which include sub-fields like physics, chemistry, mathematics, astronomy, materials science, and biology, as well as sub-disciplines such as biochemistry and biotechnology and methodologies such as computational fluid dynamics and structural analysis. These are all areas of research that can have highly data-intensive computing requirements.
- Social sciences and humanities, which includes sub-fields like economics, sociology, psychology, archeology, legal studies, linguistics, and literature. These are areas of research with typically less data-intensive computing requirements.

Natural sciences and engineering generally have more data-intensive computing requirements than social sciences and humanities; in fact, according to the Strategic Agenda Report (Deliverable D3) of the study "Development of a Supercomputing Strategy in Europe" by IDC EMEA on behalf of DG Information Society and Media of the European Commission, primary areas of usage of high performance computing include for example: weather forecasting and climate modeling, astrophysical simulation, quantum chemistry, plasma physics, molecular nanotechnology, proteomics and toxicology and national and regional scale economic modeling. As a result, for the future evolution of demand for cloud computing services, specific usage scenarios could be considered.

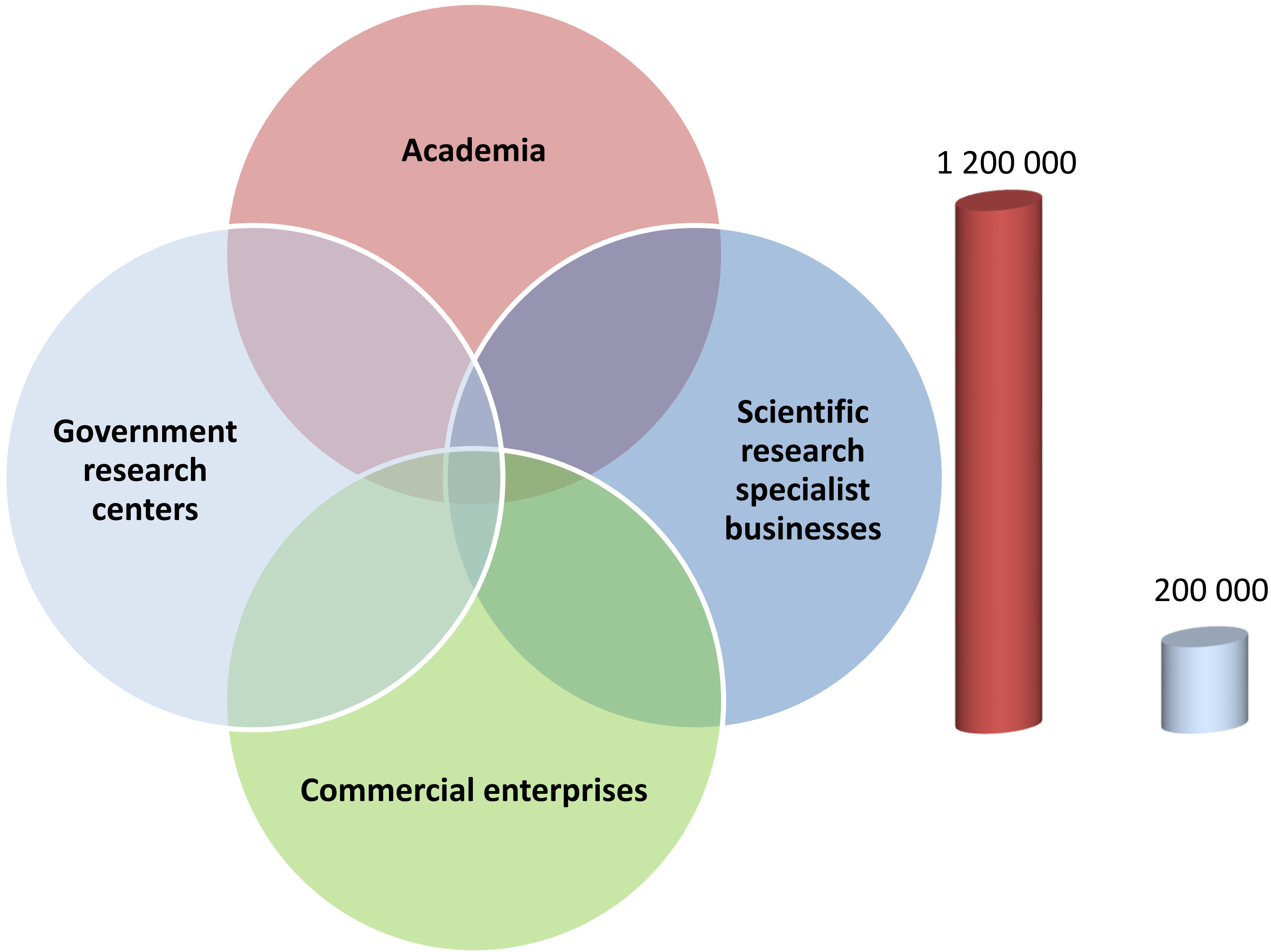
3.2.3. The main Stakeholders involved in research

The stakeholders that actively participate in research activities (see figure 1) can be classified as follows:

- Academic research centers: Eurostat estimates that there were approximately 1.2 million researchers in universities and tertiary education institutions (ISCED97 levels 5 and 6) across the 27 member states of the European Union, in 2009.
- Government research centers: Eurostat estimates that there were approximately 200,000 researchers in government research institutions, such as CNR in Italy and CNRS in France, across the 27 member states of the European Union, in 2009. Government agencies that fund scientific research through grant programs, but do not perform research activities are not included in this count.
- Commercial enterprises that conduct research and development for commercial purposes: Eurostat estimates that there were approximately 600,000 researchers in commercial enterprises across the 27 member states of the European Union in 2010.
- Scientific research businesses that are specialised in the provisioning of scientific research services: Eurostat estimates that there were approximately 90,000 researchers in commercial enterprises across the 27 member states of the European Union in 2009.
- Amateur, non-professional scientists that are particularly active in domains such as astronomy and climate change.

It is important to consider that many research projects are carried out by teams of researchers coming from different groups of stakeholders, so there is limited need to consider separate usage scenario of e-Science computing demand for the four categories, but the population of researchers provides a basic indicator to estimate computing capacity demand.

Figure 1 Stakeholders in scientific research



Source: IDC, 2012

3.2.4. The main communities

Based on this analysis, we can identify three main "layers" or groups of communities involved in scientific research:

- A "core" scientific community, spanning the tens of thousands of scientists active in the European research projects and using European e-Infrastructures. This will involve the scientists using the distributed computing platforms (EGI - European Grid Initiative or e-Science grids), and the HPC initiatives (PRACE, DEISA). This is the community involved with large scale "Big Science" projects, best able to generate a critical mass of demand.
- An extended research and higher education community, including also the university professors and students that use shared infrastructure to collaborate on research (for example the GÉANT2 high-bandwidth, academic Internet serving Europe's research and education community, which serves over 30 million researchers with a multi-domain topology spanning 34 European countries and links to a number of other world regions).
- An open and wider research community, including the non- professional scientists.

This segmentation is useful to quantify the population of researchers, but the boundaries between the different communities are progressively blurring compared to the past and do not correspond any more to sharply different tools and technologies. For example, the "core" scientific community led by the large research institutions such as CERN or ESA used to be the only ones involved in "big science" projects and almost the only ones requiring the use of extremely expensive research instruments and computational resources such as HPC.

This is no more the case: the diffusion of data-intensive computing (Big Data) together with the use of data-generating digital tools in all science domains (including the social sciences) is driving demand for distributed computing, networking and data resources from all researchers. In addition, the demand for HPC-based research has grown beyond science to many industrial/commercial domains, particularly HPC-based modeling and simulation research.

This means that there is a continuum of demand for computing resources across all the user communities: for example core community researchers may be equally likely to be involved in large scale research projects as in small science projects (the long tail of science) with minor computing power requirements, and therefore they may use HPC, or grids, or commercial cloud services, depending on their technical requirements for the type of research they are doing in a specific moment.

3.2.5. Distributed Computing in Europe and Clouds

Significant scientific advances are increasingly achieved through complex sets of computations and data analyses. These computations may comprise thousands of steps, where each step may integrate diverse models and data sources developed by different groups. Scientific workflows have emerged as a paradigm for representing and managing complex distributed scientific computations and therefore accelerate the pace of scientific progress.

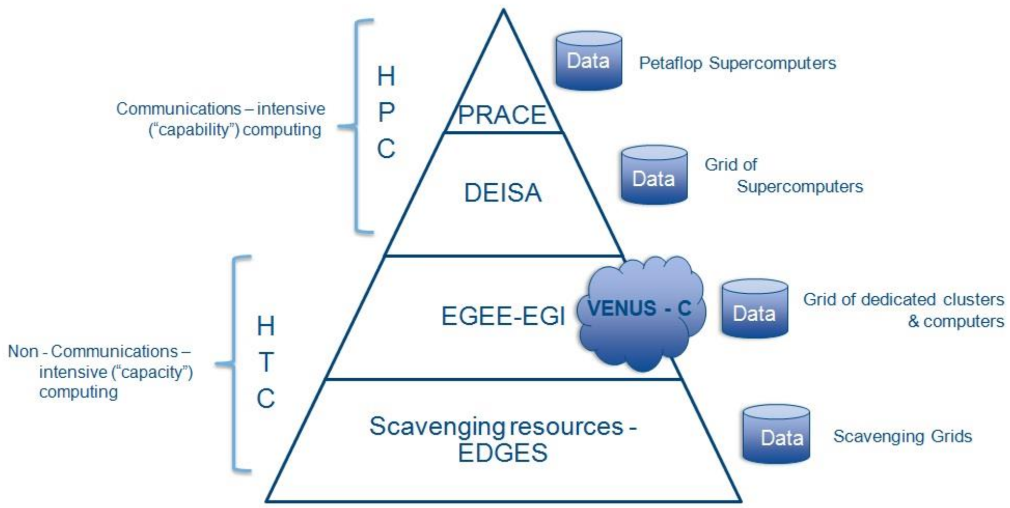
Currently, the offering of distributed computing resources in Europe is based on the e-infrastructures identified in the following Figure 2. This figure is based on the following key definitions:

High Performance Computing (HPC) *is focused on the cluster and supercomputer designed to perform the maximum number of operations per second, and make use of special architectures to achieve this goal. A key characteristic HPC machines share is a low-latency interconnection, such as InfiniBand, which makes it possible to share data very rapidly between large numbers of processors working on the same problem.*

The aim of High Throughput Computing (HTC) *is to increase research productivity of commodity CPUs, memory and networking by maximising throughput over a long period. Access to a large pool of machines, each one modest in its capabilities, can allow researchers to perform tasks such as parameter sweeps much more rapidly than being limited to a single workstation[24].*

Again, the way these differentiations are shown in the figure is broadly accurate, but does not imply that the computing systems are not capable of both types of computing. For example, big, standalone supercomputers such as the PRACE computers are used mostly for capability jobs (challenging problems), while the EGI, etc., are used mostly for capacity jobs. IDC studies for the U.S. Government have shown that even 10-person engineering businesses with the smaller HPC systems typically run a mix of capacity and capability jobs.

Figure 2 Distributed Computing Ecosystem in Europe



Source: VENUS-C 2012[25]

Today's complex, heavily data dependent scientific applications require workflow systems that support dynamic event-driven analyses, handle streaming data, accommodate

---

[24] "Cost analysis of cloud computing for research", 22-02-2012, Final report to ESPRC and JISC, par.2.4.2 page 16

[25] EDGeS = FP7 project connecting desktop grids with EGEE. Scavenging resources: authorized usage of volatile computer resources (CPU (Cycle Stealing), but also storage and communication) by a program NOT administered by the manager of a computer. This definition excludes viruses, worms, ...

interaction with users, provide intelligent assistance and collaborative support for on-going workflow design, and enable result sharing across collaborations.

Initial evaluations of the use of clouds for e-Science have drawn the following conclusions:

- There is a growing set of HPC science and engineering applications that are able to run effectively on a public cloud;
- There are applications that are better suited to a public cloud, principally because of the cloud's on-demand elasticity;
- There are applications where a cloud is a poor replacement for a supercomputer.

Evidence shows that cloud is most suitable for science in the following scenarios:

- Highly parallel applications.
- Integration of distributed sensors.
- Science gateways and portals.
- Workflow federating clouds and classic HPC.
- Batch-oriented commercial and science data analytics that can use MapReduce/Hadoop or its iterative variants.

New disciplines such as Astroinformatics, Matinformatics (real-time chemical analysis), Systems biology, Meta-genomics, Computational history, computational linguistics, etc. are the driving force for research computing. Most of these data intensive applications are loosely coupled and are ideally suited for using clouds.

3.2.6. The level of openness of research

Finally, we need to consider another dimension concerning the openness of research, which influences the ways in which research results are circulated and shared (and from the point of view of e-infrastructures, how data protection should be handled):

- Open research refers to the non-classified and non-proprietary research, which is characterized by transparent peer review collaboration processes and usage of government-funded shared infrastructures.
- Classified and private-sector research refers for example to some military projects and industrial researchers who are not pursuing peer-reviewed, open science for reasons of national security or commercial secrecy.; In some cases, private-sector companies pay for limited access to government-supported data center resources.

For the purposes of this study, IDC defines advanced research as including the following:

- Open (non-proprietary) scientific research of sufficient size and complexity to require the use of high performance computing resources. (Note: Most advanced researchers access HPC systems via laptop or desktop computers.) IDC concentrated on the "core" scientific community of researchers from multiple natural science domains, such as bio-life sciences chemistry, physics, weather/climate, combined disciplines such as biochemistry, and biophysics, as well as regional and national centers/ institutes that conduct economics research.
- Pre-competitive industrial research of sufficient size and complexity to require the use of high performance computing resources. At least some portions of the findings are intended to be published or otherwise openly shared. Even the largest, tier 1 industrial firms typically cannot justify purchasing large HPC systems and therefore seek access to large systems at national HPC facilities for conducting the most advanced pre-competitive (and competitive) research.

This study does not include industrial researchers who are not pursuing peer-reviewed, open science because historically governments have not seen it as their job to support industrial research unless it is pre-competitive. The study also excludes classified government usage because its volume is not publicly known and because IDC believes little if any of it will be run on clouds designed for open science.

3.3. Cloud Computing and Science 2.0

Beyond the traditional classification of open and classified science, there are further trends driving towards greater openness and transparency of the research activities, which create new demand for communication, interaction and sharing between the research community and the wider world of amateur scientists or simply the interested public. In other words, the boundaries of the scientific community are blurring. This has implications for the potential demand of e-infrastructures and the role of clouds.

From a more general perspective, the potential benefits of the use of clouds in the scientific environment should not be seen as a black-and-white, yes or no situation. Rather, cloud computing fits into a continuum spectrum of collaborative computing services for research and science. At one end of the spectrum there is HPC, with large academic and other public science institutions that historically have owned their computing resources and concentrated on large-scale research projects devouring computing power. A less-demanding subset of this category is represented by the majority of university and research centres, which often struggle to have access to sufficient shared computing infrastructures. At the other end of the spectrum, there is a huge number of individual scientists and researchers, extending to amateurs, who posses only individual computing resources. But this scenario, where the main demand segments used to be quite separated, is changing: there is an increasing demand for computing power by all researchers, driven by the advent of big data science as well as the changing process of scientific research which becomes more and more open. Thanks to new collaboration and communication tools, science becomes more than ever before a shared endeavor. This is what is starting to be called science 2.0.

There is also the possibility to tap into consumer resources (tablets, laptops, etc.) to use their virtualised resources to conduct analysis with less stringent requirements in terms of performance and security. In this context, cloud computing may satisfy new emerging needs, bridging between the traditional e-Science infrastructures and the new ones.

The large-scale remote and open collaboration of scientists, citizen and private-sector R&D researchers combined with making openly available not only final research results but also underlying data, software, bibliographies and annotations as well as the growing amount of data availability results in a demand for new virtual collaboration and storage spaces.

Increasing openness of scientific process as shown in the table below, augments the need for connectivity. The collaboration is not happening only between research labs but on an individual level and the granularity of the sharing process is decreasing from a fully-fledged research article to a single data set or a blog post with an initial research idea. This drives also the demand for new collaboration software that also works in the cloud.

Mendeley combines the strengths of a social networking tool with the functionality of an open reference manager. Figshare give researchers an easy tool to share and store their research data by an efficient cloud-based service. Altmetric, the winner of Elsevier's AppsforScience contest, traces the social media visibility of a given scientific article as well as its popularity in online reference managers and from mainstream news sources.

Table 6 Increasing Openness of the Scientific Research Process

Output Time	Bibliography	Data	First Analysis, working notes	Draft paper	Article	Comment on other people's work
Traditionally	Not public	Not Public	Not Public	Not public	Public	Internal, public only through articles
Emerging trend	Public	Public	Public	Public	Public	Public by all means and at all stages of work

Source: Burgelman, Jean-Claude, David Osimo, and Marc. Bogdanowicz. 2010. Science 2.0 (change will happen…), First Monday 15, no.7

As mentioned, the collaboration is not reduced to academia; non-professional scientists play an increasing role not only in providing data with the use of sensor, but also in analysing data. Zoouniverse, a collection of high-performance web applications working in a cloud environment supports several citizen science projects. Currently twelve projects ranging from astronomy (Galaxy Zoo continuations) by meteorology (Old Weather) to archaeology (Ancient Lives) and medicine (Cell Slider) are active with more than 700, 000 people participating in this online endeavor. The earliest prominent example of citizen science was the SETI Project (Search for Extraterrestrial Intelligence) sharing computing power and attracted 5MLN users.

Finally, science becomes more and more data-intensive. Not only by sharing and opening up the previously locked resources but also thanks to new technological tools for data collection (such as Internet logs, sensor data, mobile calls records or health data). Those large-scale datasets cannot be stored, captured, managed and analysed by the mean of conventional database software.

This move towards science 2.0 creates new opportunities for cloud computing infrastructures able to link together the "hardcore" research centres, individual scientists, new software providers and citizens. It can provide an open, scalable and flexible environment for mass and large-scale collaboration, sharing of research processes and results. In addition, this trend has its equivalent in open government policies and may become one of the key synergies for clouds for science and government.

3.4. Drivers and Barriers of Cloud Deployment in e-Science

Despite much discussion on cloud computing limitations, demand for cloud services in e-Science is growing rapidly, responding to the evolution of the scientific and research environment. The main drivers of demand are summarised in the following table 7. Cloud services provide access to larger (capacity) and different types of computational resources (capability) than are available on premise, enabling long-distance collaboration and information-sharing, and offering a separate environment for developmental projects and beta testing. A relevant driver is also the policy push from national strategies and availability of funding, promoting cloud for reasons of efficiency and costs savings. On the other hand, the list of barriers is long and consistent[26]. The highest barriers relate with the need to evolve organizational and practical procedures within user organizations. The insufficient knowledge of how best to exploit clouds, and of the specifics of the cost-benefits balance within the research and scientific environments, underlines the need for further awareness raising efforts and for the development of business cases and good practices.

[26] A more detailed analysis with a comparison of the drivers and barriers of cloud deployment in e-Government is presented in the Interim report deliverable of this study.

The evolution of standards and the improvement of contractual conditions are also necessary to enable wider adoption.

From the point of view of e-infrastructures suppliers, the policy drivers and the need to respond to increasing demand for computing resources, with new characteristics, without additional capital investments, are also relevant. In addition, the commoditization of virtualized technologies is pushing networking providers for science and research such as NRENs to diversify towards service provision such as cloud services.

Table 7 Main Drivers and Barriers of Cloud adoption in e-Science

Type	Drivers	Barriers
Policy/ Regulatory	National Policies and funding supporting cloud adoption	Immaturity of procurement policy unsuited to move from CAPex to OPex models, to select and manage commercial providers Unclear contractual conditions, particularly for liability and accountability of service providers. Lack of universally accepted SLA Lack of national or ministerial commitment for the actual adoption of cloud services Compliance with national data protection and privacy regulation, uncertainty about data location
Economic	Satisfying computing demand for sporadic peak usage, oscillatory demand, plateau of resources without additional capital investments Cost savings	Inertia of traditional business and funding models Uncertainty about cost-benefit and sustainability assessments (also hidden costs of traditional infrastructures)
Social/ Organizational	Increased availability in multiple location (e.g. out of home laboratories) Agility and ease of access avoiding long peer-evaluation screening Reduced need of skilled IT specialists, possibility for scientists to focus on core competencies	Uncertainty about organizational models of cloud deployment (supply side) Cultural resistance to change (shown by JISC studies in the UK for example) by scientists and IT managers (demand side) Lack of experience and training in cloud services management and deployment
Technology	Appropriate resources for middle to lower intensity computational scientific tasks Easy scalability Better and cheaper data management, backup and disaster recovery particularly for Big Data applications	Scope of applicability: highly parallel applications, embarrassingly parallel applications, integration of distributed sensors, science gateways and portals, batch-oriented commercial and science data analytics that can use MapReduce/ Hadoop or its iterative variants Risks of vendors lock-in due to proprietary standards

		Lack of interoperability between clouds, proliferation of standards Uncertainty about data safety, security and access control

Source: IDC, Trust IT: Clouds for science and public authorities 2012

3.5. Drivers and Barriers of pan-European cloud e-infrastructures in Europe

The drivers and barriers identified above are relevant also for the demand of a pan-European cloud e-infrastructure in Europe, but with some variations and additional issues due to the international dimension and the specific value added of a EU-level system (see Tables 8 and 9).

A key driver of demand is the EU policy strategy to implement the ERA and to insure the "fifth freedom" of knowledge and data mobility in the EU single market, responding to emerging demand for open, flexible and scalable computing capacity that national resources cannot satisfy. This calls for a pan-European infrastructure, able to avoid any potential "digital divide" between the large/small countries and between Big Science/Small and Mediums Science (SMS) projects insuring equal access to computing resources across the EU. For example, the current PRACE infrastructure includes HPC centers located in a small number of EU MS. While in principle these centers are open to all researchers (based on a peer-review access mechanism), according to IDC's study[27], their location in a different country does create practical and organizational constraints for researchers. The availability of cloud services by a pan-European e-infrastructure could go a long way to solve these problems.

There is also a search for economies of scale and scope in the development of e-infrastructures, due to the need to rationalize public spending and develop more sustainable business models, based on service fees as well as public funding. Again, some of these economies may be reached also by a few NRENs operating together (as for example SURF-Co next is aiming for in partnership with the UK). But economies of scale and scope will be more relevant if the whole of the 27EU are involved in a cooperative way.

The need to face global competition and the broadening of user constituencies beyond traditional research and higher education institutions and national boundaries is naturally pushing e-infrastructure providers towards international development strategies. But there is also a strong user pull. Big science users want more freedom of choice of services and providers, unlimited by national boundaries; the extended scientific and research community demands immediate, flexible and scalable computing resources at affordable prices, avoiding when not necessary the long and complex process of peer review; the science 2.0 community is naturally global and does not even take into consideration national boundaries.

Finally, demand for better interoperability and standards is naturally expressed at EU level and will require a EU-wide effort to insure equal access to all interested researchers.

There are however formidable barriers: from the immaturity of procurement policies, to complex cross-border contractual issues, to the need to develop new organizational and business models accommodating common services and cross-border operations. This is compounded by the cultural resistance by scientists and IT managers, the lack of experience in managing cross-border services rather than networks, and the need to develop new skills sets. It is clear that no single business model or monolithic approach can satisfy this complex set of issues: all the stakeholders agree that hybrid provisioning

---

[27] IDC, "A strategic Agenda for European Leadership in Supercomputing: HPC 2020" 2011, on behalf of the European Commission

scenarios, a plurality of business models and a move towards federating resources are the best way to go.

Table 8 Main Drivers of pan-European cloud e-infrastructures for science and research

<table>
<tr><th>Type</th><th>Supply-side</th><th>Demand-side</th></tr>
<tr><td rowspan="2">Policy / Regulatory</td><td colspan="2">EC and national policy strategies for EU e-infrastructures and for the ERA, to enable excellent science and the "fifth freedom" of knowledge and data mobility in the EU

Need to overcome the "digital divide" in the EU between the large/small countries and between Big Science/Small and Mediums Science (SMS) projects insuring equal access to computing resources across the EU

Compliance with EU wide regulation on data protection, sensitive data location, data privacy and copyright protection through EU wide e-infrastructures and services</td></tr>
<tr><td>Need to face global competition outside the EU and insure sustainability of e-infrastructures</td><td>Potential economies of scale and scope enabling fragmented research communities to collaborate and achieve critical mass (e.g. parallel software development)</td></tr>
<tr><td>Economic</td><td>Search for new sustainable business models, for economies of scale and scope beyond national boundaries, cost efficiencies

Competition by global public cloud providers (Amazon, Microsoft)</td><td>New demand for open, flexible and scalable computing services that national resources cannot satisfy (including Big Data)</td></tr>
<tr><td>Social/ Organizational</td><td>Considerable expertise in building cross-border e-infrastructures and federating computing resources, as shown by Géant, EGI, PRACE</td><td>Broadening of user constituencies beyond national borders and traditional HEI (science 2.0)

Scientists community increasing demand for cross-border / EU wide scientific collaboration</td></tr>
<tr><td>Technology</td><td>The commoditization of virtualized technologies, pushing networking providers for science and research such as NREN to diversify towards service provision such as cloud services</td><td>Demand for better interoperability and open standards at EU level</td></tr>
</table>

Source: IDC, Trust IT: Clouds for science and public authorities 2012

Table 9 Main Barriers to pan-European/ cross-border cloud e-infrastructures for science and research

Type	Supply-side	Demand-side
Policy / Regulatory	Persisting national differences in the applicable laws, data privacy and copyright protection, data location, particularly for sensitive data, which may affect the reliability and compliance of cross-border cloud services; Complex cross-border contractual issues about cloud providers accountability, liability, fair pricing, even tax payment	Immaturity of procurement policy unsuited to move from CAPex to OPex models, to select and manage commercial providers
Economic	Need to balance the sophisticated, secure environment of NREN and EGI with the flexibility, openness and cost-effectiveness of commercial service providers	Uncertain business case: there is still insufficient information about the cost-benefits and sustainability of cloud services provision and adoption
Social/ Organizational	Need to develop new business and organizational models for cross-border service provision; Lack of experience and training in cross-border cloud services management and deployment Practical language barriers	Cultural resistance to change (shown by JISC studies in the UK for example) by scientists, IT managers, traditional e-infrastructure managers Practical language barriers
Technology	Need to develop common services, such as AAI services (Authentication, Authorisation and Identification); Need to solve cross-border data and application interoperability and portability issues Proliferation of standards	Need to solve cross-border data and application interoperability and portability issues Fear of lock-in by service providers

Source: IDC, Trust IT: Clouds for science and public authorities 2012

3.6. Potential synergies between e-Science and e-Government clouds

A key aspect of the cloud strategy for the public sector concerns the potential synergies between e-Science and e-Government cloud infrastructures at the EU level. This is not a simple issue. The analysis of current developments of cloud in Europe has not found striking examples of shared infrastructures or co-tenancy of science and government workloads.

According to our research, the reasons for this situation are the following.

Different user requirements. The user constituencies for e-Science and e-Government are historically different and present different demand dynamics. E-Government workloads are typically akin to business operations in the private sector. These workloads consist mainly of record keeping, payments, budget tracking, document storage, and other transactions needed to support the day-to-day operations of governments. Most of these transactions can be computed in a fraction of one second, making most government workloads "embarrassingly parallel" and able to run on computers and networks with no special communications capabilities. Contrast this with advanced scientific and industrial research problems, which typically take hours, days, or weeks to run, and which nearly always benefit from special, high-capability computers and networks.

Historically, business operations in the government and private sectors have been subject to a provisioning mentality, whereby each end-user receives a desktop or laptop computer, each group of 10 users might receive a more powerful desk-side computer, each group of 100 users might receive an even more powerful departmental computer, and so on. The provisioning mentality assumes that the desktop/laptop computer is adequate to support most of the work of most end users.

In sharp contrast to this, advanced scientific and industrial research typically is subject to an enabling mentality. This mentality assumes that if end users are given access to more powerful computational resources, this will enable them to achieve greater advances.

Put another way, the provisioning mentality assumes that the number of transactions to be computed may be very large, but it is finite. The capability mentality, on the other hand, assumes that advanced research problems have an insatiable appetite for computing power and can always benefit from more capability. Research problems can almost always be run at larger scale, at higher resolution, with more elements included, or covering more possible scenarios. Equipping a single e-infrastructure to support both more-demanding e-Science and less-demanding e-Government workloads would mean over-building it for the e-Government requirements.

For these reasons, government and private sector organizations that conduct both business operations and advanced research almost never do so using the same computers and networks. IT and HPC data centers are sometimes located next to each other to take common advantage of power and cooling facilities, but seldom if ever is there co-tenancy on the same computing, networking, and storage infrastructures. The HPC data center needs to be administered with an enabling mentality that is fundamentally alien to the provisioning mentality that has been successful on the business operations side. This important difference extends beyond the data centers to distributed computing infrastructures such as grids, private clouds, and public clouds.

The Missions Are Different. There is an even more fundamental difference between e-Science and e-Government than the just-described technical requirements they impose on computing resources and e-infrastructures. While government data and documents typically must remain confidential, from personal tax records to national security information, science is an essentially open, non-proprietary, collaborative activity. Hence, the security needs, policies, and directives applicable to e-Government are very different from those needed to accelerate progress in advanced scientific and industrial research.

The Governance is Siloed. The separation of missions between e-Government and e-Science is also reflected in separate budget appropriation and governance structures. Typically, government departments are accountable for the operation of a certain program, such as revenue collection, or road maintenance, or welfare payments, for which they budget capital and operating expenditure on a annual basis. Science is generally financed through grant allocations (sourced from public or private funds) that are used to cover the entire life-cycle of a research project as if they were capital expenses related to that project.

3.6.1. Perspectives of change

Even within the current scenario, there are drivers of change which may create areas of convergence between e-Science and e-Government demand in the near and medium future. They are the following.

- **The diffusion of "Big Data" e-Government applications**, that is, the growing number of jobs that require complex mathematical models and algorithms to be applied to very large data sets, and therefore may drive demand of computing resources not simply transaction oriented but characterized by high capability as those designed for e-Science. They include national security (for example the identification of potential terrorists in near-real time); fraud and error detection in national and regional health care and pension systems (already, Italy and several other nations have acquired supercomputers for this purpose).
- **Computational applications and services for public health and environmental sciences**. In these fields, health, weather, climate and geological analyses are carried out for scientific purposes by academia, and for risk management purposes by government agencies, such as civilian and military meteorological offices, geological institutes, tourism offices, infrastructure maintenance agencies, and first responders. There is a clear potential for collaboration between research and government actors in the development and exploitation of these applications, where cloud infrastructures and services are proving to be particularly well-suited. This is shown for example by the Raincloud and WildForest case studies presented in the previous chapter.
- **Search for more sustainable business models in e-Science**. While funding patterns are presently very different between e-Government and e-Science, public authorities have a clear incentive to "open up" the user community for e-Science infrastructures to enable sustainable business models offering pay-as-you go services also to emerging demand from government users.

In summary, the development of cloud e-infrastructures covering all the needs of both e-Government and e-Science users seems a very unlikely scenario (it would fall under the pitfalls of co-tenancy discussed above). On the other hand, there is potential for an open, interoperable cloud e-infrastructure designed for e-Science to be able to meet also the needs of a variety of innovative and advanced e-Government applications, based on the collaboration between research and government actors. Or maybe, more simply, some applications developed by the e-Science community could provide analysis and reports that are valuable for government decision makers, thus could be provided to them on a usage basis. This business model would also offer the scientific community an additional option for financial sustainability of their technology investments. However, for this model to be successful, scientists will have to acquire a service management culture and organizational capabilities to market, deploy, operate, bill and support services for government "customers".

4. FORECASTING DEMAND FOR THE EU SCIENCE CLOUD

4.1. The European Scientific Cloud Forecast Methodology

The following paragraphs present the IDC estimate of potential demand of e-Science clouds up to 2016, for government-supported community and public clouds e-Science infrastructures, such as Helix-Nebula. They exclude private cloud computing (resources within an organization's security perimeter).

4.1.1. Taxonomy and scope

For the purposes of this study, IDC defines advanced research as including the following:

- Open (non-proprietary) scientific research of sufficient size and complexity to require the use of high performance computing resources. IDC concentrated on the "core" scientific community of researchers from multiple natural science domains, such as bio-life sciences chemistry, physics, weather/climate, combined disciplines such as biochemistry, and biophysics, as well as regional and national centers/ institutes that conduct economics research.
- Pre-competitive industrial research of sufficient size and complexity to require the use of high performance computing resources. At least some portions of the findings are intended to be published or otherwise openly shared.
- This study does not include industrial researchers who are not pursuing peer-reviewed, open science because historically governments have not seen it as their job to support industrial research unless it is pre-competitive. The study also exclude classified government usage because its volume is not publicly known and because IDC believes little if any of it will be run on clouds designed for open science.

4.1.2. Macro-Assumptions

The following assumptions are at the basis of the forecast:

Total Number of Cores

- The total number of cores installed in technical computing systems worldwide in 2011 was 51 million. Based on Europe's 31% share of the global HPC market in 2011 IDC estimates that the total number of cores installed in the EU in 2011 was about 16 million. This number is based on IDC's in-depth, quarterly, worldwide tracking of HPC systems sold, installed, and accepted by customers during the quarter. IDC collects this supply-side information from HPC vendors around the world and further validates it against the vendors' and customers' public announcements and other sources. To obtain the larger figures (e.g., system, processor and core counts) for the worldwide installed base of HPC systems IDC sums up quarterly data for the most recent 4.5 years, which is the average installed lifetime of HPC systems. Every 18-24 months, IDC conducts an in-depth, worldwide survey of HPC user sites to collect complementary demand-side information, which can lead to adjusting the average installed lifetime of HPC systems.

Definition of Technical Computing, also Called High Performance Computing by IDC

- IDC uses the terms technical computing and high-performance computing (HPC) synonymously to encompass the entire market for computer servers used by scientists, engineers, analysts, and other groups employing computationally intensive modeling and simulation methods. Technical servers range from small systems costing less than

$5,000 to extreme-capability machines valued at hundreds of millions of dollars each. In addition to scientific and engineering applications, technical computing includes related markets/applications areas, including economic analysis, financial analysis, animation, server-based gaming, digital content creation and management, business intelligence modeling, and homeland security database applications. These areas are included in the technical computing market based on a combination of historical developments, applications types, computational intensity, and associations with traditional technical markets. The specific categories of applications are included in annex.

The Impact of Storage on Science Cloud Demand.

- Storage is not included in the present forecast. Storage is the fastest-growing segment in IDC's five-year HPC market forecast. IDC predicts that the storage segment will grow at a robust 8.9% CAGR, from $3.7 billion in 2011 to $5.6 billion in 2016. That amounts to a 51% revenue jump in five years. Storage demand ("attach rate") differs by scientific and engineering domain, and also by application without a domain. Overall, however, IDC believes that storage attach rates (the ratio of server to storage spending) are similar for advanced scientific and advanced engineering research. IDC also assumes, based on data available to data, that storage attach rates are similar for scientific work conducted on premise and in cloud environments. IDC did not produce a detailed scientific cloud storage forecast for the present study, but it might be useful to do so in the future.

Substitution Rates

- In Tables 12 and 13, figures in the rows related to forecasts are incremental to 2011, which is set at zero (0). The HPC server tracking takes into account additions, substitutions, upgrades and decommissioning of systems over time. As explained earlier, to calculate the number of HPC systems (and components such as processors, cores) at any point in time, we sum up our quarterly QView figures for the most recent 4.5 years (18 quarters), because the average installed lifetime of an HPC system has remained at 4.5 years worldwide.

Europe's share in the cloud for science market

- Europe represented 31% of worldwide spending for high performance computing (HPC) in the most recent historical year, 2011. IDC estimates that Europe represented 28% of worldwide spending for cloud-based HPC science in 2011.

4.1.3. Growth-Assumptions

- Forecasts are incremental to 2011, which is set at zero (0). Hence, the figures in these rows represent forecasted growth in demand over 2011, rather than total demand including 2011. IDC has used this method because 2011 demand for e-Science clouds is not well understood and future demand can be more reliably forecasted. IDC believes, and non-IDC studies appear to confirm, that 2011 demand for e-Science cloud use was relatively small because of limitations described in these studies, especially data security, data transfer speeds and costs).
- Some cloud-friendly scientific applications (e.g., QCD in astrophysics) could exploit almost unlimited numbers of cores and core-hours. IDC assumes that the nearly insatiable appetites of such applications will continue to be constrained by the practice of awarding large, but limited allocations of available computational resources. Today and for the near-term future, clouds will remain best suited for open science jobs and workloads that are loosely coupled ("embarrassingly parallel") and do not require substantial interprocessor communication while the jobs are running. Clouds infrastructures enhanced to run more tightly coupled jobs effectively, such as with Infiniband or other high-capability networks, would have a higher breadth of applicability for scientific jobs.

- Most science done on clouds will continue to be open science that is awarded unpaid time allocations based on peer review. A far smaller portion of utilization will be advanced industrial scientific research allocated on a paid basis without the necessity for peer review.
- Figures include both unpaid and paid use of science clouds; therefore, CPU use (unpaid + paid) will be higher than revenue numbers (paid only).
- IDC expects most science cloud usage (and revenue) during the forecast period to be in addition to, rather than in replacement of, on premise demand for HPC in science.
- Government investment will make e-Science clouds capable of addressing more types of science problems and will result in higher usage (Tables 9 and 10) than without government investment.
- There is a limit to how quickly government investment can accelerate efforts to advance the capabilities of clouds for supporting e-Science; therefore, there is a limit to how much government investment would be useful during the forecast period. Unlimited investment will not produce unlimited advances.

4.1.4. Demand of Big Data in e-Science

IDC's European Science Cloud Forecast includes our forecast for data-intensive scientific computing ("Big Data") performed in Europe-sponsored cloud environments. IDC categorises server systems according to their primary use. We categorize an HPC/HTC server system as used for "Big Data" if more than 50% of its cycles are intended to be used for this purpose (many HPC/HTC server systems are used for multiple purposes and multiple associated workloads). IDC defines the "supercomputer" category as including HPC systems sold for $500,000 (€375,000) or more each.

As shown in the fig IDC forecasts that this market will grow rapidly from a relatively small starting point to approach $1 billion (750M euro) by 2015. Of this total, IDC estimates that 29% (about 217M euro) will occur in Europe, and approximately 62% of this amount (135M euro) will be devoted to science rather than industry/commerce. Hence, although Big Data in science is projected to grow rapidly, its small starting point today means that by 2016 it will remain a single-digit portion of the overall HPC/HTC server market.

The constraints that make high-throughput (highly parallel) science workloads more amenable to public cloud computing than communications-intensive workloads apply to Big Data as well. That is to say, high-throughput (highly parallel) Big Data science jobs are a better fit for contemporary cloud architectures than is the case for communications-intensive Big Data science jobs. As a corollary premise, improvements to cloud architectures that would make them more suitable for communications-intensive science jobs would also make them more suitable for communications-intensive Big Data science jobs.

Table 10 IDC Worldwide Data Intensive (Big Data) Focused HPC Server Revenues ($ Millions)

	2009	2010	2011	2012	2013	2014	2015	CAGR '10-'15
WW HPC Server Sales	8,637	9,504	10,034	10,564	11,397	12,371	13,485	7.20%
Big Data Worloads	535	603	655	708	786	881	989	10.40%
Big Data in HPC Portion	6.2%	6.3%	6.5%	6.7%	6.9%	7.1%	7.3%	3.0%

Source: IDC HPC 2012

4.1.5. Forecast

The following tables show the forecast scenario of the demand for government-supported community and public clouds e-Science infrastructures, such as Helix-Nebula, from 2011 to 2016.

Table 11 presents an estimate of the potential demand, measured as the number of computing cores used in open science, as defined above, regardless if it is satisfied by cloud or non-cloud services. The first row forecasts the demand for open science by the core scientific community assuming the persistence of existing mechanisms that allow to access high-performance computing resources, for example through peer review processes of specific requests. The second row, estimates the pent-up (otherwise latent) demand, which is represented by all of those requests to use high-performance computing that nowadays are turned down, because of limited resources. As it is demonstrated by the data, the amount of pent-up demand is approximately four times the current level of demand. This is because:

- According to IDC high-performance computing research, approximately three-fourth of requests are turned down;
- There are some science domains, such as astro-physics, that could express unlimited demand that is, however, non technically and financially sustainable;
- The forecast assumptions include the estimates of growth of Big Data applications, as one of the factors driving the demand for e-Science, as explained above.

The estimate of pent-up demand is very important (as it will be shown in the following forecasts) because it shows that improvements in capacity and capability of e-infrastructures are likely to generate a strong response by the community of researchers.

Table 11 Forecasts of Europe computing demand from the "core" scientific community

	2011	2012	2013	2014	2015	2016
Number of cores used in open science*	20,784,682	27,158,207	35,486,143	46,367,802	60,586,271	79,164,766
Number of cores needed for additional (pent-up) open science demand	83,138,727	108,632,827	141,944,572	185,471,208	242,345,083	316,659,063

* excludes unaddressed pent-up demand

Source: IDC HPC 2012

The following tables estimate the level of forecast demand of cloud services, based on the portion of computing cores, which will be used for cloud services up to 2016, based on two scenarios:

- The first scenario estimates the growth of existing demand, extrapolating current trends;
- The second scenario estimates the potential pent-up, unexpressed demand and the rate at which it could be satisfied.

The other main factor considered is the level of capacity and capability of e-infrastructures. From this point of view, we consider two alternative trajectories of demand forecasts:

- In the first case the capacity and capability of European e-Infrastructures grows extrapolating present trends;
- In the second case, there is additional government investment, which increases both the capacity and the capabilities of existing e-infrastructures. In this second case, the satisfaction of pent-up demand is higher because a higher portion of computing cores can migrate to the cloud.

This second demand forecast recognizes that improvements in capacity and capability, thanks to higher investments in e-infrastructures, will address pent-up demand more fully than capacity improvements alone. Capability improvements will make it feasible to run certain types of scientific jobs on the cloud that could not be handled well by today's cloud architectures, even with greater capacities.

The first scenario forecast is presented in Table 12 as follows:

- The first row estimates that approximately 2.5 million cores in use in 2011 are for use-cases that could run in the cloud, given today's cloud technology attributes (e.g. performance, interoperability, security). The first row then forecasts this demand out to the year 2016.
- The second row estimates the percentage of those use case that are actually expected to migrate to the cloud, under the scenario of investments currently planned for GÉANT, Helix Nebula, EGI-inSPIRE and so forth.
- The third row calculates the actual number of cores corresponding to the percentage of use cases expected to migrate to the cloud, under the scenario described above;
- The fourth row, assumes that additional government investments are made, , in Géant, Helix-Nebula, EGI-inSPIRE and similar programs above the levels currently planned, thus enabling government supported science clouds to expand capacity and improve technical attributes so that cloud services can accommodate more use cases. The following rows present the percentage of cores likely to be used for the cloud in this second scenario, and their total number. In fact, assuming investments happen between now and 2014, 2015 and 2016 are expected to show differences between the "current funding" scenario and the "additional funding" scenario, as the percentage of demand that is expected to migrate to the cloud is respectively 40% and 45%, in 2015, and 45% and 55%, in 2016.

Table 12 Forecast of Europe computing demand from the "core" scientific community: cloud friendly portion of existing demand

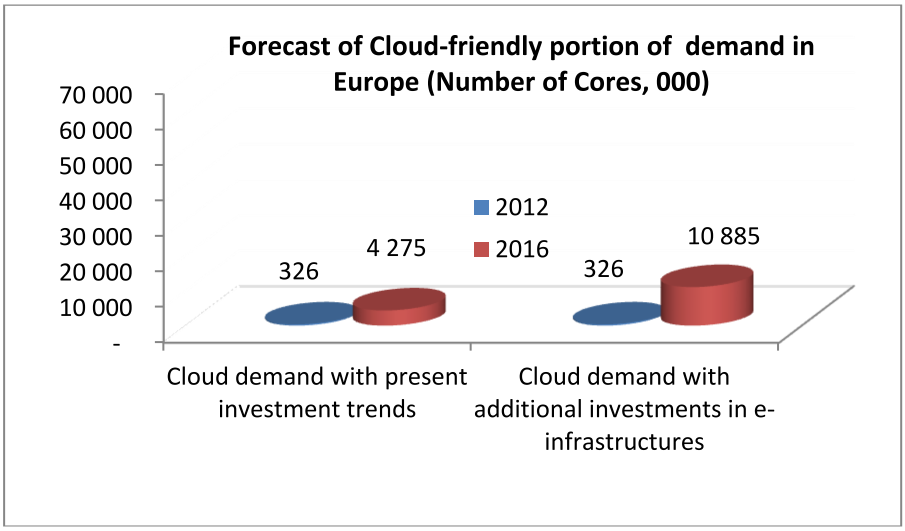
	2011	2012	2013	2014	2015	2016
Number of cores needed for cloud-friendly portion of open science	2,494,162	3,258,985	4,258,337	5,564,136	7,270,352	9,499,772
Percent of cloud-friendly open science likely to be run on the cloud		10%	20%	30%	40%	45%
Resulting number of cores needed for percent likely to be run on the cloud	-	325,898	851,667	1,669,241	2,908,141	4,274,897
If ADDITIONAL investments are made in government-supported science clouds						

	2011	2012	2013	2014	2015	2016
Number of cores needed for cloud-friendly portion of open science, with additional investments	2,494,162	3,258,985	4,968,060	7,882,526	12,117,254	19,791,191
Percent of cloud-friendly open science likely to be run on the cloud, with additional investments		15%	30%	50%	65%	80%
Resulting number of cores needed for percent likely to be run on the cloud, with additional investments	-	325,898	993,612	2,364,758	5,452,764	10,885,155

Source: IDC 2012

The summary results of Table 12 are presented in Figure 3.

Figure 3 Forecast of Cloud-Friendly portion of demand in Europe (Number of Cores, 000)



Source: IDC 2012

However, we need also to consider also the portion of pent-up demand, which could be run in the cloud. This is presented in Table 13 and Figure 4.

- The first row estimates that approximately 9.9 million cores of pent-up demand in 2011 are for use-cases that could run in the cloud, given today's technology attributes (e.g. performance, interoperability, and security).
- The second row estimates the percentage of those use cases that are actually expected to migrate to the cloud, under the scenario of investments currently planned for GÉANT, Helix Nebula, EGI-inSPIRE and so forth. In the case of pent-up demand, the analysis assumes that the take up of cloud starts earlier, because of higher urgency to tap into resources that are not available today. In fact the percentage of demand that is

expected to migrate to the cloud is already at 15% in 2012 (table 13), against 10% (table 12) for the demand that is satisfied by currently used cores.

- The third row calculates the actual number of cores of pent-up demand that are expected to move to government-supported cloud services, as explained in the previous paragraph.
- The fourth row, assumes that additional investments in Géant, Helix-Nebula, EGI-inSPIRE and similar programs are made, above the levels currently planned, thus enabling government supported science clouds to expand capacity and improve technical attributes. This allows cloud services offered by e-infrastructures to accommodate more use cases.
- The following rows present the percentage of cores likely to be used for the cloud in this second scenario, and their total number. In this forecast there are significant differences between the "no additional investments" scenario and the "additional investments" scenario, as the percentage of demand that is expected to migrate to the cloud increases from 60% under current funding to 80% with additional funding by the year 2016.

Table 13 Forecast of Europe computing demand from the "core" scientific community: cloud friendly portion of pent-up demand

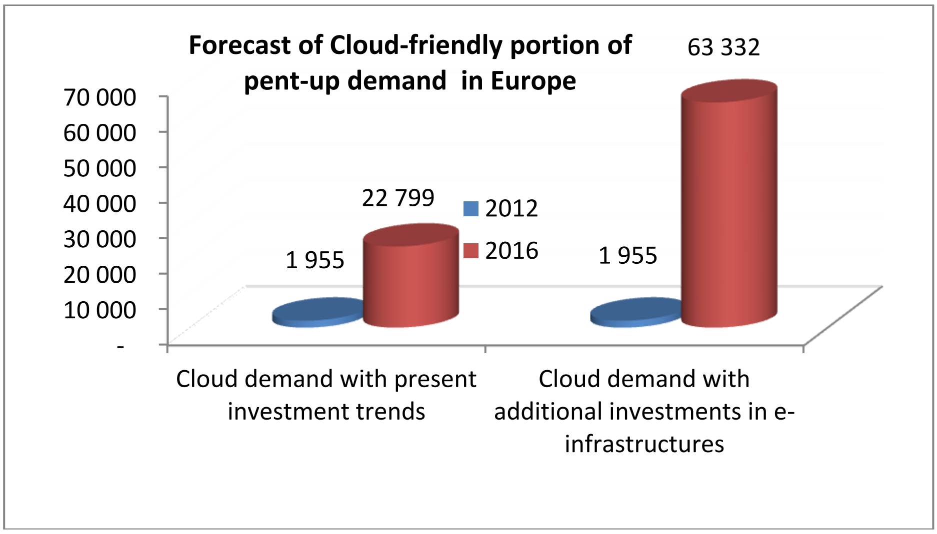
	2011	2012	2013	2014	2015	2016
Number of cores that are cloud-friendly out of the additional cores for pent-up demand	9,976,647	13,035,939	17,033,349	22,256,545	29,081,410	37,999,088
Percent of cloud-friendly open science likely to be run on the cloud, if NO ADDITIONAL investments are made		15%	30%	40%	50%	60%
Resulting number of cores needed for percent likely to be run on the cloud		1,955,391	5,110,005	8,902,618	14,540,705	22,799,453
If ADDITIONAL investments are made in government-supported science clouds						
Number of cores needed for cloud-friendly portion of pent-up demand, with additional investments	9,976,647	13,035,939	19,872,240	31,530,105	48,469,017	79,164,766
Percent of cloud-friendly open science/ pent-up demand likely to be run on the cloud, with		15%	30%	50%	65%	80%

	2011	2012	2013	2014	2015	2016
additional investments						
Resulting number of cores needed for percent likely to be run on the cloud		1,955,391	5,961,672	15,765,053	31,504,861	63,331,813

Source: IDC 2012

The summary results of Table 13 are presented in Figure 4.

Figure 4 Forecast of Cloud-friendly portion of pent-up demand in Europe (Number of cores, 000)



Source: IDC 2012

4.1.6. Summary Forecast Scenarios and System Value

In conclusion, the forecast scenario for the period 2011-2016, in case of continuing investment trends, estimate a demand of 27 million computing cores potentially used for cloud for science in the year 2016, including the growth of the existing demand and a share of pent-up demand. In case of additional government investments, improving both capacity and technical capability of cloud services offered by e-infrastructures, the potential demand for cloud could grow to 74 million cores by the year 2016 (Table 14).

In other words, the additional government investments in the period to 2016 could drive an increase of use of cloud for science almost 3 times higher than the use estimated in the forecast based on current investment trends.

Table 14 Total Forecast Scenarios - Impact of Additional Government Investments on potential Cloud demand

Number of Computing Cores - Existing demand + pent-up demand)

Note: Forecast numbers are incremental over 2011, which is set as 0

	2011	2012	2013	2014	2015	2016
Cloud Forecast Without Additional Government Investments	-	2,281,289	5,961,672	10,571,859	17,448,846	27,074,350
Cloud Forecast With Additional Government Investments	-	2,281,289	6,955,284	18,129,811	36,957,625	74,216,968
Increase Due to Government Funding (e-Science Cloud Usage)	-	-	993,612	7,557,952	19,508,779	47,142,618
% Increase			17%	71%	112%	174%

Source: IDC 2012

IDC has also calculated the average costs of the computing systems needed to meet the demand calculated in both forecast scenarios (table 15).

The system value per core is calculated as the value of HPC systems (according to IDC definition) divided by the number of cores in the systems. This is greater than the value (cost) of the cores alone. System value per core has proven to be a more useful metric for the global HPC buyer community.

Based on IDC data and estimates, the average system value per core is projected to decrease from 243.8 € to 48.4 € in 2016. This results on a total system value of 1.3 Billion € in 2016 for the 27 Million cloud cores needed in the first forecast scenario. In the case of the second scenario, driven by additional government investments, the total system value in 2016 will be of approximately 3.6 Billion €, that is almost 3 times higher.

Table 15 Total Forecast Scenarios - Total System Value to meet demand (Mill €)

Number of Computing Cores - Existing demand + pent-up demand)

Note: Forecast numbers are incremental over 2011, which is set as 0

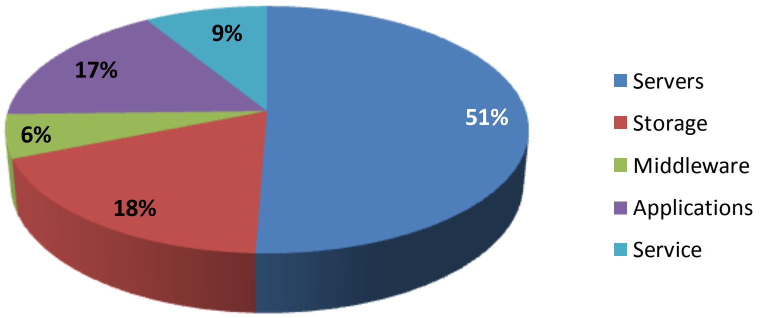
	2011	2012	2013	2014	2015	2016
System value per core (€)	243.8	213.7	147.3	101.6	70.1	48.4
Cloud Forecast Without Additional Government Investments	-	2,281,289	5,961,672	10,571,859	17,448,846	27,074,350
Total System Value to meet demand		488	878	1,074	1,224	1,309
Cloud Forecast With Additional Government Investments	-	2,281,289	6,955,284	18,129,811	36,957,625	74,216,968
Total System Value		488	1,025	1,843	2,592	3,589

	2011	2012	2013	2014	2015	2016
to meet demand						

Source: IDC 2012

Obviously, the total system value in the year 2016 is only one of the components of the investments needed to develop pan-European cloud infrastructures. Typically, servers represent about half of the total cost of scientific and technical computing systems (see Figure 5 below) while storage, middleware, applications and service represent the other half.

Figure 5 - Average Weight of Cost Components in HPC Systems (based on 2012 spending data)



Source: IDC HPC 2013

4.1.7. Concluding remarks

The forecasts presented in the previous paragraphs, even though solidly based in the depth of IDC's data about the technical and scientific computing market must be considered only as estimates of the main trends of the potential demand of clouds for science in Europe. As discussed at length in this report, this is a highly complex market with a multidimensional demand and user population. The results of this analysis lead to the following conclusions:

- The potential demand of cloud services for science and research is relevant and likely to grow rapidly in the near future, even if not all research applications are suitable for cloud.
- The main driver of demand growth is likely to be the very high unsatisfied demand, which according to IDC may represent up to 3 times the existing, visible demand. The flexibility of cloud services make them particularly suited to respond to sporadic peak demand, oscillatory demand, short-term demand, and requests by researchers and scientists from smaller research centres, or not engaged in "Big science" projects. These researchers are likely to be frustrated by the slow peer review process, traditionally used to grant access to publicly paid computing resources.
- The two forecast scenarios highlight how sensitive potential demand is not only to the capacity, but also to the technical capability of e-infrastructures. This requires the investment of considerable resources into the extension and upgrade of e-infrastructures for cloud extended to the entire EU, since the pent-up demand is distributed across the EU27, as is the emerging demand from small research centers and researchers engaged in smaller projects (sometimes called SMS, Small and Medium Science, in contrast to the Big Science projects typical of physics and other natural sciences requiring very high investments for research).

- The forecast scenarios describe an emerging demand which is not tied to a specific country and/or domestic suppliers. Not only the research community is by nature international, there are also economies of scale and scope in the provision of such services at the European level. As discussed in chapter 3.5, this calls for a pan-European infrastructure, able to avoid any potential "digital divide" between the large/small countries and between Big Science/Small and Medium Science (SMS) projects insuring equal access to computing resources across the EU.

5. ANALISYS OF GOVERNANCE AND FUNDING MODELS

This chapter analyses and delineates the governance, funding and organizational models that when applied to cloud infrastructure provisioning, ensure efficiency and effectiveness of implementations. It examines the different actors that exist in the implementation and use value chain. These models and actors are the basis of the service provisioning scenarios for e-Science clouds provided in the later section of this chapter. The analysis examines the various roles that governments, researchers, the EC and other public and private stakeholders can play in those provisioning scenarios. The categorization is intended to demonstrate the dynamics of implementations while it should be understood that implementations are in all cases a combination of the scenarios presented here. The configuration of actors involved also varies from implementation to implementation.

5.1. Governance, Funding and Organizational Attributes

Every combination of circumstances creates particular conditions of the organizational structure chosen for a cloud implementation and no implementation adheres to a given model. To illustrate the forces in play we considered the following attributes of governance, funding and organizational models:

- **Governance models** are to be intended as the structures and processes that assign decision-making responsibilities for the efficient and effective alignment of demand and supply of services. Governance decision-making processes and structures are the funding and investment management programs, portfolio and project management, change management, enterprise architecture development, sourcing and escalation mechanisms resolving service management issues. Governance is of the utmost importance in an environment that still suffers from unbalances between ever-growing demand and supply that does not always match the requirements. Highlighting this fact, the 2011 IDC study for the European Commission[28] confirmed the need for European industry to have access to large-scale HPC systems to perform pre-competitive advanced research, as well as competitive advanced research. At the same time, IDC's analysis confirmed that some PRACE centers, as well as HPC national centers in other areas of the world, are over-subscribed and would find it difficult to reserve any substantial amounts of time on their HPC systems for major new users. Revealing the complexity of the equation, in 2011, the Chinese government deployed seven substantial public clouds for advanced research, to provide Chinese industry with unpaid access to cloud. The program, which was intended to last for two or three years, was met with limited demand, demonstrating the challenge of matching supply and demand[29]. Nevertheless, in the past two years, Chinese investment in cloud infrastructure to support advanced research has soared. China Daily said cloud computing added $15 billion to China's IT industry in 2012. IDC estimates that government spending on cloud computing infrastructure will grow to $1 billion annually by 2016.
- **Funding models** refer to the mechanisms used to finance costs of developing, managing, maintaining and replacing services. Typical funding challenges in government and science include:
  - o Ensuring full coverage of the cost for the provider, while reducing the risk that a profit or margin based mechanisms reduces the incentives to invest in continuous service improvement.
  - o Making sure users avoid "over-consumption", which often happens when the resources are not charged to the end user. In this case capacity is quickly saturated with non-critical workloads. However it also includes the opposite or "under-

[28] IDC, "A strategic Agenda for European Leadership in Supercomputing: HPC 2020" 2011, on behalf of the European Commission

[29] Insights from IDC discussions with Chinese government officials and HPC center directors, 2011-2012

consumption”, which happens when excessive levels of service are provided driving costs too high. In this case administrators look for alternative, affordable resources that might offer just "good-enough" performance.
    - Balancing the regulatory compliance and audit costs among providers, brokers, developers and consumers.
    - Encouraging all stakeholders, in particular the developers, to contribute to innovation that can fill the gap of services not commercially available. Researchers are immersed in an inherently innovative organizational culture, but public sector budget constraints could stifle innovation if for example small institutions have to cover the cost of piloting and testing new applications, or porting their services to new shared infrastructure on a pay-as-you-go basis.
    - Avoiding over-engineering investment management and cost-recovery models, which can generate excessive administrative costs. In many cases a simple per-user based chargeback model is sufficient, while 100% accurate activity-based schemes increase complexity and provide only marginal improvements in terms of demand management and transparency.
    - Overcoming organizational, cultural and legal barriers. As this study mentioned in the analysis of drivers and barriers for cloud adoption, in the education community concerns have been raised about the ‘pay-as-you-go’ approach, not only because of the burden on individual researchers using their credit cards, but also claiming back the costs and covering higher costs accrued. Additionally a consistent approach to including computing costs as part of eligible research costs is lacking across Europe.
    - Including the pay-as-you-go model as a legitimate funding model in government and academia where the legal and normative framework currently mandates that research grant funding are spent on capital investments for specific projects.
    - Ensuring that the flexibility of tapping into alternative sources of funding (private, public, and public-private) does not hamper the freedom of researchers to conduct their scientific studies independently.
- **Organizational models** refer to how the services are operated: that is who is responsible for providing services to the users. There are three operating models that can be used as reference at a macro-level:
    - Centralized: all services are delivered by a single organization.
    - Federated: some services are offered centrally, and some services are offered by dedicated organizations distributed within individual users.
    - Decentralized: every group of users, in the case of e-Science, every unit of researchers, has its own dedicated set of providers, developers and so forth.

Governance, funding and organizational models are tightly linked and if addressed in isolation can generate inefficient and ineffective outcomes.

5.2. Actors

Every implementation of e-Science clouds will, of course, be unique. Every application will have different requirements, budget and intended usage. Every group of professionals implementing an e-Science infrastructure will necessarily have a unique set of figures which may include:

- E-Science service carrier: the carrier acts as an intermediary that provides connectivity and transport of services between consumers and cloud providers.
- E-Science service provider: the provider delivers the service to the consumer; the providers' task may vary depending on the type of service (IaaS, PaaS, SaaS), but they generally include installing, configuring, managing and maintaining the systems and operating the processes necessary to deliver the service, such as helpdesk and billing.
- E-Science service developer: the developer creates, publishes and monitors the services, typically "line-of-business" applications that are delivered directly to users. The developer can be considered a consumer of IaaS and PaaS services. In the case of e-Science, service developers can include specialized software companies, scientists and non-professional researchers.

- E-Science service broker: the broker matches services between providers and users, for example by providing a gateway that federates identity management, or by offering interfaces and connectors.
- E-Science service buyer: the buyer is responsible for the procurement process; increasingly in the public sector, the buyer is not the same as the user, because national or regional procurement bodies offer specialist purchasing services (tendering, pre-qualification, certification, negotiation, etc.) to aggregate bargaining power of multiple organizations, for example through framework contracts and blanket purchase agreements that end-users can leverage to speed up and limit the risk of contracting with providers.
- E-Science service consumer: is the end user or enterprise that actually uses the service, whether it is Software, Platform or Infrastructure as a Service.
- E-Science service regulator: the regulator is in charge of defining laws and policy and funding guidelines; in the realm of cloud computing this usually include privacy and security, including information assurance, certification and audit processes, and open standards for interoperability and portability.

It is clear that any scenario or model may include any of a series of actors that may be active in the models that follow.

5.3. E-Science Grids for Clouds

5.3.1. Definition

E-Science clouds are based on a federation of known, identifiable HPC and technical/scientific computing systems and related resources that are networked together (loosely coupled) across separate, collaborating organizations (e.g., separate centers within a nation or region). This assumes that a portion of the existing computing resources in organizations such as PRACE is set aside and offered to potential users, with a cloud delivery mode, that is pay-as-you go, instead of through peer review, on demand and elastically allocated based on need.

The visibility of systems on the grid gives rise to the description of grids as "transparent." This is in contrast to the description of clouds, especially public clouds, as "opaque," because the specific computers, storage, and other resources you are using cannot easily be identified. A signed agreement typically specifies the access and usage terms for the grid-networked HPC resources. We realize that grids also exist that link together HPC sites within the same public- or private sector organization (e.g., campus grids), but those are far less relevant for Europe as a whole. HPC grids are used primarily for batch-oriented computing on known resources of fixed, inelastic size and type-meaning that they typically are not designed to accommodate applications with a wide variety or changing requirements.

5.3.2. Governance

One of the main governance challenge of e-Science grids is the balancing of requirements of research centers that own HPC resources and want to make sure they satisfy first the needs of their internal scientists, before they make spare capacity available on the grid and the requirements of consumers and developers that in their research centers have limited computing resources and want cheap access. The roles of service carriers and brokers of European research networks, such as GÉANT, are essential to ensure technical optimization of workloads across the network. EC and member state funding agencies can play a key complementary role to define the portfolio management guidelines that ensure a better business optimization of workloads, based on the mission and timeline of scientific progress; for example a life-saving drug-discovery project near completion could be given priority to use computing cycles over a regional economic modeling on that is starting. The combination of technical and business prioritization processes can provide an efficient and effective clearinghouse for the grid.

5.3.3. Funding

In the interest of developing a better performing grid the EC and member states should focus funding of:

- Innovation actions aimed at providing better network performance;
- Actions aimed to improve interoperability and portability of grids.

These are generic technology investments that are hard to allocate to a specific research project or research center and even harder to justify as depreciation costs in a pay-as-you-go chargeback model. On the other hand, operating expenses that the carrier/broker managing the grid pays, can be paid back by the users with a chargeback mechanism. Adjustment to a per-user or per-usage chargeback model should be based on the size and role of the organization; for example private enterprises using the resources for competitive purposes could be charged an extra price that is used to fund access for small academic centers that have limited resources to pay for the computing capacity they use for piloting and testing, or to pay for limited access to HPC for non-professional scientists that get it as an award for developing new applications.

The SURFnet case study, presented previously, offers a useful example of this model. SURFconext is publicly funded by the Dutch Ministry of Education and Research. It is intended to deploy services responding to confirmed needs where no current offer is available on the market. The service is maintained by public funds until it is deemed commercially sustainable, at which point it is transferred - via tender - to commercial partners that commercialize it. The services are then funded by higher education and research, which pays fees for using the services. Fees include the fiber connection for Internet and depend on the size of the institutions requesting the service.

5.3.4. Organization

A grid is by definition a federated model, where the carrier and broker operate shared network infrastructure services centrally; for example federated authentication based on the SAML standard are in use within Higher Education e.g. “SURFfederatie” in the Netherlands, “InCommon” in the U.S. and “UK Access federation” in the UK. In the case of SURFnet the central unit also provides the SURFteams service to organize group management for inter-institutional online collaboration.

Computing capacity providers, application developers and researchers and other end-users operate in a decentralized manner. As a result, particularly for developers, it is important to identify mechanisms that facilitate the re-usability and portability of applications across computing platforms and networks. This central role for application dissemination and aggregation is typical of a broker and could prove beneficial in cloud application marketplaces.

Carriers and brokers of HPC grids should also ensure they optimize the combination of skills, good practices and motivation to continuously improve service management capabilities to align with HPC provider, developer, and consumer requirements. Carriers and brokers can also play a centralized buyer role for communication services and equipment and system management and middleware software.

The following table summarizes the main attributes of this provisioning scenario.

Table 16 - Main Attributes of Clouds Provisioning Scenario 1

E-Science Grids Scenario		
Attributes	**Assessment**	**Motivation**
Parallel, Problem solving performance	Very High	Maximum in capability computing, massively parallel system, excellent speed and flexibility of the interconnect
Open access / availability	Medium	Availability may be conditioned by owner's computation load
Federated sites' ability to control service levels	High	Reduced computational diversity and known hardware characteristics permit detailed control of SLAs
Resource transparency	Medium	Limited remote site user control over policy or security
Resource elasticity	Low	Limited flexibility in terms of portion of computational allocation, system configuration and available applications
Interoperability with other e-Science systems	High	Standardised data and metadata modeling formats, domain application convergence
Effectiveness of governance and funding mechanisms to encourage resource sharing	Low	Focus on internal demand
Efficiency / cost of running individual research cycles	Low	Not suited to one-off set up and run
Fixed costs	Very High	Cost distribution and allocation difficult, Systems costs very expensive

Assessment Scale: Very High - High - Medium-Low- None

Source: IDC 2013

5.4. Private or Community Clouds

5.4.1. Definition

IDC defines a Private Cloud as one or more computer systems and related resources within the firewall (security perimeter) of a single organization that the organization's authorized users can access on demand, usually for special needs, with an accounting made of their usage. Community cloud refers to the situation where there are a limited number of stakeholders that have access to the cloud infrastructure. It is not open to the public but stakeholder users may come from a number of organizations. These organizations are using high performance and high throughput computing for scientific and engineering tasks and experiments. IDC defines High-Throughput Computing (HTC) as the use of many computing resources over long periods of time to accomplish a set of loosely coupled computational tasks as opposed to High Performance Computing which places the emphasis on the cluster and supercomputers designed to perform the maximum number of operations per second, and make use of special architectures to achieve this goal.

Typically the structures were born to satisfy exclusively in-house computing needs but open to intra departmental, organizational computing needs to increase peak capacity or reduce costs.

The capabilities of private science and engineering clouds can vary greatly. In some cases, they are fully equal extensions of an organization's technical computing data center resources, but with separate rules for obtaining access. In other cases, private clouds provide types of resources that are unavailable in the data center. A variation of private clouds is community clouds, where there is a limited-membership user group.

5.4.2. Governance

Private cloud can experience governance challenges that are similar to those of HPC grids, when it comes to prioritizing among the requirements of developers and consumers that are part of the core group of researchers that funded the initial investment and other communities of scientists. The prioritization can exacerbate further if an HTC/HPC private cloud turns into a community cloud that offers services outside of the boundaries of the academic or government research institutions that owns the assets to a limited group of other researchers – a case that can materialize for example if the Helix-Nebula program gains traction as a HTC/HPC cloud provider across Europe. A possible solution to such conundrum is the creation of a separate entity responsible for operating the HTC/HPC service and offering it as a quasi-commercial enterprise to both the users that made the initial investment and those outside of the core group. This is a suitable solution when two macro-conditions are met. First of all the "outside" consumers (or buyers, or developers) must absorb a share of computing capacity demand that contributes significantly to the saturation, or even generates demand for scaling up computing resources. Secondly, there must be in the research center that created the HTC/HPC private cloud or through the creation of a shared service of all users, or through an external partner (a commercial enterprise, a public private partnership) the availability of financial resources and management competencies to take over the operation.

5.4.3. Funding

As this study shows, the first step to make cloud architectures suitable for HTC/HPC is to improve the technical performance when running parallel computing. This is an opportunity for the EC to make available target grant funding to make sure the European HTC/HPC hardware and, particularly, software engineering enterprises stay ahead of the competition in developing parallel computing capabilities for the cloud. Clearly the EC and member state government investments should be considered as seed grant or loan funding, but the hardware and software industry should take its own risks, provided there is a big enough, open and transparent market for the new technology.

The governance challenges analyzed in the previous section are very similar to the HTC/HPC Cloud challenges. When the HTC/HPC services are used only by one research institution, all of the capital investment for deploying the systems and the operating expenses for running it can be covered by that institution budget. When the cloud opens up to “group” developers, buyers and users, it progressively becomes a community cloud. As it does so the government or academic institution that originally invested it is not willing to bear all of the capital costs for equipment maintenance until the infrastructure becomes self sustainable due to dues from the community cloud users. To get around this problem the national government and the the larger European Community could provide funding as seed money to overcome the initial capital intensive phases of the building the community cloud. Analysis however must be done to ensure that services provided do not have market alternatives so as to avoid distorting market competitiveness. The initial step will be to apply a chargeback for operating costs and then to add a margin that will be accumulated into a working capital fund for upgrades and replacement of assets. If a separate entity responsible for the community cloud service is created, their funding model should be based on full cost recovery.

5.4.4. Organization

Private clouds are organized with a centralized operating model, where the provider makes all of the key decisions.

The following table summarizes the main attributes of this provisioning scenario.

Table 17 - Main Attributes of Clouds Provisioning Scenario 2

E-Science Private Clouds		
Attributes	**Assessment**	**Motivation**
Parallel, Problem solving performance	High	Good capacity computing, less effective capability computing, good speed and flexibility of the interconnect
Open access / availability	Medium	Availability may be conditioned by original owner's computation load
Federated sites' ability to control service levels	Medium	Requirement diversity is expected from "community" negatively affecting "standardization" of SLAs
Resource transparency at remote sites and users' control on security and policies	High	All users are aware of the computational capabilities, locations, and configurations
Resource elasticity (computing power, use cases and type of software)	Medium	Some flexibility in terms of portion of computational allocation, system configuration and available applications
Interoperability with other e-Science systems	Medium	Many different computing loads and software closed community security regimes
Effectiveness of governance and funding mechanisms to encourage resource sharing	Low	Focus on "community" demand
Efficiency / cost of running individual research cycles	Medium	Ability to schedule and lower or higher priority tasks
Fixed costs	Medium	Some cost distribution and allocation, originator bears systems costs initially

Assessment Scale: Very High - High - Medium-Low- None

Source: IDC 2013

5.5. Public e-Science Cloud

IDC defines a public science and engineering cloud as one or more technical computing systems (typically clusters) and related resources that are made publicly available, usually on a self-configurable, pay-as-you-go basis. These resources are often virtualized and exist outside the firewalls (security perimeters) of user organizations and are not subject to users' security measures, policies, and directives. These are generally HTC systems that are publicly available. Typically, they are provisioned at the SaaS and PaaS level and are configured on-the fly and used on a pay-as-you-go basis. These resources are often virtualized and exist outside the firewalls of user organizations. They are not subject to users' security, policies, and directives.

5.5.1. Governance

In the case of public cloud, governance mechanisms are almost entirely substituted by market mechanisms. It is the provider that is responsible for all investment decisions to ensure service levels, users (and developers in case of IaaS and PaaS) can only apply the two classical weapons of customers if they are not satisfied: "voice" or "choice". With "voice" customers can complain about the performance, relatively to the price they pay and the agreed terms and condition. With "choice", customers can select an alternative provider that offers a better price/performance ratio.

5.5.2. Funding

As in the case of Private Cloud, some public grant funding should be invested to develop cloud architectures suitable for running parallel computing.

The member states and the European Commission can decide to offer grants funding to commercial cloud players to attract them to invest in cloud computing facilities in their countries; however, IDC research suggests that harmonization of privacy, security, technical standards across the community can be a more powerful driver of investment for global players.

The consumption of public cloud services is expected to be paid for on a per-user or per-use basis, as commercial providers and brokers are maturing their billing mechanisms very rapidly. Therefore, an important change that needs to take place to enable using public clouds is in the laws and policies that prevent many researchers, particularly in public sector, to buy computing capacity under the operating expense budget.

5.5.3. Organization

In the case of public, commercial clouds, government purchasing authorities, or consortia of academic institutions, such as CINECA in Italy, can act as centralized buyers to speed up the procurement process and leverage bargaining power to get better pricing. For these central purchasing mechanisms to work effectively, it is important that joint committees of public cloud providers and end-users are put in place to agree on architectural roadmaps that provide common ground for a set of standard security, interoperability and functional attribute of the service. In some cases, the central procurement bodies could create marketplaces to facilitate access to services on an on-demand basis, as the UK government did with the CloudStore (see case study analyzed above).

Regardless of the centralization of some parts of the procurement process, users of cloud services, that is developers and consumers in the case of e-Science, will make most of their choices in a decentralized manner, thus must acquire new capabilities. In fact, instead of being technical experts in installing, configuring and managing on premise systems, they will have to become experts in interfacing cloud services with legacy, managing contracts with different SLAs and federating identity and access management across non-natively interoperable systems.

The EC, for example through ENISA, and member states can play a particularly important role relatively to interoperability and security standards and certification processes. If standard information assurance processes are not in place, procurement processes will be much longer and riskier and the creation of central procurement bodies or marketplaces will experience limited take up, as indicated by the UK government CloudStore example. In addition, if interoperability standards are not disseminated, the risk of duplication is very high. For example the RainCloud workflow system is interoperable with other European infrastructures although this has not yet been tested and is not a pre-requisite of the RainCloud project; in particular fine grain interoperability includes thin interoperable intermediate representations meaning that every workflow system can translate this common representation which is supported by the underlying execution system that are run. This can be considered as a new standard that has been developed.

The following table summarizes the main attributes of this provisioning scenario.

Table 18 - Main Attributes of Clouds Provisioning Scenario 3

E-Science Public Clouds		
Attributes	**Assessment**	**Motivation**
Parallel, Problem solving performance	Low	Limited capacity and capability computing, interconnect suitable to limited I/O requirements
Open access / availability	High	Open access and freely available
Federated sites' ability to control service levels	Low	SLAs are standardised and negotiation by pre-built contracts
Resource transparency at remote sites and users' control on security and policies	Low	Users have limited awareness of the computational capabilities, locations, and system configurations
Resource elasticity (computing power, use cases and type of software)	Medium	Flexibility in terms of portion of computational allocation in IaaS level, while system configuration and available applications in SaaS PaaS tend to be standardised
Interoperability with other e-Science systems	Low	Custom programming interfaces, not generally interoperable with GRID HPC infrastructures
Effectiveness of governance and funding mechanisms to encourage resource sharing	High	Allows economic shift from CAPEX to OPEX model, some problems in recognition of eligibility of expenses
Efficiency / cost of running individual research cycles	Very High	Cost per cycle significantly reduced, abridged application deploy times, increased hardware usage
Fixed costs	None	No capital expenditures or fixed costs required

Source: IDC 2013

5.6. Concluding considerations

The provisioning scenarios identified above are categorical and intended to illustrate categorical differences between the three main e-Science Grids, private and public provisioning models available to pubic administrations and scientific organisations requiring computational capacity and capability. The scenarios are based on three fundamentally different types of e-infrastructures. The first, which is an extension of the GRID HPC computing infrastructures, provides optimal levels of performance to tackle large computational problems in the shortest amount of time possible taking advantage of massively parallel architectures as well as the speed and flexibility of the interconnect which is possible in these types of cluster configurations. These systems are transparent and well known to the stakeholders who deploy A limited number of dedicated applications to complete computational and I/O intensive tasks. On the downside these systems are capital intensive with few institutions sharing initial investment and uncertain return and sustainability as a cloud model. The allocation of resources to users is hindered by a preference to the workload of those who implemented the systems. The excess system capacity that these owners would like to offer to potential users is made available only after internal demand has been satisfied. This makes these systems unreliable (in availability) for users who are not “owners” of the system.

The second types of private cloud systems interconnect independent systems inside a private network. They are proficient at large numbers of sequential jobs that can be individually scheduled on many different computing resources within private administrative boundaries. The systems are transparent for those within the network and capabilities can be identified and jobs scheduled with a good degree of elasticity. Like the former investment scenario these private e-Science clouds are capital intensive to start and as the community is opened to decrease costs, the availability and elasticity of computational resources declines.

The final scenario depicts the completely open public eScience cloud. These systems are very accessible and therefore have a high level of availability for users. The resources are generally very elastic and can be ramped up or decommissioned with little effort. The total lack of a priori investment and the fact that the user only pays for the exact resources commissioned and used makes them economically attractive. The downfall is of course in capability and transparency. Although there are high performance public clouds generally speaking these infrastructures lack tightly coupled, high performing infrastructure needed to run highly parallel complex scientific applications that require extremely low latency. The lack of visibility of the complete network and infrastructure where the applications are running also creates problems for policy and security limitations required by some scientific applications.

In reality, "hybrid" provisioning scenarios are more likely to be dominating the EU future developments. For example, the experience of the Helix Nebula project is moving towards the development of a standards based federated cloud architecture to enable an open platform for science innovation, based on the federation of publicly funded, community grids (such as the EGI) and commercial providers. While the HN project started out as a "community" provisioning scenario open to the research centres of the communities involved, it is moving towards open access based also on commercial services (that is, open to anybody who can pay, even if with some differences compared to the completely commercial platforms such as Amazon's).

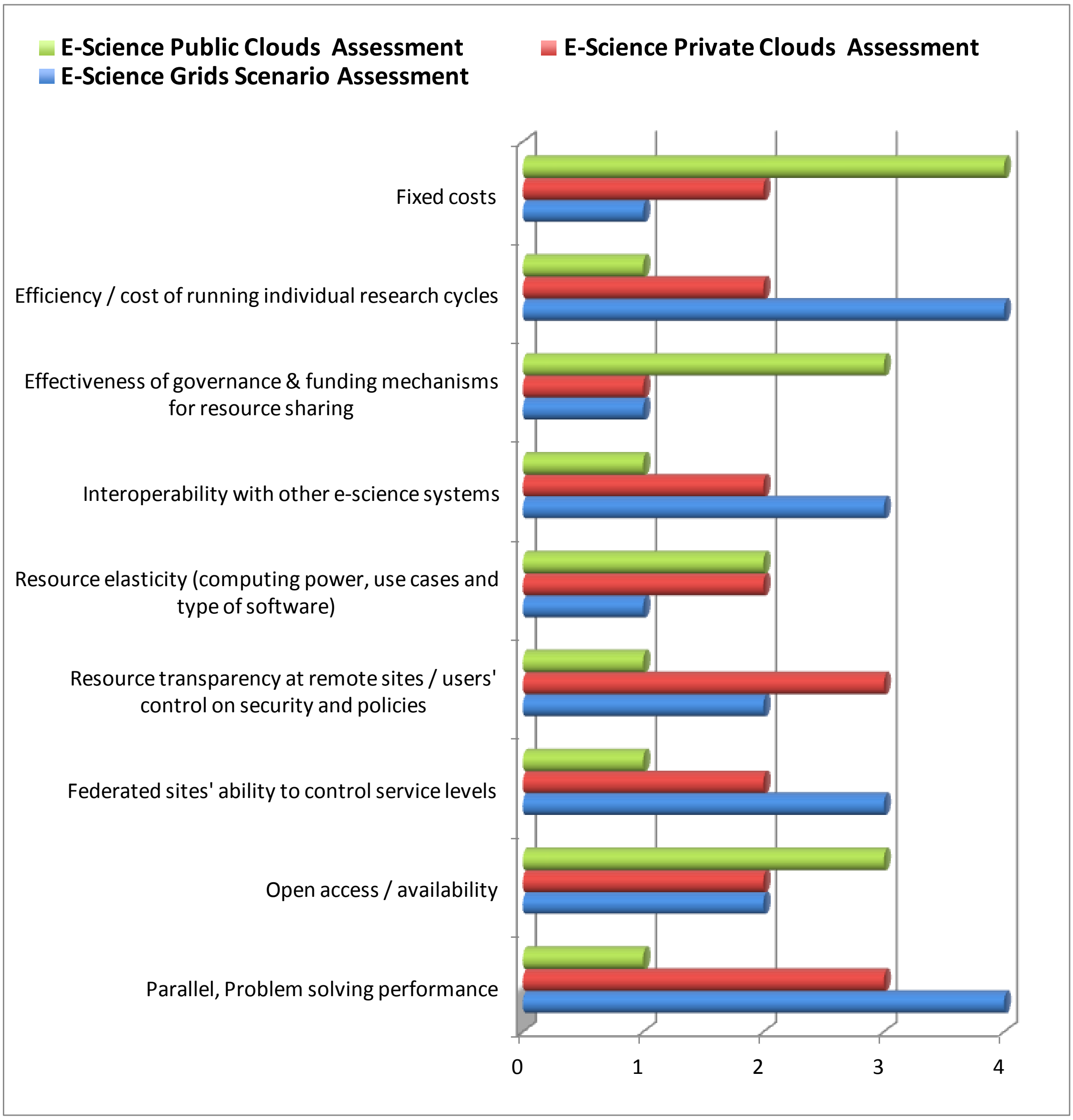
A comparative view of the attributes of the provisioning scenarios is presented in the following Table 19 and Figure 6.

Table 19 - Summary Assessment of Attributes of E-Science Cloud Provisioning Scenarios

Attributes	E-Science Grids Scenario	E-Science Private Clouds	E-Science Public Clouds
	Assessment	Assessment	Assessment
Parallel, Problem solving performance	High	High	Low
Open access/ availability	Medium	Medium	High
Federated sites' ability to control service levels	High	Medium	Low
Resource transparency at remote sites and users' control on security and policies	Medium	High	Low
Resource elasticity (computing power, use cases and type of software)	Low	Medium	Me
Interoperability with other e-Science systems	High	Medium	Low
Effectiveness of governance and funding mechanisms to encourage resource sharing	Low	Low	High
Efficiency / cost of running individual research cycles	High	Medium	Low
Fixed costs	Low	Medium	High

Source: IDC 2013

Figure 6 Summary view of E-Science Cloud Provisioning Scenarios



Source: IDC 2013

6. CONCLUSIONS AND POLICY RECOMMENDATIONS

6.1. Conclusions

The European policy strategy for the implementation of the ERA and the pursuit of excellence in science as one of the key priorities of Horizon 2020 reflect the increasing relevance of scientific and research activities for growth and innovation. The development of e-infrastructures across Europe is becoming an enabling condition for the "fifth freedom" of knowledge and data mobility in the EU, responding to emerging demand for open, flexible and scalable computing capacity that national resources cannot satisfy. Science is becoming ever more pervasive, global and open. This means that access and use of e-infrastructures in pursuit of innovation, education and excellence in the sciences, arts and humanities is becoming a public right in a society that wants to be more sustainable, smarter and more inclusive. Europe's ambition is also to become more adaptive, innovative, motivated and a driver of entrepreneurial culture. Again, the development of open and collaborative e-infrastructures is a key tool to enable industrial competitiveness, overcoming the "valley of death" between research and the market, where European innovations too often fade away to irrelevance.

As the demand for e-Science grows beyond the traditional boundaries of national research networks and big science projects, there is a clear potential for cloud computing infrastructures and services to fill the gap between traditional offerings and emerging demand.

European investments in e-infrastructures and a wide range of pilot projects have already provided early demonstrations of the potential of cloud to transform science, address big data challenges and enable collaboration across a much wider research community, as documented by this report. This broader community numbers in the thousands for each traditional supercomputer user and is important to ensure sustainable e-infrastructure. There are also clear examples of researchers pursuing new lines of research, creating start-ups and growing businesses that we need to drive the economy. Clouds are well suited to respond to peak usage or oscillatory demand for computing power, as well as to a range of emerging applications combining research and governmental risk management activities for example in the healthcare and environmental protection fields. Therefore the existence of a strong potential demand for e-Science clouds, including not only physical resources, but also data sources, services leading right the way through to computation, is clearly proven. However, demand cannot be satisfied only or mainly by public cloud commercial offerings for science, which exhibit several limitations in terms of actual capability, lack of transparency, compliance with regulation, and sometimes even higher pricing than traditional DCI.

While market drivers pushing the development of e-Science clouds exist, there are considerable barriers and challenges to be overcome in order to achieve EU-wide cloud infrastructures, avoiding potential digital divides between MS, duplication of investments at national level, and fighting the proliferation of standards and lack of interoperability, which may reduce the potential benefits of cloud services.

The need for pan-European cloud e-infrastructures

There is a natural pan-European dimension for cloud e-infrastructures, given the transnational nature of science and research, and the strong demand by users for freedom of choice beyond national boundaries. The two forecast scenarios highlight how sensitive potential demand is not only to the capacity, but also to the technical capability of e-infrastructures. This requires the investment of considerable resources into the extension and upgrade of e-infrastructures for cloud extended to the entire EU, since the pent-up demand is distributed across the EU27, as is the emerging demand from small research centers and researchers engaged in smaller projects (sometimes called SMS, Small and

Medium Science, in contrast to the Big Science projects typical of physics and other natural sciences requiring very high investments for research). A pan-European infrastructure for cloud will be able to avoid any potential "digital divide" between the large/small countries and between Big Science/Small and Medium Science (SMS) projects insuring equal access to computing resources across the EU.

From the point of view of the need for investments, no single MS can compete in advanced research with the U.S., Japan and now also China. EU wide e-Science infrastructures are necessary to support this advanced research on a pan-European scale, as well as on regional and cross-border scales within Europe. For example, most of the European FET flagship projects (Graphene, Blue Brain), funded with 1 billion euro each, and already involves scientists from most of the MS. The EU-wide infrastructure would greatly benefit these projects by supporting open (non-classified), pre-competitive science and industrial research.

However, given the current fragmented scenario, there is a need for a coherent set of policies, standards and services supporting the development of pan-European cloud infrastructures to achieve economies of scale and scope. Without such an approach, much of the potential demand may risk remaining unsatisfied. There is, however, still considerable uncertainty about the best business models supporting these infrastructures.

Potential synergies between e-Government and e-Science Clouds

This study has explored in depth the possibility to pool funding and resources in order to develop common cloud e-infrastructures for the public sector. However we were not able to find examples of shared infrastructures, or co-tenancy of science and government workloads, while instead we found many fundamental differences in the usage patterns of IT services by e-Government and e-Science. There are different user requirements, different technical requirements, funding patterns and organizational models. On the other hand, there is potential for an open, interoperable cloud e-infrastructure designed for e-Science to be able to meet also the needs of a variety of innovative and advanced e-Government applications, based on the collaboration between research and government actors, particularly concerning the diffusion of Big Data in e-Government applications and computational applications for public health and environmental sciences. This business model would also offer the scientific community an additional option for financial sustainability of their technology investments. However, for this model to be successful, scientists will have to acquire a service management culture and organizational capabilities to market, deploy, operate, bill and support services for government "customers".

Emerging Initiatives and Funding Models

Currently, e-infrastructures for research have been dependant for the most part on public funding, with each country in Europe organizing their own resources and facilities independently. Within the current European general purpose e-infrastructure, to provide a level of organisation this has meant each nation organising themselves to providing a single national level contact point that is then able to leverage existing resources from different public research centres and universities. Those resources have been acquired and maintained through public funds aimed at supporting different user communities or building up general e-Science infrastructures. Governing units are normally lightweight and in most cases with limited budget. The fees for sustaining central services that different types of e-infrastructure such as AAA, service monitoring, discovery etc. are also supported directly through public funds. This approach has several complexities, such as the lack of ownership of the resources by the coordinating centres, and the high fragmentation of resource providers, which increases notably the operation costs. It also currently makes it difficult for other providers to enter the market at a pan European level as it is unclear how they would interact with national or international groups.

The EGI.eu Federated Task Force and the Helix Nebula project are both high profile initiatives working to solve these problems.

Helix Nebula has established a growing public private partnership of 30 commercial cloud providers (suppliers) and publicly funded research organisations (users). Three high-profile flagships sponsored by CERN (high energy physics), EMBL (life sciences) and ESA (earth science) have been deployed and extensively tested across a series of cloud service suppliers. These commitments behind these initial flagships have created a critical mass that attracts suppliers to the initiative, to work together and make investments. Links have been established with DANTE and a number of NRENs so that the commercial data centres around Europe have been accessed by the user organisations via the GÉANT network. According to HN proponents, these deployments and tests have revealed a series of gaps in the current set of offerings on the cloud market and the appreciation that the best means of promoting Europe's leadership is to create an open standards based multi-vendor federated market which will allow the diversity of Europe's suppliers to compete with global leaders such as Amazon, Rackspace and Google.

Based on the experience gathered from these "proof of concept" deployments, Helix Nebula's architecture group, led by a series of cloud-savvy SMEs, have defined standards based federated cloud architecture to enable an open platform for science innovation. EGI.eu is contributing to the development of the architecture so that the EGI publicly funded e-infrastructure can be interfaced with Helix Nebula.

Flagship applications from more research disciplines that will stretch the functionality and impact of Helix Nebula have been identified for deployment during 2013.

An initial analysis of the procurement methods of the users and suppliers has been performed and a number of candidate business models highlighted that could ensure the sustainability of the initiative. The transparency of pricing of services is contributing to a more effective market and allows the users to complete a factual comparison of the cost of cloud services compared to the use of in-house resources. The public-private governance model has been expanded by refining the roles of the suppliers to an array of activities that will contribute to expanding the initiative into an ecosystem of services include consultancy, training etc.

Within this context, the main challenges identified by this study for the evolution of flexible, scalable and quality cloud services in Europe are the following.

6.1.1. Technical challenges

- **Promoting the federation of cloud services** based on open standards requires considerable efforts not only for the technical requirements but also for the implementation of all the relevant standards identified, their testing and validation[30].
- **Building innovative and integrated e-infrastructure services, aimed at diversifying service offers for different communities as much as possible**. This implies detailing the legal, fiscal and engineering issues; defining common IT security policies; dealing with organizational and procedural changes, analyzing opportunities and risks for new IT solutions; developing final business case for long-term advantages. The integration of services, including the in-house integration of services at EU level, in order to achieve more complete integration should be underpinned by the implementation of open standards. The development of demand for these services will require encouraging wider usage of computing resources by user communities from the entire e-infrastructure ecosystem. Another approach is to develop a federated set of services in Europe.
- **Need to fill the technical gaps in the offering**. There is yet a lack of maturity of SLAs on performance of development, testing and production environment, business

[30] https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups

continuity, data access, migration and helpdesk for cloud services and how they apply to different usage scenarios.

- **Need to advance the capabilities of these clouds**, especially faster communications to support scientific work that is less embarrassingly parallel and more tightly coupled.
- **Lack of commitment by commercial cloud providers for the development of applications and services needed by e-Science, due to uncertainty in the assessment of the potential demand and the risk-benefits balance**. Some companies may be reluctant to develop specific types of software due to increases in both risks and costs. They may also lack a unit of measure to define comparable costs for new services.

6.1.2. Market and Business challenges

- **Development of new business and funding models**. User communities need to pay for the e-infrastructure services they consume. This means they must have choice and purchase the services they need. On the supply side, e-infrastructure providers need to compete for innovation money and generate revenue from their customer base. This remains a considerable challenge. Providers of such services also need to champion services that ensure a competitive edge over large commercial cloud providers. It is also important to streamline the processes of gathering requirements and aporaching new user communities. Funding models must be easy to administer, because otherwise the cost of fee for service mechanisms outweighs the benefits. They must also be adaptable, because if in the short term cost recovery might be enough for operating costs, in the longer term, full cost recovery including CAPEX might be necessary.
- **Need to move beyond the CAPex funding model in research funding**. Currently, Grids suit large, well organized collaborations with large (CAPex) investments in Data centers. Clouds suit a new model when computing is treated more as an operational (OPex) cost. Cloud computing does not fit well with the typical science grant budgets[31]. There are two main challenges: fostering a change in funding policies and raising awareness of the many opportunities of using cloud computing for research.
- **Need for user-centric approaches, possibly creating a market for e-research**. The main challenge lies in re-directing the e-infrastructure provision strategy focusing it on user needs. A possible approach to do so could be the creation of a "data and e-research market" in Europe both for scientific and social science applications. This approach sees the technology (grid, cloud, combined approaches) mainly as a tool that glues together data and user communities across diverse fields. The aim is to develop an environment where research data can be generated more quickly through the cloud, shared more effectively and rapidly through an open process, and thus generate knowledge that has value. One possible model to do so is the creation of application store-like services charging a small fee, payable to the data/software provider with a high flux of transactions, plus a consultation fee. The analysis of e-infrastructure requirements should grow from technical requirements to end-user services and user-friendly tools and applications. In this way, the e-infrastructure provider could play a better role as facilitator in meeting the objectives of scientists with different IT skills.
- **Need for more agile provisioning models for science**. An important component of the evolution towards user-centric approaches concerns the way to allocate access to scarce computing resources. Traditional peer review is an effective model for determining which scientific projects are granted access to HPC resources. But the peer review process can take months to complete. This does not exploit the elastic ability of clouds to respond to more immediate needs of scientists and other researchers for additional capacity or special capabilities. Peer review bodies should consider defining circumstances in which researchers can gain more rapid access to cloud-based resources.

---

[31] VENUS-C questionnaire with end-users showed that the funding for the use of cloud as part of a research project is not always allowed and that different policies exist in EU countries, www.venus-c.eu (deliverable D3.10).

- **Solve legal and compliance issues around contracts and SLAs**. These issues concern service providers' accountability, liability, compliance with data and privacy protection regulation, both national and cross-border. They are well known and are being addressed by the European Cloud strategy. However, the e-Science environment has specific requirements which need to be addressed (for example, because of the different balance of requirements in terms of open access to data and knowledge but also intellectual property issues). The emerging scenario characterized by a mix of public-private offerings requires a new skill set for public sector stakeholders, who must take responsibility for their legal and compliance issues and become able to deal with a new range of suppliers. This requires a new skills set for public sector bodies to fully understand what is possible in terms of purchasing and maintaining effective and efficient Cloud based services, dealing with a less well-known supply chain.
- **Promote Open Access to data**. The noble principle is to make the results of publicly funded research available to all interested researchers. But the real challenge is to find a financially sustainable model to implement this and still provide some return on research investments. A possible solution would be to provide value-added analytical and search services that make using data easier. **e-Infrastructures are a strategic resource that need open policies and open access to ensure that they remain the platforms for Innovation**. Within Europe we are too fragmented and provincial and that leads to complexity, cost and inefficiency we can no longer afford. Legal frameworks need to be consistent and uniform. Policies are needed to harmonize cross-border data flows, including clear regulations on personal and sensitive data.

Europe is too slow at taking advantage of the knowledge it creates. New and imaginative ways of using the highest quality knowledge in the world need to be found. **Not only is open government data opening up new possibilities in areas such as environmental data but there are many opportunities to develop new applications** as part of the drive to address societal challenges and accept the risks involved.

6.2. Recommendations

Given the multiple market and technical challenges already discussed, the European Commission has an important role to play to promote the development of pan-European cloud e-infrastructures, insuring the availability of cloud-based quality services to research and science, as well as the public sector.

There is one overarching general recommendation, building on the consensus opinion by stakeholders interviewed in this study and present at the final workshop:

The EC should promote and sustain the spontaneous movement towards the integration and federation of clouds at the EU level, avoiding the risk to develop top-down infrastructures totally dependent on public funding and unable to adapt to the multidimensional characteristics of demand. This will allow supporting the provision of cloud services for science and research across Europe, filling the gap between the actual offering of computing resources and the emerging, pent-up demand.

The EC should however make sure that the provision of cloud services for science and research covers the whole of the EU27, avoiding any potential "digital divide" between the large/small countries and between Big Science/Small and Medium Science (SMS) projects, insuring equal access to clouds for all researchers.

The development of pan-European cloud e-infrastructures and services will support the achievement of the main challenges of the Horizon 2020 Programme as follows:

- **Achieving excellence in science**: the availability of cloud services for science and research across the EU27 will fill the gap between the actual offering of computing resources and the emerging and pent-up demand by researchers across the EU.
- **Meeting social challenges**: the development of an ecosystem of value-added cloud services addressed to research and the public sector, driven by demand and based on

user-centric approaches, will help the emergence of e-Science-as-a-service business models, provide an open, scalable and flexible environment for large-scale collaboration between scientists and citizens in an open science/ science 2.0 perspective and for increasing collaboration between science and government.

- **Promoting Industrial Leadership and Competitive Industries**: this will be achieved by the promotion of interoperability through open standards for the provision of cloud services across the EU, responding to emerging research and industry demand needs; on the other hand by supporting the transition of European e-infrastructures and cloud service providers towards more sustainable business models, where public funding supports innovation and investments, rather than operational expenditures. This should include the development of the business case for EU cloud e-infrastructures through the achievement of economies of scale and scope.

In order to implement these recommendations, the key elements of an EC strategy in this area should be the following:

1. ***Use EC funding and initiatives to promote the integration and federation of clouds and enable the migration from e-infrastructures towards a European marketplace of connectivity and cloud services for e-Research***
2. ***Promote and extend the use of clouds across multiple scientific domains and the development of a cloud services ecosystem, in order to narrow the gap between the supply and user communities and overcome cultural and resistance barriers to cloud.***
3. ***Support the consistent, comprehensive and business-case oriented analysis of cloud computing costs compared to other computing resources, requiring full costs assessment in all public funded projects***
4. ***Promote the transformation of the business models and organizational structure of e-infrastructure providers***
5. ***Create the next-generation of cloud enthusiasts, supporting the change of mindsets the development of the new skill sets needed for new clouds services and e-infrastructures***
6. ***Promote the development of innovative SMEs developing cloud-based services, also leveraging spin-offs and start-ups***
7. ***PRACE to start offering cloud services with a pay-as-you-go model***

These recommendations are presented more in detail in the following paragraphs.

6.2.1. Promote the integration and federation of clouds

Use EC funding and initiatives to promote the integration and federation of clouds and enable the migration from e-infrastructures towards a European marketplace of connectivity and cloud services for e-Research.

All the ingredients needed to create a European marketplace are potentially already in place. Europe should aim to create real value-added rather than aim to imitate or compete head-on with large, mainly US-based cloud providers. Speed of execution is very important. A real market exchange should be built with interoperability of services. Further, EC funding should be directed towards the creation of a single market for e-research procurement. The following are the main actions that should be implemented:

- Promote and support cross-border interoperability of cloud products and services using open standards. Ensure that interoperability remains an important priority to broaden choice based on the best services (including quality-price ratios), capable of effectively satisfying the different needs of the research user-based.
- Review public procurement policies for scientific and research projects to enable the transition from CAPex to Opex. There are still constraints in the funding models of research which privilege the purchase of hardware and computing resources, while they do not allow using funding for subscribing external services such as cloud computing services.

- Continue to support efforts to federate resources and services across multiple cloud providers, with a compendium of virtual machines (platforms) to be available for user communities and link together on European e-infrastructures. From a practical standpoint this means that e-infrastructures, such as Nebula and EGI, need to change from providers to brokers, at least partially. This is a significant change from the point of view of cultural motivation and professional competences for people who are used to manage servers and networks.
- Small companies are more agile and often have better development skills than large enterprises, making them ideal candidates to push new services to market, as is being demonstrated by Helix Nebula. Future work with industry should ensure greater participation of SMEs with research organisations, offering an excellent training grounds for entrepreneurs as they are very well suited to start-ups.
- Monitor and support OpenStack[32] as an open collaboration model that is driving real market exchange, creating a large market with little investment. CERN has joined OpenStack because of its strong hold across the world.
- The exclusive use of open source remains debatable according to e-Science stakeholders. Funding should therefore consider solutions that lead to better cloud services for research rather than open source versus proprietary approach.

Impact if achieved: *an open, interoperable e-infrastructure that is driven by market demand and highly focused on end-user requirements with federated resources. Some requirements could be generic, e.g. interoperability, low-cost service for customers, high added value through aggregation/visualisation, metadata catalogue*.

6.2.2. Promote the use of clouds across multiple scientific domains

Promote and extend the use of clouds across multiple scientific domains, supporting the development of an ecosystem of value-added cloud services driven by demand, based on user-centric approaches, narrowing the gap between the user and supplier communities.

The supply and user communities are still too far apart. There needs to be a real and profound shift towards user-driven and objective driven approaches. Facilitate the wider usage of cloud computing from big data challenges to smaller scale computing with appropriate training and support with a strong user-centric focus and points of access.

- Use seed funding to support research groups and SMEs in adapting applications for the cloud and test new deployments, with clear business plans. Seed funds have already proved successful to support research groups and SMEs in adapting applications for the cloud and test new deployments. This type of small-scale funding should continue to be allocated with the aim of onboarding new users from different disciplines based on clear evaluation criteria and contractual terms.
- Do not fund or support proofs of concept that have little or no market value. Make clear asset definition and business models an integrated part of e-infrastructure funding, including new application and service developments and migration to federated resources.
- Allocate part of EC funding for e-infrastructures to community building around tailored services, which commercial providers like Amazon have not yet done. E-infrastructure funding could also be used to provide new services specifically suited to research infrastructures, which gather large amounts of multi-disciplinary data.
- Monitor and support the work of the Global NRENs CEO Forum (13 NRENs) as an important step towards understanding the progress on tackling cross-border barriers. Future funding should consider support of the resulting benefits and opportunities as a way to support wider usage of the cloud through new service provision.
- Possible short-term priorities, including small-scale funding include:

[32] www.openstack.org/

- o Services and tools (including AppStore with some pay-per-user services) for the earth science community, spanning many different research roles (e.g. in synergy with ESA-ESRIN).
- o Climate change modelling, including the feasibility of running existing grid applications in the cloud and expanding the current range of applications.
- o Human diseases leveraging occurrence and other relevant data, and the creation of virtual labs and tools for the research community.

Impact if achieved: *The development of this cloud ecosystem will contribute to narrow the gap between the supply and user communities and overcome cultural and resistance barriers to cloud. This change of perspective, from technology-push to user pull, from cloud to e-Science-as-a-service would impact the way requirements are gathered and how the e-infrastructure is engineered. This will also provide an open, scalable and flexible environment for large-scale collaboration between scientists and citizens in an open science/ science 2.0 perspective. This will also provide the basis for an increasing collaboration between science and government, for example supporting the development of computational applications and services for public health and environmental sciences, as well as for "Big Data" e-Government applications in national security, fraud and error detection in healthcare and pension systems.*

6.2.3. Require full costs assessments in Research projects

Support the consistent, comprehensive and business-case oriented analysis of cloud computing costs compared to other computing resources.

Some important, recent work was done in this area. However, more work is needed to ensure more complete analyses and clarity around hidden costs, e.g. costs not disclosed by third parties and the Total Cost of Ownership (TCO) of a campus cluster). Space and power are usually buried in the university's expense overall infrastructure bills. The systems administration for small clusters is often done by students, so it appears to be free of cost. Further, the main priority of researchers is to gain benefits from new capabilities (a key driver of cloud adoption. More needs to be done to ensure researchers in collaboration with their institution and/or funding agency make informed decisions for computing resources based on a cost-impact model[33].

This activity should be seen as complementary to the EC's on-going engagement with relevant standardization forums (notably the ITU, ETSI and IEC) and international initiatives (notably the GeSI/Carbon Trust/WBCSD) in the drive towards an energy-efficient, low-carbon economy)[34].

This should include the following actions:

- **Make cost calculation a requirement for all public-funded research at national and EU level**. Costs should be analyzed more widely and on a regular basis given that the costs of cloud computing are constantly changing.
- **Ensure more transparency around the costs of all types of science, including supercomputing**. This should help to identify and measure the consequences of cost avoidance, cost sharing, and the trade-offs of the transition from CAPEX to OPEX.
- **Develop cloud usage business cases, also leveraging the NREN's experience**. Examples of ways in which NRENs can save money and generate revenue are important in fostering mainstream adoption of cloud as a new service paradigm. Success stories need to be shared more widely, highlighting different possible approaches with champions playing a leading role. Closer cooperation between e-infrastructure providers and NRENs would help identify potential new service areas that can leverage the expertise gained through e-infrastructures.

---

[33] For example, the Impact Calculator developed by JISC in the UK, http://www.jisc.ac.uk/news/stories/2011/02/impactcalculator.aspx

[34] http://ec.europa.eu/information_society/activities/cloudcomputing/docs/com/swd_com_cloud.pdf, p. 6.

Impact if achieved: *changing mindsets and creating greater awareness of costs will lead to a more responsible and smarter use of resources. It will also encourage funding agencies to define better funding allocation and academic/research institutions to measure the socio-economic impact of their research activities beyond the availability of new capabilities. National governments and the EC could also adjust their funding accordingly and justify evidence-based funding decisions*.

6.2.4. Promote new business models

Promote the transformation of the business models and organizational structure of e-infrastructure providers by:

- Supporting the change of organizational and management models towards user-centered business strategies
- Encouraging them to focus on new business models where sustainability of e-infrastructures does not depend exclusively on public funding
- Supporting them to optimize the potential for broad adoption throughout e-Science by focusing more on user needs and satisfaction
- Strongly encouraging them to invest in innovation, rather than trailing after the commercial sector (also using innovation prizes and other community oriented practices)

6.2.5. Create the next generation of cloud enthusiasts

Support the change of mindsets and the development of the skills needed for the provision of sustainable cloud services and e-infrastructures.

The issue is not just open-source versus commercial products and services, but also how to get smart people hooked up in both the commercial and research science settings to best practices, appropriate standards, and best methods for software development. This is also linked to changing mindsets as part of the drive to foster excellence in science through the wider adoption of cloud in the wider distributed computing ecosystem.

- It is crucial that this be conducted in a way that optimizes the potential for broad adoption throughout e-Science and enterprise, rather than producing isolated narrow towers of privileged funding or surrendering the reins of innovation to the pure commercial sector.
- **Inject funds into new skills development for the use of cloud** in new fields/application areas as well as the development of business skills.
- **Develop new business-oriented and user-centered skills**. Understanding the market is fundamental. Ensuring that users are encouraged to become part of the design and decision processes are key. This is a culture change from the more traditional “build it and they will come” approach.
- **Launch Innovation Prizes**. Innovation is increasingly a public opportunity. Innovation prizes for public contributions from research are an effective means and should be more widely pursued.
- **Encourage champions skilled at communicating benefits** to the wider general public and demonstrate justification of tax payers' money in supporting research in an “open” spirit.
- **Ensure developers are closer to the market** and end-user communities.

Impact if achieved: *European e-infrastructures are a place where developers, providers and end-users work together more collaboratively for the common good. Closer synergies will ensure that new services meet real-world needs with acceptable business models. Funding will be channeled to areas where it is most needed to drive higher level innovative research. Greater awareness of the general public will mean that support of future research is increasingly evidence based.*

6.2.6. Promote the development of innovative SMEs in cloud services

Promote the development of innovative SMEs developing cloud-based services, also leveraging spin-offs and start-ups capability to develop the ecosystem of new cloud-based value added services.

A series of pilot research projects in the e-infrastructures arena has demonstrated some of the benefits resulting from industry and SMEs participation to the development of innovative cloud-based applications for the public sector. This includes for example Microsoft participation to VENUS-C; ATOS, T-Systems and CloudSigma participation in Helix Nebula); the potential of CESGA's Cloud Radiotherapy planning for Spanish hospitals[35]. Further, cloud has already demonstrated its value to commercial settings through small-scale funding support. The development of new applications and services can help build an ecosystem in areas like health services, energy supply and agriculture, where Europe already has a leading position. The time is now ripe for European e-infrastructure stakeholders to strengthen synergies with industry and the public sector, including new ways to use open data and develop new services through a two-tier approach.

- **Identify and support new cloud-enabled research and services** targeting new value-added public services (e.g. education, health, energy, agriculture and food supply). The focus should be on new specialized public services, rather than on core administration services, where killer applications already have a strong market hold.
- **Provide funding for cloud projects in partnership with SMEs**. Research organizations are excellent training grounds for entrepreneurs because they are very well matched to the start-up culture. Funded projects should include partnerships with SMEs that are more agile in providing solutions and plans to identify new potential start-ups supported by business models for new services spinning out of the e-infrastructure project
- **Promote special business support schemes** (including industry support schemes) and/or CSAs in Horizon 2020 to involve SMEs

Impact if achieved: *European e-infrastructures are a place where funding is injected to support the development of new value-added services addressing societal challenges. More effective knowledge transfer between research and enterprise leads to new public-private partnerships, new spin-outs leveraging innovation, which also reduces the need for continuous funding.*

6.2.7. PRACE to start offering cloud services with a pay-as-you-go model

Continue strong support for PRACE and promote the addition of cloud capabilities and at least partially "pay-as-you-go" access models to HPC centers, to extend their utilization and best exploit their resources.

- We believe that users of the PRACE network would benefit from having cloud capabilities added, and from having a portion of the compute cycles on PRACE supercomputers made available to designated projects via cloud computing. The goal would be to enable approved projects to be assigned not only a fixed number of CPU-hours for their projects as now happens, but also the possibility of additional time, without waiting, if and when the need arises. Adding this on-demand, elastic capability should help the PRACE network and its users to respond more quickly to unexpected opportunities and other developments that can occur in course of research projects. Adding cloud capabilities should also give the PRACE network more flexibility to address its users' future requirements and to align with other research networks within and

[35] http://www.venus-c.eu/Content/UserScenarios.aspx?id=485ad817-364c-42d6-8f6d-103978248ffe.

beyond Europe. (PRACE has already begun, for instance, to collaborate with the XSEDE initiative in the United States to support a co-hosted HPC Summer School.)

Impact if achieved: *This change would improve the efficiency and effectiveness of the use of PRACE resources, contribute to the full availability of cloud services across Europe, and contribute to the achievement of excellence in science*.

MAIN REFERENCES

- IDC, "A strategic Agenda for European Leadership in Supercomputing: HPC 2020" 2011, on behalf of the European Commission http://www.hpcuserforum.com/eu/downloads/SR03S10.15.2010.pdf
- IDC "Quantitative Estimates of the Demand for Cloud Computing in Europe and the Likely Barriers to Take-up", 2012, on behalf of DG CONNECT, http://ec.europa.eu/information_society/activities/cloudcomputing/docs/ quantitative_estimates.pdf
- IDC, Best Practices: Some Confidence and Some Hesitation — Surprising Results from Our Government Cloud and Enterprise Architecture Survey , Shawn P. McCarthy, July, 2012 Document # GI236031
- "Advances in Clouds - Research in Future Cloud Computing" L. Schubert, K. Jeffery & Theilmann, W. (eds.) (2012), Cordis (Online), BE. European Commission. http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf
- "Cloud Computing for research and science: a holistic overview, policy, and recommendations", final version 30 October 2012, e-IRG task force paper, http://www.e-irg.eu/publications/e-irg-task-force-reports.html
- "Cloud Computing Use cases" White Paper, version 4 - SLA, by the Cloud Computing Use Cases Group, 2010
- "Computing e-Infrastructure cost calculation at national and European level", D2.2, by AUEB-RC. July 2012, e-Fiscal project (Financial Study for Sustainable Computing e-Infrastructures)
- "Cost analysis of cloud computing for research", 22-02-2012, Final report to ESPRC and JISC, http://www.jisc.ac.uk/media/documents/programmes/ research_infrastructure/costcloudresearch.pdf
- "Dante Strategy 2012-2015: More for all" http://www.dante.net/ server/show/nav.2549
- "European Grid Infrastructure Annual Report, 2011-2012", http:// www.egi.eu/export/sites/egi/news-and-media/EGI_AnnualReport2012.pdf
- "Future Sustainability Strategies", VENUS-C Final Deliverable 3-10, 10-7-2010, http://www.venus-c.eu/Content/Publications.aspx?id=e519a779-3df8-4adf-82a9-397f03659a69
- "Inspiring Future e-Infrastructures in Europe and Beyond - Proposal for a new e-IRG strategy", e-IRG, 30 June 2012, http://www.e-irg.eu/publications/e-irg-strategy.html
- "Knowledge without borders", Géant 2020 as the European Communications Commons, report of the Géant experts group, October 2011
- "Magellan Final Report", December 2011 30, (http://science.energy.gov /~/media/ascr/pdf/program-documents/docs/ Magellan_Final_Report.pdf)
- "Moving to the Cloud" White Paper, version 1, February 2011, by the Cloud Computing Use Cases Group, http://cloudusecases.org/Moving_to_the_ Cloud.pdf
- "16-Point Action plan for Cloud Computing in Europe" presented by EuroCloud, http://www.eurocloud.org/
- "Report on Support for Cloud Computing In response to America Competes Reauthorization Act of 2010", NSF (National Science Foundation), Section 524, report to Congress, USA
- “Seeking new horizons: EGI’s role in 2020” strategic plan by Steven Newhouse and EGI-InSpire team, April 2012.
- "SIENA European Roadmap on Grid and Cloud Standards for e-Science and Beyond" http://www.sienainitiative.eu/Repository/Filescaricati/ 8ee3587a-f255-4e5c-aed4-9c2dc7b626f6.pdf
- "Strategic Plan for a Scientific Cloud Computing infrastructure for Europe" CERN-ESA, open 08082011, http://www.helix-nebula.eu/ index.php/ uploads/ file/81/11/CERN-OPEN-2011-036.pdf.html
- "The adoption of cloud services", September 2012, ASPIRE (A Study on the Prospects of the Internet for Research and Education). Topic report version 4, http://www.terena.org/activities/aspire/

- "The strategy of the Commons: modeling the annual cost of successful ICT services for European Research", Matti Heikkurinen, Sandra Cohen, Fotis Karagiannis, Kashif Ikbal, Eoin Brazil, Sergio Andreozzi, e-Fiscal project paper presented to e-Challenges e.2012 conference
- "Towards a Global Cloud ecosystem - draft Roadmap Recommendations" by the Cloud Expert Group, Report by Keith G. Jeffery, Lutz Schubert, Burkhard Neidecker-Lutz, 11.11.2012 V0.2

GLOSSARY

CCUCDG – Cloud Computing Use Cases Group	The Cloud Computing Use Cases Group is an open web community of over 1400 members started from the Open Cloud Manifesto environment but now including many more actors. (http://cloudusecases.org.) It has published widely quoted reports on use cases of clouds.
DAE - Digital Agenda for Europe	One of the seven flagship initiatives of the Europe 2020 Strategy set to deliver sustainable economic and social benefits from a digital single market based on fast and ultra-fast Internet and interoperable applications by 2020.
DANTE (Delivery of Advanced Network Technology to Europe)	DANTE's purpose is to plan, build and operate pan-European research networks. It was set up, and is owned, by a group of National Research and Education Networks (NRENs). It was established in 1993 and has since played a pivotal role in four consecutive generations of pan-European research network: EuropaNET, TEN-34, TEN-155 and now GÉANT. It is a limited liability company and a "Not for Profit" organisation based in Cambridge. DANTE currently has an annual turnover of approximately 50 million Euros, of which around half comes from European Commission project funding. It has 'Research Association' status in the UK. http://www.dante.net/server/show/nav.2549.
DMTF – Distributed Management Task Force	DMTF enables more effective management of millions of IT systems worldwide by bringing the IT industry together to collaborate on the development, validation and promotion of systems management standards. The group spans the industry with 160 member companies and organizations, and more than 4,000 active participants crossing 43 countries. The DMTF board of directors is led by 17 innovative, industry-leading technology companies. They include Advanced Micro Devices (AMD); Broadcom Corporation; CA, Inc.; Cisco; Citrix Systems, Inc.; EMC; Fujitsu; HP; Huawei; IBM; Intel Corporation; Microsoft Corporation; NetApp; Oracle; RedHat; SunGard and VMware, Inc.
EDGeS	FP7 Project targeting user communities that require large computing power not available or accessible in current scientific e-Infrastructures. In order to support the specific needs of these scientific and other communities the consortium will interconnect the largest European Service Grid infrastructure (EGEE) with existing Desktop Grid (DG) systems in a strong partnership with the EGEE consortium. Service Grids (SG) are more flexible and can accommodate a broader variety of applications than Desktop Grids, however, their setup and maintenance require more efforts, highly skilled IT specialist, and dedicated resources. http://www.edges-grid.eu/web/edges/4
EGI European Grid Initiative	European Grid Initiative (http://www.egi.eu/) A federation of shared computing, storage and data resources from national and intergovernmental resource providers that delivers sustainable, integrated and secure distributed computing services to European researchers and their international partners.
e-IRG e-Infrastructures Reflection Group	The e-Infrastructure Reflection Group was founded to define and recommend best practices for the pan-European electronic infrastructure efforts. It consists of official government delegates from all the EU countries. The e-IRG produces white papers, roadmaps and recommendations, and analyses the future foundations of the European Knowledge Society
EIRO - European	A legal organisation and member of the EIRO Forum that has extensive expertise in the areas of basic research and the management of large,

Intergovernmental Research Organisation	international infrastructures, facilities and research programmes.
ERA - European Research Area	The area that brings together all of the European Union's (EU) resources to better coordinate research and innovation activities at the level of both the Member States and the Union. The area also aims to achieve a major ambition of the EU: to arrive at a truly common research policy.
ERIC - European Research Infrastructure Consortium	A European legal instrument adopted by the Council in 2009 to facilitate the establishment and operation of European research infrastructures on a non-economic basis. ERIC endows research infrastructures with a legal personality recognised in all Member States. To date, two European research infrastructures have been established as ERIC: SHARE and CLARIN. Six other applications have been received at the Commission, and six more are expected to be received by end-2012 or in 2013.
ESFRI - European Strategy Forum on Research Infrastructures	Created in 2002 by EU Member States and the European Commission to develop the scientific integration of Europe, and to strengthen its international outreach, ESFRI has become an increasingly important forum to advise ministries and funding agencies. Some associated countries (Albania, Croatia, Iceland, Israel, Liechtenstein, Montenegro, Norway, Serbia, Switzerland and Turkey) also participate in ESFRI.
GÉANT	GÉANT is the pan-European research network including the 32 European NREN, the Trans-European Research and Education Networking Association (TERENA), plus an additional four Associate NRENs include: Austria (ACOnet), Belgium BELnet), Bulgaria (BREN), Croatia (CARNet), Cyprus (CYNET), Czech Republic (CESNET), Estonia (EENet), France (RENATER), Germany (DFN), Greece (GRNET), Hungary (NIIF), Ireland (HEAnet), Israel (IUCC), Italy (GARR), Latvia (SigmaNet), Lithuania (LITNET), Luxembourg (RESTENA), Macedonia (MARNet), Malta (University of Malta), Montenegro (MRnet), Netherlands (SURFnet), Nordic region (includes Sweden, Finland, Denmark, Norway and Iceland) (NORDUnet), Poland (PSNC), Portugal (FCCN), Romania (RoEduNet), Serbia (AMRES), Slovakia (SANET), Slovenia (ARNES), Spain (RedIris), Switzerland (SWITCH), Turkey (ULAKBIM), and the UK (JANET). GÉANT is co-funded by the European Commission under the 7th Framework Programme. GÉANT is one of the world's largest research and education networks, comprising 25 European POPs, 12,000km of dark fibre on 18 routes, and 50,000km network infrastructure on 44 routes. GÉANT also has an incredibly diverse international footprint, serving 40 million end-users, in more than 8,000 institutions, across 40 European countries. Network availability is proven "better than carrier class", with an availability rate up to 99.999 %.
Helix Nebula (HN)	Helix Nebula is a new, pioneering partnership between leading IT providers and three of Europe's biggest research centres, CERN, EMBL and ESA, charting a course towards sustainable cloud services for the research communities - the Science Cloud. http://helix-nebula.eu/
HPC	High Performance Computing
HTC	High Throughput Computing
NGI - National Grid Infrastructure	The national federation resources, which is coordinated through a single point of contact that has an exclusive mandate to represent its national grid community in all matters falling within the scope of EGI.
Open Cloud	A manifesto launched by a group of industry actors to promote the open

Manifesto	cloud principles. http://www.opencloudmanifesto.org/
PRACE - Partnership for Advanced computing in Europe	PRACE is a European strategic approach to high-performance computing. It concentrates resources distributed in a limited number of world-class top-tier centres in a single infrastructure, forming a scientific computing network. PRACE provides access to distributed pan-European world class high performance computing and data management resources and services located in Germany, France, Spain and Italy. http://www.prace-ri.eu/HPC-access)
SHARE-ERIC - Survey of Health, Ageing and Retirement in Europe	SHARE-ERIC is a data infrastructure for the socio-economic analysis of on-going changes due to population ageing. SHARE-ERIC is the upgrade into a long-term research infrastructure of a multidisciplinary and cross-national database of micro-data of about 45,000 Europeans aged 50 or over. Construction cost: €23 million
The OpenStack Foundation http://www.openstack.org/ foundation/	The OpenStack Foundation promotes the development, distribution and adoption of the OpenStack cloud operating system. As the independent home for OpenStack, the Foundation has already attracted more than 5,600 individual members from 87 countries and 850 different organizations, secured more than $10 million in funding and is ready to fulfill the OpenStack mission of becoming the ubiquitous cloud computing platform. The goal of the OpenStack Foundation is to serve developers, users, and the entire ecosystem by providing a set of shared resources to grow the footprint of public and private OpenStack clouds, enable technology vendors targeting the platform and assist developers in producing the best cloud software in the industry.
VRC - Virtual Research Community	A group of large-scale research collaborations, or a number of separate Virtual Organisations (VOs) grouped according to research domain or computational technique. The group shares information and experience in achieving their goals through the usage of an e-Infrastructure (e.g., best practices, applications, training material).

FINAL WORKSHOP REPORT

Scenario & Key Findings

This report presents the main results of an open, one-day workshop organised by IDC and Trust IT on behalf of the DG CONNECT, which took place on November 26th 2012 at the European Commission premises in Avenue de Beaulieu, Brussels. The workshop gathered over 30 experts of e-Science and e-Government stakeholders to discuss the findings in the study "Cloud for science and public authorities" (SMART 2011/0055), the aim of which is to contribute to develop the EC Cloud Computing Strategy for science and public authorities.

The workshop presentations can be downloaded from the website
http://www.idcitalia.com/events/eucloud2012

Summary of results

The workshop presented the key findings of the study about the current and prospective development of cloud computing infrastructures for e-Science and e-Government in Europe and in the world. A key report of the study is a quali-quantitative mapping of cloud policies and initiatives in EU27 plus 13 International country profiles. According to this analysis, government cloud policies tend to be more prescriptive and top-down, while those for science allow more freedom of initiative to university and research centres. Overall, there are far more policies than actual investments, so the utilization of cloud services is still in the early development phase.

A number of government and science stakeholders presented their case studies of cloud adoption, reviewing benefits and highlighting the potential (few) synergies between government and science, Very different use profiles emerged clearly, leading to the observations that synergies so far are difficult to find, particularly concerning shared infrastructures. However, leading initiatives in both e-Government and e-Science are placing greater emphasis on user-centric and joined-up services as an outward looking benefit. Sharing of services may therefore become more likely.

A constructive discussion followed the presentation by IDC of potential demand and provisioning scenarios of pan-European cloud infrastructures for science. The need for ***Integration & Federation of clouds*** was among the key themes recurring throughout the day. Everybody agreed that there is considerable new demand emerging for e-Science services, which could be satisfied by cloud services and infrastructures including, not only physical resources, but also data sources, services leading right the way through to computation. However, given the current fragmented scenario, there is a need for a coherent set of policies, standards and services supporting the development of a pan-European infrastructure to achieve some economies of scale and scope. Without such an approach, much of the potential demand may risk remaining unsatisfied. There is, however, still considerable uncertainty about the best business models supporting these infrastructures.

The **legal and compliance issues around Contracts and SLAs** require a new skill set for public sector bodies to fully understand what is possible in terms of purchasing and maintaining effective and efficient Cloud based services. A key requirement of public stakeholders is to fully understand who the supplier is and what their supply chain is. Data controller & data processing agreements are now formal legal documents already in place in Germany and in Italy, which must be compiled to use cloud services, just to name two examples.

Today we are seeing that the supply and user **communities are still too far apart**. There is a strong tendency of companies and corporate entities involved in this area to regard the large-scale computational, network and storage needs of science only as a market, instead of seeing the potential for two-way communication that can be harnessed to produce rapid acceleration of the potential on both sides for product development.

The issue is not just open-source versus commercial products and services, but also how to get smart people hooked up in both the commercial and research science settings to the **best practices, best standards, and best methods for software development**. It is crucial that this be conducted in a way that optimizes the potential for broad adoption throughout e-Science and industry, rather than producing isolated narrow towers of privileged funding or surrendering the reins of innovation to the pure commercial sector.

The work of EGI Federated Cloud Task Force and the Helix Nebula project (presented by their representatives) is showing that a hybrid (public/private) cloud is attractive to some research communities. "Private" includes not only "in-house" but also GÉANT via DANTE/EGI/PRACE. In contrast, ESFRI projects have generally been slow to leverage distributed computing but could represent a large user community. The workshop discussion identified a number of ways in which to engage these communities. For this to happen, a number of participants see the need for closer interaction between the supply and user side with a focus on user-centric approaches.

Another result of the discussion was the confirmation, by several stakeholders, that a wide range of cloud computing applications including those of commercial public clouds are valuable for e-Science, especially for "Small Sciences" (though there are of course limits for heavy-duty e-Science applications).

There are European cloud activities already demonstrating that interoperability, portability and reversibility issues, including SLA management, can be solved in the cloud (such as the cloud federation research by CompatibleOne, OpenNebula & OpenStack in the Enterprise domain and by both EGI and Helix Nebula in the e-Science domain). These issues are not considered technical problems and can be solved today. This approach could potentially be extended to support the needs of e-Government. Therefore, pursuing this approach could lead to the development of a compliance service, able to provide a certification (a badge of authority?) to both providers and user communities on the types of available and supported services.

Case Studies Overview & Main Findings

The adoption of cloud services in Europe in the research domain is growing, but includes many examples of users' independent initiatives rather than framed within national strategies. There are similar trends in government, with a mix of bottom-up adoption patterns and top-down national strategies slowly emerging. It is too early to fully understand the economic benefits and cost savings of these initiatives, but they provide interesting evidence about provisioning scenarios and use cases. The following ones were presented at the workshop.

1. **Raincloud (Austria) – hybrid cloud for weather forecasting**

- University pilot project testing clouds for research and for spin-off services, as an alternative to grid & supercomputing. It is an example of small computational needs used to deliver the Weather Forecast-as-a-Service to regional governments. The case study demonstrates one of the current "sweet spots" of the cloud and appears to be a more effective business model than the grids.
- Cost savings can be redirected by the university to investments in real science & jobs for research. The ultimate goal is to define a business model for the institution's use of the cloud across different departments.
- The scenario demonstrates three use cases:
  - Scientists developing applications used to make the weather forecast calculations.
  - Scientists using the applications with weather and climate data from the local Institute of Meteorology, which cannot be outsourced (scientific use).
  - Local government (as customer) tapping into the resources provided by the university for its daily alerts (e.g. Avalanche Bulletin).

A similar, but not identical case is found in Greece. Through the VENUS-C project, the National Disaster Laboratory at the University of Aegean has developed fire forecasting, simulation and management service, Wildfire, which is able to predict different scenarios including wind direction. Wildfire, which is used mainly during the summer months, runs in a hybrid cloud (MS Azure and OpenNebula) and is being deployed by national fire services. Its cost-effectiveness (around €600 for the summer period) is also appealing for a small university.

2. **SURFconext (Netherlands) – clouds for research & education**

This case study demonstrates strategic "build and buy" approaches by an EU NREN, encouraging a discussion on «buy instead of build» IaaS & SaaS.

- **Buy**: purchasing at preferential rates generic, successful "killer applications" from commercial providers with potential uptake by the entire Dutch academic population (1 million users). The service negotiation extends to students and teaching staff as well as researchers.
- **Build**: SURFconext identifies opportunities to develop new applications and services, for example water resource management and university hospitals.
- **Tender**: tendering of solutions and services to SMEs after deployment.
- **Cooperation & training support**: negotiating fees for training support to University of Manchester to extend the deployment of OPENconext, which SURFconext has developed.

An important lesson on the best way to drive adoption is the creation of demos and showcases to facilitate understanding of potential benefits. Actors like SURFconext can play an important champion role for this. Examples of other NREN services include the HEAnet (Ireland) business model for storage; EduStorage charges €275 per TB per year with billing aggregated per client, governed by HEAnet's Acceptable Usage and available to Edugate federation members. Janet (UK) also provides a brokerage cloud service by directly negotiating preferential rates with service providers.

3. **UK G-Cloud – Cloud AppStore**.

The UK cloud strategy is part of a drive to move away from expensive bespoke software and dominance of a small number of providers in public sector procurement.

- Reducing the financial burden is a key driver, purchasing commodity wherever possible.
- The Cloud AppStore (launched late February 2012) is a beta version. The aim is to work towards greater agility, ensuring it is user-friendly and enables "easy in, easy out" options.
- The AppStore is proving to be an important driver of creativity, innovation and competitiveness.
- One of the biggest barriers is tackling the "cultural change" in public procurement. The Cabinet Office team is therefore working closely with the buyers, providing guidance to them, seeding and creating champions of authorities willing to buy and pushing the boundaries.
- Data remains an issue. The AppStore contains mainly data that is already in the public domain, whereas more restrictive approaches are needed for national security data, HR data and financial data.

4. **goBerlin (Germany) – trusted cloud for SMEs & public authorities**

- Germany's data protection regulations are among the most restrictive in Europe. Restrictions are higher at national and regional level, whereas municipalities are using services such as MS email.
- This case study is part of national strategy to foster mainstream adoption of cloud in the public sector by focusing on major barriers as security, trust, privacy ("Trusted Cloud" Technology Programme, considered to be among the leading examples[36]). The aim is also to develop trust certification and overcome silos in government.
- Developing innovative applications for citizens, industry and administration as a single-access point, including life-event apps for re-locating.

---

[36] Commission Staff Working Document, Unleashing the Potential of Cloud Computing in Europe, p.10

- A Marketplace run on a joint datacentre in Berlin and offering Iaas, Paas, Saas (for life situations). The most important challenges to face concern security issues, how to build trust, compliance and data protection.
- Business models include tendering the cloud infrastructure, selling apps and new services.

5. **Cheltenham Borough Council (UK) – shared services**

- Implemented a cloud-based shared document management service in cooperation with 6 other local administrations in the area. It represents a move from shared but separate systems towards integrated systems, leveraging a private cloud solution offering scalability so as to increase the number of services or partners moving forward.
- Good example of the bottom-up adoption of cloud services by governments most frequently found across Europe.
- Budget cuts for operations (around 22-23%) are a key driver towards cloud adoption, as well as an opportunity to make back-office efficiencies to maintain public services, including housing, environmental health and refuse collection. The use of cloud is currently expected to generate an RoI of £620,000 in 4 years.
- Each council retains ownership of the data.
- The approach falls short of mandating specific solutions.
- One of the service providers, Huddle, offers an "adoption success team" to foster buy-in and facilitate uptake.

6. **Emilia Romagna – Regional Public Administration**

- Regional government has created a pool of 412 regional PA stakeholders supporting the governance of regional ICT strategies and policies. Innovation is also driven through cooperation with universities and private enterprises. A regional test bed is used to develop and test new applications.
- Cloud management platform will be used for: self-service provisioning portal; IT service portal; infrastructure authority; charge back system; capacity/IT resources management; life-cycle management; orchestrator/IT process automation; external cloud connector.
- Delivery models: vendor-operated private managed cloud; vendor owned and operated hosted private cloud; shared cloud services with scalability, pay-as-you-go, support and network options.
- Business models: defining public-private investment policies compliant with Italian and EU laws; using regional government properties/assets (e.g. building, public areas, broadband Lepida network); part of Lepida assets used by private partners to run their own business; private partner finances and/or constructs cloud data centre in public areas enabling Lepida to run services towards its shareholders; defining RoI for Lepida in 5-7 years.
- Confirms drivers for Public Administration are typically: cost savings; better scalability; complexity reduction; more flexibility; efficient use of under-utilized capacity or IT investments; advanced technology; improved security; Green IT; switch from CapEx to OpEx; more (core) business focus.

Workshop – EC Policy

The EC introduced the workshop by presenting the path along which the EU Commission is stimulating Cloud adoption at the EU Level. The European strategy highlighted by Head of Unit e-Infrastructure (Unit C1) DG CONNECT, **Konstantinos Glinos**, introduces the announcement made by the European Commission last September of the Strategy for Cloud Computing.

Among the important issues that merit immediate attention is the understanding behind the term of a science cloud for Europe and the need for one. What is "European"? Is it enough to have a national set of services? Should we rather focus on interoperability?

The current study relates very closely to Action 53 of the Digital Agenda and it is important to shape the policy for H2020 over the next few months as an imminent measure. A key objective lies in understanding what already exists in government and science.

Potential demand and provisioning scenarios

Massimiliano Claps, **IDC** gave a presentation on "Cloud-based e-Science: demand scenarios, provisioning scenarios, and recommendations for an EU strategy". The starting point is the definition of a taxonomy for the demand of e-Science, which is articulated in 3 main user communities: the core scientific community, with the highest needs for computing power and HPC services; an extended research and higher education community (including the social sciences), driving the increasing demand of computing power; and the open and emerging science 2.0 community involving also the amateur, non professional scientists. These user communities have potentially different needs and it is not easy to estimate their potential demand. To do so, IDC has leveraged its in-depth knowledge of the HPC environment serving the core scientific community and extrapolated the pent-up demand observed in this world (where approximately ¾ of requests for computing time are turned down). This and a few other assumptions resulted in two alternative scenarios of potential demand projected to 2016.

The demand estimates were very much discussed, mainly because they were too heavily based on HPC-environment assumptions (even if IDC's definition of HPC is rather broad and includes a wide share of the scientific computing market). There were several suggestions that the model estimates should be deeply revised and integrated, taking into account other data sources particularly on HTC (high throughput computing). Action to gain relevant feedback in this direction has since been undertaken.

The provisioning scenarios designed by IDC highlighted the potential dynamics between public cloud infrastructures for e-Science, private clouds and community clouds. They were also discussed, but it was agreed that they did reflect many of the emerging market issues, even if it was requested to better clarify the definitions of HPC and HTC grids and their reciprocal scope. The presentation led to the discussion about the scope of the community cloud versus public cloud models and how they could satisfy the emerging demand. According to IDC, though, these are complementary, not alternative provisioning scenarios, and the issues to be solved concern their reciprocal balance, the funding sources and the business and organizational models.

eFiscal initial findings

The business model aspects were taken up again by the presentation of the e-Fiscal project initial findings.

- Initial findings show that European HPC/HTC e-Infrastructures are cost-effective where they have relatively high utilisation rates and depreciation rates. Operational expenditures are dominating the costs (70%), while personnel costs are roughly half of the total costs.
- It is important to note that comparisons are not at all easy and use case views need to complement the overall picture (i.e. specific applications examined in both environments).
    - A key recommendation is to evaluate public cloud challenges in-field during H2020, including central procurement of commercial resources, allocation and monitoring of resources (metering, effective use etc.), legal compliance, interoperability, prevention of vendor lock-in, governance and user satisfaction.
- Expected trend is a combined approach to distributed computing with both grid and cloud, evolving the current EGI environment into a virtualised service-oriented computing e-infrastructure with a public cloud mainly for smaller computing requirements ("the long tail") with a centralised pool of resources procured centrally from commercial public clouds.
- A centralised pool of cloud resources offers the following advantages:

- o Better economies of scale/prices at EU level, co-funded by EC.
- o Better stimulation of a cloud market for research at the EU level.
- o Better stimulation of interoperable, standard and recoverable (no SPoF) solutions from multiple vendors. Such an approach could help Europe advance towards interoperability and ultimately move towards standard cloud stacks and interfaces.
- o Better central control of legal, financial and policy issues.

Steven Newhouse presented the EGI perspectives

It is important to take on board the societal challenges that we are facing. The role of scientific excellence lies chiefly in analysing and solving complex problems, as opposed to eGovernment key focus areas such as the hosting of web services and consolidation of data centres. The objective of the federated cloud technologies is to avoid overlap in market offers and ensure researchers have access to uniform computing resources.

- There is much potential for the development of applications. An important policy measure would be to support the development of the right skill sets needed to drive adoption in the shift away from "investments" to "services".
- A number of characteristics specific to grid were also highlighted, for example, grid is synonymous with the secure sharing of resources, while HPC, especially in Europe has little to do with secure shared resources.
- Roles in grid: NGIs are both service providers and consolidator; researchers can also be buyer and supplier; PhD students often provide technical developments and support on a voluntary basis, which distorts the costs involved. There is both technological and financial overlapping.

Bob Jones presented the Helix Nebula perspectives

Helix Nebula is not building cloud computing services itself as these are provided by members of the partnership such as Atos, Cloud Sigma and T-System. Nebula aims at supporting IT requirements of European researchers and scientists.

- One of the challenges is that sometimes companies do not have a clear idea of how to sell a cloud service in a comparable way because they lack a unit of measure on which to apply a basic price level.
- Making specific requests for specific types of software increases risks and costs for businesses acting as service providers.
- Move beyond the distinction between private and public cloud to focus on a newer distinction between publicly owned cloud and commercially offered cloud.
- Helix Nebula has signed an agreement with DANTE to provide IP connectivity for the pilot phase for free. DANTE will also enable the testing of the hybrid model.
- In Helix Nebula, the small companies are driving developments also on behalf of the larger companies because they are more agile as an organisation and often have better development skills.
- Helix Nebula is not expected to use a single business model.
- Federated e-Identity is very important and has to allow for a third party.

Bob Jones concludes quoting the agreement signed with Dante (for the trial phase) and inviting all participants to attend the event Helix Nebula is fixing for next 16th January in which they will present some use cases.There are two possible options for collaboration between GÉANT and Helix Nebula:

- **Option 1**: DANTE becomes a member of Helix Nebula and provides IP connectivity for the pilot phase for free;
- **Option 2**: DANTE receives money for the connectivity services offered.

It was agreed that option 1 will be adopted in order to test the deployment of the Helix nebula flagship use cases.

Interactive Discussion – Main points

The participants discussed the best way to engage the developer community, including both open source and commercial products, in activities that highlight the strengths and features of new standards in ways that will promote both coherence and adoption.

- **Scientific Excellence** is a key objective of EC policy, which we must not lose sight of. Peer-reviewed approaches play an important role in identifying and supporting high-level scientific excellence. However, the peer-reviewed model cannot be the only way to provide access to computing resources for research, or there is a risk that it becomes a bottleneck. Further, peer-reviewed research tends to focus on theoretical research and complex simulations, whereas the use of the cloud is demonstrating promise for research in several disciplines, faster time to publish results and potential for spin-out services and new businesses with sustainable business models, as outlined below.
- **New emerging business models**: there are many emerging examples of smaller scale needs that can drive the development of new services and spin-outs.
    - Engineering Hub (e-HUB)[37] is a new business proposition that has emerged from three VENUS-C use cases from architecture and civil engineering: 3D visualisation rendering (GreenPrefab, an Italian start-up); structural analysis of buildings (Technical University of Valencia and a new start-up company) and eco-efficiency analysis of buildings (Danish Royal Academy and E3Lab, Milan, San Francisco and currently expanding to the UK). Various business models are being developed, including pay-per-use for the architectural design of prefab buildings. It is also important to note that the Royal Danish Academy is a VENUS-C pilot that has received seed funds (< €25,000) to adapt and deploy JE+ software through an Open Call process.
- The value of adopting a **user-centric approach** has been demonstrated by VENUS-C. A key point raised during the workshop was on re-directing the strategy towards users and their objectives with the possible creation of a “data market” in Europe both for scientific and social science applications. This approach sees the technology (grid, cloud, combined approaches) mainly as a tool that glues together data and user communities across diverse fields. The aim is to generate an environment where knowledge can be extracted more easily and therefore create value more rapidly. One model proposed is to create an application store-like service charging a small fee payable to the data/software provider with a high flux of transactions plus a consultation fee. The requirements on the e-infrastructure should be designed to meet these objectives:
    - Some requirements could be generic, e.g. interoperability, low-cost service for customers, high added value through aggregation/visualisation, metadata catalogue.
    - Some could be community specific.
    - This change of perspective, from technology-push to user pull, from cloud to e-Science-as-a-service would impact the way requirements are gathered and the e-infrastructure is engineered.
    - EC funding could also be partly steered towards community building around tailored services, which commercial providers like Amazon have not yet done. It could also be closely linked to research infrastructures, which gather large amounts of multi-disciplinary data.
- **NRENs and DANTE Strategy**: The DANTE strategy calls for concerted efforts across its member NRENs to drive positive change around new service paradigms, including cloud technologies and the delivery of services supporting global research. The strategy underscores the importance of adopting solutions suited to the diversified landscape of European NRENs with value add that is also specific to each NREN partner and to the community as a whole. Diversified approaches have emerged from both the Study and Workshop with business models already developed in countries like Ireland, the Netherlands and the UK. GRNET has adopted a different approach. It has developed an in-house cloud solution OKEANOS, which has a proprietary data centre and a cloud in

[37] http://www.hub-e.com/services

house solution, which is funded by government. However, e-Science services are in-house and not currently commercial.

- **Towards integrated services.** The workshop suggested different approaches to the integration of services, including the in-house integration of services at EU level, in order to achieve more complete integration underpinned by the implementation of open standards. Another perspective focuses on mandating the integration of user communities as part of EC funding of e-infrastructures and to aim to diversify service offers as much as possible. Another suggestion is to create a federated set of services in Europe.
- **A European marketplace.** All the ingredients needed to create an European marketplace are already in place. Europe should aim to create real value-add rather than aim to imitate or compete head-on with Amazon. Speed of execution is very important. A real market exchange should be built with interoperability of services. Further, EC funding should be directed towards the creation of a single market for procurement.
  - Make interoperability a key business driver for any resource and any service.
  - OpenStack is an example of a real market exchange, creating a large market with little investment. CERN has joined OpenStack because of its strong hold across the world.
  - The exclusive use of open source remains debatable according to the participants collectively.
- **Costs.** There is more work to be done in this area, particularly as cost assessment generally has a short-term value given the fluctuations in cloud price offers. There should be more transparency around the costs of all types of science, including supercomputing. In some instances, we are dealing with cost avoidance, cost sharing, addressing something new via OPex.

In the final session of the workshop, the participants are invited to a roundtable on their suggested recommendations to the EC about the development of a pan-European cloud for science. These are their comments.

- **Steven Newhouse**: he focuses on the integration of services achievable by Europe. Integration should be an activity that takes place in house in order to be complete, providing open standards;
- **David Wallom**: in his opinion, the user communities should create an integrated infrastructure, to be funded by EC. The service has to be more diverse;
- **Mirco Mazzucato**: he sustains the necessity for open software this should to be sustainable at European level;
- **Zajzon Bodó**: he asks the participants to discuss upon where are more interesting innovative opportunities for Europe and where Europe can be competitive;
- **Jeanne Pierre Laisne**: by answering to the EC Officer, he states that all the ingredients for Europe to create a European marketplace are already available. Europe should develop the market for services and avoid competition with Amazon. The main problem for Europe, in his opinion, is speed of execution. A real market exchange should be built and interoperability of services is something that can be produced;
- **Bob Jones**: a Federated Set of Services should be created by Europe;
- **Fotis Karagiannis**: EC funding should be directed towards the creation of a single market for procurement.

EC Conclusions

Finally, **Luis Carlos Busquets**, EC DG CONNECT, takes the floor to state the main call to action and conclude the workshop. The EC Officer points that the Commission main objective is to achieve “excellent science”, which means, the creation of a first-rate environment where the best scientists come to Europe to do research. Europe needs research infrastructures (such as those by CERN, ESA), supporting the creation of a single online European research area by 2020, where data and computing power are needed). Cloud computing is here to stay so it is important that all stakeholders play a part in driving it forward. Helix Nebula is considered to be one of the responses to Action 53 of the DAE. It is also important to deploy cloud capabilities at all the levels (IaaS, PaaS, SaaS), drive standardisation and reach out to other communities such as the long tail of science. The ultimate goal is to arrive at Science-as-a-Service.

The workshop ends on time at 17.00.

Agenda and Participants

Morning session

- Introduction and Welcome - *Konstantinos Glinos, HoU e-Infrastructures (Unit C1) European Commission, DG CONNECT*
- Keynote presentation: Key Findings of Clouds for science and public authorities in Europe and the world - *Gabriella Cattaneo, IDC European Government Consulting*
- Q&A
- Case studies of clouds in the e-Science and e-Government environment – Chaired by: *Silvana Muscella, Trust IT*

1. SurfContext: *Silvana Muscella, Trust IT*
2. Raincloud: The use of a public cloud for weather forecasting for Tyrol public services: *Radu Prodan, Associate Professor, Institute for Computer Science University of Innsbruck*
3. Cloud in Emilia Romagna: Building a regional cloud infrastructure for the public sector: *Giuliano Franceschi, DG Lepida, Emilia Romagna, Italia*
4. The UK G-Cloud strategy and the Cloud Appstore: *Nicola Westmore, Deputy Programme Director - G-Cloud and Common Infrastructure Resource Pool Manager Efficiency and Reform Group Cabinet Office*
5. goBerlin – A Trusted Cloud for SMEs and Public Authorities: *Linda Strick, Fraunhofer FOKUS, Germany*
6. Cheltenham Borough Council, UK - Shared Services: *Christopher Cox, Programme Manager, Cheltenham Borough Council*

- Cloud-based e-Science: demand scenarios, provisioning scenarios, and recommendations for an EU strategy - *Massimiliano Claps, Director, IDC Government Insights*
- Q & A

Afternoon Session

- How the EGI Cloud Federated Project sits in the European Cloud Strategy - *Steven Newhouse, EGI-InSPIRE Project Director and EGI.eu Director*
- Helix Nebula – *Bob Jones, CERN & Helix-Nebula The Science Cloud Coordinator*
- Cloud drivers, barriers and potential synergies between e-Science and e-Government clouds - Roundtable discussion with participants: led by *Gabriella Cattaneo, IDC European Government Consulting*:

1. *Fotis Karagiannis*, eFiscal coordinator & e-IRG Task Force on cloud Computing
2. *Nikos Athanasis*, University of the Aegean & Wildfire

- Discussion on Recommendations to develop Cloud infrastructures and services for e-Science in Europe – Moderators: *Richard Lloyd Stevens*, *Silvana Muscella*

1. *Nikos Athanasis,* University of the Aegean & Wildfire
2. *Fotis Karagiannis,* eFiscal coordinator & e-IRG Task Force on cloud Computing
3. *Vangelis Floris,* GRNET in-house cloud solution OKEANOS
4. *Bob Jones,* CERN & Helix-Nebula The Science Cloud Coordinator
5. *Jean-Pierre Laisne,* CompatibleOne

6. *Mirco Mazzucato,* Director of the Italian Grid Infrastructure – IGI
7. *Richard Lloyd Stevens,* independent consultant
8. *Giuliano Franceschi,* DG Lepida
9. *Pierre-Philippe Matthieu,* ESA – ESRIN
10. *Daniel Viner,* Huddle
11. *Kathryn Birch,* Huddle
12. *David Bradshaw,* Research Manager European Public Cloud services, IDC EMEA
13. *Steven Newhouse,* EGI-InSPIRE Project Director and EGI.eu Director
14. *Per Oster, Director,* Research Environments at CSC - IT Center for Science
15. *Radu Prodan, Associate Professor,* Institute for Computer Science University of Innsbruck
16. *Adil Soussi Nachit*, Account Manager, SPF Finances, Belgium
17. *Linda Strick, Fraunhofer* FOKUS, Germany
18. *Katarzyna Szkuta,* Policy Analyst, Tech4i2
19. *David Wallom,* Technical Director, National e-infrastructure Services, UK
20. *Nicola Westmore,* Deputy Programme Director - G-Cloud and Common Infrastructure
21. *Christopher Cox,* Programme Manager, Cheltenham Borough Council

- Principal Calls to Actions and Conclusions: *Gabriella Cattaneo, IDC European Government Consulting*, *Luis Carlos Busquets, EC DG Connect*

End Workshop

European Commission

Cloud for science and Public authorities - Final Report
Luxembourg, Publications Office of the European Union

2013 – 103 pages

ISBN 978-92-79-31142-0
DOI: 10.2759/25446

KK-02-13-232-EN-N

DOI: 10.2759/25446 ISBN 978-92-79-31142-0