THE FUTURE OF CLOUD COMPUTING

OPPORTUNITIES FOR EUROPEAN CLOUD COMPUTING BEYOND 2010

Expert Group Report

Public Version 1.0

Rapporteur for this Report: Lutz Schubert [USTUTT-HLRS],

Editors: Keith Jeffery [ERCIM], Burkhard Neidecker-Lutz [SAP Research]

EXECUTIVE SUMMARY

Though the concept of "clouds" is not new, it is undisputable that they have proven a major commercial success over recent years and will play a large part in the ICT domain over the next 10 years or more, as future systems will exploit the capabilities of managed services and resource provisioning further. Clouds are of particular commercial interest not only with the growing tendency to outsource IT so as to reduce management overhead and to extend existing, limited IT infrastructures, but even more importantly, they reduce the entrance barrier for new service providers to offer their respective capabilities to a wide market with a minimum of entry costs and infrastructure requirements – in fact, the special capabilities of cloud infrastructures allow providers to experiment with novel service types whilst reducing the risk of wasting resources; .

Cloud systems are not to be misunderstood as just another form of resource provisioning infrastructure and in fact, as this report shows, multiple opportunities arise from the principles for cloud infrastructures that will enable further types of applications, reduced development and provisioning time of different services. Cloud computing has particular characteristics that distinguish it from classical resource and service provisioning environments:

(1) it is (more-or-less) infinitely scalable; (2) it provides one or more of an infrastructure for platforms, a platform for applications or applications (via services) themselves; (3) thus clouds can be used for every purpose from disaster recovery/business continuity through to a fully outsourced ICT service for an organisation; (4) clouds shift the costs for a business opportunity from CAPEX to OPEX which allows finer control of expenditure and avoids costly asset acquisition and maintenance reducing the entry threshold barrier; (5) currently the major cloud providers had already invested in large scale infrastructure and now offer a cloud service to exploit it; (6) as a consequence the cloud offerings are heterogeneous and without agreed interfaces; (7) cloud providers essentially provide datacentres for outsourcing; (8) there are concerns over security if a business places its valuable knowledge, information and data on an external service; (9) there are concerns over availability and business continuity – with some recent examples of failures; (10) there are concerns over data shipping over anticipated broadband speeds.

The concept of cloud computing is linked intimately with those of IaaS (Infrastructure as a Service); PaaS (Platform as a Service), SaaS (Software as a Service) and collectively *aaS (Everything as a Service) all of which imply a service-oriented architecture.

Open Research Issues

CLOUD TECHNOLOGIES AND MODELS HAVE NOT YET REACHED THEIR FULL POTENTIAL AND MANY OF THE CAPABILITIES ASSOCIATED WITH CLOUDS ARE NOT YET DEVELOPED AND RESEARCHED TO A DEGREE THAT ALLOWS THEIR EXPLOITATION TO THE FULL DEGREE, RESPECTIVELY MEETING ALL REQUIREMENTS UNDER ALL POTENTIAL CIRCUMSTANCES OF USAGE.

Many aspects are still in an experimental stage where the long-term impact on provisioning and usage is as yet unknown. Furthermore, plenty of as yet unforeseen challenges arise from exploiting the cloud capabilities to their full potential, involving in particular aspects deriving from the large degree of scalability and heterogeneity of the underlying resources. We can thereby distinguish between *technological* gaps on the one hand, that need to be closed in order to realize cloud infrastructures that fulfil the specific cloud characteristics and *non-technological* issues on the other hand that in particular reduce uptake and viability of cloud systems:

To the *technological* aspects belong in particular issues related to (1) scale and elastic scalability, which is not only currently restricted to horizontal scale out, but also inefficient as it tends to resource over usage due to limited scale down capabilities and full replication of instances rather

than only of essential segments. (2) Trust, security and privacy always pose issues in any internet provided service, but due to the specific nature of clouds, additional aspects related e.g. to multi-tenancy arise and control over data location etc. arise. What is more, clouds simplify malicious use of resources, e.g. for hacking purposes, but also for sensitive calculations (such as weapon design) etc. (3) Handling data in clouds is still complicated - in particular as data size and diversity grows, pure replication is no viable approach, leading to consistency and efficiency issues. Also, the lacking control over data location and missing provenance poses security and legalistic issues. (4) Programming models are currently not aligned to highly scalable applications and thus do not exploit the capabilities of clouds, whilst they should also simplify development. Along the same line, developers, providers and users should be able to control and restrict distribution and scaling behaviour. This relates to (5) systems development and management which is currently still executed mostly manually, thus contributing to substantial efficiency and bottleneck issues.

On the other hand, *non-technological* issues play a major role in realizing these technological aspects and in ensuring viability of the infrastructures in the first instance. To these belong in particular (1) economic aspects which cover knowledge about when, why, how to use which cloud system how this impacts on the original infrastructure (provider) –long-term experience is lacking in all these areas; and (2) legalistic issues which come as a consequence from the dynamic (location) handling of the clouds, their scalability and the partially unclear legislative issues in the internet. This covers in particular issues related to intellectual property rights and data protection. In addition, (3) aspects related to green IT need to be elaborated further, as the cloud offers principally "green capabilities" by reducing unnecessary power consumption, given that good scaling behaviour and good economic models are in place.

Europe and Clouds

Notwithstanding common beliefs, clouds are not a phenomenon entirely imported from abroad. This report will elaborate the main opportunities for European industry *and* research to be pursued with respect to the specific capabilities and remaining gaps.

This document provides a detailed analysis of Europe's position with respect to cloud provisioning, and how this affects in particular future research and development in this area. The report is based on a series of workshops involving experts from different areas related to cloud technologies.

EUROPE'S MAIN OPPORTUNITIES TO PARTICIPATE IN THE "CLOUD MOVEMENT" CONSIST IN PARTICULAR IN ASPECTS RELATED TO EXTENDING AND COMPLETING THE CAPABILITIES OF CURRENT CLOUD SYSTEMS, WHEREBY THE LONG-TERM GOAL CONSISTS IN REALIZING META-SCALABLE CLOUD SYSTEMS AND SERVICES. THE COMPLEXITY TO REALIZE THE OPPORTUNITIES DIRECTLY DEPENDS ON THE COMPLEXITY TO PERFORM THE UNDERLYING RESEARCH WORK AND OF THE CURRENT DEVELOPMENT STATUS.

In more detail, the identified opportunities are: (1) Provisioning and further development of Cloud infrastructures, where in particular telecommunication companies are expected to provide offerings; (2) Provisioning and advancing cloud platforms, which the telecommunication industry might see as a business opportunity, as well as large IT companies with business in Europe and even large non-IT businesses with hardware not fully utilised. (3) Enhanced service provisioning and development of meta-services: Europe could and should develop a 'free market for IT services' to match those for movement of goods, services, capital, and skills. Again telecommunication industry could supplement their services as ISPs with extended cloud capabilities; (4) provision of consultancy to assist businesses to migrate to, and utilise effectively, clouds. This implies also provision of a toolset to assist in analysis and migration.

Recommendations Overview

Due to the strong commercial nature of cloud systems, both technological and non-technological aspects are involved in cloud provisioning. Since both areas still have major gaps, the recommendations are not restricted to purely technological issues, but also cover non-technological aspects related in particular to the economical and legalistic side of cloud systems.

Europe is in a strong position to address both these areas: technologically due to its excellent background in many of the key research and development aspects related to cloud systems, such as GRIDs and Service Oriented Architectures, and non-technologically due to Europe's position as a united body. Europe also has a strong market position with many of major contributors from different field originate from Europe.

The recommendations towards research and development communities and bodies as expressed in this report hence do address a wide scope of outstanding issues, ranging from specific research and development topics over general policies to legalistic aspects which currently pose a major obstacle towards wide uptake of cloud infrastructures:

Main Recommendations

Recommendation 1: The EC should stimulate research and technological development in the area of Cloud Computing

	Cloud computing poses a variety of challenges to conventional advanced ICT, mostly due to the fact of the unprecedented scale and heterogeneity of the required infrastructure. This demands a rethinking of even current advanced ICT solutions. Plenty of research issues remain to be addressed in the context of cloud provisioning. Europe should exploit the available expertise and results from areas such as Grid, Service Oriented Architectures and e-infrastructure to help realizing the next generation of services on cloud systems. Particular research topics to be addressed further are: (1) Elastic scalability; (2) Cloud (systems) development and management; (3) Data management; (4) Programming models and resource control; (5) trust, security and privacy.

Recommendation 2: The EC together with Member States should set up the right regulatory framework to facilitate the uptake of Cloud computing

	Cloud systems are mostly in an experimental stage – to fully exploit their capabilities in particular from a commercial side, the according impact, dependencies, requirements etc. need to be evaluated carefully. Accordingly, research efforts need to be vested not only into technological aspects of *realizing* cloud systems, but also into the aspects related to commercial and business aspects, in particular involving economical and legalistic concerns. Accordingly, business consultants, legal researchers, governmental bodies etc. should be encouraged to participate in investigating the particular circumstances of cloud provisioning. Obviously, technologies thereby need to recognize results from these areas, just as economical and legalistic views need to acknowledge the technological capabilities and restrictions. In summary, the specific issues are: (1) Economical aspects; (2) Legalistic issues; (3) Green IT

Additional Recommendations

Additional Recommendation 1: The EU needs large scale research and experimentation test beds

	A major obstacle for European research communities to develop and test effective large scale cloud systems consists in the lack of available resource infrastructures of a size that allow experimentation and testing. Such infrastructure test beds could be provided through joint, collaborative efforts between existing resource owners and public, as well as non-public research bodies, e.g. through public-private partnerships or through fostering existing research communities building up on public e-infrastructures etc.

Additional Recommendation 2: The EC together with industrial and public stakeholders should develop joint programmes encourage expert collaboration groups

	The development of future cloud infrastructures and in particular of meta-clouds necessitates the collaboration of experts from various backgrounds related to cloud systems, as can be implicitly seen from the recommendations above. This would include research and development, academia and industry equally. To encourage such collaboration, the need for interoperable meta-clouds needs to be promoted more clearly.

Additional Recommendation 3: The EC should encourage the development and production of (a) CLOUD interoperation standards (b) an open source reference implementation

	The development of standards and a reference implementation would assist European SMEs in particular in ensuring their products and service offerings in the cloud environment have the widest possible market and customer acceptability. The standards should encourage all suppliers to be able to interoperate; the reference implementation is to allow plug-tests to prove standards compliance.

Additional Recommendation 4: The EC should promote the European leadership position in software through commercially relevant open source approaches

	Maintaining an open source approach for research results and cloud infrastructure support tools ensures uptake and simplifies adaptation to different environments. The European open source movement should thereby work strongly together with industry to support commercial cloud based service provisioning.

LIST OF EXPERTS

Prashant Barot [Oracle], Francis Behr [Syntec], Peter Bosch [Alcatel Lucent], Ivona Brandic [Vienna University of Technology], Brigitte Cardinael [France Telecom], Thierry Coupaye [France Telecom Orange Labs], Richard Davies [Elastichosts], David De Roure [University of Southampton], Philippe Dobbelaere [Alcatel Lucent], Andreas Ebert [Microsoft], Aake Edlund [KTH], Guido Falkenberg [Software AG], Jürgen Falkner [Fraunhofer], William Fellows [The 451 Group], Friedrich Ferstl [SUN], Ioannis Fikouras [Ericsson], Mike Fisher [British Telecom], Behrend Freese [Zimory], Alfred Geiger [T-Systems], Juanjo Hierro [Telefonica I&D], Giles Hoghen [ENISA], Keith Jeffery [ERCIM], Ricardo Jimenez-Peris [UPM], Ruby Krishnaswamy [France Telecom Orange Labs], Frank Leymann [University of Stuttgart - IAAS], Ignacio Llorente [UCM], Monica Marinucci [Oracle], Joan Masso [Gridsystems], Cyril Meunier [IDC], Christine Morin [INRIA], Sebastian Müller [Google], Burkhard Neidecker-Lutz [SAP], Mathieu Poujol [PAC], Thierry Priol [INRIA], Harald Schöning [Software AG], Lutz Schubert [High Performance Computing Center Stuttgart], Dave Snelling [Fujitsu Labs Europe], Paul Strong [eBay], Werner Teppe [Amadeus], Clemems Thole [Fraunhofer], Dora Varvarigou [NTUA], Stefan Wesner [High Performance Computing Center Stuttgart], Per Willars [Ericsson], Yaron Wolfsthal [IBM], Hans Wortmann [University of Groningen].