International Communication Union
ITU
Arab Regional Office

Arab League Educationaln Cultural
and Scientific Organization
ALECSO

Guidelines to improve the use of the Cloud Computing Technology in Education in Arab Countries

December, 2016

The Arab League Education, Culture and Science Organization (ALECSO) is a Tunis-based specialized institution working under the umbrella of the League of Arab States. It is essentially concerned with the development and coordination of the activities related to education, culture and science in the Arab world.

The International Telecommunication Union (ITU) is the United Nations specialized agency for Information and Communication Technologies – ICTs. The ITU Arab Regional Office in Egypt provides the support to ITU Members in the Arab Countries under several activities related to ICT development and deployment of telecommunication/ICT infrastructure and services.

This study was commissioned as part of the project 'Promoting Cloud Computing Technology in Education' carried out by ALECSO organization and ITU Arab Regional Office.

ISBN: 978 – 9973 – 15 – 381 - 4

Edited by:

- Prof. Mohamed Jemni and Dr. Mohamed Koutheair KHRIBI, Arab League Educational, Cultural and Scientific Organization - ALECSO.
- Mr. Slaheddine Maaref, International Telecommunication Union ITU Arab Regional Office.

Prepared by:

- Prof. Farouk Kamoun, University of Manouba, Tunisia
- Prof. Christophe Cerin, University of Paris 13, France
- Eng. Karim Chine, United Kingdom
- Dr. Heithem Abbes, University of ElManar, Tunisia
- Dr. Wajdi Louati, University of Sfax, Tunisia

Published by

Arab League Educational, Cultural and Scientific Organization - ALECSO
Street of Mohamed Ali Akid, Centre Urbain Nord, BP 1120, Cité Khadra 1003, Republic of Tunisia

Tel. : +216 70 01 39 07
Fax : +216 71 94 86 61
Web : www.alecso.org

International Telecommunication Union ITU
Regional Office for the Arab States
Smart Village, Giza, Egypt

Tel. : +20 2 353 71777
Fax : +20 2 353 71888
Web : http://www.itu.int

JEMNI, Mohamed

Guidelines to improve the use of the Cloud Computing Technology in Education in Arab Countries / ed. By Mohamed Jemni, Mohamed Koutheair Khribi, Slaheddine Maaref; prep. By Farouk Kamoun, Christophe Cerin, Karim Chine…[et al.] . – Tunis: Arab League Educational, Cultural and Scientific Organization; Cairo: International telecommunication Union ITU. Regional Office for the Arab States, 2016. – 139 p.

ISBN : 978 – 9973-15-381-4

001/05/2017/وق

Outline

Executive Summary

Background

Ten years ago, Amazon unleashed a technology that sparked a revolution in IT. What became called "cloud computing" turned out to be a paradigm shift in the way we use computers. Today, the cloud is transforming Industry and impacting education and research more than any other technology has ever did. The current figures about cloud adoption leave no doubts: The cloud is here to stay.

The trajectory towards cloud pervasiveness has been all but linear. In research and education, the last few years have been years of transformation. Individual educational institutions all over the globe started embracing various types of cloud technologies and embedding them as core capabilities of their teaching/learning frameworks. National and international funding bodies elaborated strategies and work programs to support projects aiming at fostering the use of the cloud in research and education. After years of building capacities and e-Infrastructures fully relying on Grid technologies, both the European Commission and the American National Science Foundation (NSF) kick started a slow but steady shift towards making the cloud the corner stone of their national or trans-national federated research infrastructures.

This study aims at:

- Raising awareness about cloud computing technology and related benefits and advantages especially for the education field;
- Developing specific guidelines for migrating to the cloud which take into account several parameters and national contexts for the Arab countries;
- Delivering a cloud migration policy for decision makers;
- Defining a roadmap to develop and deploy cloud infrastructures and platforms for education and research.

The report is structured as follows:

- The first chapter deals with the state of the art of cloud adoption and cloud technologies; It explains the concepts, introduces some initiatives and surveys focusing on cloud computing in education and research and sketches technical elements of the cloud ecosystem in order to characterize the main tools used in production around the world;
- Chapter 2 classifies the Arab countries according to their level of development in the ICT field; It also presents the methodology adopted to conduct a survey on cloud computing in the Arab countries as well as the survey's results;
- Chapter 3 outlines the policy recommendations taking into account the findings related to current state of cloud adoption worldwide, particularly in educational and research as well as the outcome of the survey and the state of ICT in the Arab world;

- Chapter 4 provides a practical guide to help educational institutions and information technology (IT) managers plan and operate application and workload migration to the cloud. It provides a detailed roadmap that covers a list of technical steps for migration.

State of the Art in Cloud Adoption and Cloud Technologies

The Main global trends in the cloud market and the state of adoption are summarized as follows:

- The cloud infrastructure will be available as affordable services part of the ambient environment, and as a commodity;
- The public cloud market has experienced a tremendous growth over the past few years. The Saas segment will continue to outrun the IaaS and PaaS segment with about 60% of total market in 2018;
- Cloud applications will account for 90% of worldwide mobile data traffic by 2019, compared to 81% at the end of 2014;
- Major providers in public cloud are AWS, Google, Microsoft Azure and Rackspace. AWS dominates the public cloud.

RightScale's recent survey (2015) on a sample of 930 organizations belonging to various sectors of activities shows that:

- Cloud is a given and hybrid cloud is the preferred strategy of the respondents.
- Scalability, faster access to infrastructure, and availability are the top three benefits experienced using cloud ,whereas "security" and "lack of cloud resources and expertise" are the top 2 challenges.
- Central IT teams are increasingly offering self-service portals as a central hub to broker cloud services to the enterprise

The NSF-XSEDE cloud adoption survey (2012-2013) conducted in higher education and research institutions across a wide variety of scientific areas and the humanities, arts, and social sciences, found that:

- The cloud benefits identified by the survey participants were: flexible pricing model, lower costs, compute elasticity, data elasticity, Software as a Service, Education as a Service, broader use, scientific workflows, rapid prototyping, and data analysis;
- The cloud challenges identified included the learning curve, the variability in bandwidth, the lack of private/public cloud interoperability, the security, the data movement and cloud computing cost and the funding availability.

Given the importance of the cloud phenomenon, the European Commission and the National Science Foundation (NSF) launched major initiatives in cloud computing. The EU adopted strategy in 2012 for "Unleashing the potential of cloud computing in Europe". It also adopted cloud computing among the research priorities in the H2020 research and innovation funding program.
In August 2014, The NSF has announced the funding of several projects to create cloud computing testbeds "to enable a new future for cloud computing". Its aim at **transforming** the current US research environment by supporting research infrastructure, enabling transformative research at the frontiers of computing.

In addition to these initiatives, several universities have adopted and actually migrated a certain number of applications such as distance learning and / or scientific research applications to the public cloud.
Among the universities that have migrated educational applications to the public cloud, the following examples are noteworthy:

- Anhanguera one of the largest universities in Brazil and in the world supporting distance learning, Anhanguera is one of the largest users of Moodle, an e-learning platform for collaborative learning, migrated Moodle to Amazon AWS ;
- The Khan Academy posts a vast collection of free educational online. Students answer some 1.5 million practice questions per school day all served through Google App Engine.

- The Tokyo University of Technology (TUT) has opted for a full shift of its ICT environment to the Azure cloud.

Among the universities that have migrated scientific applications to the public cloud, the following examples are worth mentioning:

- The Algorithms, Machine, and People (AMP) Lab at the University of California Berkeley leverages AWS to quickly scale the compute resources needed to apply analysis algorithms to genomic data;
- The Collaboration project MSSNG Autism Speaks adopted Google to migrate and scale up its database to hold information from the whole genomes of 10,000 individuals, making it the world's largest single repository of autism-related DNA sequencing data.

In addition to these experiences with the public cloud, educational institutions have adopted private or community cloud models. Two case studies are to be mentioned:

- The University of Hawaii at Manoa (UHM) College of Education (COE) located on several remote islands adopted the open source cloud platform OpenStack to virtualize its infrastructure;
- In 2015 the University Sorbonne Paris Cité (USPC) decided to renovate and to federate its research infrastructure through the launch of a community cloud based on the OpenNebula open source technology.

State of the ICT and cloud computing in education in the Arab World

The overall assessment of the state of ICT in the arab World is measured according to the ITU ICT Development Index(IDI) and the Internet World Stats.

IDI is a composite index that combines 11 indicators into one benchmark measure that can be used to monitor and compare developments in information and communication technology (ICT) between countries and over time. The IDI is divided into the three sub-indices: (i) An access sub-index, (ii) A use sub-index, (iii)A skills sub-index. IDI is constructed as the weighted average of the normalized values of those indicators, from 0 to 10. A score that is close to 10 implies a very good ICT development level while a score close to 0 indicates a poor development.

It's to be noted that IDI rankings for 2010 and 2015 consider only 16 out of the 22 Arab countries.
The 16 Arab countries are ranked from 27 to150 out of 162 classified countries. They all achieved some progress as compared to 2010. They may be divided into three classes: the first one which consists of 7 countries is above the world's average of 5.03; two countries of this class even exceed the average of the most advanced countries (7.35); the second class is made up of four countries that are close to the world's average with an IDI varying between 4.75 and 4.40. The last class consists of 5 countries with an IDI varying between 3.71 and 2.07.

According to the World Stats Institute, the average rate of Internet penetration for the 22 Arab countries is equal to 44.8% and is slightly below the world's average of 46.4%. With regard to Internet penetration we note that 3 groups emerge: one group of 10 countries whose scores are below the world average, a group of 7 countries which exceed the European average (73.5%) and a group of 5 is in the middle. This classification confirms the outcome of the analysis based on the IDI index.

Itis also an indication that a great deal of effort is to be accomplished in most Arab countries to reach an adequate ICT infrastructure development to enable cloud deployment.

To collect data on cloud adoption in educational and research institutions in the Arab countries, a survey based on an online questionnaire conceived by the study team was launched by ALECSO and publicized via its official channels throughout its members states. The survey was aimed at government departments, engineering schools, faculties and research centers, and educational data centers.

The objective of the survey was (i) to make an inventory of ICT use in Arab institutions (ii) to understand how those institutions perceive the benefit and applicability of the cloud to their use cases and whether and how they plan to use it (iii) to understand the barriers and risks they currently apprehend.

Forty responses provided by institutions from 10 Arab countries were usable. Out of the 40 respondents, 5 have already adopted cloud computing at the production level which the prerequisite to be designated as "cloud adopter". The 35 other respondents are either in the implementation stage (28%), or in the trial stage (12%), or in the discussion stage (28%) or are not involved (20%) which categorizes them as "non cloud adopters".

The survey results revealed the following trends:

- **A clear move toward cloud adoption**: 52% of surveyed institutions are either in the production, implementation or test phase;

- **A positive perception of cloud benefits by non cloud adopters :** the most important benefits perceived by the highest percentage of respondents are "disaster recovery capabilities" followed by "hardware cost saving" and "reliable data storage";
- **Type of cloud for non cloud adopters**:
    - 40% of institutions prefer to migrate to a private cloud,
    - 20% hybrid cloud
    - Only 9% to a public cloud
- **Applications to migrate to the cloud for non cloud adopters**:
  « collaborative application » with 46% of respondents, followed by storage service (34%) and scientific applications (29%) are the preferred applications during the first year. They are followed by «learning management system» for (26% the first year and 26% the second year);
- **Perceived barriers for cloud adoption for non cloud adopters**: Among the most significant barriers are "security issues", "Integration with Existing systems", "data protection and / or privacy Concerns";
- **Support needed to migrate to the cloud**: the highest support demands by "non cloud adopters**"** are related to « cost benefit analysis » and « security and privacy ».Most cloud adopters (4/5) engaged in change management to overcome cloud adoption barriers;
- **Cloud adopters expressed their satisfaction with cloud utilization.**

It is to be noted that the results of the survey should be considered with caution since the survey sample is not representative enough. They give however some clear indication as to the growing cloud adoption in many Arab countries.

To better validate the outlined trend, the working group has completed its research using available documents and material published on the net mostly by the major cloud providers in order to search for Arab educational institutions that have adopted the cloud. The literature review revealed that:

- Morocco seems to be the leader in Africa in cloud computing adoption. Twelve universities adopted « Google Apps For Education » as a service for their students, faculty members and administrative staff;
- Morocco has also decided to use the cloud in vocational training: it adopted Microsoft Office 365 to provide training in 327 training institutes across the country for 500,000 students in 35 fields of studies;
- In Tunisia, four higher education institutions adopted public clouds
- In Jordan two universities have adopted the cloud;
- In Saudi Arabia, two universities have also adopted the cloud;
- In addition, a few institutions in Qatar and Sudan adopted the cloud.

The review of the literature and the survey's results show that cloud adoption is a movement that is gradually spreading within the higher education institutions of the Arab world. Therefore, it is timely to reflect on the strategy and practical measures to be implemented to support the Arab educational institutions and help them successfully migrate to the cloud.

Cloud adoption: Policy guidelines

Policy recommendations presented are based on the findings related to cloud adoption in the world and particularly in educational institutions. They also refer to the state of infrastructure and ICT development in the Arab world as well as to the rate of cloud adoption or intentions and preferences with regard to cloud migration for a sample of educational institutions

The development of the key policies presented below is also the result of a literature review and of intensive exchanges and debates between the experts of the project.

Four key policies are proposed. These policies are structured according to major stakeholder's profiles: ministerial departments, universities and educational institutions.

Key policy one: high quality network

Arab countries whose ICT Development Index scores are average or weak, need to accelerate the development of their network infrastructure in order to foster and succeed in migrating to the cloud. The first key policy principle is: ***Give to investment in a high quality network (intranet+access to INTERNET) the highest priority***. To this end:

- **Ministries** in charge of higher education and research should negotiate contracts' frameworks and SLAs at national level with Telecom operators to provide a best of breed network access to all institutions regardless of their locations and size. Bandwidth should be appropriately sized and should be upgraded on a regular basis.
  It is essential to coordinate with the different ministries in order to implement a framework of incentives promoting universal access to the internet. Students, teachers and researchers should be able to have access to a reliable internet connection not only on campus but also at home (anywhere anytime).
- **Universities/ higher education institutions** have to allocate the right budgets to deploy a redundant/reliable high-quality intranet, properly sized and supported.
  Wireless access should be provided and campus-wide coverage should be optimal and permanently monitored and supported. Wired connections should be available in all the rooms where learning and practical activities requiring maximum reliability and bandwidth are taking place.

It is recommended to negotiate with hardware providers to lower the barrier for acquiring laptops and negotiate with Telecom operators to lower the barrier for gaining 3G/4G internet access on mobile devices to all students, teachers and researchers.

Key Policy 2: Always Public cloud first

Public Cloud computing offers computing capacity and storage as well as a wide range of services and scientific applications that can be used as needed. It offers flexibility, efficiency and enables users to benefit from developments made by suppliers. Services are accessible via the Internet everywhere all the time.

The policy to be adopted consists, therefore, in encouraging the use of public cloud as the first choice wherever possible and hence enabling educational institutions fast access to advanced IT and catching up with international practice: ***Always public cloud first***is the second key policy

It is recommended to consider Always SaaS first and Use Pubic PaaS/IaaS instead of local infrastructures. Software-as-a-Service should be considered first, it entirely delegates all the unwanted complexity to a service providers and empowers users through seamless access to tailored and effective Web User interfaces. PaaS and IaaS (public if possible or private otherwise) should then be considered as the way to go when it comes to the provisioning of custom applications and capabilities, they expose more complexity but can be harnessed thanks to APIs and automation frameworks. Existing on-premise applications that (i) have critical constraints or (ii) wouldn’t benefit from public or private clouds or (iii) can’t be cloudified should continue operating on-premises.

- **Ministries** in charge of higher education and research should negotiate contracts’ frameworks with the key global and local cloud providers. Engage in ambitious partnerships with those players (such as Google, Microsoft, Amazon, etc.) in order to make the use of cloud in education part of a larger framework of cooperation, promotion of innovation and adoption of cutting-edge IT solutions throughout their countries.
  Ministries should also put in place communication strategies contributing to a mindset shift about pervasiveness of cloud solutions.
- **Universities/ higher education institutions** should review all services and applications in use, benchmark existing SaaS alternatives and migrate if a mature and satisfactory solution exists which also respects the different constraints and legal rules of the institution.
  For PaaS/IaaS migration, IT departments should train their members on technologies such as containers and clouds automation APIs and tools. They should expose self-service portals to the end user and automate user interaction with tools and applications as much as possible.

Key policy 3: “Cloudify” the existing local infrastructures and applications at institutional level

To speed up cloud migration, intermediate solutions may be adopted to improve the management of the existing IT infrastructure without disrupting the operation of the applications used by educational institutions. It is recommended to "***cloudify” the existing local infrastructures and applications at institutional level***. To this end, 2 options are available:

a- Use containers for applications deployment;

b- Use a mature open source cloud toolkit (OpenStack, OpenNebula). Eventually with virtual machines running Docker Engines (or a similar container engine). Avoid proprietary virtualization technologies and use open source hypervisors (KVM, Xen) wherever possible.

Option (a) alone represents an easier option than (b)

- **Ministry** in charge of higher education and research should put in place a program for pilot projects to show case the cloudification using containers and cloud toolkits of typical institutions' infrastructures. It is also necessary to publish and disseminate results, know-how and lessons learned.
- **University/higher education institution**'s IT departments should train their members on technologies such as containers, virtualization technologies and clouds toolkits. Pilot projects must be put in place to help making informed decisions about the specific container orchestration technology and cloud toolkit that would best fit the needs of the institution.

Key policy 4: Adopt a cloud friendly governance model for IT

The migration to the cloud requires a change in IT governance models and practices. First, the offer to be made available to users must meet two main requirements: ease of use and flexibility. Secondly the establishment of an intermediary entity should facilitate relationships between suppliers and customers. It is recommended to ***adopt a cloud friendly governance model for IT***. Hence it is advisable to:

a- Allow end users (Teachers, students, researchers, staff) to make informed choices about the specific services they may want to experiment with or use on the long run. Redefine the role of the IT department to become a service and support provider without limiting the scope of initiatives that users may want to take;

b- Create an entity at national, regional or institutional level that deals with brokerage and procurement in a centralized way.

That entity's major missions would be to:

(i) Negotiate the SLAs and the general terms and conditions with both Network providers and Global/national public clouds providers when possible/applicable. Make sure the terms and conditions are compatible with all the legal frameworks to which the institutions are subject;

(ii) Centralize the procurement of public cloud resources;

(iii) Operate the cloud broker and the self-service portal;

(iv) Take responsibility for a federated/consistent data strategy and data management plan;

(v) Act as a single point of information for IT/data security and privacy policies/rules.

- **Ministry** in charge of higher education and research should
  - o study the different options for the intermediate entity and short-list those which are compatible with existing rules and procurement processes;

    - arbitrate between public and private entity options and if the latter option is chosen, select a company based on the legal rules in place;
    - otherwise, arbitrate between the creation of a new public entity and the delegation of the new role to an existing one that has the operational capacity for that. Make sure that the entity is governed in a way that gives to the end users/institutions power to contribute to the strategical decision making.
- **Universities/ higher education institutions**: If the Ministry decides not to operate an entity at national or regional level, the university should (i) create that entity or (ii) attribute its role to a team with in the IT department or (iii) contract with a company.
  In case of (i) or (ii), Universities/ higher education institutions must make sure that
    - all the required human resources and skills are available to meet the intermediation challenges at technical and legal levels;
    - the entity is open to end users' contributions to the strategical decision making.

Applying the above policy choices requires to overcome many **barriers to cloud adoption**. Among the most significant barriers that have been identified are: "security issues", "Integration with existing systems", "data protection and privacy concerns". To be able to meet the challenges of cloud migration, institutions need support. The main support demands requested by institutions are related to " cost benefit analysis " and " security and privacy ";

Capacity building is key to a successful transition to cloud. **I**nstitutions must make investment in the governing structures, organization processes, people and their skills required to make cloud technology an essential element in how the organization services are managed. IT managers need to understand what capabilities are offered and how they can be combined with internal resources, and develop a plan to leverage these combined resources:

Cultural change is also essential and it cannot be achieved without a substantial educational effort. Cloud experimentation by IT specialists and users would greatly facilitate culture change; workshops to share the knowhow and experiences cross-institutions would also be needed. One needs to keep in mind that outsourcing arrangement or a technical platform are not enough. Institutions must plan for cloud computing as a **strategic choice** and the elaborated strategies should be adapted to the institution's specific environment and use cases. Cloud migration should be used as a vector for transformation and improvement. For instance, new curricula need to be addressed in parallel with the deployment of cloud technologies. The curricula may concern the teaching of cloud technologies-related skills or may leverage the capabilities of the cloud to provide more effective education. This applies to all data science related curricula for example.

Cloud migration: Implementation and guidelines

The objective is to provide a practical guide to help educational institutions and information technology (IT) managers to plan and operate applications and workloads migration to the cloud. It provides a detailed roadmap that covers a list of technical steps for migration.

Three scenarios for migrating applications, services and workloads to Cloud Computing infrastructures are considered:

- **SaaS Migration**: Consists in replacing in-house applications with new SaaS applications;

- **IaaS and PaaS Migration**: deals with migrating in-house applications and workloads to private or public IaaS/PaaS Clouds;
- **Cloud Service Brokerage for IaaS and PaaS Migration**: describes the Usage of cloud brokers to manage the delivery of cloud services and negotiate relationships between educational institutions and private or public Clouds providers.

A migration use case is also presented for each scenari:

SaaS Migration: Many schools, colleges and universities have moved their email, collaboration and communication services to the cloud. The objective is to completely replace in-house applications and services with new SaaS capabilities. Email, Video Conferencing, Storage, Social Network, Office tools are all eligible applications for migration. Both Google and Microsoft offer such educational tools in SaaS mode for free in many countries. The Amazon Educate global initiative provides students and educators with AWS credits for use in courses and projects. The Amazon's cloud powers various innovative SaaS offerings for eLearning. Echo360 for example is an AWS-backed active learning platform, its SaaS classroom capabilities replace several classroom learning technologies that are neither scalable nor designed to support the diversity of modern faculty/student interaction.

Domain administrators in charge of moving their students, faculty and staff members to Google Apps for Education or to Office 365 for Education or to AWS will find in the report references to guides with a step-by-step outline for completing the technical aspects of the deployment. The outlines include relevant help center's articles and videos for the three cloud providers.

Migration to an IaaS /PaaS: The major technical steps for migrating in-house applications and workloads to public or private IaaS/PaaS Clouds are: (i) Elaborate a cloud migration project plan (ii)Set up the cloud environment (iii)Set up the Applications in the cloud (iii)Set up a pre-migration prototype (iv)move to the production cloud.
These steps are outlined in the report, their detailed description is beyond the scope of this executive summary

Two migration use cases are presented one public cloud case and one private cloud case.

The first one deals with a public cloud migration of an institutional Moodle to AWS, a typical multitier dynamic web application. Amazon provides a set of whitepapers targeted at architects and technical decision makers looking to build a cloud migration strategy. The actual migration steps described for this use case are: (i) Cloud Assessment (ii) Proof of concept (iii) Data migration (iv) Application migration (v)Co-existence phase (vi)Optimization.

The second use case deals with the set up of an open stack based private cloud environment and the migration process as implemented by ENIS (school of engineering at Sfax university). ENIS is expected to achieve the following objectives:

- Migrate business applications (e.g. office order, timetable, mailing service, web service, etc…) to the private cloud;
- Replace traditional PCs deployed in practical work classrooms with thin clients connected to a private cloud based VDI (Virtual Desktop Infrastructure);
- Offer virtual machines as a service for researchers to deploy their scientific workloads;

- To offer a self-service portal for administrators to allocate virtual servers (to run applications) and for end users (teachers, students, researchers, etc.) to access their own virtual desktops and virtual machines.

Cloud service brokerage for IaaS and PaaS Migration: Cloud service brokerage acts as a middle man between educational institutions and public cloud providers (like Amazon AWS, Google Compute/App Engine, Microsoft Azure) by aggregating and offering multiple cloud resources and services that best suit their needs. The broker represents a single interface for interacting with multiple public and private clouds. The broker may be deployed on premises, within one institution data center or in a community fashion, under the supervision of supra-institutions entity (e.g. ministry of research and education) or as a paid external service.

Cloud broker can manage multiple public clouds through their APIs (Application Programming Interfaces). In order to avoid vendor lock-in and to ensure services' portability and interoperability, it is recommended to deal with public clouds offering extensive and well documented APIs. Clouds allowing virtual artifacts to be easily exported are to be preferred.

Cloud brokers are often made accessible through self-service portal which act as a central hub within the organization. The portal allows users to seamlessly provision, access and share the IT capabilities and services they need without having the IT administrators directly involved.

Three examples of cloud brokerage platforms are presented:

RosettaHUB[www.rosettahub.com] is an innovative eScience and eLearning platform. It provides researchers, teachers and students with a streamlined experience in their day-to-day interactions with (i) clouds, HPC clusters and supercomputers which are made accessible through one easy-to-use web console(ii) data science environments, tools and libraries which are interconnected and exposed in the browser as collaborative services the Google Docs way(iii) data storage capabilities which are mapped to an easy-to-use interactive framework and next to which compute capabilities are created and remotely controlled (iv)big data frameworks such as Spark(v) Local Desktop tools such as Excel and Word which become clients to remote advanced cloud-based data processing capabilities (vi) Local applications and code which can leverage the RosettaHUB libraries to programmatically provision and use cloud resources and tools for scientific computing and data analysis (vii) peers and collaborators with whom real-time collaboration in the browser can be engaged.

Various other open source technologies can be used as brokerage building blocks for the institutional one-stop access portal such as the CloudSelect broker, the CompatibleOne broker and JellyFish. Some of them are based on open standards such as OCCI (Open Cloud Computing Interface).

Chapter 1
State of the Art in Cloud Adoption and Cloud Technologies

1.1. Introduction

Ten years ago, Amazon unleashed a technology that sparked a revolution in IT. What became called "cloud computing" turned out to be a paradigm shift in the way we use computers. Today, the cloud is transforming Industry and impacting education and research more than any other technology has ever did. The current figures about cloud adoption leave no doubts: The cloud is here to stay. The trajectory towards cloud pervasiveness has been all but linear. The pace of innovation of the technology since its early days has been tremendous. Most predictions and strategies elaborated by the different actors trying to harness the cloud and make the most out of it have been obsoleted by the technological metamorphoses. As a reaction to what Amazon, followed by Microsoft and Google have been offering, the idea of building "private clouds" to mimic those new public services on premises emerged and drained a lot of energy and time. It became clear today that the most viable approach is the hybrid one: Use the public cloud wherever possible and keep applications with strict legal or privacy constraints on premises, eventually on an infrastructure that is virtualized and automated the cloud way.

In order to streamline operations and provide both end users and IT administrators with the best experience in dealing with cloud infrastructures, avoid lock-in and standardize applications, cloud brokers are used as a layer of federation. Self-service portals complement those brokers and act as a user facing one-stop shop for the provisioning, access and sharing of all required applications.

The virtualization technologies that made it possible to cloudify the management of infrastructures are being complemented and sometime replaced with light weight containers technologies such as Docker. Those technologies are deeply impacting not only the cloud landscape but also all the IT Development/Operations arena (DevOps). Clouds of containers are enabling automation of all the processes of software delivery and infrastructure changes: the building, testing, and releasing of software and services can happen today more rapidly, more frequently, and more reliably than ever before.

In research and education, the last few years have been years of transformation. Individual educational institutions all over the globe started embracing various types of cloud technologies and embedding them as core capabilities of their teaching/learning frameworks. National and international funding bodies elaborated strategies and work programs to support projects aiming at fostering the use of the cloud in research and education. After years of building capacities and e-Infrastructures fully relying on Grid technologies, both the European Commission and the American National Science Foundation (NSF) kick started a slow but steady shift towards making the cloud the corner stone of their national or trans-national federated research infrastructures. The public cloud tsunami created politically hard to manage situations where large investments appeared to have been geared towards technologies that clearly had no future. The governance models had to be deeply adjusted. New procurement processes had to be explored. Cloud computing and the public clouds that used to be a "taboo" aren't any more as the most recent work programs published testify. In order to measure the current progress of cloud adoption by research and educational institutions, a number of surveys have been conducted among which the XSEDE survey which shed some light on various aspects of cloud usage, use cases and applicability.

The organization of this Chapter is as follows. After a general introduction of the core concepts, we provide in section 2, elements of context to setup vocabulary and to explain what we can expect from cloud technologies in a near future.

In the third section we introduce some initiatives and Surveys focusing on Cloud computing in education and research.

In the fourth section we sketch technical elements of the cloud ecosystem in order to characterize the main tools used in production around the world. For each main cloud offering, we highlighted strengths, weaknesses and reported on a few case studies where the technologies have been successfully used by higher education and research institutions.

In section five, we provide a synthesis of what is possible to do with clouds, the challenges and methodological milestones.

1.2. Elements of context

1.2.1. Cloud Services Models, SaaS, PaaS, IaaS

According to NIST, the cloud has three service Models, SaaS, PaaS, IaaS [1].

SaaS: Software as a Service

"Cloud application services, or Software as a Service (SaaS), represent the largest cloud market and are still growing quickly. SaaS uses the web to deliver applications that are managed by a third-party vendor and whose interface is accessed on the clients' side. Most SaaS applications can be run directly from a web browser without any downloads or installations required, although some require plugins.

Because of the web delivery model, SaaS eliminates the need to install and run applications on individual computers. With SaaS, it's easy for enterprises to streamline their maintenance and support, because everything can be managed by vendors: applications, runtime, data, middleware, OSes, virtualization, servers, storage and networking.

Popular SaaS offering types include email and collaboration, customer relationship management, and healthcare-related applications. Some large enterprises that are not traditionally thought of as software vendors have started building SaaS as an additional source of revenue in order to gain a competitive advantage.

SaaS Examples: Google Apps, Salesforce, Citrix GoToMeeting, Cisco WebEx, Microsoft Office 365

PaaS: Platform as a Service

Cloud platform services, or Platform as a Service (PaaS), are used for applications, and other development, while providing cloud components to software. What developers gain with PaaS is a framework they can build upon to develop or customize applications. PaaS makes the development, testing, and deployment of applications quick, simple, and cost-effective. With this technology, enterprise operations, or a third-party provider, can manage OSes, virtualization, servers, storage, networking, and the PaaS software itself. Developers, however, manage the applications.

Enterprise PaaS provides line-of-business software developers a self-service portal for managing computing infrastructure from centralized IT operations and the platforms that are installed on top of the hardware. The enterprise PaaS can be delivered through a hybrid model that uses both public IaaS and on premise infrastructure or as a pure private PaaS that only uses the latter.

Similar to the way in which you might create macros in Excel, PaaS allows you to create applications using software components that are built into the PaaS (middleware). Applications using PaaS inherit cloud characteristic such as scalability, high-availability, multi-tenancy, SaaS enablement, and more. Enterprises benefit from PaaS because it reduces the amount of coding necessary, automates business policy, and helps migrate apps to hybrid model.

PaaS Examples: Azure Web Apps, AWS Elastic Beanstalk, Google App Engine

IaaS: Infrastructure as a Service

Cloud infrastructure services, known as Infrastructure as a Service (IaaS), are self-service models for accessing, monitoring, and managing remote datacenter infrastructures, such as compute (virtualized or bare metal), storage, networking, and networking services (e.g. firewalls). Instead of having to purchase hardware outright, users can purchase IaaS based on consumption, similar to electricity or other utility billing.

Compared to SaaS and PaaS, IaaS users are responsible for managing applications, data, runtime, middleware, and OSes. Providers still manage virtualization, servers, hard drives, storage, and networking. Many IaaS providers now offer databases, messaging queues, and other services above the virtualization layer as well. Some tech analysts draw a distinction here and use the IaaS+ moniker for these other options. What users gain with IaaS is infrastructure on top of which they can install any required platform. Users are responsible for updating these if new versions are released"[1].

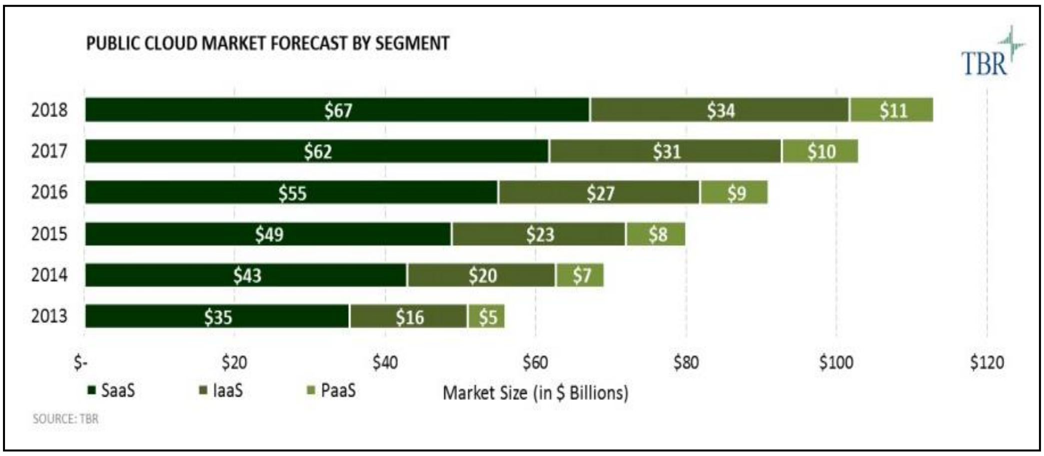
(IaaS Examples: Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE),

1.2.2. Cloud Market Evolution

Public Cloud market experienced a tremendous growth over the past few years and will continue to do so. According to TBR, the market is projected to grow from $80B in 2015 to $112B in 2018 as shown in fig.1.1. The SaaS segment will continue to outrun the IaaS and PaaS segment with about

60% of total market in 2018. Cloud applications will account for 90% of worldwide mobile data traffic by 2019, compared to 81% at the end of 2014.

Figure1.1. Public Cloud market



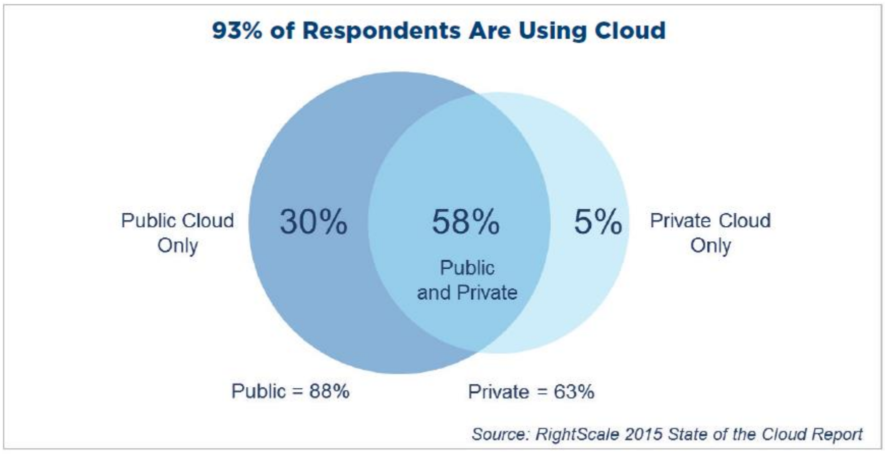
Source: [2]-Roundup Of Cloud Computing Forecasts And Market Estimates Q3 Update, 2015 – Forbes, http://www.forbes.com/sites/louiscolumbus/2015/09/27/roundup-of-cloud-computing-forecasts-and-market-estimates-q3-update-2015/print/

The market is dominated by a few major global players such as Amazon AWS, Microsoft Azure, Google, Rackspace public cloud etc.).

1.2.3. Cloud Usage Evolution

Cloud adoption is spreading throughout all segments of business, government, education, health care, entertainment etc. A recent survey conducted by RightScale[3] In January 2015 provides significant insights about strategies, rates of cloud adoption, types of usage, choices providers adopted by a sample of 930 organizations. The sample covers a broad spectrum of organizations belonging to various sectors of activities: Software (29%), Business services (4%), Advertising and Agencies (4%), Media Publishing (5%), Financial Services (6%), Education (7%), Tech Services (16%) and others (29%). Among the key results, it is worth noting that 93 percent of organizations surveyed are running applications or experimenting with IaaS infrastructure, 82 % have a hybrid cloud strategy up from 74% in 2014. The trend towards public clouds massive adoption is pointed out.

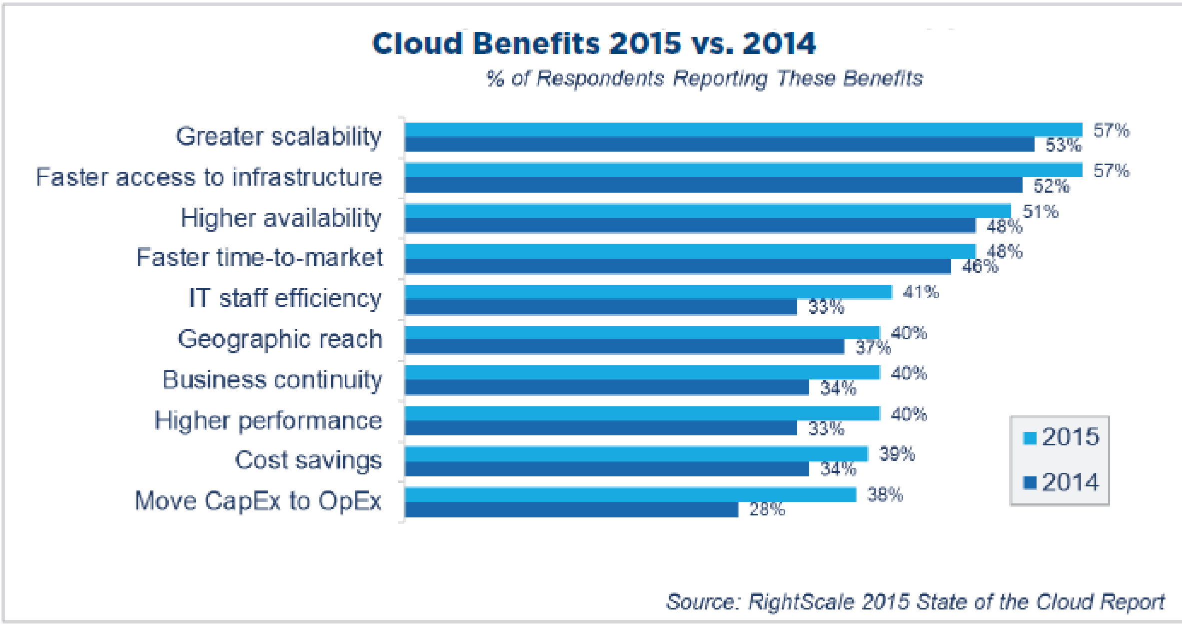
Figure 1.2 Cloud strategy



Among the cloud providers Amazon AWS has by far the largest market share with 57% of running applications, followed by Microsoft Azure (21%) , Google (13%) , Rackspace (11%) etc.

While there is a clear shift towards the use of Public clouds, a large percentage of organizations still run most of their workloads on private clouds. Applications running on the Cloud are mainly: Development and tests, Websites, Greenfield, Marketing campaigns and to a lesser extent Legacy applications.
Respondents reported that "scalability, faster access to infrastructure, and availability" are the top three benefits experienced using cloud, see Fig. 1.3

Figure 1.3 Cloud Benefit



Respondents regarded "security" and "lack of cloud resources and expertise" as the top two significant challenges with practically an equal weight.

When it comes to IT governance, IT departments are taking the lead in cloud spending decisions. They are increasingly offering self-service portals as a central hub to access cloud services. 43% of the respondents reported to have a portal, 41% are planning or developing a portal while 16% have no plan. Overall DevOps adoption is growing and reached 66%. The adoption of Docker is experiencing a huge start.

Even though the RightScale survey sample contains 7% of organizations in education, it doesn't provide any specific results for this category. However, if we refer to an older survey ran in the USA by CDW in 2013, we notice that large businesses and higher education lead on cloud adoption. Moreover the rate of adoption (organizations implementing or maintaining cloud computing) for higher education increased significantly from 34% in 2011 to 43% in 2012 [4].

1.2.4. Cloud and Data Science

Data science is a new discipline that relies on computers and mathematics, particularly statistics to extract information from data. The term big data rather characterizes the data as data/information and less as scientific knowledge extraction methods. In data science, researchers rely on data mining, statistics, signal processing, learning and data visualization. Each of these disciplines produce and exchange data. The production sites are geographically distant implying movement and accommodation of data between sites. Cloud computing is the key technology to implement data science when it provides the analysis, extracting, visualization tools as a service. Cloud computing is the key technology if it allows the data scientist to choose and to configure its own tools according to its will and without any system administrator intervention. Given the growing importance of data science, Data science curricula have been added by thousands of universities around the globe. The modern requirement of mastering data science disciplines and big data techniques and technologies gives the cloud a cornerstone position and makes its adoption by research and education institutions a necessity.

1.3. Cloud computing in education and research, initiatives and Surveys

The UK, through the e-Science program, the US through the NSF-funded cyber infrastructure and the European Union through the ICT Calls aimed to provide "the technological solution to the problem of efficiently connecting data computers, and people with the goal of enabling derivation of novel scientific theories and knowledge" [5].

The Grid [6] [7], foreseen as a major accelerator of discovery, didn't meet the expectations it had excited at its beginnings and was not adopted by the broad population of research professionals. The Grid has been a good tool for particle physicists and it has allowed them to tackle the tremendous computational challenges inherent to their field. However, as a technology and paradigm for delivering computing on demand, it doesn't work and it can't be fixed. On one hand, "the abstractions that Grids expose – to the end-user, to the deployers and to application developers – are inappropriate and they need to be higher level" [8], and on the other hand, academic Grids are inherently economically unsustainable. They can't compete with a service outsourced to the Industry whose quality and price would be driven by market forces.

The virtualization technologies and their corollary, the Infrastructure-as-a-Service (IaaS) style cloud, hold the promise to enable what the Grid failed to deliver: a sustainable environment for computational sciences that would lower the barriers for accessing federated computational resources, software tools and data; enable collaboration and resources sharing and provide the building blocks of a ubiquitous platform for traceable and reproducible computational research. Amazon Elastic Compute Cloud, Microsoft Azure and Google Compute Engine are examples of Infrastructure-as-a-Service that anyone can use today. Their considerable success announces the emergence of a new era. However, bringing that era for research and education still requires new technical and organizational frameworks to bridge the gap between the researchers and educators' current tools and practices and the tremendous capabilities that modern clouds expose. The National Science Foundation and the European Commission have put in place strategies geared toward creating the missing links and unleashing fully the cloud technologies potentialities.

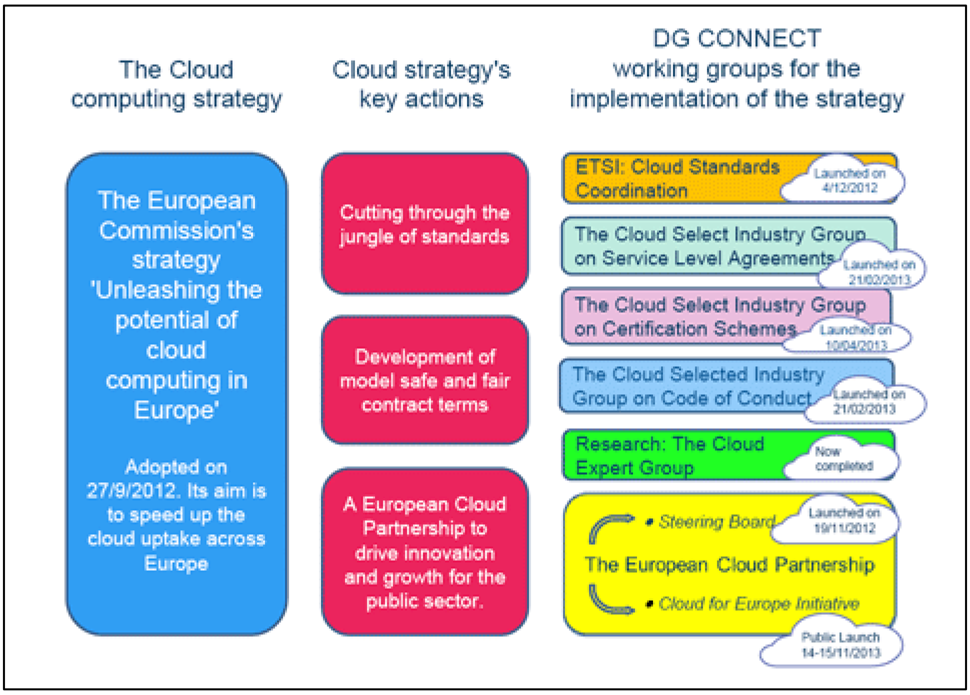
1.3.1. European Commission's initiatives in Cloud Computing

In September 2012, the European Commission adopted a strategy for "Unleashing the Potential of Cloud Computing in Europe" [9]. The strategy, see Figure 1.4, outlines actions to deliver a net gain of 2.5 million new European jobs, and an annual boost of €160 billion to the European Union GDP (around 1%), by 2020. The strategy is designed to speed up and increase the use of cloud computing across all economic sectors. This strategy is the result of an analysis of the overall policy, regulatory and technology landscapes and of a wide consultation with stakeholders, to identify ways to maximize the potential offered by the cloud.

The European Commission Directorate General for Communications Networks, Content & Technology (DG Connect) manages the Digital Agenda of the EU [9a]. It launched at the start of 2013 a web-based public Consultation with a view to defining future research priorities in Cloud Computing, Software and Services, ahead of the H2020 ICT Work Program 2014-15 [10]. The call "e-Infrastructures" in the work program that followed [11] [12] [13] stressed a number of priorities including (i) Integrating e-infrastructure resources and services across all layers (networking, computing, data, software, user interfaces), in order to provide seamless services tailored to user needs. (ii) Implementing the e-infrastructure to ride the wave of "big data" (iii) Providing support to the e-infrastructure for Open Access (iv) Implementing the e-infrastructure part of the EU strategy on High Performance Computing (HPC)

Under the flagship of FP7, the e-Infrastructures activity of the EU is part of the Research Infrastructures program, funded under the FP7 'Capacities' Specific Program. It focuses on the development and evolution of the high-capacity and high-performance communication network (GÉANT), distributed computing infrastructures (grids and clouds), supercomputer infrastructures, simulation software, scientific data infrastructures, e-Science services as well as on the adoption of e-Infrastructures by user communities. From our point of view, this program related to infrastructure is the more relevant and overlaps some of our concerns. In such calls, people may apply on the two subjects of cloud adoption and migration, from a technical point of view and maybe from all the others concerns: human resource impact, processes…

Figure 1.4 The 2012 EU Agenda



Source: https://ec.europa.eu/digital-agenda/en/european-cloud-computing-strategy

At the research level, the first Cloudway workshop was held in Italy on September 2015 [13 a]. The workshop is a forum for discussions about Cloud adoption and migration. The topics cover a broad spectrum of fields, including processes, patterns and frameworks; Empirical studies; Crosscutting concerns; Architecture evolution, adaptation and transformation; Infrastructure and data challenges; Hybrid and multi-cloud migration; Migrating to cloud using micro services; DevOps process and cloud migration. In short, it covers all the research and practical aspects for the migration toward cloud technologies. The first edition of the workshop may be considered as a very prospective one that will bring together more and more people in the future.

To sketch a little, under FP7 and H2020, a number of projects have been funded to explore the use of cloud computing in research and education in Europe among them:

1.3.1.1.VENUS-C

VENUS-C is a project funded under the European Commission's 7th Framework Programme drawing its strength from a joint co-operation between computing service providers and scientific user communities to develop, test and deploy a large, Cloud computing infrastructure for science and SMEs in Europe [13 b].

1.3.1.2. Helix Nebula

Helix Nebula is a project funded under the European Commission's 7th Framework Programme. It pioneered a partnership between big science and big business in Europe and aimed at charting the course towards the sustainable provision of cloud computing - the Science Cloud [13 c]. The partnership brings together leading IT providers and three of Europe's leading research centers, CERN, EMBL and ESA in order to provide computing capacity and services that elastically meet big science's growing demand for computing power. The CERN & HNI project Co-Coordinator published a paper [14] describing his vision for European Open Science Cloud.

1.3.1.3. GEANT

GÉANT is the pan-European data network for the research and education community. It interconnects national research and education networks (NRENs) across Europe, enabling collaboration on projects ranging from biological science to earth observation and arts & culture [14 a]. The GÉANT project combines a high-bandwidth, high-capacity 50,000 km network with a growing range of services. These allow researchers to collaborate, working together wherever they are located. Services include identity and trust, multi-domain monitoring perf SONAR MDM, dynamic circuits and roaming via the eduroam service. Together with European NRENs, GÉANT connects 50 million users in over 10,000 institutions. Through links to research networks in other regions (such as Internet2 and ESnetin in the USA, AfricaConnect in Africa, TEIN in Asia-Pacific and RedCLARA in Latin America), GÉANT enables collaboration between researchers in over half the world's countries. Co-funded by the European Commission and Europe's NRENs, the GÉANT network was built and is operated by DANTE. The GÉANT project is a collaboration between 41 partners: 38 European NRENs, DANTE, TERENA and NORDUnet (representing the five Nordic countries), and 30 Open Call project partners.

The Geant Cloud Services' [15] aim is to establish a place for NRENs to collaborate on clouds, share experiences, aggregate demand, to optimally position the NRENs to play an active role with respect to the rapidly developing cloud paradigm. The Support to Clouds activity enables NRENs to deliver cloud and mobile services to their communities with the right conditions of use, so that the benefits of the cloud can be fully realized and the attendant risks appropriately managed.

1.3.1.4.EUDAT2020

EUDAT2020[16] brings together a unique consortium of e-infrastructure providers, research infrastructure operators, and researchers from a wide range of scientific disciplines under several of the ESFRI themes, working together to address the new data challenge. In most research communities, there is a growing awareness that the “rising tide of data” will require new approaches to data management and that data preservation, access and sharing should be supported in a much better way.

1.3.1.5. EUBrazil Cloud Connect

EUBrazil Cloud Connect [17]is a project funded under the second EUBrazil coordinated call under the topic a) Cloud computing for Science. Its aim is to create an intercontinental federated eInfrastructure for scientific usage. This e-Infrastructure will join resources from different frameworks, like private clouds, supercomputing and opportunistic desktop resources to offer the community high level scientific gateways and programming models. The overarching objective of EUBrazil Cloud Connect is to drive cooperation between Europe and Brazil by strengthening the scientific and knowledge-based society as key to sustainable and equitable socioeconomic development. The core of this collaboration is defined through the scientific uses cases selected, which will require the collaboration between Brazil and Europe in the provision of data, services and expertise.

1.3.2. National Science Foundation (NSF) initiatives in cloud computing

Academic research has contributed significantly to the development of the cloud architectures, however much of the recent innovation in this domain has been driven by the industry. Academic researchers are now considering a new generation of innovative applications of cloud computing architectures that the Industry made available. In August 2014, The NSF has announced the funding of two $10 million projects “Chameleon” [18] and “CloudLab” [19] to create cloud computing testbeds. Through these two projects, the NSF’s Directorate for Computer and Information Science and Engineering (CISE) goal is to enable a new future for cloud computing. Its aims is to transform the current US research environment by supporting research infrastructure, enabling transformative research at the frontiers of computing and providing unique opportunities for current and future generations of computing researchers and educators.

1.3.2.1. Chameleon

Chameleon is a large-scale, reconfigurable experimental environment for cloud research, co-located at the University of Chicago and The University of Texas at Austin. The project’s infrastructure will consist of: 650 cloud nodes with 5 petabytes of storage with reconfigurable slices with the support both “bare-metal-access” and virtualization technologies. The project’s aims to (i)Allow testing of a range of problems ranging from machine learning and adaptive operating systems to climate simulations and flood prediction (ii)Allow the experimentation with new virtualization technologies that could improve reliability, security and performance (iii) Create a social network for researchers to share solutions and discuss new ideas (iv) Support heterogeneous computer architectures such as low power processors, general processing units (GPUs) etc.

1.3.2.2. CloudLab

CloudLab is a large-scale distributed infrastructure based at the University of Utah, Clemson University and the University of Wisconsin. The project’s infrastructure will consist of: 15,000 processing cores and in excess of 1 petabyte of storage at its three data centers with 100 gigabit-per-second connections on Internet2's advanced platform between its different sites. The project aims to (i)Provide a facility where researchers can build their own clouds and experiment with new ideas with complete control, visibility and scientific fidelity (ii) Enable new applications with direct benefit to the public in areas of national priority such as real-time disaster response or the security of private data like medical records (iii) Integrate diverse cutting-edge platforms for research that will be provided by HP, Cisco and Dell (iv) Support Open Flow (an open standard that enables researchers to run experimental protocols in campus networks) and other software-defined networking technologies.

Prior to Chameleon and CloudLab, NSF has funded and supported the XSEDE Project which brings advanced Cyber infrastructure, digital services and expertise to American Scientists and Engineers. **XSEDE is currently one of most advanced, powerful and robust collection of integrated advanced digital resources and services in the world.**

1.3.2.3. The XSEDE project and the XSEDE report on cloud adoption

The Extreme Science and Engineering Discovery Environment (XSEDE) as a project is a five-year, $121 million grant award made by the NSF to the National Center for Supercomputing Applications NCSA at the University of Illinois and its partners [20]. XSEDE is a virtual organization that provides a dynamic distributed infrastructure, support services, and technical expertise, enabling researchers, engineers, and scholars to address the most important and challenging problems facing the USA and world. XSEDE supports a growing collection of advanced computing, high-end visualization, data analysis, and other resources and services.

XSEDE is a successor to the NSF-funded TeraGrid project, which itself succeeded the NSF supercomputer center program that began in the 1980s. As a set of services, XSEDE integrates supercomputers, visualization and data analysis resources, data collections, and software into a single virtual system for enhancing the productivity of scientists, engineers, social scientists, and humanities experts. XSEDE facilitates scientific discovery by providing common authentication and security mechanisms, global namespace and file systems, remote job submission and monitoring, file transfer services, advanced support services, and a user portal designed to help researchers work as efficiently and effectively as possible.

XSEDE's suite of advanced digital services connects with other high-end facilities and campus-based resources around the US, serving as the foundation for a national computing ecosystem. Additionally, XSEDE's architecture allows open development for future digital services and enhancements.

A National Science Foundation-sponsored cloud user survey[1] was conducted from September 2012 to April 2013 by the XSEDE Cloud Integration Investigation Team to better understand how cloud is used across a wide variety of scientific areas and the humanities, arts, and social sciences. Let's now quote the main facts and results of the report [21].

Data was collected from eighty projects, representing a family of cloud users from twenty-one science and engineering disciplines and the humanities, arts, and social sciences. Quantitative dimensions of cloud usage (number of cores used peak/steady state, bandwidth in/out of the cloud, amount of data stored in the cloud, etc.) and qualitative experiences (the benefits and challenges of using the cloud) were explored.

The top three reasons found by XSEDE that explain that survey participants used the cloud are: (1) on-demand access to burst resources, (2) compute and data analysis support for high throughput scientific workflows, and (3) enhanced collaboration through the rapid deployment of research team web sites and the sharing of data.

MapReduce, a framework popularized by the Apache Hadoop implementation, was the most heavily used special feature offered by the cloud service providers, followed by access to community datasets.

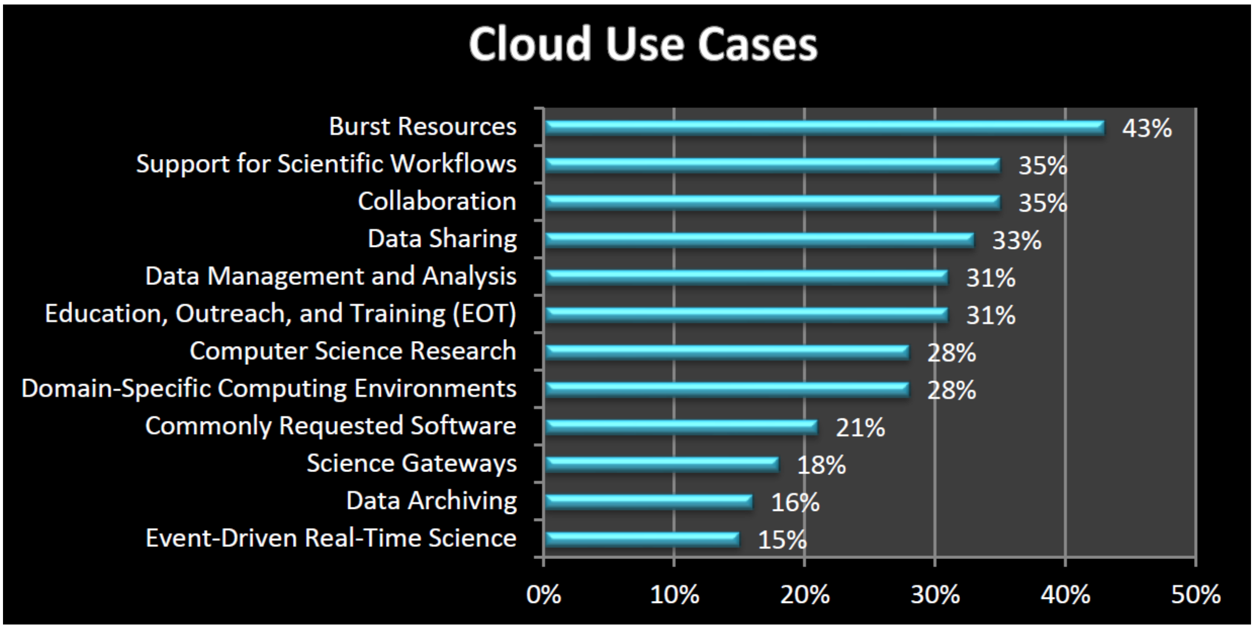
For XSEDE, application and programming models considered as good candidates for the cloud are: high throughput, embarrassingly parallel workloads; academic labs and teaching tools; domain-specific computing environments; commonly requested software; science gateways; and, real-time event-driven science.

Cloud benefits identified by the survey participants were: pay as you go, lower costs, compute elasticity, data elasticity, Software as a Service, Education as a Service, broader use, scientific workflows, rapid prototyping, and data analysis.

Cloud challenges identified in the NSF-XSEDE report included *the learning curve, virtual machine performance, variability in bandwidth, memory limits, database instability, private/public cloud interoperability, security, data movement, storage, and cloud computing cost and the funding availability*.

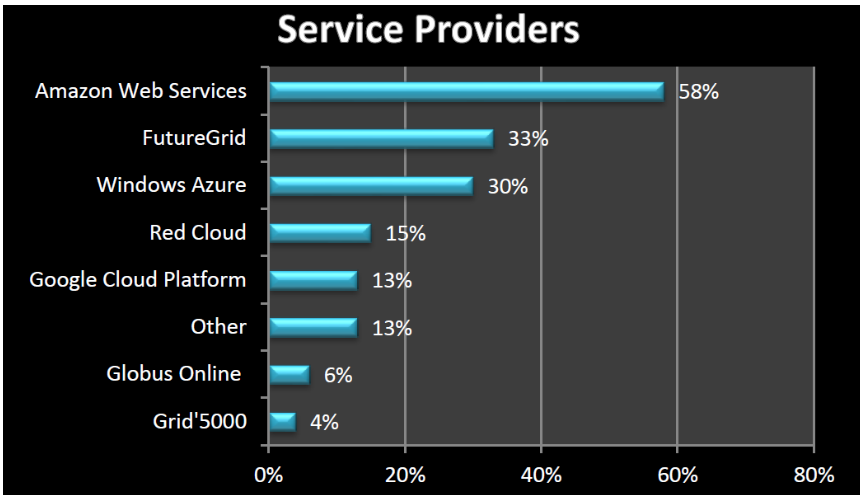
To the question "What cloud use cases are represented by your research or education project?" The participants to the XSEDE Survey [22] gave answers summarized by the graphics below on Figure Y.

Figure 1.5 XSEDE survey (illustrative example about use cases)



In order to address those use cases, different cloud services providers were used (see Figure Z).

Figure 1.6 Cloud providers studied in the XSEDE survey



The survey's results show a clear dominance of Amazon Web Services, followed by Azure and the Google Cloud Platform. Those results are in line with Gartner's latest Magic Quadrant [3] which covers all the common use cases for cloud IaaS.

1.3.3. French initiatives in Cloud Computing

Université Sorbonne Paris Cite (USPC) is one of the major French university consortium. It gathers four Parisian universities and four institutes for higher education and research:

- Sorbonne Nouvelle University (Paris 1);
- Paris Descartes University (Paris 5);

- Paris Diderot University (Paris 7);
- Paris 13 University;
- Institut de Physique du Globe de Paris (Institute of Earth Physics of Paris - IPGP);
- Institut National des Langues et Civilisations Orientales (National Institute of Eastern Languages and Civilizations - INALCO);
- Institut d'Etudes Politiques de Paris (Institute for Political Sciences - Sciences Po);
- Ecole des Hautes Etudes en Sciences Sociales (School for Social Sciences and Public Health - EHESS).

In addition to 120.000 students, USPC hosts numerous research teams most often also affiliated to the french public research institutions 'Centre de la Recherche Scientifique' (CNRS), 'Institut National de la Sante et de la Recherche Medicale' (INSERM), 'Institut de Recherche pour le Developpement' (IRD), 'Commissariat a l'Energie Atomique' (CEA). USPC puts strong emphasis on the quality of training and on research development, and gathers approximately 17.000 researchers, teachers, engineers and technicians. The USPC infrastructures dedicated to experimental sciences present strengths and weaknesses that directly arise from the wide internal diversity. Some institutes greatly call for scientific computing, have access to national super-computing facilities, and have been using cluster/grid computing for a while. This is the case for IPGP but it is far less developed for practitioners from life and health sciences, that constitute about half of the total number of researchers at USPC.

The IDV program[23]is a multidisciplinary and interdisciplinary program centered on the development and use of life imaging. It was founded at the end of 2014 for a 3-4 years period, and currently gathers approximately 200 researchers from USPC. USPC partly supports the operational expenses of the program, and the costs of some human resources, such as master and PhD students, postdocs, or support engineers.

The IDV members set up a survey targeting the business processes of the IDV members. The 200 members belong to 30 different teams from research laboratories, imaging core facilities and departments from 14 hospitals. This community is very diverse in terms of initial training, background and practice. The following disciplines are represented among the members: physics, chemistry, biology, medicine, psychology, pharmacy, mathematics & computer sciences for image processing, data sciences and humanities like law and applied ethics. We built a questionnaire with the aim to provide an overview of the various scientific activities, methods and tools used by the IDV members, at the beginning of the project. We also meant at depicting the life cycle(s) of the data used in the daily work of the community. The answers from participating members were collected using the Sphinx Online (v 3.1.2) platform, and results presented below obtained during a seventeen-week period covering questioning to final analysis.

The IDV community can roughly be divided into 2 populations: on one hand, a 'biologist population', for whom experiences, observation and statistical analyzes are central to the research process; on the other hand, a "computer scientist population", who very routinely uses theoretical calculations, models and numerical simulations. Respectively half of the IDV members, and more than one third, work in teams from 3 to 5 and 5 to 10 researchers respectively. Collaboration among scientists from different teams or even laboratories is widespread practice. More than 70% of answers indeed confirmed that they needed sharing data with other teams in their daily work.

The IDV survey[23]also focus and characterize the Image formats and software used by the IDV members and the data characteristics, storage location and data preservation (if any) managed by members. The current practices do not allow easy nor efficient data sharing among scientists and laboratories. The spatial fragmentation leads to frequently under-exploited data, and hard drive crashes often generate data loss. Only few IDV members use online solutions, i.e. clouds (around 10% of collected answers). Long term data preservation is far from current practices, with approximately one half of all the data produced by the IDV members deleted within the 10 years following production.

Note that this survey has been made possible because IDV members funded an engineer position to manage it. Following this example, USPC has decided in July 2015 to fund a position for a research engineer to disseminate the best practices with clouds, to promote usages in particular towards the population of researchers and teachers that are not coming from the computing fields. This person is also in charge of the coordination of white papers, for instance on the topic related to hosting data from the technical, ethical, legal point of views. This specific white paper also focuses on privacy, for instance for the data from patients in hospitals, and makes a room for relevant use cases in using clouds in this context.

It is to be mentioned that IDV members and other people are currently working on Data hosting problems, another aspect of the problems faced by large scale infrastructures (cluster, grids and clouds). The document in preparation[1]will deal with the problems related to data hosting especially when cloud computing technologies are used. It was written in non-technical language in order to be accessible to an audience of neophytes.

It is intended primarily for an audience of higher education and research people so that they can understand the fundamentals and be aware of the challenges. Authors assume in fact that a university, for example, needs elements to define a data hosting policy inside a private cloud. Authors offer them a complete panorama, a collection of objective and factual information that are proposed, discussed, commented on in the underlying ecosystems, public or private. Particular incision is made on the ethical and legal issues.

The document in preparation may be instructive to anyone not belonging to the world of higher education and research but wishing to get an idea about the challenges and solutions in terms of hosting and massive data management. Authors also hope that the reader will find a holistic approach and methodological elements on the topics of migration and adoption in our daily lives of cloud computing technologies

Moreover, USPC (Université Sorbonne Paris Cité) conducted in 2015-2016 [23] a similar survey to the IDV one targeting everyone at USPC. The survey was devoted to isolate trends in the practices and also to identify communities of users. About 250 people answered the survey.

Concerning the utilization of platforms for computing, analyzing or storage purposes, 62.39% of respondents use a local platform, 21.37% a platform in their institutions, 12.82% use a platform in another institution, 12.82% use a national infrastructure, 3.42% an international infrastructure.

[1]Contact information: christophe.cerin@lipn.univ-paris13.fr

Concerning the long term archiving (> 10 years), 69.62% of responders use a local system (hard disk), 19.62% an Intranet solution and 10.76% an Internet solution. 33.6% of respondents needs to store less than 100GB of data per year and 9.6% of respondents need to save more than 5TB of data per year.

The collaborative tools used by the USPC community is (per percentage of responders) Emails (80.3%) , Dropbox (69.3%), Google Drive (46.2%), Google Docs (42%), Visio-conferencing tools (33.6%), Shared agenda (28.2%), Wiki (10.9%), Social networks (9.2%). More than 30 other tools are used, among them Owncloud (5%), Git (5%), MyCoRe (3%).

The need for training is very important and was expressed by 64.8% of respondents.

Most answers were quite expected but they revealed few issues that any organization should address for a better usage of its infrastructures, for instance to fight against fragmentation of data. The survey shows the importance of starting from the current usage and supporting end-users via dedicated and specific trainings.

At least in France, the Conference des Presidents d'Universite (CPU) formulated in May 2015 ten general proposals for the digital area, among which are: *Establishing infrastructure to deal with public data produced by research, teaching and training; Organizing open and participative sciences with data repositories for research and innovation; Laying the groundwork for a new work organization that is generated by the digital Era*.

1.4.Cloud computing in education and research, overview of technologies, providers and use cases

1.4.1.Elements about the Public Cloud Ecosystem

We start by descriptions of the main public cloud vendors from the Gartner report [24] complemented with descriptions of case studies from universities and research centers.

1.4.1.1.Amazon Web Services

"Amazon Web Services (AWS), a subsidiary of Amazon.com, is a cloud-focused service provider with a very pure vision of highly automated, cost-effective IT capabilities, delivered in a flexible, on-demand manner.

1.4.1.1.1. Offerings

AWS offers Xen-virtualized multitenant and single-tenant compute, with multitenant storage, along with extensive additional IaaS and PaaS capabilities, including object storage with an integrated CDN (Amazon S3 and CloudFront) and a Docker container service (EC2 Container Service). It is willing to negotiate large-scale single-tenant and on-premises deals (such as the U.S. intelligence community cloud deal). The AWS Marketplace has an extensive selection of third-party software and services. Enterprise-grade support is extra. It has a multi-fault-domain SLA. Colocation needs are met via AWS Direct Connect.

1.4.1.1.2. Strengths

- *AWS has a diverse customer base and the broadest range of use cases, including enterprise and mission-critical applications. It is the overwhelming market share leader, with over 10* *times more cloud IaaS compute capacity in use than the aggregate total of the other 14 providers in this Magic Quadrant. This has enabled it to attract a very large technology partner ecosystem that includes software vendors that have licensed and packaged their software to run on AWS, as well as many vendors that have integrated their software with AWS capabilities. It also has an extensive network of partners that provide application development expertise, managed services, and professional services such as data center migration.*

- *AWS is a thought leader; it is extraordinarily innovative, exceptionally agile, and very responsive to the market. It has the richest array of IaaS features and PaaS-like capabilities. It continues to rapidly expand its service offerings and offer higher-level solutions. Although it is beginning to face more competition from Microsoft and Google, it retains a multiyear competitive advantage. Although it will not be the ideal fit for every need, it has become the "safe choice" in this market, appealing to customers who desire the broadest range of capabilities and long-term market leadership. It is the provider most commonly chosen for strategic adoption.*

1.4.1.1.3 Cautions

- *AWS can be a complex vendor to manage. Customers must ensure that they receive the level of sales and solution architecture engagement they need to be successful. AWS is a price leader, but it charges separately for optional items that are sometimes bundled with competing offerings; use of third-party cost management tools, such as RightScale Cloud Analytics and Cloudability, is highly recommended. AWS's support offerings are tiered based on the level of support that a customer purchases, rather than on a "relationship" or size-of-spend basis; customers need Business-tier support in order to ensure excellent support.*
- *AWS is spreading its efforts very broadly. Although many new services are highly successful, services that turn out to be of less interest to customers will not get the same depth of continued investment as more popular services. As AWS expands, it increasingly encroaches on the territory of traditional IT vendors, heightening competitive pressure and the need to invest deeply in engineering efforts. Furthermore, new capabilities often compete with products and services from AWS partners; though this is normally positive for customers, it creates ecosystem conflicts that AWS must continue to manage carefully"*[24].

1.4.1.1.4. Research and Education Case Studies

1.4.1.1.4.a The University of Berkeley [25]

The Algorithms, Machine, and People (AMP) Lab at the University of California Berkeley is a multi-disciplinary research effort designed to build scalable machine learning and data analysis technology. Among the many experiments run by the AMP Lab, one area of concentration is in the field of genomics and cancer research. Due to the vast amount of data that genome sequencing produces, the AMP Lab leverages AWS to quickly scale the compute resources needed to analyze the algorithms that are used in genomics work. As a result, researchers are able to use many machines in the cloud simultaneously to process genome data faster and more cost effectively.

1.4.1.1.4.b Anhanguera [26]

Founded in 1973, Anhanguera is one of the largest universities in Brazil and one of the largest in the world, offering over 500,000 undergraduate and postgraduate courses. The university has over 70 campuses throughout Brazil, and hundreds of offices supporting distance learning. Anhanguera University has more than 10,000 teachers and over 6,000 administrative staff.

Anhanguera is one of the largest users of Moodle, an e-learning platform for collaborative learning. When it needed to scale the platform to meet user demand, the university leveraged AWS to build a solution that would deliver applications reliably and grow with the user population.

Anhanguera decided to work with Dedalus, a premier consulting member of the Amazon Partner Network (APN), to build a cloud environment using Amazon Web Services for increased reliability and scalability. After provisioning Amazon Virtual Private Cloud (Amazon VPC) to create a closed environment within the AWS Cloud, Anhanguera and Dedalus migrated Moodle and other applications to AWS. The university uses Amazon DynamoDB for fast retrieval of e-learning content at any level of traffic. Amazon Relational Database Service (Amazon RDS) with Multi-AZ provides failover capability, improving reliability for Moodle and other applications.

1.4.1.1.4.c University of Maryland University College [27]

University of Maryland University College (UMUC) is an open-access university serving working adult students pursuing higher education through on-site and online courses. Most students of UMUC (80%) take their courses online. UMUC is leveraging a broad array of products from the AWS platform including EC2 for virtual machine instances, RDS oracle for databases and Amazon redshift for analytics to derive insight from four core datasets totaling 10 terabytes from its Learning Management System, Student Information System, Financials and Customer Relationship System (CRM). By replacing its legacy applications with a new cloud computing analytics platform, UMUC is able to analyze data much more efficiently than before and make sense of patterns to understand enrolments trends and demand for academic programs, measure student and faculty engagement in online course, analyze patterns for re-enrolment and course taking behaviors, identify at risk students that need assistance and analyses the effectiveness of its marketing campaigns.

1.4.1.1.4.d University of Western Australia [28]

The Centre for Software Practice (CSP) is a research and development body within the Faculty of Engineering, Computing and Mathematics at the University of Western Australia (UWA). Formed in 2006, the center creates practical learning opportunities for students studying software engineering at UWA and conducts research into the impact of technology on online communities, open source development, and health informatics.

In 2012, after partnering with Stanford University to obtain a massive open online course (MOOC) platform, CSP created Class2Go to deliver multiple online courses and assessments over the web. Then the center needed an infrastructure capable of managing and delivering course assets—including streaming video, online quizzes, and practical coding assessments—to several thousand online participants. The infrastructure had to be able to scale to accommodate the delivery of new courses and enable participant interaction. In addition, the CSP needed to fund Class2Go without assistance from third parties, meaning it had a limited budget to spend on infrastructure.

CSP was able to deploy Class2Go on AWS in a few hours and launch the online platform at less than 10% the cost of previous, similarly sized projects.

1.4.1.2.Microsoft

"Microsoft is a large and diversified technology vendor that is increasingly focused on delivering its software capabilities via cloud services. Its Azure business was previously strictly PaaS, but Microsoft launched Azure Infrastructure Services (which include Azure Virtual Machines and Azure Virtual Network) into general availability in April 2013, thus entering the cloud IaaS market.

1.4.1.2.1. Offerings

Microsoft Azure offers Hyper-V-virtualized multitenant compute (Virtual Machines), with multitenant storage, along with many additional IaaS and PaaS capabilities, including object storage (Blob Storage) and a CDN. The Azure Marketplace offers third-party software and services. Enterprise-grade support is extra. It has a multi-fault-domain SLA. Colocation needs are met via Azure ExpressRoute. Microsoft's SaaS offering includes Office 365 which provides all the Microsoft office tools on the cloud, therefore enabling collaboration and online and ubiquitous access. According to Microsoft's COO, one in four Microsoft's enterprise Customers adopted Office 365. Offices 365 tools include Word Online, Excel Online, People, OneNote Online, PowerPoint Online, Calendar, Excel Online and OneDrive the place where the user can save all office files.

1.4.1.2.2. Strengths

- *Microsoft Azure encompasses integrated IaaS and PaaS components that operate and feel like a unified whole. Microsoft has been rapidly rolling out new features and services, including differentiated capabilities. It has a vision of infrastructure and platform services that are not only leading stand-alone offerings, but also seamlessly extend and interoperate with on-premises Microsoft infrastructure (rooted in Hyper-V, Windows Server, Active Directory and System Center), development tools (including Visual Studio and Team Foundation Server), and applications, as well as Microsoft's SaaS offerings.*
- *Microsoft's brand, existing customer relationships, history of running global-class consumer Internet properties, deep investments in engineering, and aggressive roadmap have enabled it rapidly to attain the status of strategic cloud IaaS provider. Microsoft Azure is growing rapidly, and is in second place for market share, with more than twice as much cloud IaaS compute capacity in use as the aggregate total of the remaining providers in this Magic Quadrant (excluding market share leader AWS). Microsoft has pledged to maintain AWS-comparable pricing for the general public, and, on a practical level, customers with Microsoft Enterprise License Agreement discounts obtain a price/performance ratio that is comparable to AWS.*

1.4.1.2.3. Cautions

- *Microsoft has previously reliably met its promised time frames for introducing critical features that help Azure fulfill enterprise needs for security, availability, performance, networking flexibility and user management, but it has not finished introducing all such functionality. Customers who intend to adopt Azure strategically and migrate applications over a period of one year or more (finishing in 2016 or later) can begin to deploy some workloads now, but those with a broad range of immediate enterprise needs may encounter challenges. Furthermore, customers express concern about the global impact of many past Azure outages, which may necessitate ensuring that critical applications on Azure have a non-Azure disaster recovery solution.*
- *Microsoft's partner ecosystem is still relatively nascent. It recently launched a software marketplace and has begun aggressively recruiting managed service and professional*

services partners. However, many of these partners lack extensive experience with the Azure platform, which could compromise the quality of the solutions they deliver to customers. Furthermore, the Azure ecosystem is very dependent on existing Microsoft relationships. Although customers do run heterogeneous environments in Azure, this lessens the appeal of Azure to non-Microsoft-centric organizations."[24]

1.4.1.2.4 Research and Education Case Studies

1.4.1.2.4.a Virginia Tech [29]

DNA sequencing analysis can lead to medical and pharmaceutical breakthroughs, but it requires supercomputing resources and Big Data storage that many researchers lack. Through a grant provided by the National Science Foundation and Microsoft, computer scientists at Virginia Tech developed an on-demand, cloud-computing model using the Microsoft Azure HDInsight Service. Researchers now have easier, cost-effective access to DNA sequencing tools and resources, which could lead to faster advancements in medical research.

These cloud-enabled tools allow biologists and bio-informaticians to access their work from any device, collaborate more easily and run analysis much quicker. Thanks to the on-demand nature of cloud computing, costs of genome sequencing dropped from millions of dollars a few years ago to thousands and could potentially drop further to pennies in the near future.

1.4.1.2.4.b Tokyo University of Technology [30]

The Tokyo University of Technology (TUT) was founded in 1986, it has promoted three particular aims: training in use of technologies and expert scientific theory for the betterment of society, engagement in advanced research and passing research findings back to society and creating an ideal educational and research environment.

TUT has opted for a full shift of its ICT environment to the Azure cloud. It has combined usage of cloud services for PaaS and SaaS and effective system operations thanks to building up the school’s core database. Information entered from all systems will be stored in a newly created core database; and through a redeveloped, university-wide, wireless network, it will be possible to use data extensively for CRM (Customer Relationship Management) and for university administration.

1.4.1.2.4.c The University of California, Davis

The University of California, Davis has 33,000 students and multiple undergraduate and graduate colleges on campus, including the College of Agricultural and Environmental Sciences. To support the changing campus technology needs and build streamlined applications for the entire campus, the college deployed Windows Azure. The platform saves the college money, promotes innovative application development, fosters collaboration, provides higher availability, and scales for future use among other state universities.

Initially, the college chose an infrastructure-as-a-service solution, but then changed to Platform-as-a-Service one. The college relies on Azure Cloud Services and Windows Azure Web Sites, as well as Windows Azure SQL Database for its major projects, and Microsoft Office 365. Windows Azure Blob Storage was used for unstructured data and Windows Azure Shared Caching was deployed to ensure consistency for users. With Windows SQL Data Sync, protected data can be pushed into the cloud, and New Relic is an add-on to simplify billing.

The college created two new applications on Windows Azure: an Academic Course Evaluations application for students and a campus wide Purchasing System application for students, faculty, and staff. With Academic Course Evaluations, 33,000 students can quickly and easily assess the faculty and courses online. Nearly 10,000 faculty, staff, and students use the Purchasing System application to buy notebooks, pens, and laptops online.

1.4.1.3. Google

"Google is an Internet-centric provider of technology and services.

1.4.1.3.1. Offerings

Google Cloud Platform combines an IaaS offering (Compute Engine), a PaaS offering (App Engine) and a range of complementary services, including object storage and a Docker container service (Container Engine). Compute Engine VMs are KVM-virtualized and metered by the minute. Enterprise-grade support is extra. It has a multi-fault-domain SLA. Colocation needs are met via Google Cloud Interconnect. Google has been a precursor in providing SaaS services targeting educational institutions. Google apps is a platform for creating, sharing and collaborating on various types of documents and content. The online access makes those apps ubiquitous, students and professors can interact with those apps from anywhere, using any device. Google apps consists of Gmail, Google Chrome, Google Docs, Google Sheets, Google Hangouts, Google Drive, Google Forms, Google Slides, Google Classroom, Google Sites. Google Chrome is a web browser that makes all bookmarks and user's passwords accessible from anywhere provided the user is signed in.

Gmail is an online email service that can be accessed from anywhere with unlimited email storage. Google Calendar is an online calendar to store events, schedules with the ability to send meetings invitations, create tasks, and share the calendar with other users. Google Docs is an online document creator with the ability to share and co-edit documents. Documents are saved automatically at any time there is a document change. Google Docs is a simplified version of Microsoft Office giving free unlimited access for education. Google Sheets is similar to Excel and allows the creation of collaboration sheets which are fully collaborative and auto-saves the same way as Google Docs. Google Hangouts is a web conferencing tool that allows real-time messages, voice and video chatting. Google Drive is the place where all Google files are stored. It acts as a hard drive for the user's Google apps account. Google forms is a tool which helps create online forms that gather information and store it in the form of a Google sheet. It can be used for tests, questionnaires or surveys. Google Slides is similar to Microsoft's PowerPoint for the creation of slides that are accessible from anywhere. Google Classroom is an online tool that allows teachers to setup classrooms, assign work to students, grade the work and return it to the students. Many universities are using Google Apps. In 2013 about one third of UK universities had already adopted Google Apps.

1.4.1.3.2. Strengths

- *Google's strategy for Google Cloud Platform centers on the concept of allowing other organizations to "run like Google" by taking Google's highly innovative internal technology capabilities and exposing them as services that other companies can purchase. Consequently, although Google is a late entrant to the IaaS market, it is primarily productizing existing capabilities, rather than having to engineer those capabilities from scratch.*
- *Google has a comprehensive vision for, and extensive experience with, how cloud-native applications are developed and managed through the life cycle. It has a fluid notion of the*

boundaries between IaaS and PaaS, along with the spectrum of deployment options from VMs to containers, that will, over time, enable customers to choose their trade-offs between control and automated management. Although many customers currently choose Google for its excellent price/performance value and exceptionally fast VM provisioning, over time, Google will differentiate itself with platform and manageability features, not prices.

1.4.1.3.3. Cautions

- *Although Google has significant appeal to technology-centric businesses, it is still in the rudimentary stages of learning to engage with enterprise and midmarket customers, and needs to expand its sales, solutions engineering and support capabilities. Prospective customers report difficulties in gaining the attention of Google's sales staff and being directed toward appropriate solutions. Furthermore, Google needs to earn the trust of businesses. Google also lacks many capabilities important to businesses that want to migrate legacy workloads to the cloud. Its hybrid cloud strategy is open-source and partner-centric, focused on the ecosystem surrounding Kubernetes, its container cluster management software. Google needs to build an ecosystem around Google Cloud Platform; its partner program is nascent.*
- *Google's cloud IaaS adoption has been driven primarily by cloud-native use cases, including batch computing. Google's short-term focus is on better enabling new cloud-native applications, with less attention being paid to capabilities needed for other workloads. While the offering has been improving steadily, Google's feature release velocity has not been as fast as expected. Google's deep engineering investment could potentially advance its offering much more rapidly in the future. Google is not yet taking full advantage of its potential opportunity with Google Cloud Platform"*[24].

1.4.1.3.4 Research and education case studies

1.4.1.3.4.a Autism-speaks [31]

Autism Speaks is an autism advocacy organization in the United States that sponsors autism research and conducts awareness and outreach activities aimed at families, governments, and the public.

The Collaboration project MSSNG the organization has with Google helped Connecting biological discoveries with Google expertise in extracting value from huge amounts of information in order to advance not only autism research, but the entire field of genomic medicine. Working through Google Genomics, the MSSNG Project has access to the same technologies that power Google Search and Maps. Using these technologies, the MSSNG team is creating solutions for securely storing, processing, exploring and sharing complex biological datasets. Autism Speaks has already uploaded nearly 100 terabytes of data from more than 1,300 genomes onto Google Cloud Storage and has an additional 2,000 samples in the sequencing queue. In the end, the MSSNG database will hold information from the whole genomes of 10,000 individuals, making it the world’s largest single repository of autism-related DNA sequencing data.

An important part of the MSSNG project is to promote the Open Science model by sharing these data with the global autism research community. Until now, the transport of genomic data between collaborators involved physically shipping hard drives, a costly and time-intensive process. The MSSNG database lets the autism community instantly power research projects, by providing web-based access to genomic data from thousands of individuals, together with new online analysis tools. The MSSNG Project has the capability to dramatically accelerating breakthroughs in understanding the causes and subtypes of autism in ways that can advance diagnosis and treatment as never before

1.4.1.3.4.b The Khan Academy [32]

Based in Mountain View, California, the Khan Academy is a not-for-profit that produces and posts a vast collection of free educational online videos about math and science topics ranging from algebra and trigonometry to biology and economics. Millions of students, educators and self-learners around the world watch the videos, both on the Khan Academy's YouTube channel and on its hugely popular website (www.khanacademy.org), where students answer some 1.5 million practice questions per school day.

The Khan Academy's development team continually tweaks the site based on how visitors choose to learn. The Khan Academy chose Google App Engine as its hosting and application development platform because Google App Engine could easily house its growing collection of 2,000-plus videos, resolving the organization's overall server and maintenance issues with a single solution. Using Google App Engine freed the team to focus on the user experience and the array of content that makes the academy such a powerhouse. During the U.S. school year, the Khan Academy receives more than 3.8 million unique visits a month – all served through Google App Engine. Such a traffic is supported by the team without having any dedicated system administrators.

1.4.1.4. Rackspace

"Rackspace is an independent Web hoster with a long track record of leadership in the managed hosting market. It has numerous related businesses; some, such as SaaS email, are part of Rackspace itself, while others, such as Jungle Disk, are subsidiaries.

1.4.1.4.1 Offerings

Rackspace Public Cloud is a fully multitenant, OpenStack-based, Citrix Xen-virtualized offering; the offering also has OpenStack Ironic-based bare-metal servers (OnMetal) that are provisioned in approximately 5 minutes, and paid for per minute. Rackspace also offers three flavors of hosted private cloud: vCloud Director-based and VMware-virtualized, Microsoft Cloud OS-based and Hyper-V virtualized, and OpenStack-based and KVM-virtualized. It also offers a Rackspace-operated OpenStack private cloud on the customer's premises. Private clouds are priced on the basis of dedicated capacity. Rackspace has object storage with an integrated CDN (Cloud Files). Customers must choose either a paid support plan or managed services.

1.4.1.4.2 Strengths

- *Over the course of 2014, Rackspace successfully pivoted from its "Open Cloud Company," OpenStack-oriented strategy, and returned to its roots as "a company of experts," emphasizing its managed service expertise and superior support experience. Rackspace has a coherent vision of cloud-enabled managed services that utilize automation and a DevOps philosophy. Rackspace is no longer targeting customers that want to self-manage exclusively, except in the context of its private cloud services.*
- *Rackspace's industrialized private cloud offerings are thoughtfully constructed, more automated than most competing offerings, and operated in a fashion that allows Rackspace to deliver reliable, well-supported services at economical prices. During 2014, Rackspace embraced technology partnerships with VMware and Microsoft in order to expand its range of private cloud offerings, and began to provide managed services for third-party offerings, starting with VMware's vCloud Air. While Rackspace continues to participate in the*

OpenStack community, and has invested significantly in its OpenStack-based private cloud services, it is increasingly technology-neutral. It is now focused on the ways in which it can add value beyond providing basic infrastructure, such as offering solution templates within its OpenStack-based private cloud.

1.4.1.4.3 Cautions

- *Rackspace Public Cloud is a developer-centric offering, and has appealed primarily to small businesses seeking a replacement for low-cost mass-market hosting. Although Rackspace now delivers a solid set of basic features, it has not been able to keep up with the pace of innovation of the market leaders. Rackspace has refocused its business upon customers that need expert managed services for mission-critical needs, rather than try to compete directly for self-managed cloud IaaS against hyperscale providers that can rapidly deliver innovative capabilities at very low cost, or against established IT vendors that have much greater resources and global sales reach. It is a stronger competitor in private cloud IaaS, which does not require the same breadth of capabilities.*
- *Rackspace has made many cloud-related acquisitions to enhance its cloud capabilities and rapidly expand the number of developers it employs. However, it has not integrated these acquisitions into a cohesive whole. Many of the acquisitions actually can manage or operate with multiple cloud IaaS providers. While this potentially positions Rackspace for future multicloud management, and enables it to take advantage of the growth of competitors, Rackspace will need to create a compelling value proposition for using its own cloud IaaS offerings"[24].*

1.4.1.4.4. Research and education case studies

1.4.1.4.4.a The University of Berkeley [33]

The University of Berkley uses extensively cloud computing for both research and education. It has developed JupyterHub, an open source web application which provides interactive data science and scientific computing across a wide range of programming languages including Python and R. JupyterHub is fully deployed on Rackspace. It is used at Berkley for teaching various courses, the main advantages of using JupyterHub are: (i) No software installation is required (ii) Students start from the same baseline: same library versions, same software version etc. (iii). Students submissions adhere to the same structures: same filenames, same directory structure (iv). Students feedback/Grading can be uploaded to the Jupyter server

1.4.2.Elements about the private clouds ecosystem

We introduce now some cloud technologies as well as research and education case studies based on those technologies.

1.4.2.1. OpenStack

OpenStack is a free and open-source cloud-computing software platform. Users primarily deploy it as an infrastructure-as-a-service (IaaS) [33 a].

1.4.2.1.1. Offerings

The technology consists of a group of interrelated projects that control pools of processing, storage, and networking resources throughout a data center—which users manage through a web-based dashboard, through command-line tools, or through a RESTful API. OpenStack.org released it under the terms of the Apache License.

1.4.2.1.2. Strengths

(i)Widely accepted cloud management API (ii) Broad Ecosystem (iii) Adaptability (iv) Open source (v) Interoperability [33 b].

1.4.2.1.3. Cautions

(i) Difficulty of implementation (ii) Shortage of skills available in the market (iii) Conflicting or uncoordinated OpenStack project governance (iv) Weak spots in some OpenStack projects (v) Integration with existing infrastructure.

1.4.2.1.4. Research and education case studies

1.4.2.1.4.a University of Hawaii College of Education [34]

The University of Hawaii at Manoa (UHM) College of Education (COE) is a research college in Honolulu on the island of Oahu that supports faculty, staff, and students on several remote islands. Faculty at UHM COE wanted to tap into cloud storage and compute resources to improve theirresearch and productivity. Cloud resources could accelerate their work and their time-to-deploy services, without having to depend on the university's centralized IT department.

Rolling out an OpenStack cloud gave the college flexibility, and the IT team found that moving from server infrastructure to a virtualized infrastructure let them meet faculty demands as they arose, and fostered the spirit of innovation that already existed at the College of Education. The college's OpenStack cloud met the following use case goals: (i) Support faculty in their research by providing fast and convenient computing and storage resources (ii) Provide comprehensive file sharing and synchronization (via OwnCloud on the OpenStack cloud) to store sensitive data as part of the university's security initiative (iii) Runs the workloads of public website, intranet, and student information systems

1.4.2.2. OpenNebula

The OpenNebula [35]platform manages a data center's virtual infrastructure to build private, public and hybrid implementations of IaaS. OpenNebula is free and Open Source software developed according to the assembly of C, C++, Ruby, Java, Shell scripts programs.

1.4.2.2.1. Offerings

OpenNebula orchestrates storage, network, virtualization, monitoring, and securitytechnologies to deploy multi-tier services in Virtual Machines (VMs).

1.4.2.2.2. Strengths:

Innovative functionality for private and hybrid clouds and datacenter virtualization. Easy to use and to configure;

1.4.2.2.3. Cautions:

(i) Add-ons to connect to other systems are few (ii) strategy in relation to docker still unclear even the oneDock framework is putting OpenNebula on the right path (iii) shortage in competencies

1.4.2.2.4. Research and education case studies

1.4.2.2.4.a Research at University Sorbonne Paris Cité (USPC)

In 2015 USPC decided to renovate and to federate its infrastructure for doing research. Two clusters and one data center is now available for the community of researchers. The “Imagerie du Vivant” (IDV [23]) project is a challenging project that uses cloud computing technologies. The data center technology is based on OpenNebula and currently the 3 systems are not managed by a single system. A web portal will be available at the end of the year 2015 at least for authenticating users. The convergence of systems (cluster, grids, clouds) is among the most important questions that USPC will address in the very near future.

1.4.2.2.4.b Collaboration tool at Paris 13 university

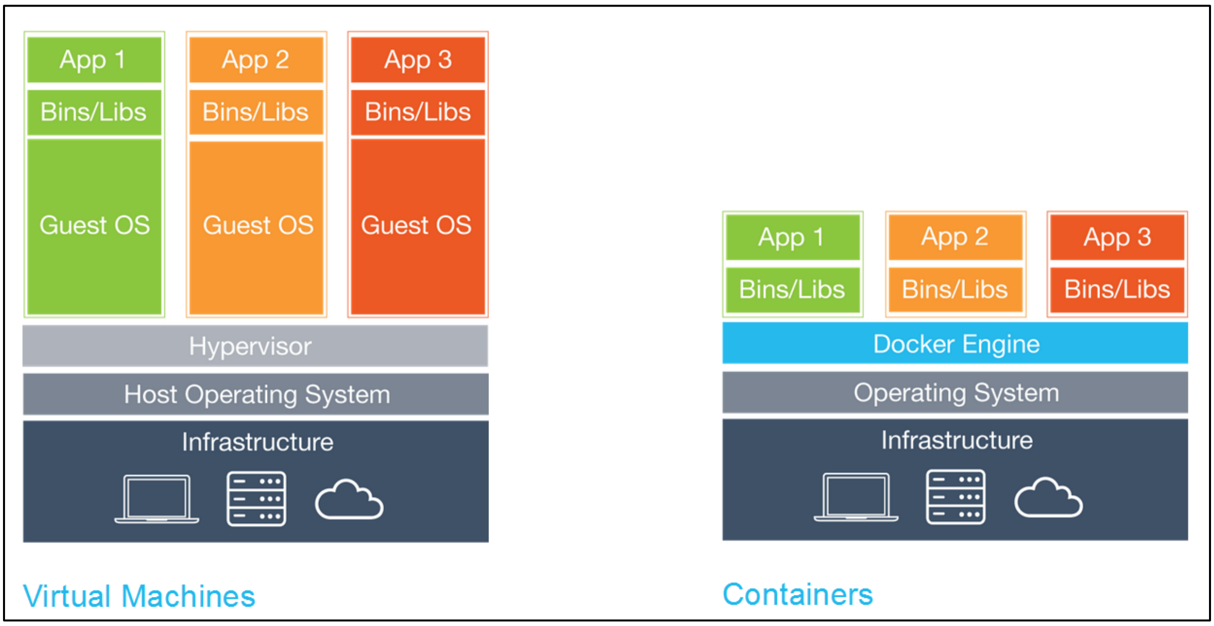
Paris 13 offers a series of universal services through a common interface and platform which is called Post-it. Post-It is the collaborative workspace at University Paris 13. It allows you to exchange ideas and documents within private groups that you create or to which you are invited to participate. Users can invite other people and manage their perimeters of sharing. Post-it lets users view posted messages and files, comment messages, post messages and files or manage the group with you. With Post-it, users can access their files from different groups and directly from the computers, tablets or mobile phones because the service is implemented as a Software as a Service meaning that only a browser is required to use the service.

1.4.2.3. Docker

Docker is an open-source program that enables a Linux application and its dependencies to be packaged as a container. Container-based virtualization isolates applications from each other on a shared operating system (OS). This approach standardizes application program delivery, allowing apps to run in any Linux environment, whether physical or virtual. Because they share the same operating system, containers are portable among different Linux distributions and are significantly smaller than virtual machine (VM) images. Docker competes with proprietary application containers such as the VMware vApp and infrastructure abstraction tools like Chef. Docker is not the only existing open source containers framework but it is by far the most popular today and it is foreseen to become an unescapable building block for the clouds of the future. Both AWS and Google Cloud offerings include Container Services: ECS for Amazon and Container Engine for Google. The Latter is powered by Kubernetes, an Open Source toolkit for managing a cluster of Linux containers as a single system. *Docker Swarm* is a native clustering for Docker and became the main competitor of Google’s Kubernetes. It turns a pool of Docker hosts into a single, virtual Docker host and serves the standard DockerAPI : any tool that already communicates with Docker can use Swarm seamlessly.
Public Container services are growing in popularity. They make it possible to easily run and manage Docker-enabled applications across a cluster of virtual machine instances. Docker became a key technology for accelerating migration to clouds and for avoiding vendors lock-in.

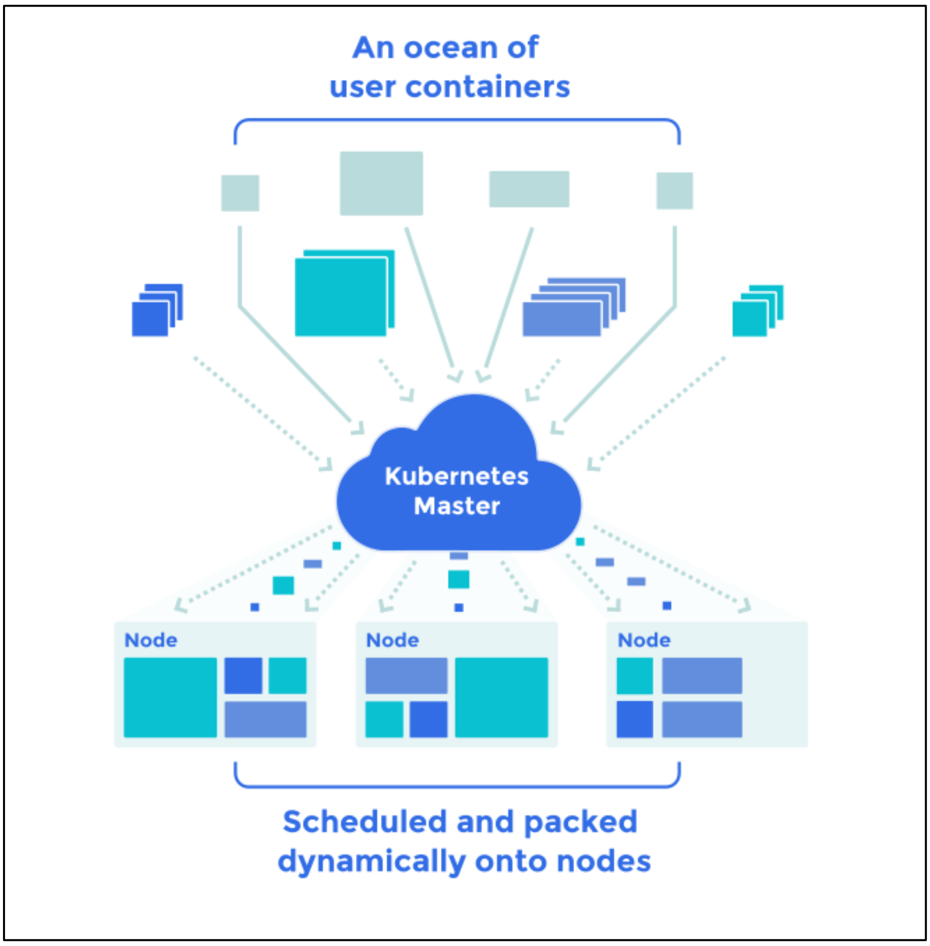
Docker accelerates Dev and simplifies Ops: DevOps [36] (a clipped compound of "development" and "operations") is a culture, movement or practice that emphasizes the collaboration and communication of both software developers and other information-technology (IT) professionals while automating the process of software delivery and infrastructure changes. It aims at establishing a culture and environment where building, testing, and releasing software, can happen rapidly, frequently, and more reliably.

Figure 1.7 Containers



Source [36]: https://en.wikipedia.org/wiki/DevOps

Figure 1.8 Container's scheduling



Source[37]:http://kubernetes.io/

1.5. Conclusion

The Main global trends in cloud market and adoption presented in this Chapter may be summarized as follows:

- the cloud infrastructure will be available as affordable services part of the ambient environment, and as a commodity
- Public Cloud market has experienced a tremendous growth over the past few years. The SaaS segment will continue to outrun the IaaS and PaaS segment with about 60% of total market in 2018.
- Cloud applications will account for 90% of worldwide mobile data traffic by 2019, compared to 81% at the end of 2014.
- Major providers in public cloud are AWS, Google, Microsoft Azure and Rackspace. AWS dominates in public cloud;

RightScale's recent survey on a sample of 930 organizations shows that:

- Cloud is a given and hybrid cloud is the preferred strategy of the respondents.
- Scalability, faster access to infrastructure, and availability are the top three benefits experienced using cloud, whereas "security" and "lack of cloud resources and expertise" are the top 2 challenges.
- Central IT teams are increasingly offering self-service portals as a central hub to broker cloud services to the enterprise

The NSF-XSEDE cloud adoption survey conducted in higher education and research institutions across a wide variety of scientific areas and the humanities, arts, and social sciences, found that:

- Cloud benefits identified by the survey participants were: pay as you go, lower costs, computes elasticity, data elasticity, Software as a Service, Education as a Service, broader use, scientific workflows, rapid prototyping, and data analysis.
- Cloud challenges identified included the learning curve, virtual machine performance, variability in bandwidth, memory limits, database instability, private/public cloud interoperability, security, data movement, storage, and cloud computing cost and the funding availability.

Given the importance of the cloud phenomenon the European Commission and the National Science Foundation (NSF) launched major initiatives in cloud computing. The EU adopted strategy in 2012 for "Unleashing the Potential of Cloud Computing in Europe". It also adopted cloud computing among the research priorities in the H2020 research and innovation funding program.
In August 2014, The NSF has announced the funding of several projects to create cloud computing testbeds "to enable a new future for cloud computing". Its aim is to transform the current US research environment by supporting research infrastructure, enabling transformative research at the frontiers of computing.

In addition to these initiatives, several universities have adopted and actually migrated a certain number of applications such as distance learning and / or scientific research applications to the public cloud.
Among the universities that have migrated educational applications to the public cloud, the following examples are noteworthy:

- Anhanguera the largest universities in Brazil and in the world supporting distance learning, and one of the largest users of Moodle, an e-learning platform for collaborative learning, migrated Moodle and other applications to AWS for increased reliability and scalability

- University of Maryland University College (UMUC), an open-access university is leveraging a broad array of products from the AWS platform by replacing its legacy applications with a new cloud computing analytics platform.
- The Centre for Software Practice (CSP) is a research and development body at the University of Western Australia (UWA). In 2012, CSP migrated its Class2Go, a massive open online course (MOOC) platform, to the cloud in order to be able to scale to accommodate the delivery of new courses and enable participant interaction
- Khan Academy posts a vast collection of free educational online. Students answer some 1.5 million practice questions per school day all served through Google App Engine.
- The Tokyo University of Technology (TUT) has opted for a full shift of its ICT environment to the Azure cloud.

Among the universities that have migrated scientific applications to the public cloud, the following examples are worth mentioning:

- The Algorithms, Machine, and People (AMP) Lab at the University of California Berkeley is a multi-disciplinary research effort designed to build scalable machine learning and data analysis technology. The AMP Lab leverages AWS to quickly scale the compute resources needed to analyze the algorithms that are used in genomics work.
- The Collaboration project MSSNG Autism Speaks adopted Google to migrate and scale up its database to hold information from the whole genomes of 10,000 individuals, making it the world's largest single repository of autism-related DNA sequencing data . The MSSNG project is to promote the Open Science model by sharing these data with the global autism research community.

In addition to these experiences with the public cloud, educational institutions have adopted private or community cloud models. Two cases are to be mentioned:

- The University of Hawaii at Manoa (UHM) College of Education (COE) located on several remote islands adopted OpenStack to virtualize its infrastructure in order meet faculty demands as they arose, and fostered the spirit of innovation
- In 2015 University Sorbonne Paris Cité (USPC) decided to renovate and to federate its infrastructure for doing research through the launch of a community cloud based on OpenNebula technology.

References

[1] IaaS, PaaS, SaaS (Explained and Compared), https://apprenda.com/library/paas/iaas-paas-saas-explained-compared/

[2]-Roundup Of Cloud Computing Forecasts And Market Estimates Q3 Update, 2015 – Forbes, http://www.forbes.com/sites/louiscolumbus/2015/09/27/roundup-of-cloud-computing-forecasts-and-market-estimates-q3-update-2015/print/

[3] RightScale 2015 State of the Cloud Report, http://www.rightscale.com/lp/2015-state-of-the-cloud-report

[4]CDW's 2013 State of The Cloud Report, http://www.cdwnewsroom.com/wp-content/uploads/2013/02/CDW_2013_State_of_The_Cloud_Report_021113_FINAL.pdf

[5] Cyber infrastructure definition on Wikipedia, http://en.wikipedia.org/wiki/Cyberinfrastructure.

[6]. I. Foster. What is the Grid? A Three Point Checklist. Grid Today, 1(6):22-25,2002.

[7]. I. Foster, C. Kesselman, J. Nick, and S. Tuecke. The physiology of the Grid.Grid Computing: Making the Global Infrastructure a Reality, 2003.

[8]. ShantenuJha, Andre Merzky, Geoffrey Fox, "Using clouds to provide grids, with higher levels of abstraction and explicit support for usage modes", http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0529:FIN:EN:PDF.

[9] Unleashing the Potential of Cloud Computing in Europe, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0529:FIN:EN:PDF

[9 a] Digital Agenda: New strategy to drive European business and government productivity via cloud computing, http://europa.eu/rapid/press-release_IP-12-1025_en.htm?locale=en

[10]Report on the public consultation for H2020 Work Programme 2014-15: Cloud Computing, Software and Services
http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=2172

[11] HORIZON 2020 WORK PROGRAMME 2014–2015,
https://ec.europa.eu/research/participants/portal/doc/call/h2020/common/1617601-part_1_introduction_v2.0_en.pdf

[12] HORIZON 2020 WORK PROGRAMME 2014–2015, European research infrastructures (including e-Infrastructures),
http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-infrastructures_v2.0_en.pdf

[13]HORIZON 2020 WORK PROGRAMME 2014–2015, General Annexes, https://ec.europa.eu/research/participants/portal/doc/call/h2020/common/1617621-part_19_general_annexes_v.2.0_en.pdf

[13 a] CloudWay 2015, https://sites.google.com/site/cloudway2015/

[13 b] Venus-C web site, http://www.venus-c.eu/Pages/Home.aspx

[13 c]Helix Nebula web site, http://www.helix-nebula.eu

[14]Towards the European Open Science Cloud, https://zenodo.org/record/16001/files/EUSciCloud-March2015.pdf

[14 a]GEANT project's web site, http://www.geant.org

[15]GEANT Cloud Services web site, http://services.geant.net/clouds/About_Us/Pages/Home.aspx

[16] EUDAT project's web site, http://eudat.eu/

[17] EUBrazil Cloud Connect project's web site, http://www.eubrazilcloudconnect.eu/

[18] Chameleon project's web site, https://www.chameleoncloud.org/

[19]CloudLab project's web site, https://www.cloudlab.us/

[20] XSEDE project's web site, https://www.xsede.org/

[21]XSEDE Cloud Survey Report, http://www.cac.cornell.edu/technologies/XSEDECloudSurveyReport.pdf

[22] XSEDE Cloud Survey Report http://www.cac.cornell.edu/technologies/XSEDECloudSurveyReport.pdf

[23] Imageries du Vivant project's web site, http://idv.parisdescartes.fr

[24] Gartner: Magic Quadrant for Cloud Infrastructure as a Service, Worldwide report, Lydia Leong et al, published 18 May, 2015, http://www.gartner.com/technology/reprints.do?id=1-2G2O5FC&ct=150519&st=sb

[25]AWS Case Study: University of California Berkeley AMP Lab's Genomics Research Project https://aws.amazon.com/solutions/case-studies/university-of-california-berkeley-amp-lab-genomics-research/

[26] AWS Case Study: Anhanguera, https://aws.amazon.com/solutions/case-studies/anhanguera/

[27] AWS Case Study: University of Maryland University College, https://aws.amazon.com/solutions/case-studies/university-of-maryland-university-college/

[28] AWS Case Study: University of Western Australia https://aws.amazon.com/solutions/case-studies/university-of-western-australia/

[29] University Transforms Life Sciences Research with Big Data Solution in the Cloud, https://azure.microsoft.com/en-us/case-studies/customer-stories-virginiatech/

[30] Tokyo University of Technology Implements Futuristic Environment Using Cloud-Based Solutions, https://customers.microsoft.com/Pages/CustomerStory.aspx?recid=863

[31] Harnessing big genomic data to find the missing answers of Autism Spectrum Disorder, https://cloud.google.com/customers/autism-speaks/

[32] The Khan Academy Scales and Simplifies with Google App Engine, https://cloud.google.com/customers/khan-academy/

[33] Deploying JupyterHub for Education, Jessica Hamrick, https://developer.rackspace.com/blog/deploying-jupyterhub-for-education/

[33 a] OpenStack definition on Wikipedia, https://en.wikipedia.org/wiki/OpenStack

[33b] Is OpenStack a Success? Alan Waite, http://blogs.gartner.com/alan-waite/2015/04/27/openstack-success/

[34]University of Hawaii College of Education harnesses OpenStack to support faculty research via a secure private cloud
https://www.mirantis.com/blog/case-study-university-of-hawaii-openstack-private-cloud/

[35] OpenNebula's web site, http://opennebula.org

[36] DevOps definition on Wikipedia, https://en.wikipedia.org/wiki/DevOps

[37] OpenNebulaDocker Driver and Datastore, http://opennebula.org/opennebula-docker-driver-and-datastore/

Chapter2
State of ICT and
cloud computing in education in the Arab World

2.1. Introduction

This chapter is organized in four parts: the first part positions the Arab countries according to their level of development in the ICT field; the second part presents the methodology adopted to conduct a survey on cloud computing in the Arab countries; the third part presents and discusses the results of the survey. The chapter ends with a general conclusion.

Arab countries considered in this chapter are those members of the Arab League of (22). These countries and their respective populations are shown in Table 1.

Tab. 2.1. Population by country

Countries	Population (2015 Est.)	Countries	Population (2015 Est.)
Algeria	39,542,166	**Morocco**	33,322,699
Bahrain	1,346,613	**Oman**	3,286,936
Comoros	780,971	**Palestine** (West Bank)	2,785,366
Djibouti	828,324	**Qatar**	2,194,817
Egypt	88,487,396	**Saudi Arabia**	27,752,316
Iraq	33,309,836	**Somalia**	10,616,380
Jordan	6,623,279	**Sudan**	36,108,853
Kuwait	3,996,899	**Syria**	22,878,524
Lebanon	4,151,234	**Tunisia**	11,037,225
Libya	6,411,776	**United Arab Emirates**	9,445,624
Mauritania	3,596,702	**Yemen**	26,737,317

Source [1]:Internet Worl Stats http://www.internetworldstats.com/stats.htm

2.2. ICT status in the Arab world

The objective of this paragraph is to provide an overall assessment of the state of ICT in the arab World as measured according to the ITU ICT Development Index and the Internet World Stats.

This assessment is justifed by the fact that cloud deployments require a certain level ICT readiness and infrastructure development .

2.2.1. The ITU ICT Development Index

"*The ICT Development Index (IDI) is a composite index that combines 11 indicators into one benchmark measure that can be used to monitor and compare developments in information and communication technology (ICT) between countries and over time.*"[2]

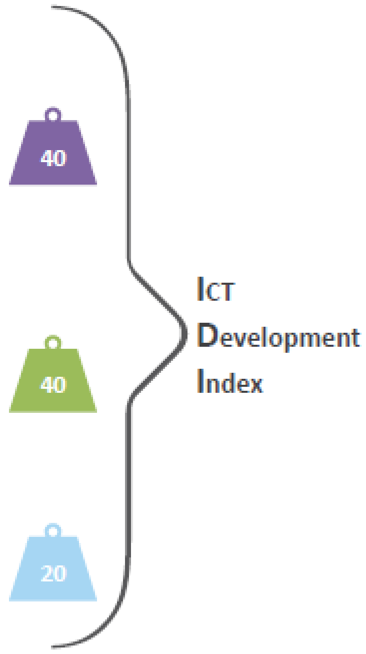
"The IDI is divided into the following three sub-indices":

- *"Access sub-index: This sub-index captures ICT readiness, and includes five infrastructure and access indicators"*
- *"Use sub-index: This sub-index captures ICT intensity, and includes three intensity and usage indicators"*
- *"Skills sub-index: This sub-index seeks to capture capabilities or skills which are important for ICTs. It includes three proxy indicators*"

The set of indicators is summarized in the following figure 2.1.

Figure 2.1. ICT Development Index

ICT access	Reference value	(%)
1. Fixed-telephone subscriptions per 100 inhabitants	60	20
2. Mobile-cellular telephone subscriptions per 100 inhabitants	120	20
3. International Internet bandwith (bit/s) per internet user	962'216*	20
4. Percentage of households with a computer	100	20
5. Percentage of households with Internet access	100	20
ICT use	**Reference value**	**(%)**
6. Percentage of Individuals using the Internet	100	33
7. Fixed-broadband subscriptions per 100 inhabitants	60	33
8. Active mobile-broadband subscriptions per 100 inhabitants	100	33
ICT skills	**Reference value**	**(%)**
9. Adult literacy rate	100	33
10. Secondary gross enrolment ratio	100	33
11. Tertiary gross enrolment ratio	100	33



Note: *This corresponds to a log value of 5.98, which was used in the normalization step.
Source: ITU.

Source [2] Information Society Report 2015
http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf

IDI is constructed as the weighted average of the normalized values of those indicators, from 0 to 10 (see Fig. 2.1). A score that is close to 10 implies a very good ICT development level while a score close to 0 indicates poor development.

It's to be noted that IDI rankings for 2010 and 2015 consider only 16 out of the 22 Arab countries.

Tab. 2.2. IDI rankings, Arab States region, 2015

Economy	Regional rank 2015	Global rank 2015	IDI 2015	Global rank 2010	IDI 2010	Global rank change 2015-2010
Bahrain	1	27	7.63	48	5.42	21
Qatar	2	31	7.44	37	6.10	6
United Arab Emirates	3	32	7.32	49	5.38	17
Saudi Arabia	4	41	7.05	56	4.96	15
Kuwait	5	46	6.83	45	5.64	-1
Oman	6	54	6.33	68	4.41	14
Lebanon	7	56	6.29	77	4.18	21
Jordan	8	92	4.75	84	3.82	-8
Tunisia	9	93	4.73	93	3.62	0
Morocco	10	99	4.47	96	3.55	-3
Egypt	11	100	4.40	98	3.48	-2
Algeria	12	113	3.71	114	2.99	1
Syria	13	117	3.48	106	3.14	-11
Sudan	14	126	2.93	127	2.05	1
Djibouti	15	148	2.19	143	1.69	-5
Mauritania	16	150	2.07	146	1.63	-4
Average			**5.10**		**3.88**	

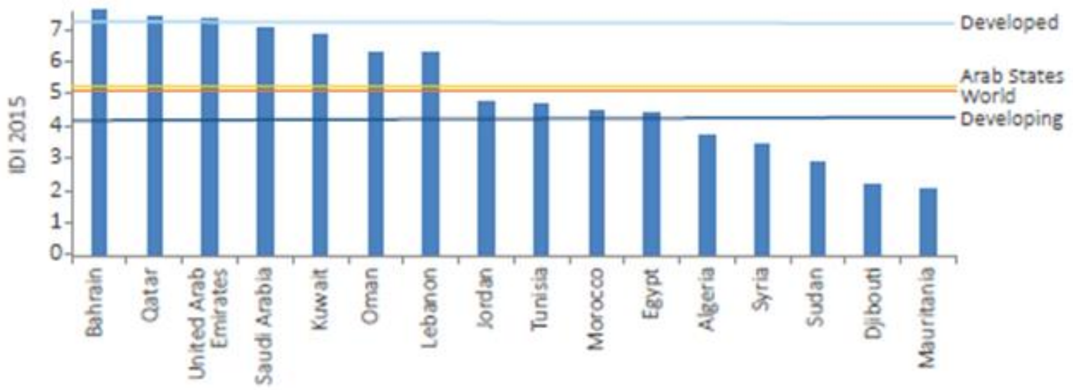
Source: ITU.

Source [2] Information Society Report 2015
http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf

The 16 Arab countries are ranked from 27 to150 out of 162 classified countries. They all achieved some progress as compared to 2010. They may be divided into three classes: the first one which consists of 7 countries is above the world average of 5.03; two countries of this class even exceed the average of the most advanced countries (7.35); the second class is made up of four countries that are close to the world average with an IDI varying between 4.75 and 4.40. The last class consists of 5 countries with an IDI has varying between 3.71 and 2.07.

The figure below shows the IDI scores of the Arab countries as well as their positioning relative to the world average, the Arab countries, the developed and developing countries averages.

Figure 2.2 IDI by country compared with global averages, 2015

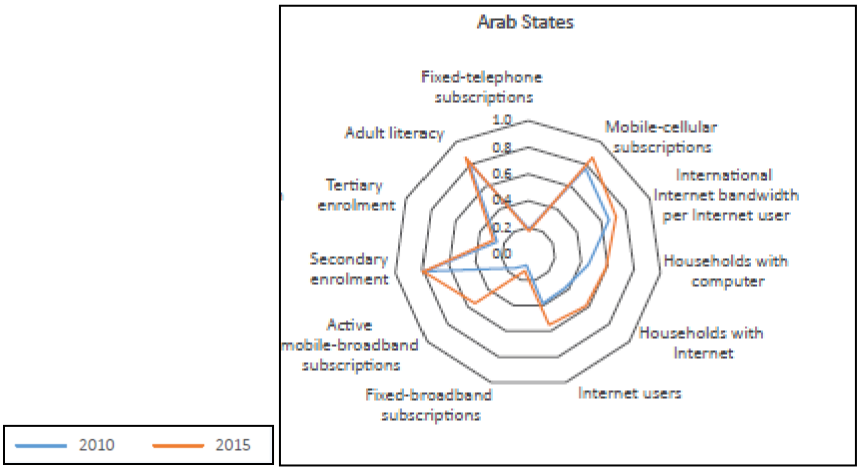


Source: ITU.

Source [2] Information Society Report 2015
http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf

The average IDI of Arab countries slightly exceeds the world average. It is worth mentioning that most of the indicators included in the IDI index for the Arab countries have evolved positively between 2010 and 2014 as shown in the figure below.

Figure 2.3 Average IDI rating for each indicator, arab region, 2010 and 2015



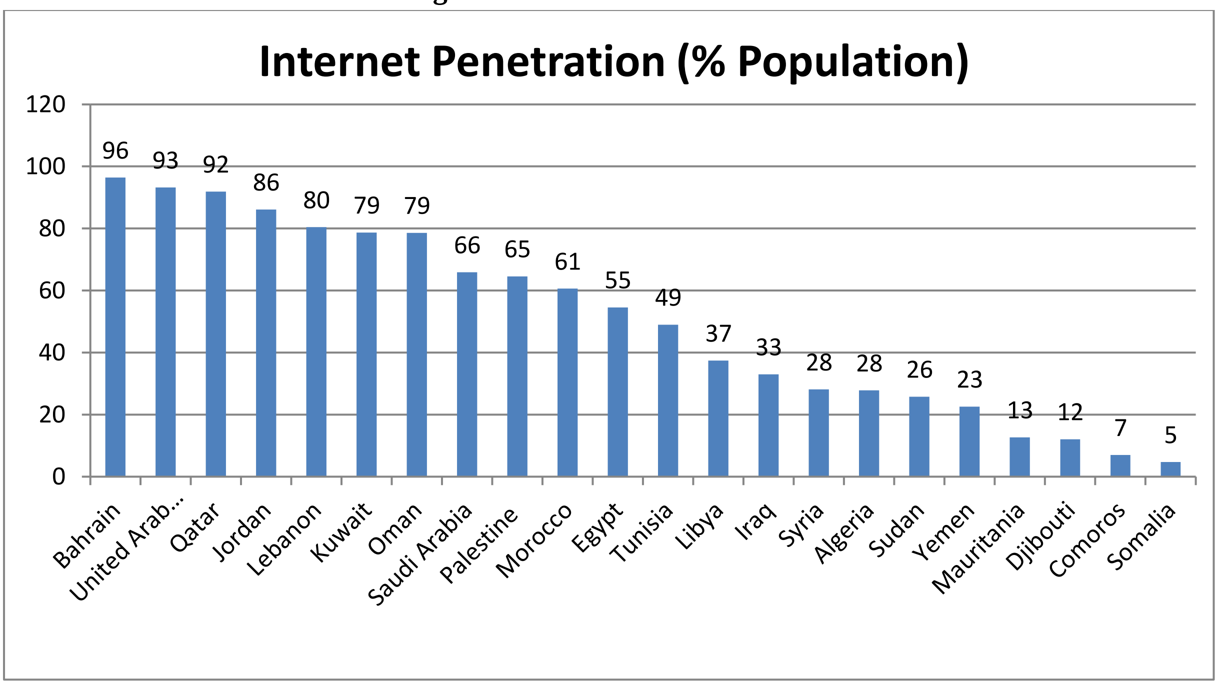
Source [2] Information Society Report 2015
http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf

The most important increases are related to: active mobile-broadband subscribers, Internet-users, households with Internet and households with computer indicators.
Recent figures in terms of Internet penetration for all Arab countries are presented in the next section.

2.2.2 The Internet Penetration according to Internet World Stats

The World Stats website regularly publishes data on the penetration and use of the Internet in the world. The situation for the Arab countries on the date of November 2015 is as follows[1]

Figure 2.4 Internet Penetration



The average for the 22 Arab countries is 44.8% and is slightly below the world average of 46.4%. It is greater than the average of Africa (28.6%), below of the average rate of the North American countries of 87.9% which corresponds to the highest regional average in the world. It may be noted that 3 groups emerge: one group 10 countries whose scores are below the world average, 7 of which exceeded the European average (73.5%) and 5 are in the middle. This classification confirms the outcome of the analysis based on the IDI index.

2.3. Survey Methodology on cloud computing in education in the Arab world

To collect data on cloud adoption in educational and research institutions in the Arab countries, a survey based on an online questionnaire conceived by the study team was launched by ALECSO via the official channels throughout the Arab countries. The survey was aimed at government departments, engineering schools, faculties and research centers, and data centers.

The objective of the survey is (i) to make an inventory of ICT use in Arab institutions (ii) to understand how those institutions perceive the benefit and applicability of the cloud to their use cases and whether and how they plan to use it (iii) to understand the barriers and risks they currently apprehend. The outcome will be used to elaborate recommendations for triggering or catalyzing Arab institutions' migration to the cloud.

The questionnaire (see annex) is structured as follows:

The first part addresses users and non-users of the cloud. It aims at collecting general data about the institution and assessing the status of its current ICT tools and infrastructure.

A second part is for non-users of the cloud. It focuses on the positioning of the institution in relation with cloud utilization which includes its current perception of the expected benefits and usability of the technology as well as its concerns

This part aims also at collecting data to characterize the types of potential cloud utilizations and to infer actions likely to lower the barriers to cloud adoption.

A third part is for cloud users. It aims to collect data on (i) the types of clouds used and their levels of use (ii) the perceived benefits and usability (iii) the challenges raised by the migration to the cloud and the solutions adopted to overcome those challenges (iv) the level of satisfaction with the cloud solutions that have been put in place (v) The actions to be taken to boost the migration to the cloud.

The Sphinx software[3]was used to allow online data collection and processing. This choice is motivated by the desire to reach a maximum number of respondents and accelerate data collection. The survey was conducted between August and November 2015.

2.4. Survey Results

Survey results will be presented in four parts:

- Sample characteristics
- ICT status
- Results for institutions that do not use the cloud
- Results for institutions that use the cloud

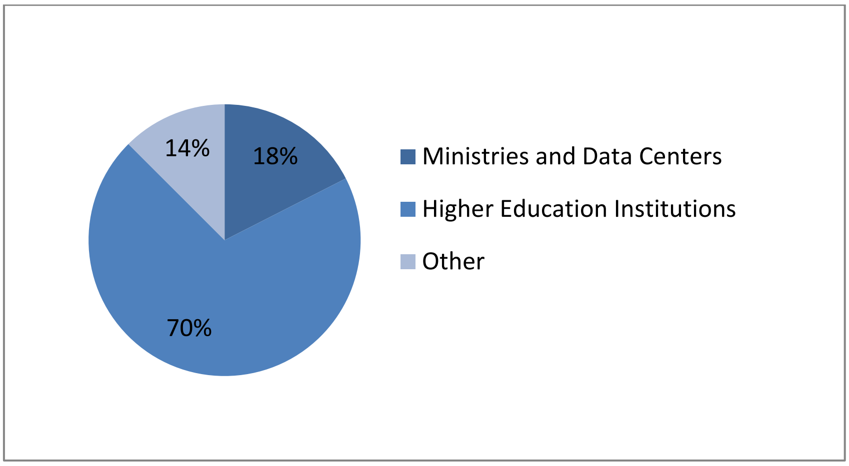
2.4.1. Sample characteristics

65 questionnaires were completed online, but only 40 were valid. 25 were rejected because of lack of information or multiple responses from the same organization or unrealistic data. 10 Arab countries participated in the survey.

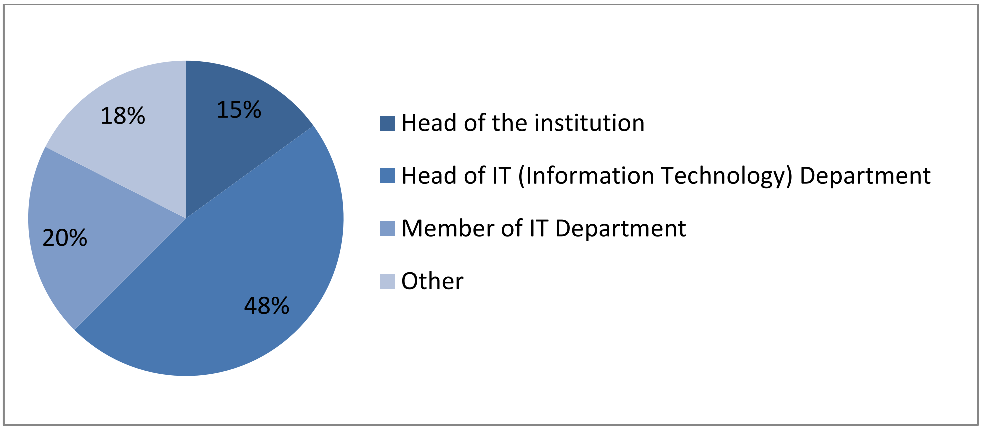
Tab. 2.3 Breakdown of respondents by country

Countries	**Number of respondents**
Algeria	1
Sudan	1
Morocco	1
Tunisia	6
United Arab Emirates	1
Jordan	11
Palestine	1
Kuwait	7
Qatar	1
Bahrain	9
n.a.	1
Total	**40**

Higher education institutions (faculties and engineering schools) (28) make up 70% of the sample while government departments and data centers 19%.

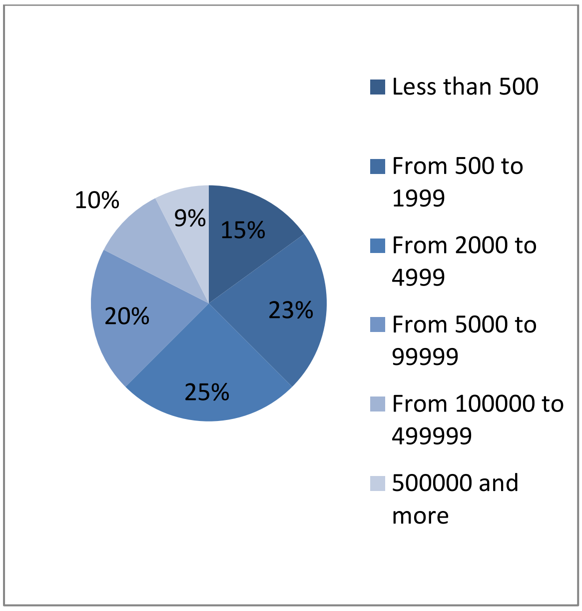


The majority of respondents consists of directors and members of IT departments (68%).

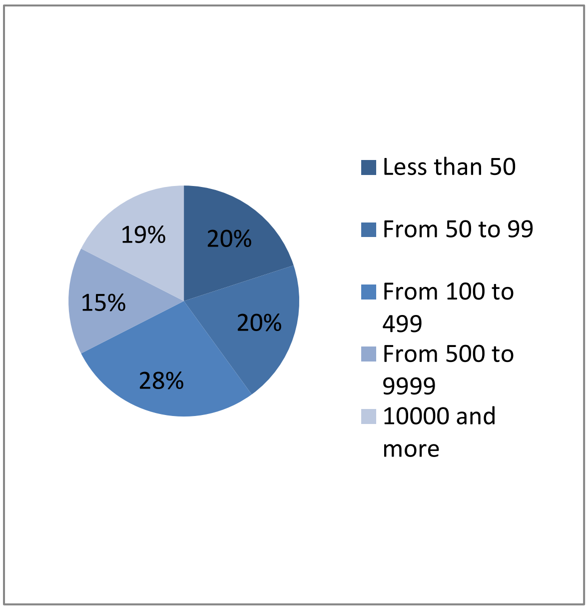


The sizes of the surveyed institutions measured by the number of students and faculty members are highly variable as shown in the figures below:

Students



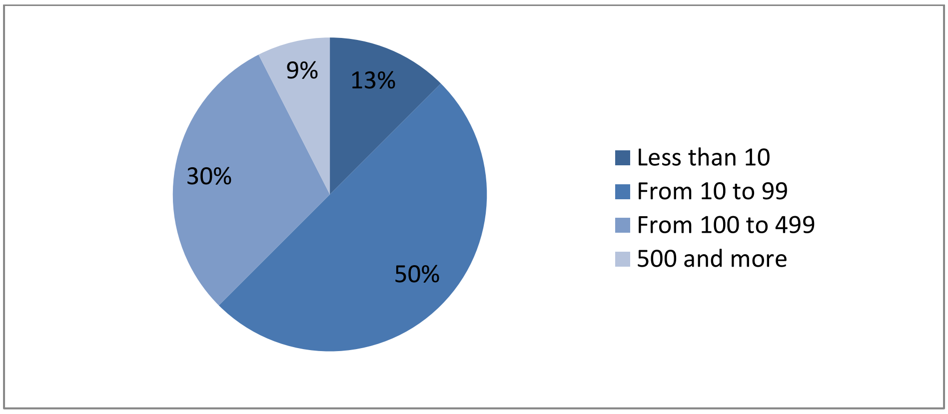
Faculty members



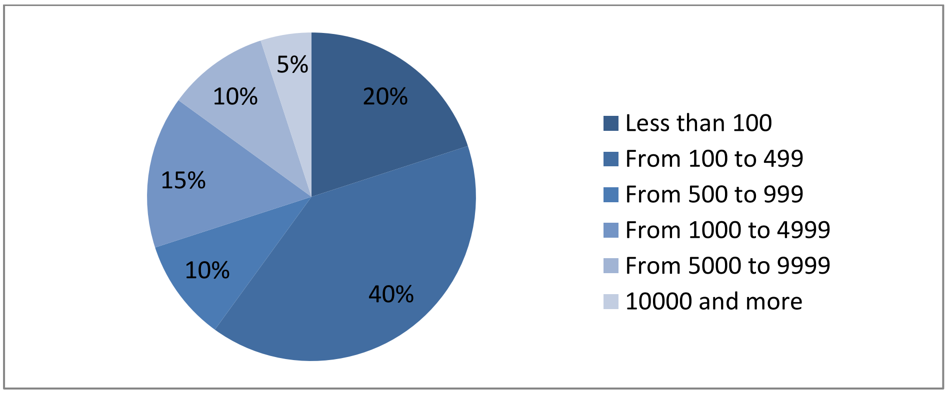
2.4.2. ICT Status

The ICT status is measured with regard to the following indicators: network bandwith, number of PCs and terminals, the existence of a data center and the types of IT applications and services used in the institution.

Speed of Internet connection in Mbit/s



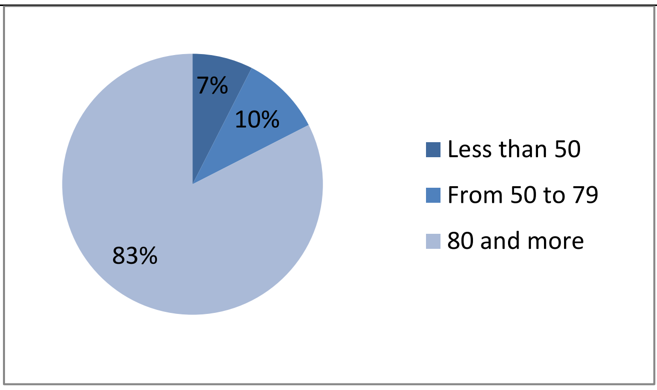
Number of desktop in institutions



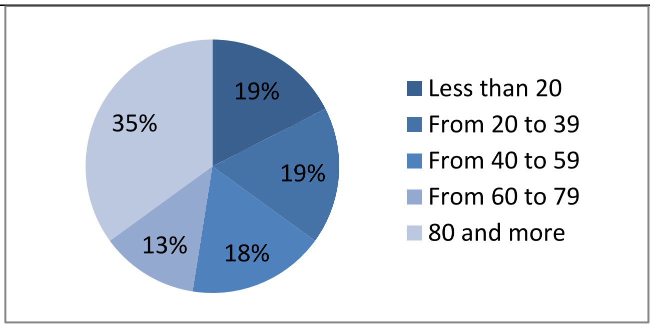
The majority of institutions surveyed (85%) own a data center.

The majority of students have an email address; however the number of students with a laptop is relatively low as shown in the following figures:

Percentage of students with email addresses

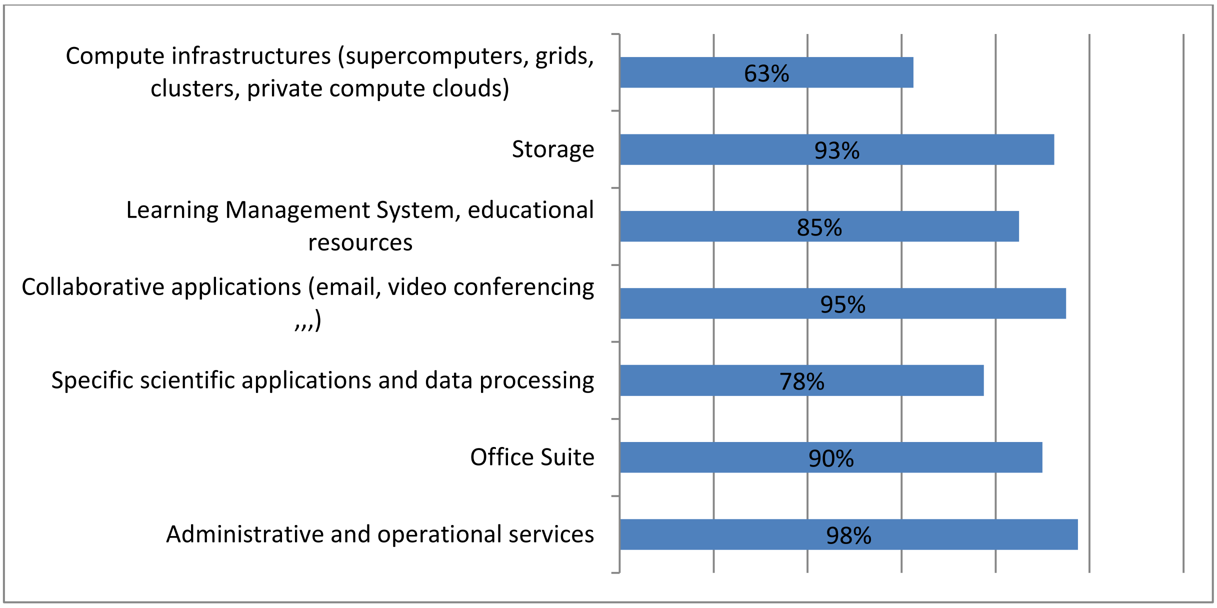


Percentage of students who own a laptop



Institutions use different types of IT applications and services with the highest concentrations on Administrative and operational services applications (98%), collaborative applications (95%), storage (93%) and office suite (90%).

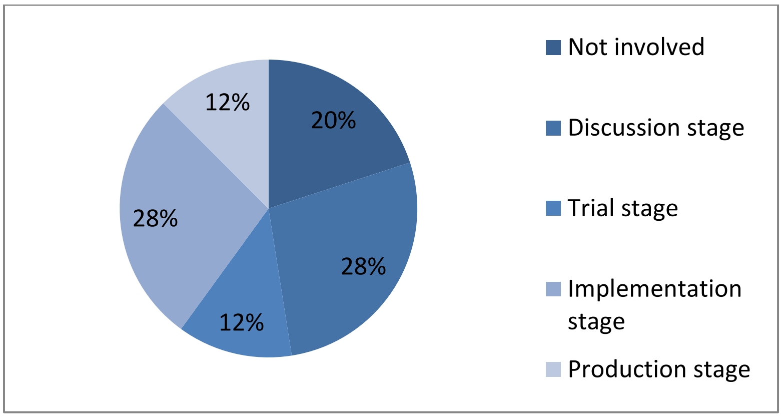
Types of IT applications and services



68% of respondents are very satisfied with the existing IT applications and services.

As regards cloud adoption 40% of the respondents indicated that they are either at production (12%) or implementation stage (28%). If we add the percentage of institutions in the trial stage then the total reaches 52%.

Cloud utilization in institutions

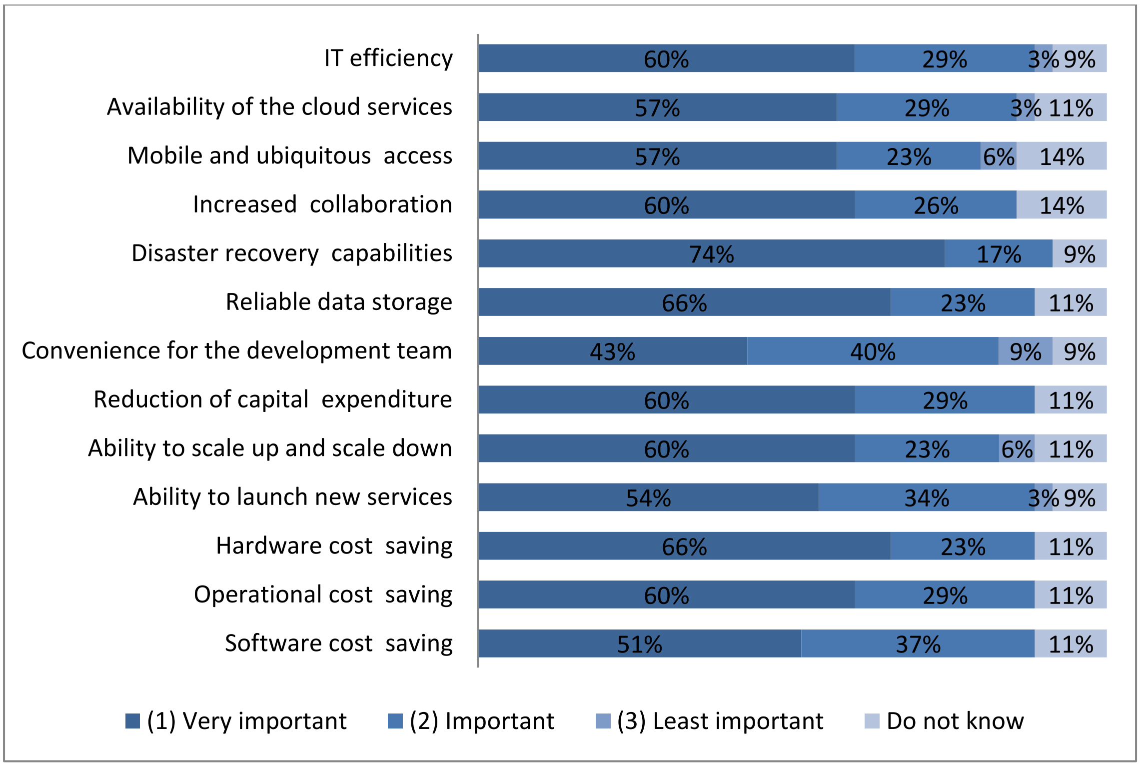


Even though the percentage of early adopters is fairly low there seems to be an interest towards cloud technology.

2.4.3. Survey results for institutions that do not use the cloud in a production stage

The survey results presented in this section are based on the perceptions of the 35 respondantswhich are not in production stage. The answers are exposed on a scale with three positions: very important, important and least important.

Perceived benefits of cloud use in education and research

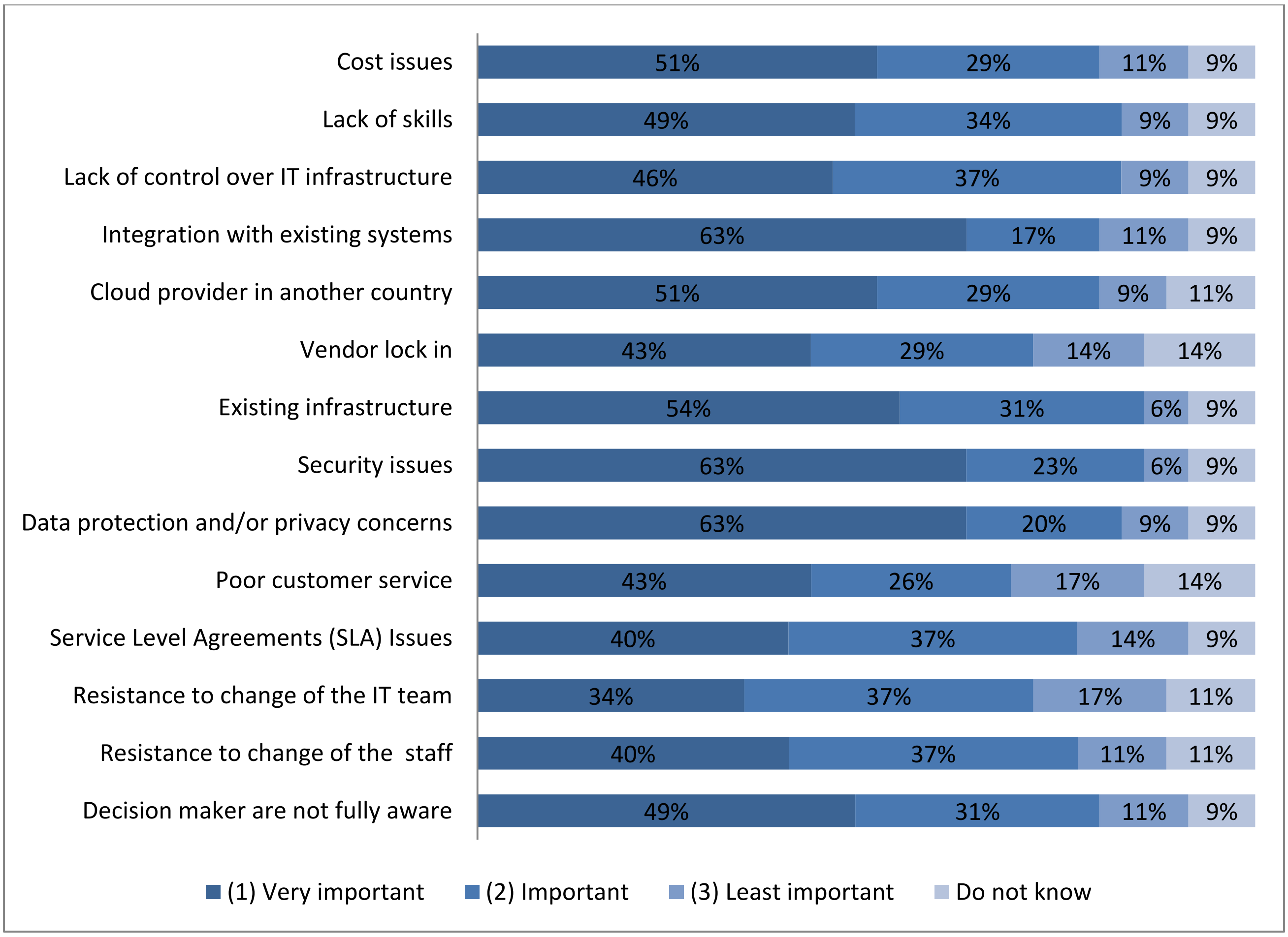


The most important benefits perceived by the highest percentage of respondents are "disaster recovery capabilities" followed by "hardware cost saving" and "reliable data storage". More than 50% of respondents consider that all the other benefits are very important except for "convenience for the development team" (43%).

The vast majority of respondents perceive the use of the cloud as easy or very easy.

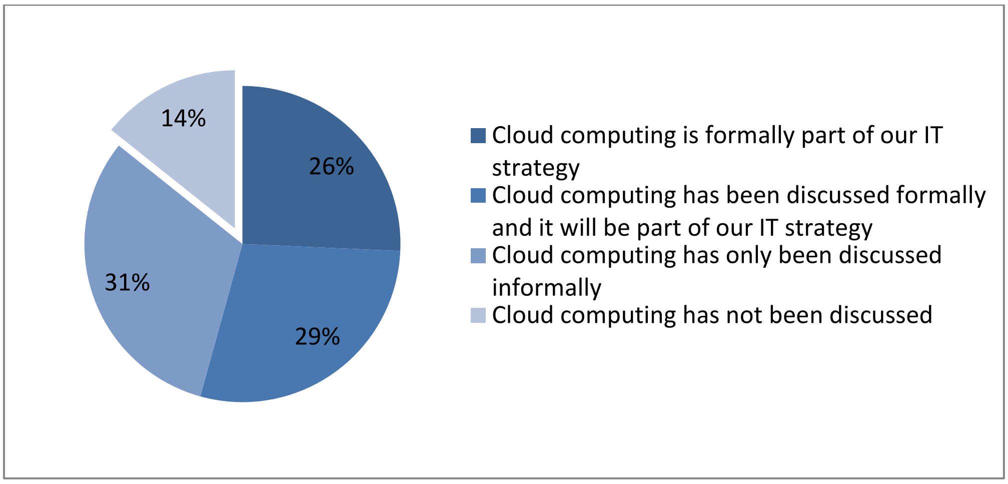
Despite this positive perception of the benefits and the ease of use of the cloud, a number of barriers to cloud adoption were identified by respondents. Among the more significant barriers we find "security issues", "Integration with Existing systems", "data protection and / or privacy Concerns". Overall, all the barriers mentioned in the questionnaire are perceived by the majority of respondents as important / very important.

Perceived barriers for cloud adoption



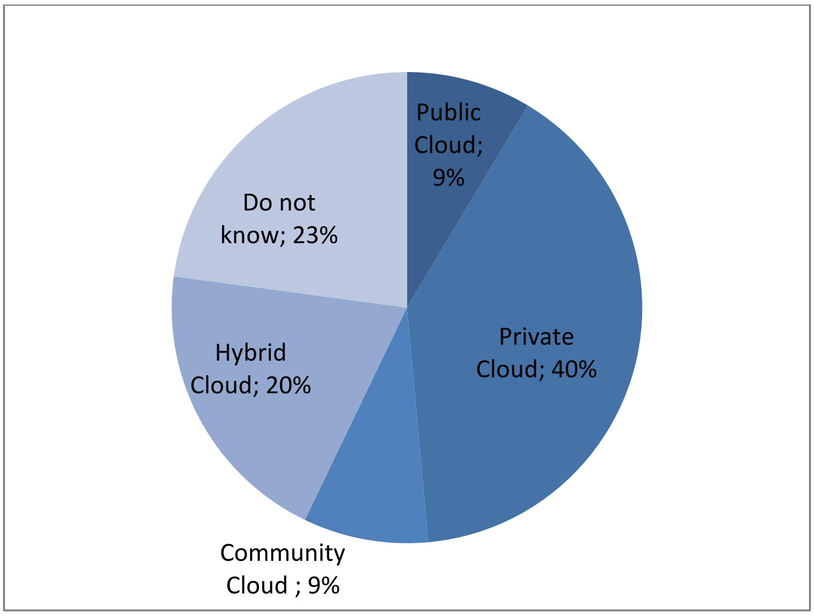
Despite these perceived barriers, 55% of respondents formally introduced cloud computing in their strategy or are in a formal discussion stage to make it part of their strategies. This confirms the trend noted above.

Intention to use the cloud



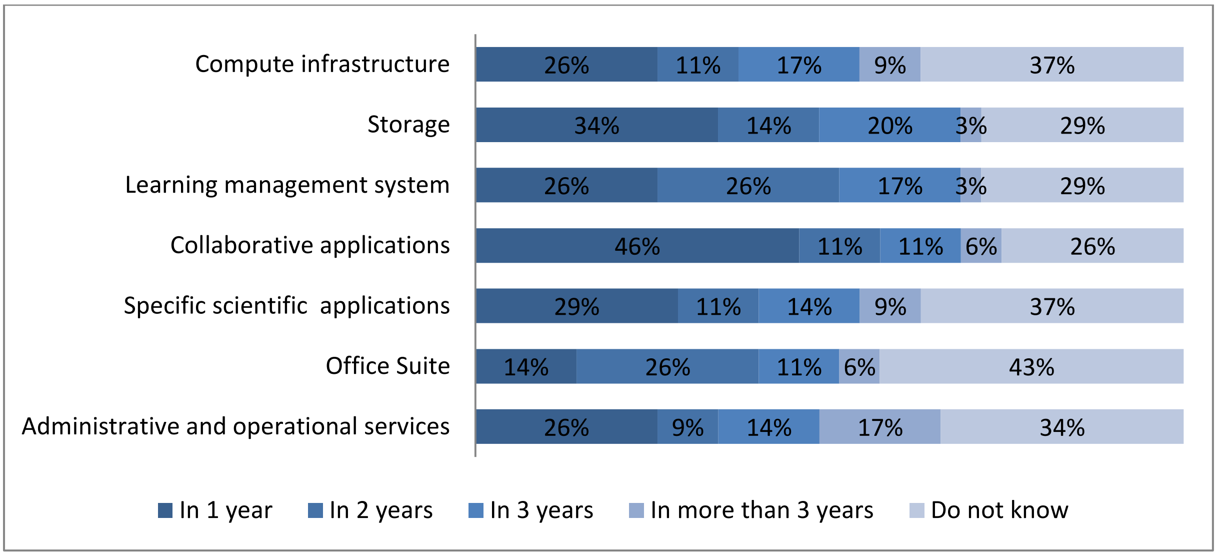
When asked if your institution intends to migrate, what would be the type of cloud of their choice, 40% of institutions prefer to migrate to a private cloud, 20% hybrid cloud and only 9% to a public cloud. This result could be explained by a lack of familiarity with the cloud.

Type of cloud in case of migration



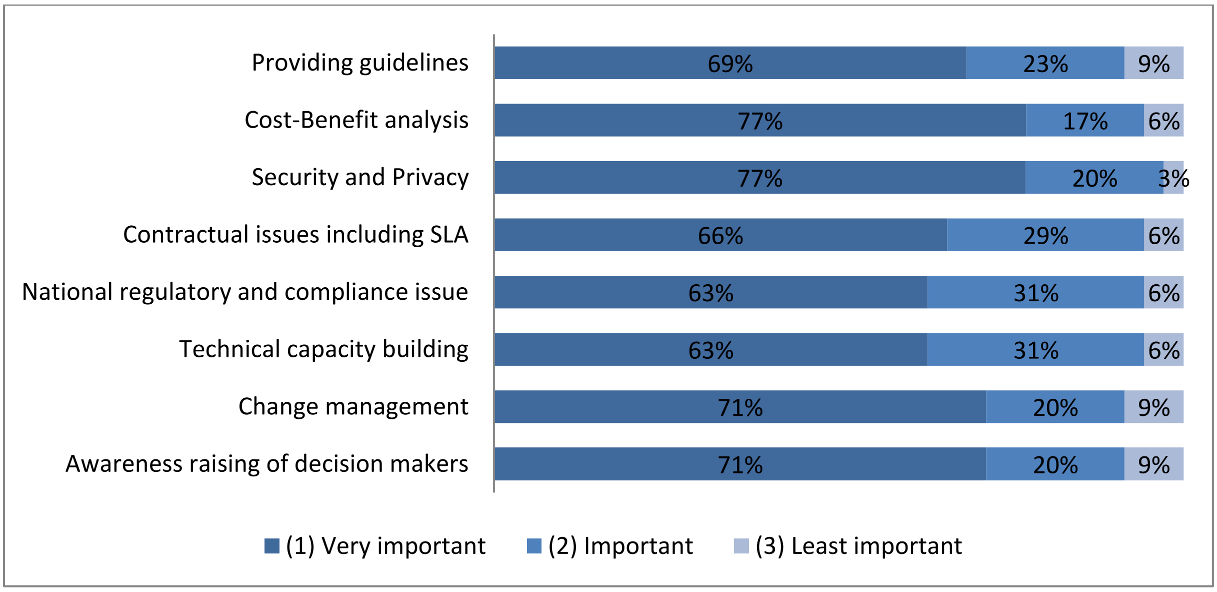
In case of migration, according to a time frame, the preferred applications to migrate to the cloud during the first year are « collaborative application » with 46% of respondents, followed by storage service (34%) and scientific applications (29%). They are followed by "learning management system" (26% the first year and 26% the second year). It is also worth mentioning that for various applications, a significant number of respondents (from 26% to 43%) didn't know what priorities to set.

Intention of migration IT applications and services to the cloud



In order to remove barriers and facilitate cloud migration, respondents expressed their needs for support as given in the following figure:

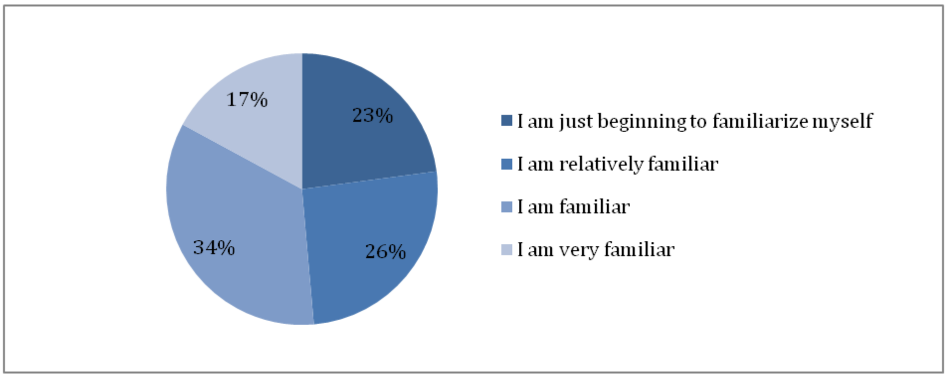
Following areas need support regarding decision to migrate to the cloud



All proposed support items received a “very important” score from more than 60% of respondents. The highest support demands are related to « cost benefit analysis » and « security and privacy ».

The following figure shows 49% all respondents are just familiar or beginning to familiarize themselves with cloud computing. Given that the majority of respondents belong to IT departments this suggests that there is a strong need for support in raising awareness and capacity building.

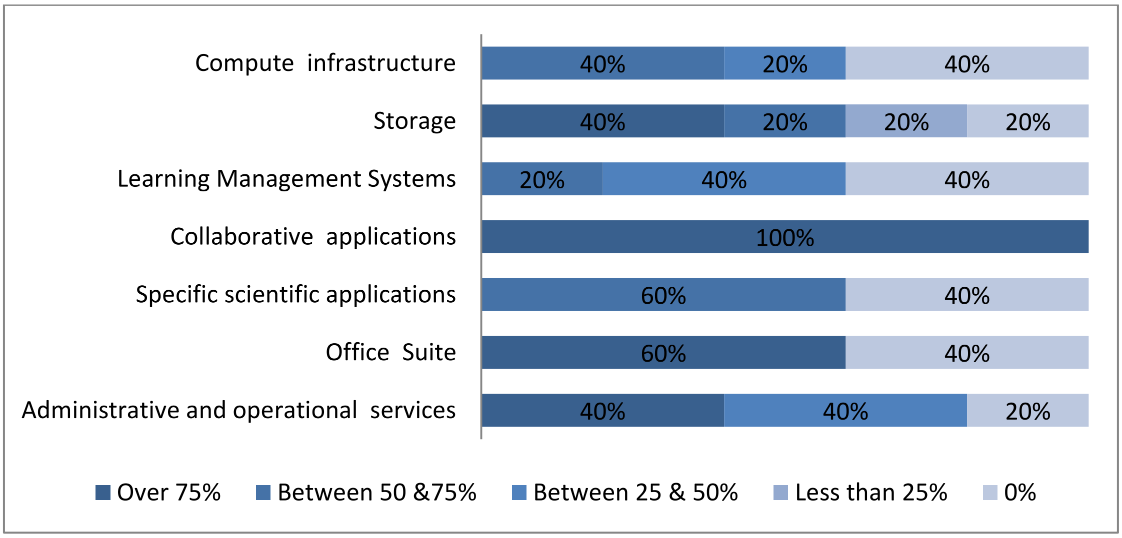
Familiarity of respondent with the cloud



2.4.4. Survey results for institutions using the cloud in a production stage

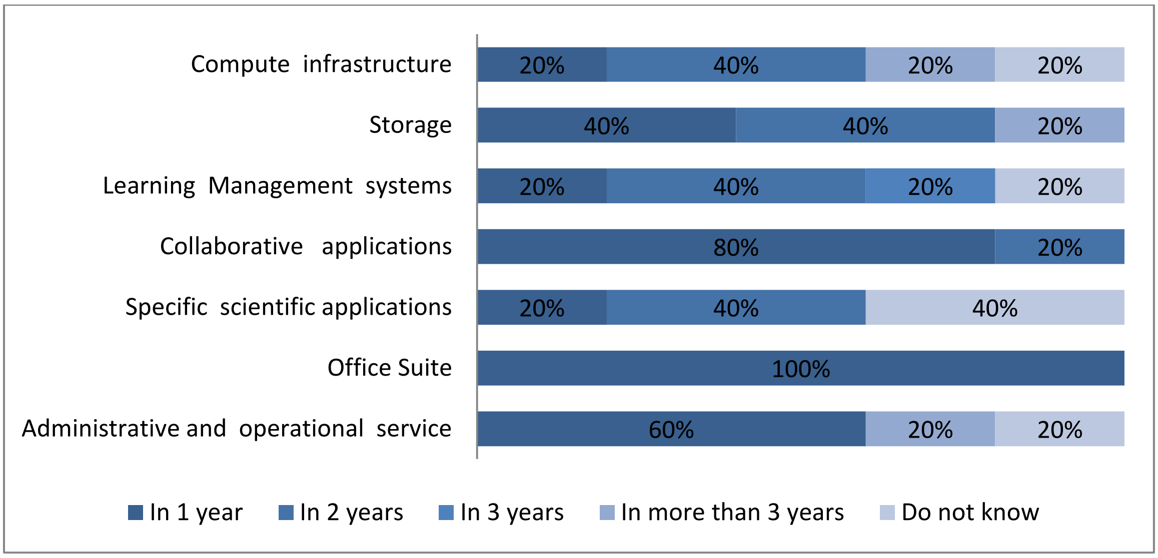
The number of institutions in the sample using the cloud is 5. Therefore, conclusions drawn from this sample must be taken with great caution.
4 out of 5 (80%) of the respondents have adopted the hybrid cloud.
The five institutions have migrated more than 75% of their collaborative applications to the cloud. Three of them have migrated more than 75% of "specific scientific" and "office suite" applications to the cloud. Only one institution has migrated more than 75% of "Learning Management Systems" application to the cloud, as shown in the figure below.

IT applications and services migrated to the cloud



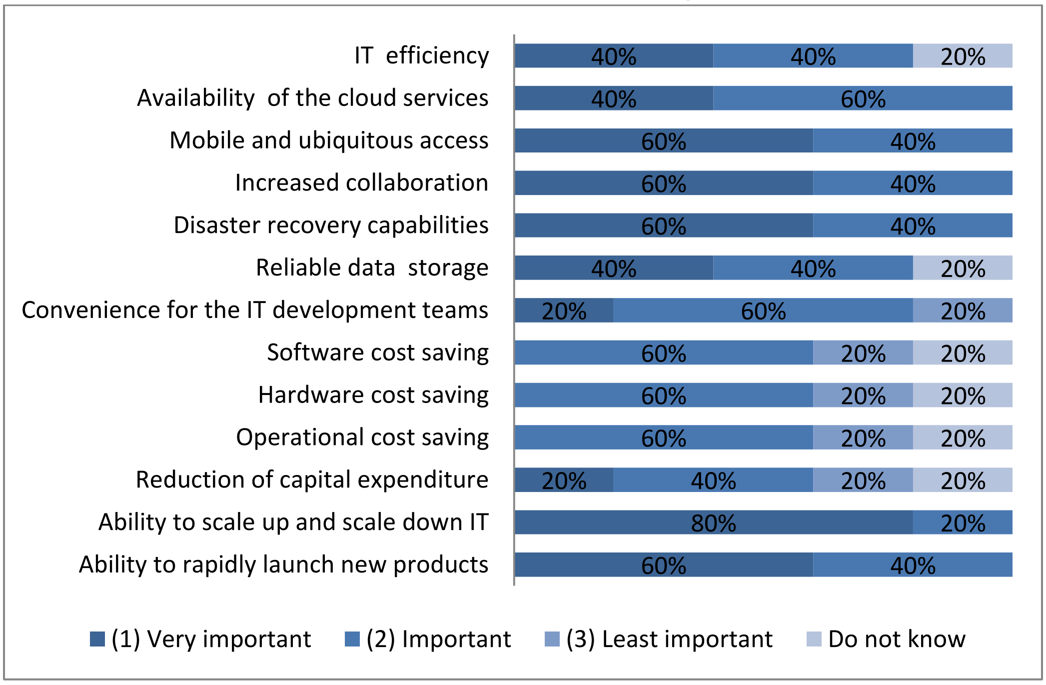
Applications and services for which the respondents intend to pursue migration or plans to migrate to the cloud are given below:

Plan to migrate to the cloud



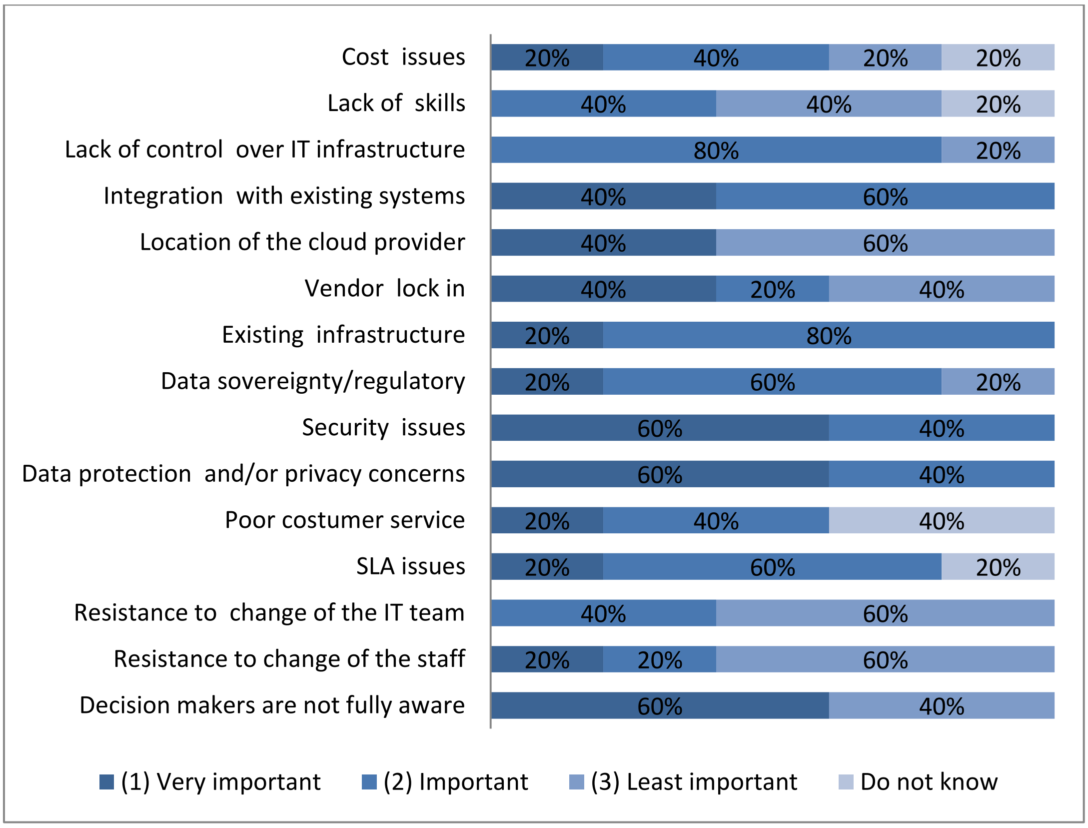
In one year, five institutions intend to migrate "office suite" applications to the cloud, and "collaborative applications" (or rather what remains of them since they have already migrated more than 75%) for 4 them.
The learning management system does not seem to be a priority since a single institution is planning its migration in a year.
Four institutions recognize that the "Ability to scale up and scale down IT" is a very important benefit. Only one institution considers that "Convenience for the IT development teams" is a very important benefit. Similarly, only one institution considers that the "Reduction of capital expenditure" benefit is very important.

Benefits that have been made in moving to the cloud



The main barriers that institutions encountered when migrating to the cloud are particularly: the “Lack of control over IT infrastructure”, “security issues”, “Data protection and/or privacy concerns”, "Decision makers are not fully aware of the benefits the technology can deliver". For 4 respondents the existing infrastructure was a major barrier in the migration process.

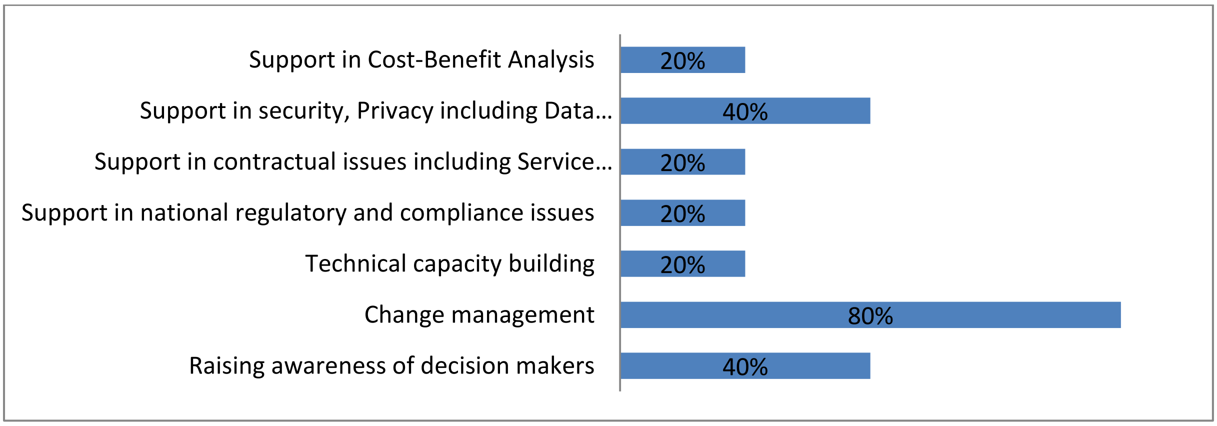
Barriers encountered when migrating to the cloud



.

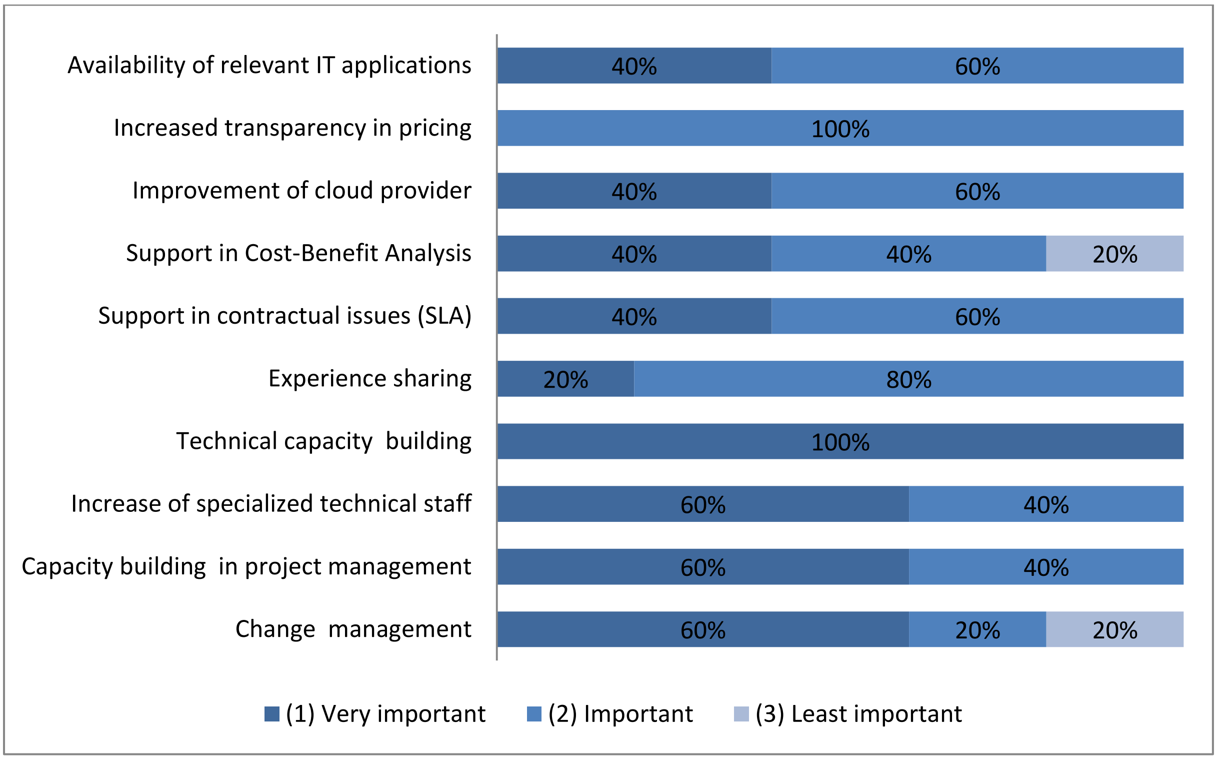
Most respondents (4) adopted change management to overcome these barriers.

How did your institution overcome these barriers?



Three out five institutions are very satisfied with the cloud, while the other two are satisfied. According to the respondents the main actions to be taken to boost cloud migration are especially the needs for increased transparency in pricing and technical capacity building.

What are the actions to be taken to boost migration of your institution to the cloud?



2.5. Conclusion

The analysis of the of ICT status in the Arab world demonstrated a great diversity between Arab countries and an average level close to the world average.

Indeed, compared to the world average of the IDI of 5.03, three classes of Arab countries have been acknowledged. The first one which consists of 7 countries is above the world average of 5.03; two countries of this class even exceed the average of the most advanced countries (7.35); the second class is made up of four countries that are close to the world average with an IDI varying between 4.75 and 4.40. The last class consists of 5 countries with an IDI varying between 3.71 and 2.07.

With regard to Internet penetration, we note that 3 groups emerge: one group 10 countries whose scores are below the world average, a group 7 of which exceeds the European average (73.5%) and a group 5 are in the middle. This classification confirms the outcome of the analysis based on the IDI index. It’s also an indication that a great deal of effort is to be accomplished in most Arab countries to reach an adequate ICT infrastructure development to enable cloud deployment.

This diversity is also reflected in the results of the survey of 40 institutions of higher education in the Arab world. Despite this diversity strong trends emerge:

- **A clear move toward cloud adoption**: 52% of surveyed institutions are either in the production, implementation or test phase;
- **A positive perception of cloud benefits by non cloud adopters (not in production phase)**: the most important benefits perceived by the highest percentage of respondents are "disaster recovery capabilities" followed by "hardware cost saving" and "reliable data storage".
- **Type of cloud for non cloud adopters**: 40% of institutions prefer to migrate to a private cloud, 20% hybrid cloud and only 9% to a public cloud.
- **Applications to migrate to the cloud for non cloud adopters**: « collaborative application » with 46% of respondents, followed by storage service (34%) and scientific applications (29%) are the preferred applications during the first year. They are followed by «learning

management system» for the first couple of years (26% the first year and 26% the second year).
- **Perceived barriers for cloud adoption for non cloud adopters** Among the most significant barriers are "security issues", "Integration with Existing systems", "data protection and / or privacy Concerns".
- **Support needed to migrate to the cloud**: the highest support demands by non **cloud adopters** are related to « cost benefit analysis » and « security and privacy ».
- Most cloud adopters (4/5) engaged in change management to overcome cloud adoption barriers.
- **Cloud adopters expressed their satisfaction with cloud utilization**

It is to be noted that the results of the survey should be considered with caution since the survey sample is not representative enough. They give however some clear indication as to the growing interest to cloud adoption in many Arab countries.
In an attempt to better validate the outlined trend, the working group has completed its research using available documents and material published on the net mostly by the major cloud providers in order to search for Arab educational institutions that have adopted the cloud. The literature review revealed that:
- Morocco seems to be among the leaders in Africa in cloud computing adoption. Indeed universities of IbnZohr (Agadir), Chouaib Doukkali (El jadida), Sidi Mohammed Ben Abdellah (Fès), Ibn Toufail (Kénitra), Cadi Ayad (Marrakech), Moulay Ismail (Meknès), Mohammed 1er (Oujda), Mohammed V de Souissi (Rabat), Abdelmalek Essadi (Tetouan/Tanger), Hassan II (AinChok- Casablanca), Sultan Moulay Slimane (BeniMellal) et Hassan II adopted « Google Apps For Education » as a service for their students, faculty members and administrative staff [4]
- Morocco has also decided to use the cloud in vocational training: OFPPT, a public institution for vocational training, adopted Microsoft Office 365. OFPPT provides training in 327 training institutes across the country for 500,000 students in 35 fields of study [5].
- In Jordan 2 universities have adopted the cloud: Jordan University of Science and Technology (Jordan) and Princess Sumaya University for Technology.
- In Saudi Arabia, two universities have also adopted the cloud: King Abdallah University of Science and Technology and King Abdulaziz University
- In Tunisia, 4 higher education institutions adopted Google apps for education: the National School of Engineering of Sfax, SESAME University, ESPRIT School of engineering and the National School of computer science in Tunis.
- In addition, institutions in Qatar and Sudan have adopted the cloud: Qatar Cloud Computing Center, Nile Center for Technology Research

To conclude, even though not much information is available, the review of the literature shows that cloud adoption is a movement that is gradually spreading within the higher education institutions of the Arab world. Therefore, it is timely to reflect on the strategy and practical measures to be implemented to support the Arab educational institutions and help them succeed migration to the cloud.

References

[1] Internet Worl Stats,http://www.internetworldstats.com/stats.htm

[2]-Measuring the Information Society Report 2015,

http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf

[3] Sphinx Logiciels,http://www.lesphinx-developpement.fr/

[4] Cloud Computing And Its Role In Education In Morocco, http://www.gjesr.com/Issues%20PDF/Archive-2015/january-2015/9-CLOUD%20COMPUTING%20AND%20ITS%20ROLE%20IN%20EDUCATION%20IN%20MOROCCO.pdf

[5] Microsoft to Provide Educational Opportunities for 500,000 Students and Trainees in Morocco Via the Cloud and New IT Academies, https://news.microsoft.com/2013/01/30/microsoft-to-provide-educational-opportunities-for-500000-students-and-trainees-in-morocco-via-the-cloud-and-new-it-academies/

Chapter 3
Cloud adoption: Policy guidelines

3.1. Introduction

Policy recommendations presented in this chapter are based on the findings related to cloud adoption in the world and particularly in educational institutions. They also refer to the state of infrastructure and ICT development in the Arab world and in particular in a sample of educational institutions as well as their rate of cloud adoption or intentions and preferences with regard to cloud migration.
The development of key policies presented below are also the result of a literature review and intensive exchanges and debates between the members of the project.

Four fundamental policy principles are proposed:

- High quality network
- Always Public cloud first
- "Cloudify" the existing local infrastructures and applications at institutional level
- Adopt a cloud friendly governance model for IT.

The four key policies are further developed according to major stakeholder's profiles: ministerial departments, universities and educational institutions.

3.2. Key policy one: high quality network

The analysis of the ICT status in the Arab world based on the ICT Development Index (IDI) demonstrated a great diversity between Arab countries. In reference to the world average of the IDI of 5.03, three classes of Arab countries have been acknowledged.

With regard to Internet penetration we note that 3 groups emerge: one group often countries whose scores are below the world's average, a group of seven countries of which exceeds the European's average (73.5%) and a group of five countries in the middle. This classification confirms the outcome of the analysis based on the IDI index. It's also an indication that a great deal of effort is to be accomplished in most Arab countries to reach an adequate ICT infrastructure development to enable cloud deployment.
Consequently, Arab countries whose scores are average or weak, need to accelerate the development of their network infrastructure in order to foster and succeed migration to the cloud. The first key policy principle is:

KEY POLICY 1: Give to investment in a high quality network (intranet+access to INTERNET) the highest priority.

It is also necessary to consider the implementation of a backup network and to negotiate contracts to include appropriate Service Level Agreements (SLAs) and suitable terms and conditions.

This key policy is further elaborated according to the 3 stakeholder's profiles as follows.

3.2.1. Main guidelines for the implementation of key policy 1 for ministries in charge of higher education and research

Ministries in charge of higher education and research should negotiate contracts' frameworks and SLAs at national level with Telecom operators to provide a best of breed network access to all institutions regardless of their locations and size. Bandwidth should be appropriately sized and should be upgraded regularly. Ideally, optic fiber and redundancy should be provisioned where possible.

It is essential to coordinate with the different ministries in order to implement a framework of incentives promoting universal access to the internet. Students, teachers and researchers should be able to have access to a reliable internet connection not only on campus but also at home (anywhere anytime).

Budgets for the internet connections implementation and support could be managed at ministry level or at university/institution level.

3.2.2. Main guidelines for the implementation of key policy 1 for universities/ higher education institutions

Universities/ higher education institutions have to allocate the right budgets to deploy a redundant/reliable high-quality intranet, properly sized and supported.
Wireless access should be provided and campus-wide coverage should be optimal and permanently monitored and supported. Wired connections should be available in all the rooms where learning and practical activities requiring maximum reliability and bandwidth are taking place.

It is recommended to negotiate with hardware providers to lower the barrier for acquiring laptops and negotiate with Telecom operator to lower the barrier for gaining 3G/4G internet access on mobile devices to all students, teachers and researchers. If successful, such an approach would make it possible to have a "bring your own device" policy campus-wide.

3.3. Key Policy 2: Always Public cloud first

The survey results of 40 higher education institutions in the Arab world (see chapter 2) showed:

- **A clear move toward cloud adoption**: 52% of surveyed institutions are either in the production, implementation or test phase;
- **A positive perception of cloud benefits by non cloud adopters (not in production phase)**: the most important benefits perceived by the highest percentage of respondents are "disaster recovery capabilities" followed by "hardware cost saving" and "reliable data storage".

However, 40% of institutions prefer to migrate to a private cloud, 20% hybrid cloud and only 9% to a public cloud. This result could be explained by a lack of familiarity with the cloud, 49% all respondents are just familiar or beginning to familiarize themselves with cloud computing. This lack of familiarity may reduce the perception of respondents regarding the importance of the cost to implement a private cloud and the risks of failure and their potential impact on users. In contrast to the private cloud, public cloud can be deployed quickly without the institutions having to carry out developments and investments in infrastructure that can support virtualization and load.

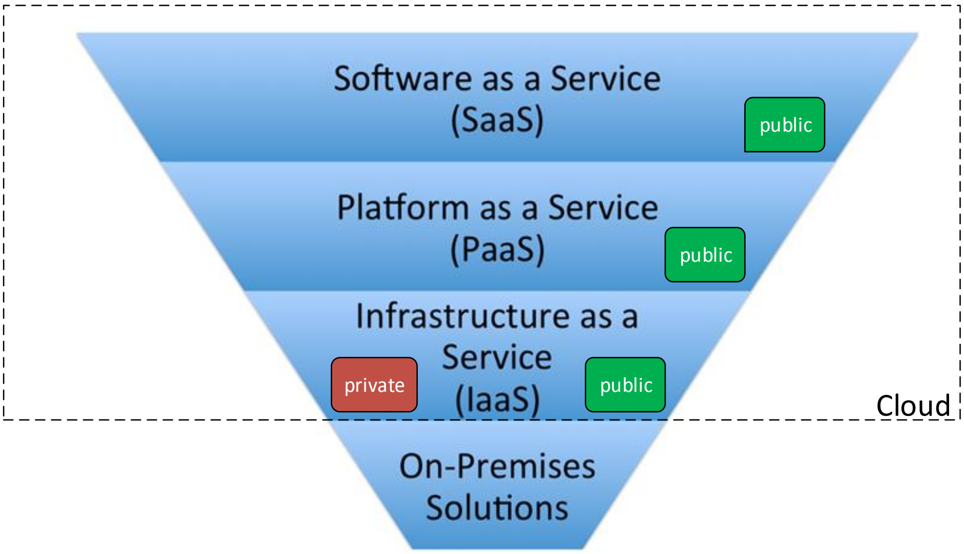
Public Cloud computing offers computing capacity and storage as well as a wide range of services and scientific applications that can be used as needed. It offers the advantage of flexibility, efficiency and enables users to benefit from developments made by suppliers. Services are accessible via the Internet everywhere all the time.

The policy to be adopted consists, therefore, in encouraging the use of public cloud as the first choice whenever it's possible and hence enabling educational institutions fast access to advanced IT and catching up with international practice

KEY POLICY 2: Always public cloud first

Always SaaS first, Use Pubic PaaS/IaaS instead of local infrastructures: As represented by fig. 3.1, Software-as-a-Service should be considered first: it entirely delegates all the unwanted complexity to a service providers and empowers users through seamless access to tailored and effective Web User interfaces. PaaS and IaaS (public if possible or private otherwise) should then be considered as the way to go when it comes to the provisioning of custom applications and capabilities, they expose more complexity but can be harnessed thanks to APIs and automation frameworks. On-premise existing application deployments that (i) have critical constraints or (ii) wouldn't benefit from public or private clouds or (iii) can't be cloudified would continue operating on-premises.

Figure 3.1 Public/private cloud Services



Source [1]: adapted from Cloud Strategy For Higher Education Building A Common Solution (2014), https://Net.Educause.Edu/Ir/Library/Pdf/Erb1413.Pdf

3.3.1.Main guidelines for the implementation of key policy 2 for ministries in charge of higher education and research

Ministries in charge of higher education and research should negotiate contracts' frameworks with the key global and local cloud providers. Engage in ambitious partnerships with those players (such as Google, Microsoft, Amazon, etc.) in order to make the use of cloud in education part of a larger framework of cooperation, promotion of innovation and adoption of cutting-edge IT solutions throughout the country.

Put in place communication strategies contributing to a mindset shift about pervasiveness of cloud solutions. Those strategies should foster the adoption of the “public cloud first” approach and catalyze the everything-as-a-service paradigm shift.

3.3.2. Main guidelines for the implementation of key policy 2 for universities/ higher education institutions

Institutions should review all services and applications in use, benchmark existing SaaS alternatives and migrate if a mature and satisfactory solution exists which also respects the different constraints and legal rules of the institution.
Review all SaaS applications used by similar institutions and adopt those which may empower students, educators and researches and address their different use cases in an innovative and effective way.
For PaaS/IaaS migration, IT departments should train their members on technologies such as containers and clouds automation APIs and tools. IT administrators should use containers to avoid lock in and maximize institution’s agility. They should use brokers to decouple applications and services from the underlying IaaS and to expose a federated layer of management of both public and on premises infrastructures. They should expose self-service portals to the end user and automate user interaction with tools and application as much as possible.

3.4. Key policy 3: “Cloudify” the existing local infrastructures and applications at institutional level

To speed up cloud migration, intermediate solutions may be adopted to improve the management of the existing IT infrastructure without disrupting the operation of the applications used by educational institutions. “Cloudification” solutions give developers the freedom to specify new environments and deploy applications in a faster and easier way. They enable to offer users new possibilities and streamline the operation of the existing infrastructure.
For these reasons, the project team proposes:

KEY POLICY3: “Cloudify” the existing local infrastructures and applications at institutional level

a- Use containers for applications deployment
Deploy Docker with a containers management platform (Kubernetes, Swarm+Mesos for example).
b- Use a mature open source cloud toolkit (OpenStack, OpenNebula). Eventually with virtual machines running Docker engines. Avoid proprietary virtualization technologies.

Option (a) alone represents an easier option than (b) since it doesn’t involve the use of hypervisors and standard/heavy weight virtualization technologies: up to 30% of the resources are consumed by the virtualization overhead hence containers allow for a more effective use of the available IT resources. Cloud toolkits (b), containers and containers orchestration frameworks (a) can be combined and can work together seamlessly offering the best of the two worlds:

"OpenStack is about simplifying the management of datacenter heterogeneity. The value of OpenStack today is that hundreds of vendors have written and are actively maintaining thousands of drivers for physical infrastructure, and OpenStack exposes them as a fabric of APIs and allows for the orchestration of bare metal and virtual infrastructure, whether it is compute, storage, or network. Kubernetes and Mesos are the underlying components of what people typically refer to as a platform as a service. They are great at managing and scaling applications wrapped in containers; OpenStack takes care of datacenter heterogeneity, but has a poor story when it comes to application and container management. You put Kubernetes or Mesos on top of OpenStack, and you get technical nirvana" [2].

3.4.1.Main guidelines for the implementation of key policy 3 for ministries in charge of higher education and research

Ministry in charge of higher education and research should put in place a program for pilot projects to show case the cloudification using containers and cloud toolkits of typical institutions' infrastructures. It is also necessary de publish and disseminate results, know-how and lessons learned.

3.4.2.Main guidelines for the implementation of key policy 3 for Universities/ higher education institutions

Universities/higher education institutions IT departments should train their members on technologies such as containers, virtualization technologies and clouds toolkits. Pilot projects must be put in place to help making informed decisions about the specific container orchestration technology and cloud toolkit that would best fit the needs of the institution.
IT managers can then decide the Dockerization of the existing applications which will then be put in production on the newly cloudified infrastructure (More technical details are provided in Chapter 4).

3.5. Key policy 4: Adopt a cloud friendly governance model for IT

The migration to the cloud requires a change in IT governance models and practices. First, the offer to be made available to users must meet two main requirements: ease of use and flexibility. Secondly the establishment of an intermediary entity should facilitate relationships between suppliers and customers.

Key policy 4: Adopt a cloud friendly governance model for IT

a- Allow end users (Teachers, students, researchers, staff) to make informed choices about the specific services they may want to experiment with or use on the long run. Redefine the role of the IT department to become a service and support provider without limiting the scope of initiatives that users may want to take.

b- Create an entity at national, regional or institutional level that deals with brokerage and procurement in a centralized way. (Delegate those responsibilities to an existing public or private entity where applicable).

That entity would:

(i) Negotiate the SLAs and the general terms and conditions with both Network providers and Global/national public clouds providers when possible/applicable. Make sure the terms and conditions are compatible with all the legal frameworks to which the institutions are subject.

(ii) Expose to individual users and member institutions terms and conditions that engage their individual responsibility in case of a non-authorized/legally reprehensible behavior.

(iii) Centralize the procurement of public cloud resources, eventually negotiate[3 & 4] cost-effective options with global or national public cloud providers (reserved instances etc.).

(iv) Operate the cloud broker and the self-service portal.

(v) negotiate software licenses on behalf of a group of institutions(where applicable) and make those licenses available as a shared pool

(vi) Make it possible for users from a group of institutions to interact/collaborate/share any cloud resources, applications or services they acquire/access through the intermediate entity.

(vii) Take responsibility for a federated/consistent data strategy and data management plan.

(viii) Act as a single point of information for IT/data security and privacy policies/rules.

3.5.1. Main guidelines for the implementation of key policy 4 for Ministry in charge of higher education and research

Ministry in charge of higher education and research should:

- study the different options for the intermediate entity and short-list those which are compatible with existing rules and procurement processes;
- arbitrate between public and private entity options and if the latter option is chosen, select a company based on the legal rules in place;
- otherwise, arbitrate between the creation of a new public entity and the delegation of the new role to an existing one that has the operational capacity for that. Make sure that the entity is governed in a way that gives to the end users/institutions power to contribute to the strategical decision making.

3.5.2. Main guidelines for the implementation of key policy 4 for Universities/ higher education institutions

If the Ministry decides not to operate an entity at national or regional level, the university should(i) create that entity or (ii) attribute its role to a team within the IT department or (iii) contract with a company.

In case of (i) or (ii), Universities/ higher education institutions must make sure that

- all the required human resources and skills are available to meet the intermediation challenges at technical and legal levels;
- the entity is open to end users' contributions to the strategical decision making.

3.4. Other issues related to migration

The success of the operationalization of the above policy choices requires to tackle the following issues.

- **Barriers for cloud adoption**: among the most significant barriers for cloud adoption for non cloud adopters are "security issues", "Integration with existing systems", "data protection and / or privacy Concerns". To overcome these barriers institutions need support. Main support demands requested by institutions are related to « cost benefit analysis » and « security and privacy ».

- **Capacity building**: institutions must make investment in the governing structures, organization processes, people and their skills required to make cloud technology an essential element in how the organization services are managed. To gain the full advantage that result from sharing technical, human and resources, institution managers need to understand the resource capabilities they possess as well as the resources that the cloud providers make available. IT managers need to understanding what capabilities are offered and how they can be combined with internal resources, and developing a plan to leverage these combined resources.

- **Culture Change:** aside from various technical issues, institutions managers fail to recognize that they have to take into consideration the cultural environment of their own company. A very common problem that institutions may encounter is the internal politics, which will tend to slow down cloud computing adoption. Many administrators will reason that they will have difficulty controlling and deploying application updates. They will resist cloud computing solutions due to fear of the unknown. The internal structure of IT departments in institutions and the legacy approaches to their process management will definitely curb the implementation of cloud computing projects. Culture change must occur. This change cannot be done only by means of speech and information; cloud experimentation by IT specialists and users would greatly facilitate culture change.

- **Strategy**: Institutions must plan for cloud computing as a strategic choice rather than as an outsourcing arrangement or a technical platform. Organizations that approach the cloud as a replacement for internal technology solutions limits its ability to think broadly about how it can support the strategic direction of the organization. A limited technical perspective will not tend to lead to the examination of the organizational structure, governing processes, organization architecture and culture.

- **Education**: New curriculum needs to be addressed in parallel with the deployment of cloud technologies. For some students the curriculum may concern how to use cloud technologies. For others, since it becomes easier to access shared pools of virtual resources, for instance computational resources, it becomes urgent to renovate the contents of courses in a direction of more practical work with the computers and (open source) software.

3.5. Conclusion

Chapter 3 dealt with the political choices to be made for migrating to the cloud. Four key policies were proposed:
- High quality network
- Always Public cloud first
- "Cloudify" the existing local infrastructures and applications at institutional level
- Adopt a cloud friendly governance model for IT.

The four policies were further developed according to the 3 major stakeholder's profiles: ministerial departments, universities and educational institutions.
Finally other issues related to migration have been presented.

In line with the proposed policies and in order to move to an operational level, the following chapter (4) will provide practical guidelines to help educational institution and information technology (IT) managers to plan and operate application and workload migration to the cloud. It provides a detailed roadmap that covers a list of technical steps for migration.

References

[1] Cloud Strategy For Higher Education Building A Common Solution (2014), https://Net.Educause.Edu/Ir/Library/Pdf/Erb1413.Pdf

[2] Will OpenStack, Kubernetes, Or Mesos Control Future Clusters? (2015), http://www.nextplatform.com/2015/08/13/will-openstack-kubernetes-or-mesos-control-future-clusters/

[3] HNSciCloud: Bridging Cloud Computing Innovation & Open Science (2016) ,http://www.helix-nebula.eu/

[4] Procurement Innovation for Cloud Services in Europe (2014), http://www.picse.eu/

Chapter 4
Cloud migration: Implementation and guidelines

4.1. Introduction

Migrating applications and workloads to the cloud starts with a clear and methodical strategy, and success is achieved only through detailed planning and strong technical execution. The objective of this document is to provide a practical guide to help educational institution and information technology (IT) managers to plan and operate application and workloads migration to cloud. It provides a detailed roadmap that covers a list of technical steps for migration.

The previous chapter 3 presented and developed a set of strategic scenarios and activities for IT managers to determine what applications can be moved (or not) to which delivery models (public, private, hybrid/brokerage).The next steps to develop a detailed technical plan that implements the migration strategy and scenarios. This is exactly the aim of the chapter which discusses and details all technical migration steps and requirements related to cloud service types (SaaS, PaaS and IaaS) and cloud deployment models.

This chapter presents first the set of educational and research services and applications candidates for migration to Cloud Computing. Three major migration scenarios are then discussed and detailed for migrating applications, services and workloads to clouds (SaaS migration, IaaS and PaaS migration, cloud service brokerage for IaaS and PaaS migration). A migration use case is presented for each scenario.

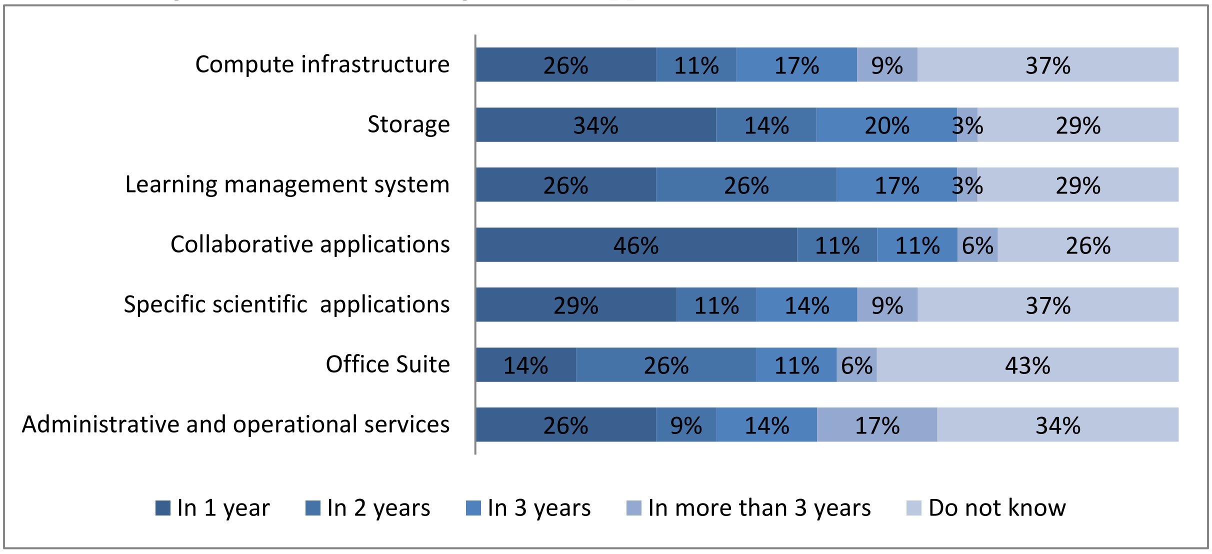
4.2. Cloud services for education and research: needs and requirements

This section relies on the situational analysis in Arab countries conducted by chapter 2. The objective is to identify the needs for educational institutions and to explore how existing cloud services can meet these requirements. It is in line with the recommendations and policies described in chapter 3 which puts emphasis on public clouds and broker-based approaches.

The results of the survey (see chapter 2) of 40 higher education institutions in the Arab world showed that the preferred applications to migrate to the cloud during the first year are « collaborative application » with 46% of respondents, followed by storage service (34%) and scientific applications (29%). They are followed by «learning management system" (26% the first year and 26% the second year).

The next sections present migration scenarios explaining how these applications listed above can be moved to various delivery models (public, private, hybrid or brokerage)and different cloud service types (Saas, PaaS and IaaS).

Figure 4.1Intention of migration IT applications and services to the cloud



4.3. Applications and Workloads' Migration Scenarios

4.3.1. Migration scenarios

This section presents four scenarios for migrating applications, services and workload to Cloud Computing:

- **SaaS Migration**: Replacing in-house applications with new SaaS models.
- **IaaS and PaaS Migration**: Migrating in-house applications and workload to (private or public) IaaS and PaaS Clouds.
- **Cloud Service Brokerage for IaaS and PaaS Migration**: Using Cloud Brokers to manage the delivery of cloud services and negotiates relationships between educational institutions and (private or public) Clouds.

These migration scenarios are detailed in the next subsections.

4.3.2. SaaS Migration

Many schools, colleges and universities have moved their email, collaboration and communication services to the cloud. The objective is to completely replace in-house applications and services (including Email, Video Conferencing, Storage, Social Network, Office tools…) with new SaaS models.

Both Google and Microsoft offer educational tools and software as a service for free in many countries. While both products--Google Apps for Education [1] and Microsoft's Office 365 Education [2] --have similarities, administrators must decide which platform to adopt with respect to the terms and conditions, regulatory issues, and the interoperability with their existing suite of applications.

The Amazon Educate [3] global initiative offers a third alternative and provides students and educators with AWS credits for use in courses and projects. Those credits make it possible to use any AWS services (for example Workspaces).

Google and Microsoft expose standard eLearning capabilities in a SaaS mode. The Amazon's cloud powers various innovative SaaS offerings for eLearning. Echo360 [4] for example is an AWS-backed active learning platform, its SaaS classroom capabilities replace several classroom learning technologies that are not scalable or designed to support the diversity of modern faculty/student interaction.

4.3.2.1. Migration to Google Apps for Education

Google Apps is a software as a service from Google that provides customizable versions of several Google products using a domain name. It offers several Web applications and office suites, including Gmail, Hangouts, Drive, Google Calendar, Docs, Sheets, News, Slides, Play, Groups, Sites, and Vault. Google Apps is 100% free, with no hidden costs. The main disadvantage of Google Apps is that users require a Google account. Permissions are required for students under the age of 13 to create Google account.

Figure 4.2 Google Apps for Education

Source [1]: https://www.google.com/edu/products/productivity-tools/

Domain administrators in charge of moving their student and/or faculty/staff populations to Google Apps for Education will find a guide with a step-by-step outline for completing the technical aspects of the deployment, including relevant help center articles and videos. Google Apps deployments typically take between 1 day and 6 weeks[5].

4.3.2.2. Migration to Office 365 for Education

Microsoft Office 365 provides schools with free suite (including mailing, storage, messaging, web conferencing, document editing…) and tools (such as OneDrive, Lync, Skype, Office Mix and OneNote). Students and faculty are able to use any browser to create documents in Word, PowerPoint, and Excel. Developers could develop new apps on top of the open production platform Office365 to satisfy institution-specific needs and fulfill the evolving needs of educators.

Figure 4.3 Microsoft Office 365 for Education

Source [2] :https://www.microsoft.com/en-us/education/

Several webinars and webcasts as well as deployment guides are available to help IT administrators deploy Office throughout their institutions [6].

4.3.2.3. Migration to AWS

4.3.2.3.1.Migration to AWS Workspaces

The AWS Educate [5] credits allow Students and educators to use for example Amazon Workspaces which is a managed desktop computing service in the cloud. Students and educators can easily provision cloud-based desktops and gain access to the documents, applications, and resources they need from any supported device including Windows and Mac computers, Chrome books, iPads, Kindle Fire tablets, and Android tablets.

A large number of resources is available for migrating to AWS Workspaces in [7].

4.3.2.4.Migration to AWS/echo360

The Echo360 Active Learning Platform Capabilities include active learning tools for student engagement, learner analytics, lecture capture and content management, all integrated into a scalable campus system that allows institutions to make better use of their LMS capabilities, supports extended learning modalities like hybrid distance learning and captures behavioral learning data to support predictive analytics.

Echo360 online Help system provides guidance to teachers, students and administrators and can be found in [8].

4.3.3. Migration to an IaaS /PaaS

There are two target service models for the migration of existing applications to Clouds– Infrastructure as a Service (IaaS) and Platform as a Service (PaaS). This section presents the different technical steps for migrating in-house applications and workload to public or private IaaS and PaaS Clouds.

4.3.3.1. Technical Steps

Cloud Migration Phase	Technical Steps

<table>
<tr>
<td>Cloud Migration Project Plan</td>
<td>
Prepare a formal project plan:

Identify the cloud migration project team including technical staff, IT managers, decision makers, ... The technical staff must have the skills to migrate applications to cloud environments. The creation of cloud images and their deployment may involve skills new to the educational institution.
Organize a Kick-off meeting to discuss migration options, priorities, technical aspects, objectives...
Conduct a technical feasibility study:

Constraints and requirements related to workloads and applications migration should be addressed such as required resources (CPU, RAM, Storage, BW…) and cost estimates.
Study the Integration issue of migrated applications: Define dependencies between the application components being migrated and those remaining on-premises. Migrating some application components to the Cloud may involve performance and configuration changes including access control, network latency and workflow process.
Check security issues:

Public cloud case: The institution is responsible for authentication and authorization at the application level. A single identity service should be set up for access control and role-based permissions. The cloud service provider is responsible for security controls at infrastructure level.
Private cloud case: The institution is responsible for authentication, authorization and security control at both application level and Infrastructure levels.

Check and audit the existing networking architecture: The migration process should consider the performance and resiliency of the intra-cloud networking (in the case of private cloud) and the access network infrastructure (in the case of public cloud). Technical stuff must look carefully at the architecture, topology and some metrics (Latency, Bandwidth..) of the LAN/WAN networks of their institution. Additional techniques such as dynamic VPNs (Virtual Private Clouds) can provide resilient and secure cloud environments.
Check compatibility of monitoring tools and explore the deployment of a unified monitoring platform.

</td>
</tr>
</table>

	○ Check if the application design, eventually with few refactoring, would allow for a scalable and elastic multi-nodes deployment. Choose to deploy on a single high capacity instance otherwise. ○ Check if the existing applications and productions servers can be migrated to containers (e.g. *Docker)* to avoid vendor lock-in and to simplify the migration process. • Define the roadmap, task assignment, performance indicators. • Choose project management tools and set up an Open Trouble Ticket System and issue resolution process
Set up the cloud environment	**<u>Cloud IaaS Case</u>** 1. **Set up and configure the cloud infrastructure:** ○ Case 1: Private IaaS ▪ Specify the hardware requirements ▪ Purchase, provision and configure the data center servers. ▪ Deploy the virtualization layer and the Cloud Manager ○ Case 2: Public IaaS ▪ Access to the self-service portal of the public cloud infrastructure ▪ Discover the set of IaaS tools, services and APIs offered by the cloud provider to configure and manage the infrastructure 2. **Set up security rules and policies** ○ Configure Firewalls, VPN routers, Security groups, etc **3. Configure the storage service and create volumes** **4. Create virtual machine images/docker container files or images and export them to the cloud.** **5. Instantiate Virtual Machines or Docker containers** ○ Create virtual machines from machine images (or docker containers from docker images) and attach them to their respective volume units **6. Configure Network** ○ Configure VLANs, VPN tunnels and gateways ○ Configure the Domain Name Service (DNS) ○ Test virtual machines connectivity **7. Configure identity services** **8.** Establishing and testing connections to the in-house directory server (LDAP, Active Directory, etc) from cloud machine instances.

	9. Install the execution environments if needed o Install and configure databases, OS, application server software and libraries on the virtual machines or containers. **<u>Cloud PaaS Case</u>** **1. Access to the public cloud platform** **2. Discover the platform environment and the software stack (such as application servers and databases) offered to support migrated applications** **3. Check if the PaaS execution environment offers the configuration and tools required by the application (database and application servers versions, monitoring tools, etc)**
Set up the Applications in the cloud	**<u>Cloud IaaS Case</u>** 1. **Install and Configure the application within a virtual machines (or a container)** o Upload the application and installation the cloud's virtual machines. o Use Docker to run applications inside containers (https://docs.docker.com/v1.8/userguide/dockerizing/) **2. Implement and test Integrations** o set up all integrations between the migrated application components and on-premises resources. **3. Configure Management and Monitoring tools** o Set up all monitoring tools and solutions to supervise the running application **<u>Cloud PaaS Case</u>** **1. Upload applications via the PaaS self-service portal** **2. Implement and test Integrations** o set up all integrations between the migrated application components and on-premises resources.
Set up a Pre-Migration Prototype	**1. Check Cloud Services Agreement** o In the case of public IaaS, ensure all contractual aspects are in place with the cloud provider. **2. Stop in-house servers** o Schedule downtime(in case of offline migration) for existing systems to move applications to the pre-migration prototype o Make complete backups of the on-premises databases and

	configuration scripts related to the applications 3. **Export all configuration files and databases to the cloud servers.** **4. Run test scripts to validate:** o Application and data migration o Connectivity from all endpoints o Authentication mechanisms **5. Start the cloud hosted applications** o Connect the on-premise system to the migrated applications **6. Test and Validation** o Ask a preselected group of test users to validate the work environments and systems are functional o Evaluate performance **7. Terminate the pre-migration prototype and restart the on-premises application**
Move to Production Cloud.	**1. Update the migration plan,** o Organize a meeting between all stakeholders to discuss ▪ All issues encountered during the pre-migration phase ▪ Migration Tasks to be added or removed ▪ Changes in resources required vs initially expected **2. If needed, Notify users about any new procedure to access the applications** **3. Final check of licensing systems, security parameters and application monitoring tools for the production cloud environment** 4. **Close the Migration Project, move to the production environment and dismiss the deprecated servers.**

4.3.3.2.Public cloud case: migration to AWS

Amazon provides a set of whitepapers targeted at architects and technical decision makers looking to build a cloud migration strategy. Here is a summary of one of the white papers provided, adapted to the use case of an institutional Moodle migration to AWS (typical multitier dynamic web application) [9]

4.3.3.2.1. Use Case description

Moodle is an open-source e-learning platform (aka Learning Management System) that is widely adopted by educational institutions to create and administer online courses. For larger student bodies and higher volumes of instruction, Moodle must be robust enough to serve thousands of learners, administrators, content builders and instructors simultaneously. Availability and scalability are key requirements as Moodle becomes a critical application for course providers. Moreover, the usage pattern of Moodle and the resources it requires to cope with the load may vary: Several teachers within institution A have decided to set up online exams using Moodle Quizzes, thousands of students may connect simultaneously during exams periods and the responsiveness of Moodle during those periods is highly critical.

4.3.3.2.2. Migration Steps

Cloud Assessment:

"On financial assessment, the technical program manager at institution A's IT department was able to map the hardware configuration of physical servers to equivalent EC2 instance types and estimate the combined storage and bandwidth requirements. The team realized that they could free up the current infrastructure for other internal projects, discontinue a tape backup maintenance contract and reduce their operating expenditures by 30%. During the technical assessment, they discovered that the entire institution technology stack was compatible with AWS and could run on Amazon EC2 Instances with Linux. They also discovered that Moodle can be configured to run at peak capacity (7 Servers) during exams periods, medium capacity (4 Servers) on weekdays and low capacity (2 Servers) on weekends.

Proof of Concept:

The IT team was skeptical about the relational database migration. To test, they decided to build a proof of concept application. During the proof of concept, the team learned the following techniques: starting, terminating and configuring Amazon EC2 virtual machine instances and Amazon RDS DB Instances, storing and retrieving Amazon S3 storage objects, and setting up elastic load balancers using the AWS Management Console. They learned a ton about AWS and saw that they have full control over the environment, and felt a lot more confident about moving to the next step. The relational database files (binary and transaction logs) were moved to Amazon RDS instances using the standard mysql import utility. The team was able to successfully test and migrate all data to a DB instance, get performance metrics using Amazon CloudWatch, and set retention policies for backups. They built migration scripts to automate the process and created awareness within the organization by organizing a brownbag session and successfully demonstrated their work to their peers.

Data Migration

Once the proof of concept was complete, the team decided to move all of the Moodle's static course material (video, audio, etc. files) into an Amazon S3 storage bucket, created a CloudFront distribution of that Amazon S3 bucket, and modified the references in web pages so that end-users get the content directly from Amazon S3 and Amazon CloudFront. With a few scripts and the AWS SDK for Java library, they were able to transfer all data from tape drives and upload it to Amazon S3.

Application Migration

During the application migration phase, the IT team launched both small and large instances for their web and PHP servers. They created AMIs (Amazon Machine Images) for each server type. AMIs were designed to boot directly from an EBS volume and fetch the required version of Moodle during launch. They modified their build and deployment scripts to use the cloud as an endpoint. Security Groups were defined to isolate web servers from the applications and database servers. Testing (functional, load, performance etc.) was performed to ensure that the systems were performing at expected levels, and that exit criteria for each component were met.

Co-existence Phase

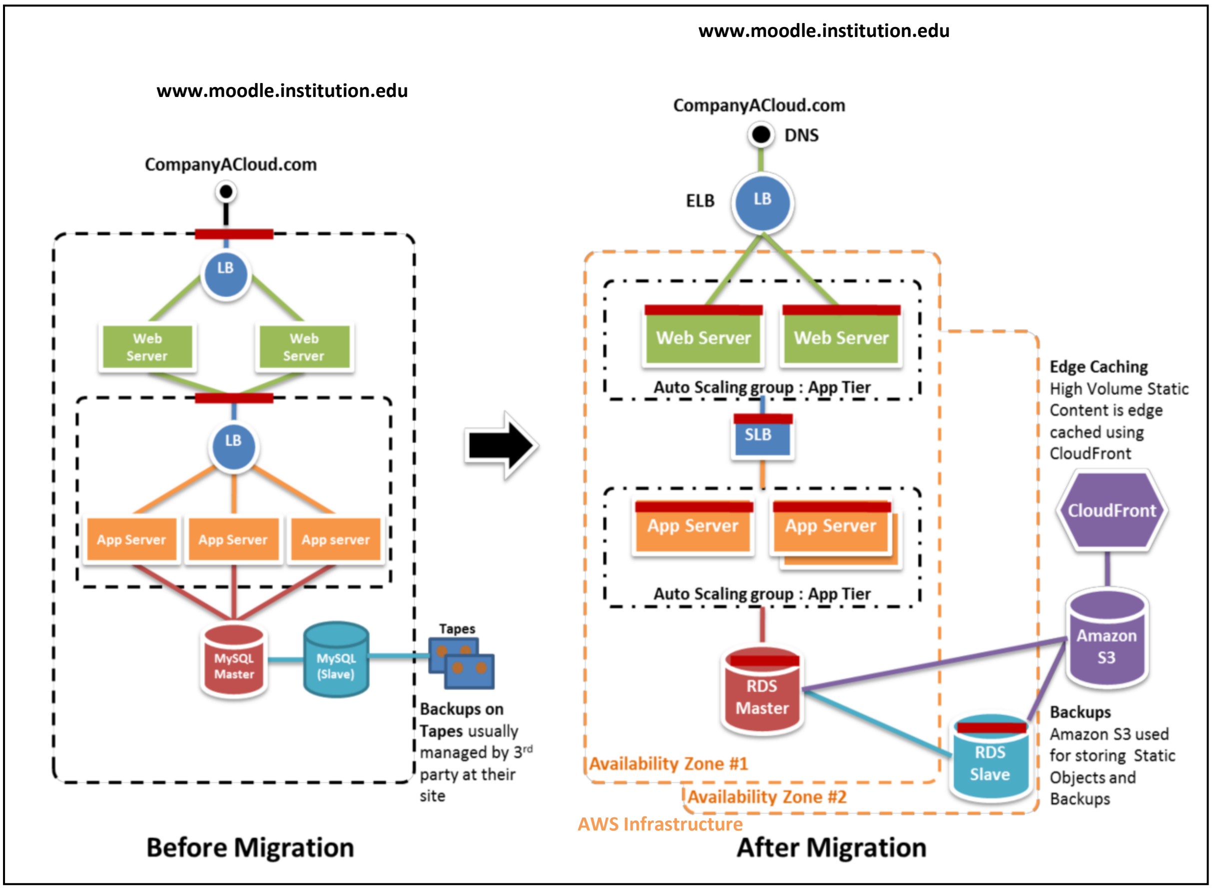
During the migration phase, the local infrastructure was not deprecated immediately. Institution A employed a hybrid migration strategy during the migration of all Moodle-related servers. The configuration of the on premise hardware load balancer was modified to send requests to the new instances in the cloud. For a short duration, the load balancer was routing traffic to the servers in the cloud as well as to the physical servers. After verifying that the servers in the cloud were performing at required levels, the physical servers were dismissed one by one, the load balancers were updated, and all of the web traffic was being served up by the EC2 instances running in the cloud.
After testing was completed, the DNS was switched to point to the cloud-based web servers and the application was fully migrated to the AWS cloud.

Optimization

During the optimization phase, the IT team analyzed their utilization patterns and realized that they could save 30% if they switched to Reserved Instances. They purchased four Reserved Instances (2 for web servers and the other 2 for the PHP servers). They built additional scripts to run their web application in 3 different "modes": weekend, weekday and exams mode.
These modes defined the minimum number of servers to run. The team also integrated Amazon CloudWatch into their existing dashboards so that they can monitor the system metrics of every instance in their cloud fleet" [9]

Mode	Web Servers	App Servers
Weekend	1	2
Weekday	2	3
Exams	5	7

Figure 4.4 Migration to AWS

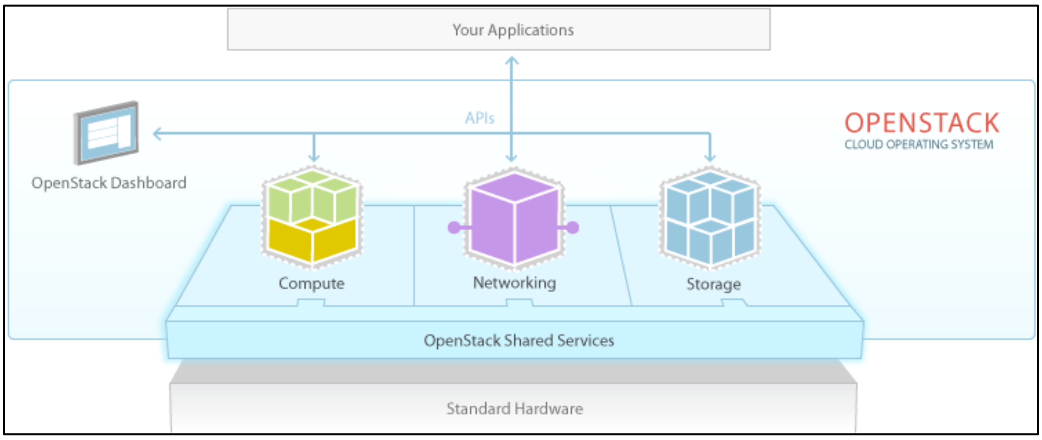


Source [9]: http://d36cz9buwru1tt.cloudfront.net/CloudMigration-scenario-wep-app.pdf

4.3.3.3.Private cloud case: Migration to OpenStack

OpenStack is a flexible and extensible open source cloud management platform for private and hybrid cloud deployment models [10]. It offers a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter. These resources are managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface. Cloud infrastructure services are exposed through ReST APIs.

Figure 4.5 OpenStack



Source [10]http://openstack.org

4.3.3.3.1. Migration steps

Choose hardware

When choosing the hardware on which the OpenStack environment will be deployed, the following characteristics should be considered:

- **CPU --** Consider the number of expected virtual machines to deploy in the cloud environment and the number of core per virtual machine. If the environment is used for heavy computational work, more powerful CPUs are required than environments used primarily for storage, for example.
- **Memory --** Depends on the amount of RAM assigned per virtual machine and the controller node.
- **Storage --** Depends on the local drive space per virtual machine, remote volumes that can be attached to a virtual machine, and object storage.
- **Networking --** Depends on the selected Network Topology (VLAN, GRE, VLANx..), the network bandwidth per virtual machine, and the network storage.

Eventually, get support in defining requirements using the OpenStack Hardware BOM Calculator [11].

Choose the Hypervisor

A hypervisor provides software to manage virtual machine access to the underlying hardware. The hypervisor creates, manages, and monitors virtual machines. OpenStack Compute supports many hypervisors to various degrees, including: KVM, LXC, QEMU, VMware ESX/ESXi, Xen, Hyper-V, etc.

KVM is the most widely adopted hypervisor in the OpenStack community followed by Xen. LXC, VMware, and Hyper-V are often used depending on the context.

Deploy OpenStack's cloud manager

To deploy OpenStack, the IT manager can rely on a number of tutorials, guidelines and automated deployment tools available on line. Some references are listed as following:

- **Reference Guideline**: Users, Operations and Administration Guides are available on [12]
- **DevStack** is widely used to quickly setup a fully operational OpenStack proof of concept infrastructure.
- **Compass**: an open source deployment manager for OpenStack. It is, to a pool of servers what the Live CD is to a single box: It bootstrapps the server pool.
- **Fuel** is an open source orchestration layer on top of Puppet, MCollective, and Cobbler. Fuel codifies Mirantis' best practices of OpenStack deployment.
- **Triple O**: installs, upgrades, and operates an OpenStack cloud on top of an existing OpenStack basic infrastructure.

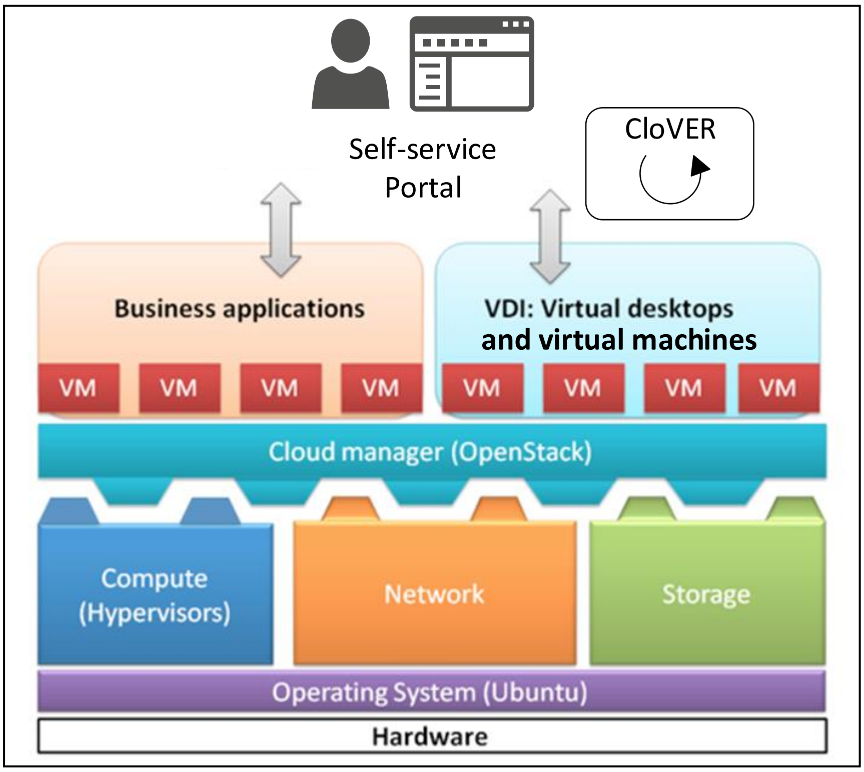
4.3.3.3.2. Migration Use case description

An educational institution (Sfax University-Tunisia) is expecting to migrate to a private cloud based on OpenStack. The objective is to:

- Migrate business applications (e.g. office order, timetable, mailing service, web service, etc…) to the private cloud
- Replace traditional PCs deployed in practical work classrooms with thin clients connected to a private cloud based VDI (Virtual Desktop Infrastructure).
- Offer virtual machines as a service for researchers to deploy their scientific workload.

The IT managers first purchased servers (datacenter) running Linux Ubuntu Server and deployed the KVM hypervisor and the OpenStack cloud manager. The private cloud offers a self-service portal for administrators to allocate virtual servers (to run applications) and for end users (teachers, students, researchers, etc.) to access their own virtual desktops and virtual machines. Then they acquired CloVER [13], a platform based on OpenStack that delivers virtual desktops to users. Virtual desktop is a computing environment where the machine's Desktop interface is delivered to a remote user over a network using some remote display protocols such as VNC, RDP, or PCoIP. CloVER offers a self-service portal to end users to create and access on demand virtual desktops from anywhere and using any device (smart phone, thin clients…). To address the needs of the different departments, a threshold of 240 simultaneous VM in operation was set and required the following configuration: (i) 12 servers with XEON E5645 processors (8 cores, 2.4 Ghz)(ii)128 Go RAM (iii) 40 To of SATA storage (iv) 10 Gigabit switch

Figure 4.6 OpenStack deployment Sfax University



4.3.4. Cloud Service Brokerage for IaaS and PaaS Migration

4.3.4.1. Overview

Cloud service brokerage provides the intermediary between educational institutions and public cloud providers (like Amazon AWS, Google Compute/App Engine, Microsoft Azure) by aggregating and offering multiple cloud resources and services that best suits their needs. Cloud broker is a component that acts as an intermediary between the cloud providers and consumers (universities/schools).

The broker represents a single interface with multiple (public and private) clouds. It makes it possible to provision, manage, monitor and use virtual machines and/or Linux containers (such as Docker's containers).

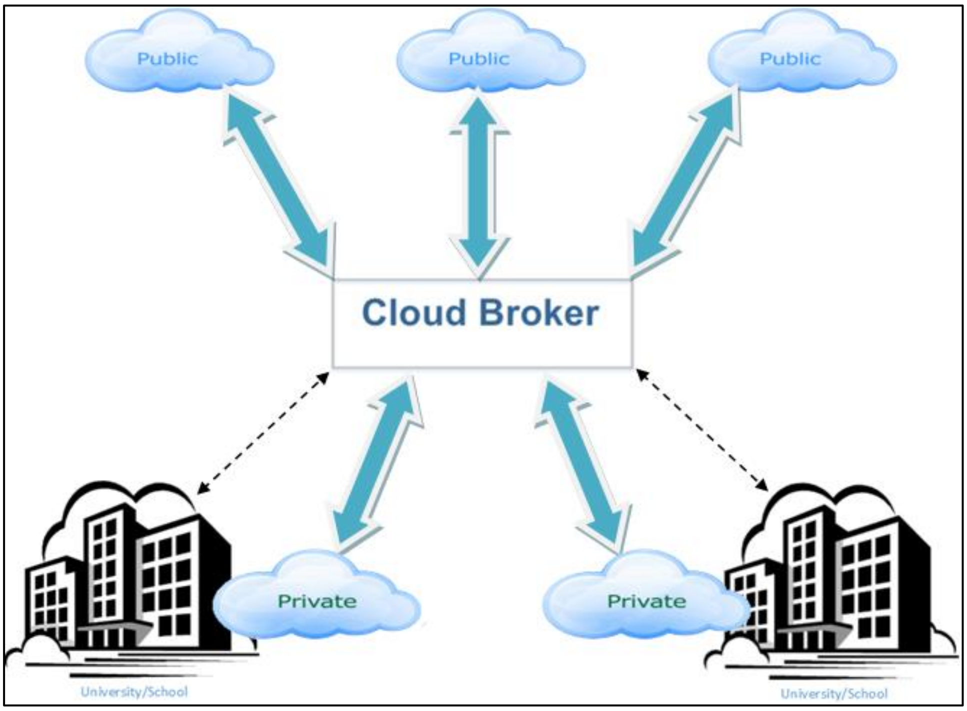
The broker may be deployed:

- **privately**: on premises, within one institution's data centre. It could be a customized third-party technology (open source or proprietary) or a technology developed and tailored to the specific needs of the institution (and leveraging existing open source libraries and toolkit for clouds federation such as jClouds [14]
- **in a community fashion**, under the supervision of supra-institutions entity (e.g. ministry of research and education). The Broker would be used as a federation layer for accessing public and private clouds and it could also act as a procurement delegation and federation entity.
- **as a paid external service** to which institutions may subscribe.

Cloud brokers would allow educational institutions to integrate their in-house applications or their hosted private cloud services with resources offered by IaaS or PaaS cloud providers. The cloud brokerage concept allows institutions to efficiently select the appropriate cloud model (private, hybrid or public) that best suits their requirements. They can also switch smoothly from a cloud model to another and also between cloud vendors while ensuring service continuity and reliability.

IaaS-style Clouds commonly expose their features and services via REST and/or SOAP APIs (application programming interfaces). Cloud broker can manage multiple public clouds using their APIs. In order to avoid vendor lock in and to ensure services' portability and interoperability, it is recommended to deal with public clouds offering extensive and well documented APIs and allowing virtual artifacts to be easily exported. It also important to extensively rely on containers technologies such as Docker.

Figure 4.7 Broker Architecture



Cloud brokers are often made accessible through self-service portals which act as a central hub within the organization and allow users to seamlessly provision, access and share the IT capabilities and services they need without having the IT administrators directly involved. The Web Interfaces provided by public clouds and private clouds toolkits are too complex to handle for regular non-IT user.

4.3.4.2.Example of cloud brokerage platforms

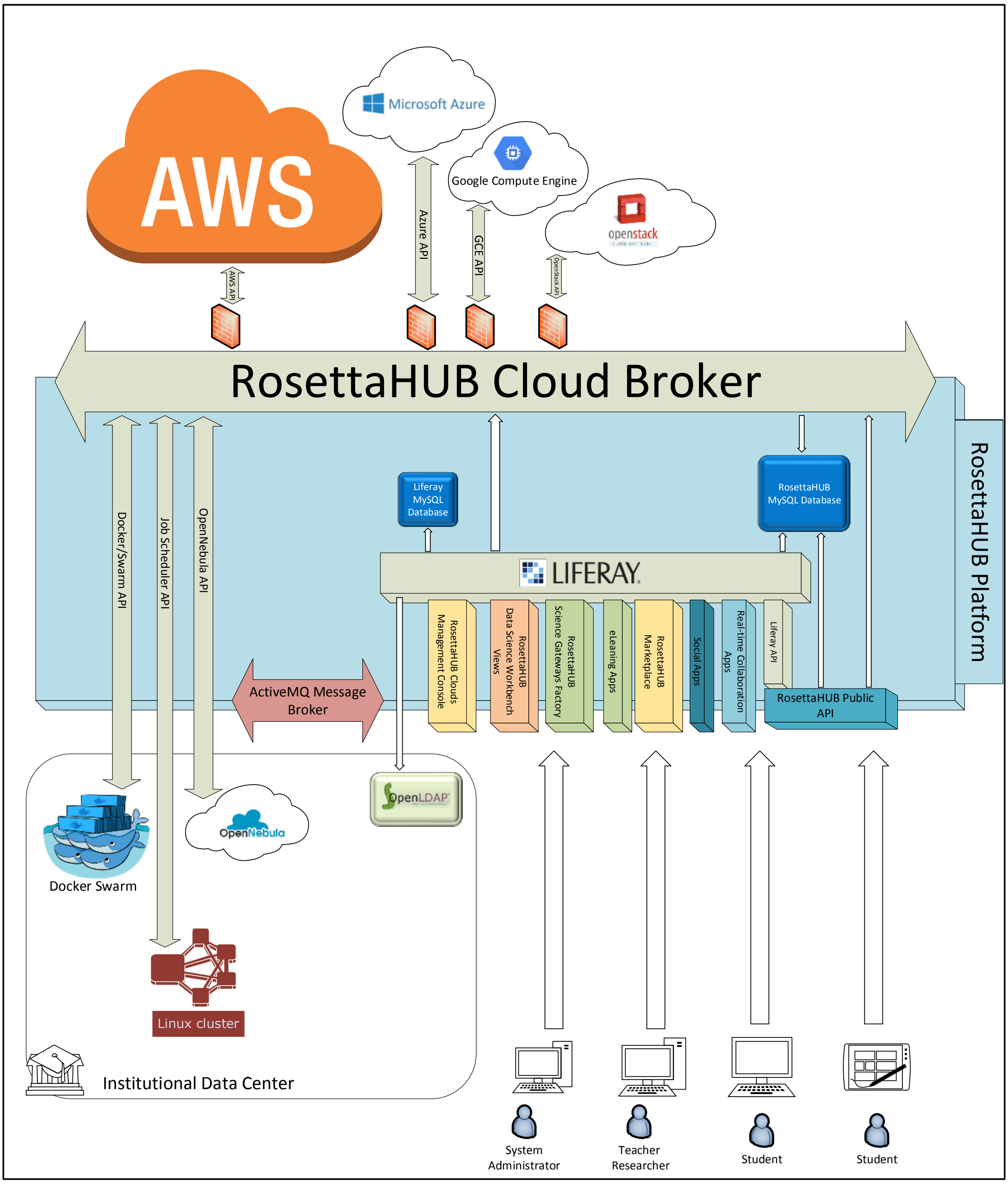
4.3.4.2.1. RosettaHub

RosettaHUB [15] is an innovative eScience and eLearning platform. It provides researchers, teachers and students with a streamlined experience in their day-to-day interactions with (i)clouds, HPC clusters and supercomputers (ii) data science environments, tools and libraries (iii) data storage capabilities (iv)big data frameworks (v) peers and collaborators

RosettaHUB relies on a set of federation concepts and frameworks that include

- A Cloud Broker, capable of provisioning and managing instances of virtual machines and Linux Containers on any public or private cloud and any Linux cluster. It uses the open APIs of the different infrastructures and exposes (i) a “universal” object model for programmable infrastructures (ii)SOAP and Restful APIs (iii) unified management web consoles. It is accessible through a self-service portal allowing researchers, teachers and students to acquire and use the resources they need and it enables many scenarios of cloud resources sharing and collaboration
- A universal “cloud-based data science engine” concept and API which integrates tools such as R, Python, Julia, SQL, Scala, Spark, Mathematica into one single consistent framework so that they all share their workspaces, variables and functions.
- A universal Web Workbench capable of exposing all the capabilities of data science engines within an extensible and real-time collaborative context.
- A framework for building and publishing custom Science Gateways, eLearning environments and interactive big data applications.
- A set of clients and Software Development Kits calling the RosettaHUB APIs and making it possible to access all the platform’s capabilities from desktop applications including MS Office.

Figure 4.8 RosettaHUB platform



Source : www.rosettahub.com

Usage scenarios

Use Case 1

The Leibniz Supercomputing Centre (Leibniz-Rechenzentrum, LRZ) [16]provides the following services to the scientific and academic communities in Munich: (i) General IT services for more than 100,000 university customers in Munich and for the Bavarian Academy of Sciences and Humanities (BAdW) (ii) A powerful communications infrastructure called the Munich Scientific Network

(Münchner Wissenschaftsnetz, MWN) (iii) A competence centre for data communication networks, archiving and backup of large amounts of data on extensive disk and automated magnetic tape storage (iv) A technical and scientific high performance Supercomputing Centre for all German universities.

RosettaHUB is forseen to be a one-stop shop for accessing the LRZ compute resources including the OpenNebula cloud and the HPC Linux Cluster as well as public clouds (AWS). Researchers, teachers and students would access the platform with their LDAP credentials and create virtual collaborative research environments inside or outside the organization firewalls. Researchers interact with the compute resources they seamlessly acquire (from the RosettaHUB self-service portal) via tools such as the RosettaHUB data science workbench, noVNC, Jupyter, RStudio, Excel, Word, etc.

Use Case 2

Within Coursera or any MOOCs SaaS platform, RosettaHUB would make it possible for students to access from their browsers the baselines of tools, scripts, data sets required for a specific set of exercises. Those tools become accessible in one-click, as a service, in a Google-Docs-like collaborative fashion. Student Aicha can share her environment with student Fatima the way a spreadsheet is shared on Google Docs and they solve the course's exercises in collaboration. Student Ali can share his environment with teacher Radhia and receive guidance interactively. Teacher Radhiacan access all the history of interaction between Ali and his virtual working environment for evaluation or pedagogy adjustment purposes.

Use Case 3

Teacher and Students create accounts on the public version or RosettaHUB [15]. Teacher Mohamed prepares a big data training involving use of Spark and Hadoop. Teacher Mohamed creates a new group and adds all his students to that group. Teacher creates and shares tokens backed by a Google Compute Engine free account with the group. Students use the tokens they have been granted to start Spark and Hadoop clusters and follow the steps provided by Teacher.

Migration aspects.

- a dedicated scalable RosettaHUB portal linked to the educational organization's specific authentication/authorization system (such as OpenLDAP) can be made available to the institution upon request.

- IT administrator register the institution's AWS account with RosettaHUB. That account is eventually provisioned with credits from AWS Educate.

- IT administrator generates tokens and grants them to teachers. Those tokens give indirect/controlled access to AWS and allow the teachers to start machines with predefined capacities on AWS.

- Teachers make use of those resources by taking control over the containers they host. Teachers create and customize virtual e-research and e-learning environments and register snapshots of those environments with RosettaHUB.

- Teachers redistribute AWS tokens (bundled with learning environments snapshots) as needed to students.

- IT administrator registers private clouds, Docker clusters and Linux clusters with RosettaHUB and grants access permissions to users according to their roles.

4.3.4.2.2. CloudSelect broker

CloudFx group has launched CloudSelect, an innovative open standards-based Cloud Services Brokerage platform. The CloudSelect platform provides value-added services based on performance monitoring, capacity management, security and availability and full stack provisioning. It enables enterprises and service providers to offer "Anything as a Service" -- including underlying IT infrastructure, application development platforms and fully configured business software applications. More information about the broker is available on line [17].

4.3.4.2.3. CompatibleOne broker

CompatibleOne is an open source cloud broker based on open standards such as OCCI (open cloud computing interface). The CompatibleOne broker allows:

- Rapid selection of the most appropriate provider
- Full automation of the provisioning and the migration process of resources
- Elimination of any form of vendor lock-in and permit full interoperability
- Monitor and control the service delivery quality of cloud providers

More information about the broker is available on line [18].

4.3.4.2.4. JellyFish broker

JellyFishis an open-source cloud broker platform, allowing enterprises to effectively manage their cloud resources. It provides a framework and platform for organizations to act as internal or external brokers of cloud services. Project Jellyfish adds a layer of core broker-type functionality to cloud management, such as advanced catalog search and compare, project-based workflows, collaboration, dashboards, quotas and service blueprinting.
More information about the broker is available on line [19].

References

[1] Google Apps for Education (GAE), https://www.google.com/edu/

[2]Microsoft's Office 365 Education; https://www.microsoft.com/en-us/education/

[3] AWS Educate home page: https://aws.amazon.com/education/awseducate/

[4] AWS Echo360 Online site, http://echo360.com/

[5] GAE Guideline, https://sites.google.com/a/googleapps.com/university-guide-to-going-google/technical-integration

[6]Office365 Guideline, https://www.microsoft.com/en-us/education/products/office-365-deployment-resources/default.aspx

[7] Amazon Workspace, http://aws.amazon.com/workspaces/resources/

[8] Echo 360, http://help.echo360.org/index.htm#t=Main_Help_Page.htm

[9]Amazon Web Services Migration Scenarios: Web Application Architecture http://d36cz9buwru1tt.cloudfront.net/CloudMigration-scenario-wep-app.pdf

[10] OpenStack, http://openstack.org

[11] BOM calculator, https://www.mirantis.com/blog/hardware-bom-calculator-controllers-switches-bells-and-whistles/

[12] Reference Architecture, http://docs.openstack.org/

[13] CloVER, http://www.procan-group.com/products/clover

[14] jclouds, The Java Multi-Cloud Toolkit web site, https://jclouds.apache.org/

[15] RosettaHUB platform's web sitehttps://www.rosettahub.com

[16] Leibniz SupercomputingCenter's web site, https://www.lrz.de/english/

[17] CloudSelect, https://www.cloudselect.com

[18] CompatibleOne, http://www.compatibleone.com/

[19] JellyFish, http://projectjellyfish.org/

Annex

Survey Questionnaire

Survey Background

Cloud computing technologies and services have been steadily evolving and spreading in a growing number of application domains. ALECSO in cooperation with ITU intends to assess cloud adoption in educational and research institutions in the Arab countries. The objective of this study is (i) to make an inventory of ICT use in Arab institutions (ii) to understand how those institutions perceive the benefit and applicability of the cloud to their use cases and whether and how they plan to use it (iii) to understand the barriers and risks they currently apprehend.

This questionnaire is structured as follows:

The first part (questions 1-12) addresses users and non-users of the cloud. It aims at collecting general data about the institution and assessing the status of its current ICT tools and infrastructure.

A second part (questions 13-22) is for non-users of the cloud. It focuses on the positioning of the institution in relation with cloud utilization which includes its current perception of the expected benefits and usability of the technology as well as its concerns This part aims also at collecting data to characterize the types of potential cloud utilizations and to infer actions likely to lower the barriers to cloud adoption.

A third part (questions 23-31) is for cloud users. It aims to collect data on (i) the types of clouds used and their levels of use (ii) the perceived benefits and usability (iii) the challenges raised by the migration to the cloud and the solutions adopted to overcome those challenges (iv) the level of satisfaction with the cloud solutions that have been put in place (v) The actions to be taken to boost the migration to the cloud.

The collected data will be used to elaborate recommendations for triggering or catalyzing Arab institutions' migration to the cloud.

The data will be anonymized and treated in full respect of the relevant data protection legislations.

ALECSO and ITU thank you for your cooperation.

General Information

What is the name of your institution (please indicate the full name)?

Please specify the country of your institution

Please provide your email address

What type of institution do you work for ? (check the appropriate box)

- ○ Ministries and Data Centers
- ○ Higher Education Institutions
- ○ Other

What is your position in your institution? (check the appropriate box)

- ○ Head of the institution
- ○ Head of IT (Information Technology) Department
- ○ Member of IT Department
- ○ Other

How many staff members and how many students does your institution count (by category)

Students	
Faculty members (lecturers)	
Researchers other than faculty members	
Administrative and technical staff other than IT	
IT Administrative and technical staff	

ICT tools and infrastructure status

What is the speed of your Internet connection in Mbit/s?

[ ]

What is the number of PCs or Terminals installed in your institution?

Number of Desktops	[ ]
Number of Laptops	[ ]
Number of Terminals	[ ]

What is the number of PCs and Terminals installed in laboratories

Number of Desktops	[ ]
Number of Laptops	[ ]
Number of Terminals	[ ]

Do you have a data center (check the appropriate box)

- ◯ No
- ◯ Yes

Number of IT administrative and technical staff members

Is remote access to your data center available

- ◯ Yes
- ◯ No

What is the percentage of students who own a laptop?

What the percentage of desktops in laboratories relative to the number of students ?

What is the percentage of students with email addresses?

Do you have a wireless local area network (check the appropriate box)

- ◯ No
- ◯ Yes

Check all that apply

- ☐ Accessible to students
- ☐ Accessible to faculty members
- ☐ Accessible to researchers
- ☐ Accessible to the administrative and technical staff

What are the types of IT applications and services used in your institution? (check all that apply)

- ☐ Administrative and operational services
- ☐ Office Suite
- ☐ Specific scientific applications and data processing
- ☐ Collaborative applications (email, video conferencing ...)
- ☐ Learning Management System, educational resources
- ☐ Storage
- ☐ Compute infrastructures (supercomputers, grids, clusters, private compute clouds)
- ☐ Other

Satisfaction level of your institution with the existing IT applications and services

What is the level of satisfaction of your institution with the existing IT applications and services? Please rank satisfaction from 1 to 3 with 1 being very satisfied.

- ◯ (1) Very satisfied
- ◯ (2) Satisfied
- ◯ (3) Not satisfied

Cloud utilization in your institution

At what stage your institution is, with regards to cloud computing? (check the appropriate box)

- ◯ Not involved
- ◯ Discussion stage
- ◯ Trial stage
- ◯ Implementation stage
- ◯ Production stage

Perceived benefits of the cloud (of cloud use in education and research)

What does your organization view as the most important benefits of the cloud? Please rank benefit from 1 to 3 with 1 being the very important (check all that apply)

	(1) Very important	(2) Important	(3) Least important	Do not know
Software cost saving	○	○	○	○
Operational cost saving	○	○	○	○
Hardware cost saving	○	○	○	○
Ability to rapidly launch new products and services	○	○	○	○
Ability to scale up and scale down IT capacity on demand	○	○	○	○
Reduction of capital expenditure	○	○	○	○
Convenience for the IT development teams	○	○	○	○
Reliable data storage	○	○	○	○
Disaster recovery capabilities	○	○	○	○
Increased collaboration	○	○	○	○
Mobile and ubiquitous access to applications and services	○	○	○	○
Availability of the cloud services round the clock	○	○	○	○
IT efficiency	○	○	○	○

Perceived ease of use of the cloud

How do you perceive the use of the cloud for the personnel of your institution? Please rank utilization from 1 to 3 with 1 being the very easy (check all that apply)

	(1) Very easy	(2) easy	(3) Not easy	Do not know
Learning how to apply cloud technology	○	○	○	○
Interaction with cloud technology	○	○	○	○

Perceived barriers for cloud adoption

What are the most significant barriers to cloud adoption in your institution? Please rank barriers from 1 to 3 with 1 being the very important (check all that apply)

	(1) Very important	(2) Important	(3) Least important	Do not know
Decision makers are not fully aware of the benefits the technology can deliver	○	○	○	○
Resistance to change of the staff (lecturers, researchers, ...)	○	○	○	○
Resistance to change of the IT team	○	○	○	○
Service Level Agreements (SLA) Issues	○	○	○	○
Poor customer service	○	○	○	○
Data protection and/or privacy concerns	○	○	○	○
Security issues	○	○	○	○
Existing infrastructure	○	○	○	○
Vendor lock in	○	○	○	○
Location of the cloud provider in another country	○	○	○	○
Integration with existing systems	○	○	○	○
Lack of control over IT infrastructure	○	○	○	○
Lack of skills	○	○	○	○
Cost issues	○	○	○	○
Other (specify)	○	○	○	○

Intention to use the cloud

Intention to use the cloud

Has cloud computing been incorporated into your institution's strategy? (check the appropriate box)

- ◯ Cloud computing is formally part of our IT strategy
- ◯ Cloud computing has been discussed formally and it will be part of our IT strategy
- ◯ Cloud computing has only been discussed informally
- ◯ Cloud computing has not been discussed

If your institution intends to migrate, what would be the type of cloud? (check the appropriate box)

- ◯ Public Cloud
- ◯ Private Cloud
- ◯ Community Cloud (shared cloud among several institutions)
- ◯ Hybrid Cloud
- ◯ Do not know

If your institution intends to use the cloud, which IT applications and services would it migrate to the cloud? (check all that apply)

	In 1 year	In 2 years	In 3 years	In more than 3 years	Do not know
Administrative and operational services	◯	◯	◯	◯	◯
Office Suite	◯	◯	◯	◯	◯
Specific scientific applications and data processing	◯	◯	◯	◯	◯
Collaborative applications (email, video conferencing ...)	◯	◯	◯	◯	◯
Learning management system, educational resources	◯	◯	◯	◯	◯
Storage	◯	◯	◯	◯	◯
Compute infrastructure	◯	◯	◯	◯	◯
Other (specify)	◯	◯	◯	◯	◯

In which of the following areas do you need support regarding your decision to migrate to the cloud? Please rank areas from 1 to 3 with 1 being the very important. (check all that apply)

	(1) Very important	(2) Important	(3) Least important
Awareness raising of decision makers	◯	◯	◯
Change management	◯	◯	◯
Technical capacity building	◯	◯	◯
National regulatory and compliance issues	◯	◯	◯
Contractual issues including Service Level Agreements	◯	◯	◯
Security, Privacy including data protection issues	◯	◯	◯
Cost-Benefit analysis	◯	◯	◯
Providing guidelines	◯	◯	◯
Other (specify)	◯	◯	◯

Familiarity of respondent with the cloud

How much confidence do you currently have in your knowledge of the cloud? (check the appropriate box)

- ◯ I am just beginning to familiarize myself
- ◯ I am relatively familiar
- ◯ I am familiar
- ◯ I am very familiar

What kind of cloud is used by your institution? (check the appropriate box)

- ◯ Public Cloud (Infrastructure operated by a third-party vendor)
- ◯ Private Cloud (Your own infrastructure or infrastructure operated solely for your company)
- ◯ Community Cloud (cloud shared among several institutions)
- ◯ Hybrid Cloud (Combination of Public and Private Cloud Infrastructure)

Cloud usage level

What are the IT applications and services that your institution has migrated to the cloud and at what level? (check all that apply)

	Over 75%	Between 50 &75%	Between 25 & 50%	Less than 25%	0%
Administrative and operational services	◯	◯	◯	◯	◯
Office Suite	◯	◯	◯	◯	◯
Specific scientific applications and data processing	◯	◯	◯	◯	◯
Collaborative applications (email, video conferencing ...)	◯	◯	◯	◯	◯
Learning Management System, educational resources	◯	◯	◯	◯	◯
Storage	◯	◯	◯	◯	◯
Compute infrastructure	◯	◯	◯	◯	◯
Other (specify)	◯	◯	◯	◯	◯

What IT applications and services for which your institution intends to pursue migration or plans to migrate to the cloud? (check all that apply)

	In 1 year	In 2 years	In 3 years	In more than 3 years	Do not know
Administrative and operational services	◯	◯	◯	◯	◯
Office Suite	◯	◯	◯	◯	◯
Specific scientific applications and data processing	◯	◯	◯	◯	◯
Collaborative applications (email, video conferencing ...)	◯	◯	◯	◯	◯
Learning Management system, educational resources	◯	◯	◯	◯	◯
Storage	◯	◯	◯	◯	◯
Compute infrastructure	◯	◯	◯	◯	◯
Other (specify)	◯	◯	◯	◯	◯

Benefits made using the cloud

Benefits made using the cloud

What are the benefits that your institution has made in moving to the cloud? Please rank benefit from 1 to 3 with 1 being the very important (check all that apply)

	(1) Very important	(2) Important	(3) Least important	Do not know
Ability to rapidly launch new products and services	○	○	○	○
Ability to scale up and scale down IT capacity on demand	○	○	○	○
Reduction of capital expenditure	○	○	○	○
Operational cost saving	○	○	○	○
Hardware cost saving	○	○	○	○
Software cost saving	○	○	○	○
Convenience for the IT development teams	○	○	○	○
Reliable data storage	○	○	○	○
Disaster recovery capabilities	○	○	○	○
Increased collaboration	○	○	○	○
Mobile and ubiquitous access to applications and services	○	○	○	○
Availability of the cloud services round the clock	○	○	○	○
IT efficiency	○	○	○	○

Ease of use of the cloud

Is the use of the cloud for the staff of your institution easy? Please rank utilization from 1 to 3 with 1 being the very easy (check all that apply)

	(1) Very easy	(2) easy	(3) Not easy	Do not know
Learning how to apply cloud technology	○	○	○	○
Interaction with cloud technology	○	○	○	○

Barriers faced in migrating to the cloud

What are the barriers that your organization has encountered when migrating to the cloud? Please rank barriers from 1 to 3 with 1 being the very important (check all that apply)

	(1) Very important	(2) Important	(3) Least important	Do not know
Decision makers are not fully aware of the benefits the technology can deliver	○	○	○	○
Resistance to change of the staff	○	○	○	○
Resistance to change of the IT team	○	○	○	○
Service Level Agreements (SLA) issues	○	○	○	○
Poor costumer service	○	○	○	○
Data protection and/or privacy concerns	○	○	○	○
Security issues	○	○	○	○
Data sovereignty / regulatory and compliance	○	○	○	○
Existing infrastructure	○	○	○	○
Vendor lock in	○	○	○	○
Location of the cloud provider in another country	○	○	○	○
Integration with existing systems	○	○	○	○
Lack of control over IT infrastructure	○	○	○	○
Lack of skills	○	○	○	○
Cost issues	○	○	○	○
Other (specify)	○	○	○	○

How barriers have been lifted?

How did your institution overcome these barriers? (check all that apply)

- ☐ Raising awareness of decision makers
- ☐ Change management
- ☐ Technical capacity building
- ☐ Support in national regulatory and compliance issues
- ☐ Support in contractual issues including Service Level Agreements
- ☐ Support in security, Privacy including Data Protection Issues
- ☐ Support in Cost-Benefit Analysis
- ☐ Other (specify)

What is the level of satisfaction of your institution with existing cloud applications and services. Please rank satisfaction from 1 to 3 with 1 being very satisfied. (check the appropriate box)

- ○ (1) Very satisfied
- ○ (2) Satisfied
- ○ (3) Not satisfied

What are the actions to be taken to boost migration of your institution to the cloud. Please rank actions from 1 to 5 with 1 being most important. (check all that apply)

	(1) Very important	(2) Important	(3) Least important
Change management	○	○	○
Capacity building in project management	○	○	○
Increase of specialized technical staff	○	○	○
Technical capacity building	○	○	○
Experience sharing between institutions	○	○	○
Support in contractual issues including Service Level Agreements	○	○	○
support in Cost-Benefit Analysis	○	○	○
Improvement of cloud provider services	○	○	○
Increased transparency in provider's pricing models	○	○	○
Availability of relevant IT applications	○	○	○
Other (specify)	○	○	○

ISBN : 4- 381 – 15 – 9973 – 978