Instituto Superior de Engenharia do Porto

PRIVATE CLOUD COMPUTING PLATFORMS

Analysis and implementation in a Higher Education Institution

Nuno Miguel Correia do Novo Pinhal

Dissertação para obtenção de grau de Mestre em

Engenharia Informática

Área de Especialização em

Arquiteturas, Sistemas e Redes

Orientador: Doutor Nuno Alexandre Magalhães Pereira

Juri:

Presidente:

Doutor José António Reis Tavares, Professor Adjunto, ISEP

Vogais:

Doutor Alexandre Manuel Tavares Bragança, Professor Adjunto, ISEP

Doutor Nuno Alexandre Magalhães Pereira, Professor Adjunto, ISEP

Porto, outubro 2012

Acknowledgments

I would like to express my acknowledgment to Adjunct Professor Nuno Alexandre Magalhães Pereira of *Instituto Superior de Engenharia do Porto do Instituto Politécnico do Porto* (ISEP-IPP) for having made possible the fulfillment of this work and have always been available.

I am also grateful to my colleagues at Information Technology Service of the *Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto* (ICBAS-UP), for their support in each of their specialized areas making possible with team commitment and collaboration, to make this work viable for testing and implementing with this type of technology.

Last but not least, my family for their support, motivation and understanding.

Abstract

The constant evolution of the Internet and its increasing use and subsequent entailing to private and public activities, resulting in a strong impact on their survival, originates an emerging technology. Through cloud computing, it is possible to abstract users from the lower layers to the business, focusing only on what is most important to manage and with the advantage of being able to grow (or degrades) resources as needed.

The paradigm of cloud arises from the necessity of optimization of IT resources evolving in an emergent and rapidly expanding and technology.

In this regard, after a study of the most common cloud platforms and the tactic of the current implementation of the technologies applied at the Institute of Biomedical Sciences of Abel Salazar and Faculty of Pharmacy of Oporto University a proposed evolution is suggested in order adorn certain requirements in the context of cloud computing.

Keywords: cloud computing, private cloud computing, xen server, microsoft terminal servers, microsoft application virtualization, open stack, open virtual desktop, hyper-v

Resumo

A constante evolução da Internet e a sua crescente utilização e consequente vinculação às atividades privadas e públicas, traduzindo-se num forte impacto à sua sobrevivência, origina uma tecnologia emergente. Através de *cloud computing*, é possível abstrair os utilizadores das camadas inferiores ao negócio, focalizando apenas no que realmente é mais importante de gerir e ainda com a vantagem de poder crescer (ou diminuir) os recursos conforme as necessidades correntes.

Os recursos das TI evoluíram consideravelmente na última década tendo despoletado toda uma nova consciencialização de otimização, originando o paradigma da computação em nuvem.

Neste sentido, após um estudo das plataformas de *cloud* mais comuns, é abordado um *case study* das tecnologias implementadas no Instituto de Ciências Biomédicas de Abel Salazar e Faculdade de Farmácia da Universidade do Porto seguido de uma sugestão de implementação de algumas plataformas de *cloud* a fim de adereçar determinados requisitos do *case study*.

Distribuições produzidas especificamente para a implementação de nuvens privadas encontram-se hoje em dia disponíveis e cujas configurações estão amplamente simplificadas. No entanto para que seja viável uma arquitetura bem implementada, quer a nível de *hardware*, rede, segurança eficiência e eficácia, é pertinente considerar a infraestrutura necessária como um todo. Um estudo multidisciplinar aprofundado sobre todos os temas adjacentes a esta tecnologia está intrinsecamente ligado à arquitetura de um sistema de nuvem, sob pena de se obter um sistema deficitário. É necessário um olhar mais abrangente, para além do equipamento necessário e do *software* utilizado, que pondere efetivamente os custos de implementação tendo em conta também os recursos humanos especializados nas diversas áreas envolvidas.

A construção de um novo centro de dados, fruto da junção dos edifícios do Instituto de Ciências Biomédicas de Abel Salazar e da Faculdade de Farmácia da Universidade do Porto, possibilitou a partilha de recursos tecnológicos. Tendo em conta a infraestrutura existente, completamente escalável, e assente numa abordagem de crescimento e de virtualização, considera-se a implementação de uma nuvem privada já que os recursos existentes são perfeitamente adaptáveis a esta realidade emergente.

A tecnologia de virtualização adotada, bem como o respetivo *hardware* (armazenamento e processamento) foi pensado numa implementação baseada no XEN Server, e considerando que existe heterogeneidade no parque dos servidores e tendo em conta a ideologia das tecnologias disponíveis (aberta e proprietária) é estudada uma abordagem distinta à implementação existente baseada na Microsoft.

Dada a natureza da instituição, e dependendo dos recursos necessários e abordagem a tomar, no desenvolvimento de uma nuvem privada, poderá ser levado em conta a integração com nuvens públicas (por exemplo Google Apps), sendo que as possíveis soluções a adotar poderão ser baseadas em tecnologias abertas e/ou pagas (ou ambas).

Este trabalho tem como objetivo, em última instância, o desígnio de verificar as tecnologias utilizadas atualmente e identificar potenciais soluções para que em conjunto com a infraestrutura atual, disponibilizar um serviço de nuvem privada.

O trabalho inicia-se com uma explicação concisa do conceito de nuvem, comparando com outras formas de computação, expondo as suas características, revendo a sua história, explicando as suas camadas, modelos de implementação e arquiteturas. Em seguida, no capítulo do estado da arte, são abordadas as principais plataformas de computação em nuvem focando o Microsoft Azure, Google Apps, Cloud Foundry, Delta Cloud e Open Stack. São também abordadas outras plataformas que emergem fornecendo assim um olhar mais amplo para as soluções tecnológicas atuais disponíveis. Após o estado da arte, é abordado um estudo de um caso em particular, a implementação do cenário de TI do novo edifício das duas unidades orgânicas da Universidade do Porto, o Instituto de Ciências Biomédicas Abel Salazar e a Faculdade de Farmácia e sua arquitetura de nuvem privada utilizando recursos partilhados. O estudo do caso é seguido de uma sugestão de evolução da implementação, utilizando tecnologias de computação em nuvem de forma a cumprir com os requisitos necessários e integrar e agilizar a infraestrutura existente.

Palavras-chave: cloud computing, private cloud computing, xen server, microsoft terminal servers, microsoft application virtualization, open stack, open virtual desktop, hyper-v

Contents

List of Figures

List of Tables

Chapter 1

Preamble

This chapter is an introductory note to contextualize the theme of the thesis. It is covered the motivation, the presentation of where the study was carried out, its reality and its requirements and defined a way of approach to the topic concerned.

1.1 Motivation

Not long ago the ICBAS data center had it's servers stacked on top of each other coexisting in the same room as the IT staff. Gradually better conditions of hardware, isolation and technology began to arise. As consequence of a study in virtualization technologies and the development of the project/internship of DEI-ISEP dedicated to the implementation of a high availability cluster of virtualized servers [Pinhal, 2008], XEN hypervisor was adopted for a new virtualization platform with three new servers and storage. This scenario was sufficient for the existing needs at ICBAS and it was still capable of being scaled.

In the beginning of the transition process for new building housing ICBAS|FF-UP, it was necessary in conjunction with the logistics process, the assurance of the business continuity at the old building and the implementation of a potentially scalable technology infrastructure at the new building that would enable a smooth migration procedure with the least possible impact to the users. The new data center shared between two organic units of the University of Porto, ICBAS and FF was made available within the stipulated schedule, as well as all the hardware needed to implement a new architecture. The purpose of the new system is essentially the creation of a flexible solution to support the existing requirements as well as the new ones that may be triggered after the beginning of school's semester activities. In order to take advantage of the new data center's full potential, it is pertinent to embark on the emerging technology of cloud computing. However, originally, due to the short timing of the transition process, this was not accomplished, to address of the issue of cloud computing. The actual system configuration, although innovative in certain aspects, to meet the existing requirements, it remained with the virtualization infrastructure used in the old building essentially to address legacy issues.

Considering the nature of the higher education institutions involved in this process, and to ensure the quality and their business continuity, transforms this problem into a pertinent technological challenge.

1.2 Presentation

Considering that the Institute of Biomedical Sciences of Abel Salazar and the Faculty of Pharmacy of the Porto University are, for all intents and purposes, a Public Institution, whose legal status was recently changed to Foundation, this section contains a broad presentation of the organizational units involved in the case study and it is also addressed a brief framework on the technology titled in the perspective of the public administration on which it is contextualized.

1.2.1 Public Administration

The importance of cloud in the context of public policies has a particular emphasis on the center of the Public Administration (or cloud government) that is increasingly a reality in the coming years, with the theme reflected in key policy guidelines. The strategy Europe2020 [Commission, 2012] with its Digital Agenda, puts a strong focus on this issue. European Action Plan for Electronic Government 2011-2015.

In Portugal, it also gives a prominent role to the development of cloud computing is one of the areas mentioned in the "Agenda Digital 2015" [Português, 2012] (development of public services using the potential of cloud) and the "Plano de Racionalização e Redução de Custos nas TIC" (IT Plan of Rationalization and Cost Reduction) needs further work and investment made in recent years in Portugal in the area of eGovernment.

The purpose of cloud computing in Public Administration is to improve the quality of public services to society (citizens and businesses), best use of existing resources in the goverment (rationalization), reducing costs and gaining efficiencies. For example the use of cloud to stimulate areas such as scientific research [Truong and Dustdar, 2010], benefiting society as a whole.

The GRID initiative [UMIC, 2012], under development by FCCN [1] / UMIC [2], is a good example by taking advantage of existing technological infrastructure in institutions of higher education and research, including Web servers, databases, etc..

The main benefits of using cloud in Public Administration, in very general terms, can contribute to:

- improve services to citizens and to businesses;
- more efficiently and use of existing technological resources (economies of scale);
- cost reduction;
- use of shared resources and shared services;
- reduction of energy consumption;

[1] Fundação para a Computação Científica Nacional (Foundation for National Scientific Computation)
[2] Agência para a Sociedade do Conhecimento, IP (Agency for the Knowledge Society, Public Institution)

- increase of the productivity;
- common monitoring systems;
- stimulus to public procurement and ecommerce;

Collaboration between public and private sector in the field of cloud computing will be increasingly a reality in the future (private cloud, public cloud or hybrid). It should take into account in this respect the analysis of possible business models, which make this kind of sustainable partnerships, interoperability, data security and data protection. [Loutas et al., 2010]

In a relationship of this nature, consideration should be given rigorous review and quantified of what benefits the cloud can bring to the Government, aiming at the public services that will be enhanced, identifying how can you monetize the infrastructure existing technology in public administration and what the viable business models that might be adopted.

The creation of a Government Cloud, aimed at rationalizing the infrastructure of public administration, theoretically, we could have won at the level of technological infrastructure maintenance (backups, security, etc..), Rationalization of resources (sharing servers, databases, applications , etc..), lower spending on software licenses (many of the applications in cloud computing are free, or just paid depending on the type of use), higher levels of productivity, as the networking and collaboration may encourage exchange ideas, knowledge and stimulate innovation.

Through the creation of a Government Cloud (G-Cloud), one can rationalize the infrastructure of the Public Administration. [Maio, 2011] The IT park is vast and the Public Administration can be better utilized, continuing the process of dematerialization and virtualization under way with the SIMPLEX program, and now with the "IT Plan to Reduce Costs in Public Administration". As an example, the "GRID National Initiative" (Higher Education) allows the sharing of resources, providing high computational capabilities at the expense of distributing processing tasks in coordinated and efficient systems throughout various Research and Development institutions. This approach to the higher education has been proven to be the awakening to a new successful paradigm to adopt. [Sultan, 2010]

Resistance by the Administration regarding the implementation of cloud computing systems is manifested mainly due to little concrete knowledge of the potential of cloud and a few cases at the international level that are eligible to serve as an example. Questions still open: Security? Data Protection? Most appropriate type of cloud? Direct Costs and cost-benefit analysis? However there are still outstanding issues, including security, data protection, the most appropriate type of cloud, direct costs and cost-benefit analysis. [Mutavdžic', 2010]

Portugal is at the vanguard of the G-Cloud as already in eGovernment at European level. It is expected that contributions to the rise of new business models that inspire the development opportunities around the cloud can be brought by the "Digital Agenda 2015". There are still some challenges coming to take into account, as the legal framework [Yang and Chen, 2010] (national, community, international, data protection, security ...), the organizational framework (governance), the definition of services (sovereignty; sensitive data ..), as create value for AP and for society (open innovation in AP), infrastructure (RNG; data centers) and accesses (devices).

Clinical research or nanotechnology can be stimulated using the technologies of networking and distributed computing. Health, Education, eGovernment The good results achieved by Portugal in recent years in eGovernment may have enhanced the opportunities around the cloud, also generating new business opportunities for technology companies. In the private sector, the cloud will be more related to efficiency gains and cost savings.

1.2.2 University of Porto

The University of Porto is, at the present, one of the largest education and research institution in Portugal. It has about 31.000 students, 2.300 teachers and researchers along 1.700 administrative staff attend its 15 schools and 69 scientific research units, spread across 3 university campuses located in the city of Porto.

The University of Porto has 14 faculties and a business school offering a wide range of subject areas at all levels of higher education currently offering 700 training programs (from degrees to continuous professional training).

Teaching staff comprising 1.860 full-time equivalent teachers (71% of which PhDs) ensures a high quality of training that makes the University of Porto the most sought-after University in Portugal and it's the choice for candidates with the highest grades.

Currently, nearly 31.000 students (6.500 postgraduates) attend the 273 degrees, masters and doctoral courses of the University of Porto. Every year, around 2.000 international students choose this university to complete their higher education.

The University of Porto has 69 research units and is responsible for more than 20% of Portuguese articles indexed each year in ISI Web of Science, it has some of the most productive and prestigious Portuguese Research and Development centers and international evaluations rated half of the research units with "Excellent" or "Very Good".

Stimulating economic value to the scientific production and the establishment of partnerships with the Portuguese industry leaders is one of the aims pursued by the University that has already resulted in several innovations with proven success in global markets.

ICBAS - School of Health and Life Sciences

The Institute of Biomedical Sciences Abel Salazar, hereinafter ICBAS, is a unit of the University of Porto with a structure of school and university center for teaching, scientific research, culture and service to the community.

The scientific fields of ICBAS lie within the Fundamental and Applied Biology, particularly in the areas of Health, Environment, Animal Production, Processing and Food Processing and Quality Control. So, essentially technological or techno professional, formations ICBAS emphasize preparation for professional practice of human medicine, veterinary medicine, the sciences of the aquatic environment, biochemistry and bioengineering, or offer specialization of 2nd and 3rd cycles in these areas and other related.

The preparation of students requires highly qualified teachers and professionally demanding, so the ICBAS works in close collaboration with associated structures, such as the Hospital Centre of Oporto (Hospital Geral de Santo António Júlio Dinis Maternity Hospital and Maria Pia), Magalhães Hospital Lemos, Portuguese Institute of Oncology (IPO), Joaquim Urbano Hospital, Centre Hospitalier de Vila Nova de Gaia Hospital Center of the Northeast - Braganza Hospital and Health Centers of the ARS North, where students receive part of their training.

In addition to the central campus that has the campus ICBAS U. Port, this unit boasts organic facilities on the campus of Agrarian Vairão (ICAV) - in Vila do Conde - where in addition to the teaching of veterinary science, clinical research is practiced in this area, through the Center for Animal Reproduction Vairão (CRAV) and Clinical Center for Equine Vairão (CCEV).

Alongside the ICBAS collaborates with leading research institutions associated with the U. Port, such as the Associated Laboratories, Institute of Molecular and Cell Biology (IMCB), the Institute of Molecular Pathology and Immunology (IPATIMUP) and the Interdisciplinary Centre for Marine and Environmental Research (CIIMAR), and two research units based on your space base : the UMIB (Unit for Multidisciplinary Investigation in Biomedicine - anatomy, pharmacology, molecular biology and cytology) and UNIFAI (Unit for Research and Training in Adults and Elderly).

It should be noted also that this partnership is extended to national and foreign public entities, including the Research Institute of Fisheries and Marine (IPIMAR), the Coastal Station of Aguda (ELA), the Faculties of Science and Engineering, U. Porto. And universities and research institutes from France, Spain, Germany, UK, Sweden, Netherlands, Romania, Thailand, United States, Chile, Argentina and Brazil, among others.

The ICBAS also offers specialized services to the Community, such as Investigations and diagnosis collaborating with the Hospitals of St. Anthony, St. John IPO Porto and Lisbon, Institute of Medical Genetics, Maria Pia Hospital and Institute of Legal Medicine of Porto.

- Clinic and surgery of companion animals (consultations, treatments ...)
- Surgery and clinic of large animals
- Support the implementation of aquaculture projects
- Diagnosis of Chromosomal Instability (DIC)
- Support for the seafood industry
- Analysis of water quality
- Studies in the areas of ecotoxicology and environmental
- Education and training programs for in specific areas of intervention

Concerning the human resources, ICBAS has about 260 teachers, of whom 132 are PhDs and 27 Scholars career, which provide education from the 1st to the 3rd study cycles, actions of training, and the pursuit of research from the basic to the applied.

Researchers of international renown are part of the ICBAS human resources, developing significant work in multiple areas of science such as anatomy, cell biology, pharmacology, immunology, pathology, psychology, toxicology and chemistry, constantly generating numerous articles in international journals.

The ICBAS owns a Biomedical Simulation Center, related with the clinical teaching which includes simulators of various diseases, some of which are unique in the country, such as some simulators of cardiovascular disease.

Students and teachers have in their disposal a Center for Biomedical Simulation that is accessible to clinical cycle and is equipped with simulation models of basic life support, immediate life support, reviewing the background of the eyes and ears, digital rectal examination, pathology cardio-respiratory, intubation of patients, obstetrics and gynecology, among others. The ICBAS gives a strong emphasis to higher qualification by offering postgraduate students and the general public, Master's degrees in various branches of health sciences and life as well as PhDs in different areas.

To respond efficiently to its ambitious educational project, the ICBAS has now the new building completed and ready to use. The process of monetization and sharing between the Faculty of Pharmacy ICBAS and made possible the existence of a large health Polo, in the broadest sense of the term, in close collaboration with the Center Hospital of Porto, increasing the valences of the Hospital of Saint Anthony, and Júlio Dinis Maternity Hospital Maria Pia, that all may benefit in the near future.

ICBAS, as part of its openness to society, gives secondary schools and their students in particular the opportunity for guided tours of their laboratories, anatomical museum and consultation with numerous works of Alberto Saavedra your library. The extracurricular activities underpinned by ICBAS by AEICBAS and other associations (such as the prestigious Biomedical Coral) are many and enriching, contributing to the reputation of this institution.

Faculty of Pharmacy of the University of Porto

The Faculty of Pharmacy of the University of Porto (FFUP) was founded in 1921, and its origins go back to 1836, when the old School of Pharmacy was established.

The FFUP provides education, research and development, while engaged with academic excellence in the field of Pharmaceutical Sciences. In addition to this, the FFUP is also a center of creation, communication and diffusion of culture, science and technology, in several domains such as health, chemical and biological sciences, whilst serving Man and respecting His rights.

The FFUP pursuits the following goals:

- To teach students human, ethical, cultural, scientific and technical education;
- To teach students subject disciplines, which are important for their Scientific and technical education, bearing in mind the development of specific competences;
- To organize courses within the FFUP, or in partnership with other Faculties of the University or with external institutions;
- To promote extracurricular teaching sessions, as well as professional formation;

- To support pedagogic, scientific and technical mobility with national and international institutions;
- To develop products and community services in a worthy reciprocal way.

1.3 Requirements

Considering the planned relocation, housing two schools on the same campus, a new architecture need to be planned from scratch. Although the two colleges have distinct characteristics, there are certain places where resources are common, including classrooms, library and computer rooms for students. This new requirement (shared resources) could easily be solved with a trust relationship between the ICBAS domain and the FF domain, however, the FF did not have a domain, given its size and limited resources. The requirements, in a higher level of abstraction are the establishment of two separate domains where administrative maintenance operations are performed and a common domain that can be used and administered by either of the above. In each of the domains (ICBAS and FF) it's expected to be contemplated entities, resources, and isolation of services, departments and respective laboratories (if applicable) and students. Although the FF had not, so far, a computer facilities comparable to ICBAS, it is intended that this be integrated and to adopt part of the work methodologies exist in ICBAS, thus easing the work of two teams IT. Embarking on a new technological platform, must be as transparent as possible for the end users, not impairing either the lective activity or the services, ensuring business continuity transversely.

1.3.1 Platform

The platform used is mostly Windows, with all the necessary software and technology support services, departments and respective laboratories are based on Microsoft Windows technologies. However, there are some services that have use Linux, the platform for server virtualization, electronic mail, VPN, monitoring (internal and external) and web. All of the clients use either Microsoft Windows Windows XP (who are the target of an effort to update) or Windows 7.

The Microsoft Campus Agreement (MCA) is a contract signed annually between the University of Porto and Microsoft. Through this contract is made the licensing of products known as primary and servers. The licensing of primary products contrary to what one might suppose is not the number of machines involved, but according to the number of employees. The accounting of these is given in the concept of FTE (Full-time equivalent), ie, according to the definition given by Microsoft: people who work for the institution 200 or more hours / year and to use computers in carrying out their activities. The MCA is renewed annually and runs from April 1 of the calendar year corresponding to 31 March the following calendar year. The primary products cannot be purchased separately and must be subscribed in as many units as the number of FTE reported. The server products are licensed on as many units as given annually by each organizational unit and agency U.PORTO. Regarding the desktop operating system, it should be noted that the license only covers upgrades, requiring that the computer first having a version of Microsoft's desktop operating system or Macintosh properly licensed (license typically OEM - Original

Equipment Manufacturer). It is emphasized that the acquisition of a new computer, if the MS Windows operating system comes installed, the supplier is obliged to give the customer the relevant license and this should always be required. Under the conditions of licensing, primary products subscribed by each organizational unit or agency may be used on all computers in the same. The license also includes the possibility to use the institution's staff on their personal computers, for example at home, the primary products that the unit or agency is licensed (Work at Home Rights).

Primary products:

- Updates for Windows Enterprise
- Office Professional Plus :
  - Word
  - Excel
  - PowerPoint
  - Outlook with Business Contact Manager
  - Access
  - InfoPath
  - Publisher
  - SharePoint Workspace
  - Lync
  - Office Web Apps
  - OneNote
- Exchange Server Ent
- Microsoft Office for Mac (Professional Edition)
- Core CAL [3] Licences (Forefront EndPoint Protection, System Center Configuration Manager Client ML, Windows Server, Exchange Server Standard, SharePoint Server Standard, Lync Server Standard)
- Remote Desktop Services Client Access Licenses
- Microsoft Project Professional (as requested)
- Microsoft Visio Premium (as requested)

Server Products:

- Windows Server Std

[3] Client Access Licenses

- Windows Server Ent
- Exchange Server Ent
- System Center 2012
- SQL Server 2012
- Sharepoint Server
- Exchange Server
- ForeFront TMG Standard
- Other server products, as requested.

1.3.2 Organizational

Both ICBAS as FF have common organizational characteristics that allow the design of an architecture based on these pillars, bearing in mind other existing asymmetries. The commonalities are described below:

- Students;
- Management
- Divisions and it's services;
- Departments and it's laboratories;
- Community service:
  - Laboratory of veterinary pathology
  - Veterinary clinic;
  - ...
- Partner organizations;

1.3.3 Campus

The new location includes five new buildings, including one exclusively for the administrative services of the ICBAS, one other that contains the library and amphitheaters (common) and three other buildings with the same number of floors which house the departments, laboratories, offices and classrooms of both organizational units as represented in Figure 1.1.

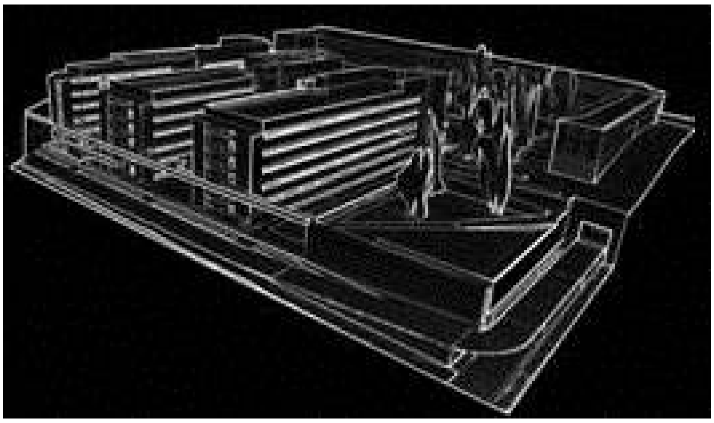

Figure 1.1: ICBAS|FF-UP Plant

1.4 Approach

The approach to this challenge, it boils down mainly to the study and discrimination of the technologies, assess what has been achieved in the particular case of ICBAS. After evaluating viable cloud computing alternatives that fit the existing requirements (or that can be potentially scalable) and suggest one or more solutions.

1.5 Outline

This document is divided into six consecutive chapters in which in this first chapter is done contextualization of the framework. In the Preamble chapter is made a description of the motivation that led for this work to take place, afterwards the subject of the work is contextualized within the public administration and a brief description of the institution and its functional and organizational framework is presented. It is made a study of requirements regarding the platform, organizational structure and is shown a vision of the campus in focus to encompass the whole problem. After discussed the mentioned subjects a methodology of approach for the resolution of the work is enunciated.

The subject of cloud computing is studied in the next chapter, addressing some of the most pertinent issues as an introduction to the technology inherent to this work. This includes a comparison with other technologies, the characteristics, a summary of the history, the architecture and deployment models. It is also made a comparison between the two most common models of implementation and a brief reference to the necessary infrastructure, the data center.

The study of the state of the art is disseminated with cloud computing technologies, which addressed some of the most used and brought by different initiatives originated by world renowned companies and other organizations and institutions. In this chapter several cloud computing technologies are covered such as Microsoft Azure, Google Apps, Cloud Foundry, the Delta Cloud, Eucalyptus, Open Stack and XEN Cloud Platform.

The technological implementation at ICBAS|FF-UP is studied and its approach to successfully migrate the existing infrastructure causing the least impact possible on the users and providing the means to maximize the new resources. The technologies, architecture, hardware, network and organization are analyzed.

A proposal for the evolution to the technologies used is followed in order to implement a private cloud solution without neglecting the requirements initially defined. Some technologies that can be adapted and integrated with the existing system are proposed, as StackOps, the Ulteo, Hyper-V.

The work ends with a chapter where the conclusion concerns the reflection on all the previous chapters and their respective themes, trying to foresee a future, hopefully near, in which it is possible to undertake the proposed development.

Chapter 2

Cloud computing

Of all the existing definitions around the term cloud computing, the one I consider most comprehensive to encompass all its aspects is used by Gartner that defines cloud computing as an elastic computational model where it's IT-related capabilities are delivered "as a service" using Internet technologies to multiple external customers. [Stevens and Pettey, 2008]

Cloud computing integrates both hardware and software systems existing in the data center and applications that are made available as services over the Internet. The services provided are often called software as a service (SaaS) and the existing hardware and it's software within the data center are known as a cloud. When the cloud is made available to the general public, in a pay-as-you-use manner, it is refered as a public cloud or as a service sold (utility computing). Data centers within an organization or business that are not available to the public, are defined as private clouds. [Armbrust et al., 2009]

Cloud computing delivers computation, software, data access, and storage services that do not need end-user knowledge of the physical location and configuration of the system that provides the services. [Marston et al., 2011] Equivalents to this concept can be strained with the electricity grid, wherein end-users consume power without needing to understand the infrastructure or component devices required to provide the service. [Zhang et al., 2010]

The concept of cloud computing satisfies a continuous need of IT by increasing the capacity or adding capabilities on demand without investing in new infrastructure, training new personnel, or licensing new software. Cloud computing incorporates any subscription-based or pay-per-use service that, in real time over the Internet, thus extending IT's existing capabilities. [Sultan, 2011]

Cloud computing defines a new supplement, consumption, and delivery model for IT services based on Internet protocols, and it typically associates provisioning of dynamically scalable and frequently virtualized resources. It is a result and consequence of the ease-of-access to remote computing sites provided by the Internet. Simply by the form of web-based tools or applications that users can access and use through a web browser as if the resources (programs) were locally installed on their own computers.

Cloud computing providers deliver applications via the internet, which are accessed from a web browser, while the business software and data are stored on remote servers. While the computing resources are consolidated at a remote data center location, legacy applications (line of business applications that until now have been dominant in thin client Windows computing) are delivered via a screen-sharing technology. In other cases, entire business applications have been developed using web-based technologies such as AJAX.

Most cloud computing infrastructures comprise of services delivered through shared data-centers and seeming as a single point of access for consumers' computing needs. Commercial contributions may be required to meet service-level agreements (SLAs), but particular terms are less often negotiated by smaller companies. [Qian et al., 2009]

In diverse different perspectives, the core of private cloud computing is the intersection of the fast growing cloud-computing trend and the maturing virtualization trend. [Bittman, 2010a]

It is also a natural evolution for enterprises that are heavily virtualized [Bittman, 2011e], private cloud computing is emerging from the hype and has turned into very real implementations. It's expected in 2011 and future years to be an especially big year for proofs of concept and implementation. [Bittman and Scott, 2011]

2.1 Comparison

Cloud computing is a technological evolution of virtualization platforms and the distributed paradigm, inhering some features related to this technologies and thus sharing similar characteristics with the following:

- Autonomic computing designates distributed computational systems capable of adapting to unpredictable changes while hiding intrinsic complexity to operators and users, hence being self-managed.

- Client-server model computing states generally that any distributed application that differentiates between service providers (servers) and service requesters (clients).

- Grid computing is a method of distributed and parallel computing, where a massive virtual computer, composed of a cluster of networked, loosely coupled computers is acting in concert to perform very large tasks.

- Mainframe computer are powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as census, industry and consumer statistics, financial transaction processing, and enterprise resource planning.

- Utility computing represents the packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity.

- Peer-to-peer is distributed architecture not needing for central coordination, where users are at the same time both consumers and suppliers resources (in divergence to the traditional client-server model).
- Service-oriented computing refers to software-as-a-service.

2.2 Characteristics

The main benefits of private cloud computing fall into three specific areas: economics, QoS and agility. In each different implementation scenario values more or less the mentioned above characteristics. [Bittman, 2011c]

Cloud computing reveals the following key characteristics:

- Agility improves with users capability to re-provision technological infrastructure resources.
- Application programming interface (API) accessibility to software that allows machines to interact with cloud software in the same way the user interface enables interaction between humans and computers. REST-based APIs are typically used for cloud computing systems.
- Cost is demanded to be reduced and in a public cloud delivery model capital cost is converted to operational cost. This is supposed to lower barriers to entry, as infrastructure is usually provided by a third-party and does not need to be purchased for one-time or occasional intensive computing tasks. Pricing on a utility computing basis is fine-grained with usage-based options and fewer IT skills are required for implementation (in-house).
- Location and device independence allow users to access systems using a web browser regardless of their location or what device they are using (e.g. mobile phone, PC). Users can connect from anywhere, as infrastructure is off-site (usually provided by a third-party) and accessed via the Internet.
- Multi-tenancy allows sharing of resources and costs across a large pool of users therefore enabling:
  - Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
  - The peak-load capacity increases without the need of users to engineer for highest possible load-levels
  - Efficiency and utilization improvements for systems that are often only 10-20% utilized.
- Reliability is enhanced where multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.

- Elasticity and scalability via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis near real-time, without users having to engineer for peak loads.
- Resources and system performance is monitored and consistent and loosely coupled architectures are constructed using web services as the system interface.
- Security could is enhanced due to centralization of data, improved security-focused resources, etc., but apprehensions can persist about the loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than under traditional systems, in part because providers are able to dedicate resources to solving security issues that many customers cannot afford. Nevertheless, the complexity of security is significantly increased when data is distributed over a wider area or greater number of devices and in multi-tenant systems that are being shared by unrelated users. Furthermore, user access to security audit logs may be difficult or impossible. Private cloud installations are in part encouraged by the users desire to preserve control over the infrastructure and evade losing control of information security.
- The maintenance of cloud computing applications is less complex since they do not need to be installed on every user's computer.

2.3 History

The term "cloud" is used as a metaphor for the Internet representation that originated on the cloud drawing used in the past to represent the telephone network, and later to illustrate the Internet in diagrams of computer network as an abstraction of the underlying infrastructure it represents.

With the widespread adoption of virtualization, cloud computing is a natural evolution, service-oriented architecture, autonomic, and utility computing. The details are abstracted from end-users, who no longer have need for expertise in, or control over, the technology infrastructure "in the cloud" that supports them. [Weiss and Chuba, 2011]

The underlying concept of cloud computing dates back to the 1960s, when John McCarthy lectured that computation could someday be organized as a public utility. Most of all the modern-day characteristics of cloud computing (elastic provision, provided as a utility, online, illusion of infinite supply), the comparison to the electricity industry and the use of public, private, government, and community forms, were thoroughly explored in Douglas Parkhill's 1966 book, The Challenge of the Computer Utility. [Parkhill, 1966]

The actual term "cloud" derives from telephony in that telecommunications companies, who until the 1990s offered primarily dedicated point-to-point data circuits, began offering Virtual Private Network (VPN) services with equivalent quality of service but at a much inferior cost. By switching traffic to balance utilization as they saw fit, they were able to make use of their overall network bandwidth more effectively. The cloud symbol was used to designate the demarcation point between that which was the responsibility of the provider and that which was the responsibility of the user. This boundary is extended with cloud computing to cover servers as well as the network infrastructure.

After the dot-com bubble, cloud computing development was influenced by Amazon that played a key role by modernizing their data centers, which, like most computer networks, were using as little as 10% of their capacity at any one given time, just to leave room for occasional spikes. The finding of the fact that the new cloud architecture resulted in significant internal efficiency improvements whereby new features could be added in faster and easier manner, Amazon started the development of a new product with the effort to provide cloud computing to external customers, and launched Amazon Web Service (AWS) on a utility computing basis in 2006.

Eucalyptus, in early 2008, became the first open-source, AWS API-compatible platform for deploying private clouds [Tomas, 2011]. OpenNebula, in early 2008, enhanced in the RESERVOIR European Commission-funded project, became the first open-source software for deploying private and hybrid clouds, and for the federation of clouds. In the same year, efforts were concentrated on providing QoS guarantees (a requirement to use real-time interactive applications) to cloud-based infrastructures (such as a multimedia cloud [Zhu et al., 2011]), in the framework of the IRMOS European Commission-funded project, resulting in a real-time cloud environment. Gartner previewed, by mid-2008, an opportunity for cloud computing to profile the relationship among consumers of IT services, those who use IT services and those who sell them and concluded that organizations are inclined to switch from company-owned hardware and software resources to per-use service-based models so that the projected shift to cloud computing will result in dramatic growth in IT products in some areas and major reductions in other areas.

In July 2010, OpenStack was announced, attracting nearly 100 partner companies and over a thousand code contributions in its first year, making it the fastest-growing free and open source software project in history.

Private cloud computing is one of the technologies on the rise to the Peak of Inflated Expectations during 2011. This technology is one of the several analyzed that has broad-ranging impact across the business. It is an emerging technology in Gartner's Hype Cycle with numerous expectations on its shoulders. [Fenn and LeHong, 2011]

2.4 Architecture

Software engineering best practices have rebalanced over the years among monolithic, modular, and object- and service-oriented. The typical centralization versus decentralization problem has played out recurrently as its weight influenced among the "big iron" and "big client" paradigm. The interpretation of this problem, together with the rise of cloud computing and the propagation of powerful new computing devices, leads to the client-cloud application paradigm, which inherits from and improves client/server architecture. [Knipp et al., 2011]

The architecture of the software systems involved in the delivery of the cloud computing model, characteristically involves multiple cloud components communicating with each other over a loose coupling mechanism such as a messaging queue as represented in Figure 2.1. [Teneyuca, 2011]

The Intercloud is considered to be an extension of the Internet "network of networks" on which it is

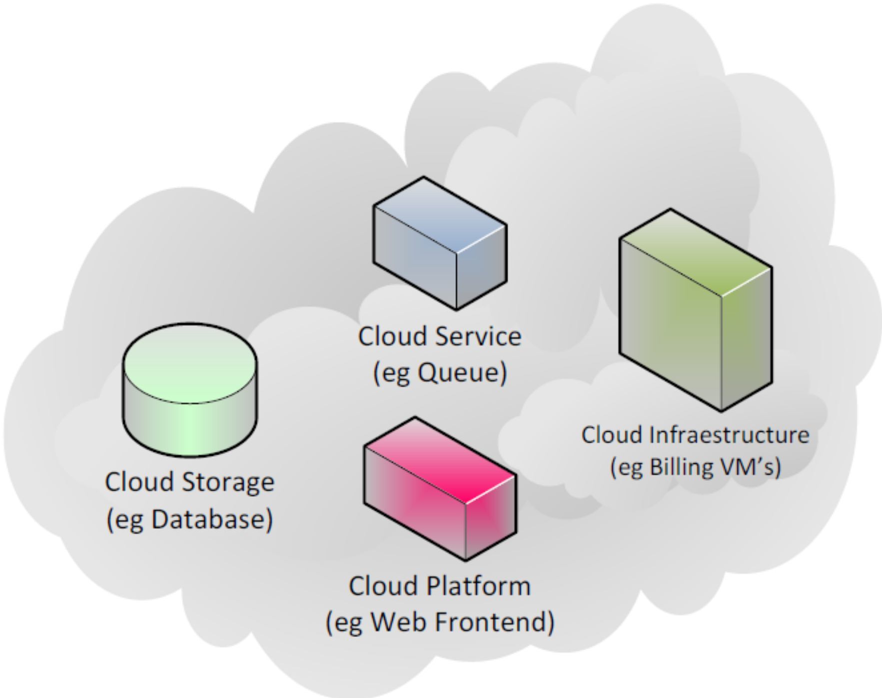

Figure 2.1: Cloud computing architecture

based and may be represented as an interconnected global "cloud of clouds".

Despite the massive evolution around this technology, there is no standard and specifications yet for cloud computing, however some progression in creating a Cloud Computing Open Architecture (CCOA) is already being drafted. [Zhang and Zhou, 2009]

Applying engineering disciplines to cloud computing arises in Cloud engineering. It brings a systematic approach to the high level concerns of commercialization, standardization, and governance in conceiving, developing, operating and maintaining cloud computing systems. It is a multidisciplinary process encompassing contributions from diverse areas such as systems, web, software, information, performance, security, platform, risk, and quality engineering.

Cloud computing is generically associated with the provision of three services (among others) to its users, such as application, platform and infrastructure.

The main premise behind the choice of a pyramid to represent layers is to think about building a structure, where each layer is built upon the next, potentially, creating a larger whole. Although each layer can be somewhat dependent upon each other and directly related, they don't require interdependence. In reality, each layer can, and does exist on its own. It is possible, for example, to build a Cloud Application on top of a Cloud Platform or Cloud Infrastructure, but the building process primarily works from the ground up. The reverse is not possible (e.g., build a cloud platform on top of a cloud application). There are other ways to describe this hierarchy as simple layers or interconnecting circles, but those don't necessarily express the strength of the structure, or infrastructure in this case. The cloud pyramid shown encompasses this idea in a simple visual representation in Figure 2.2.

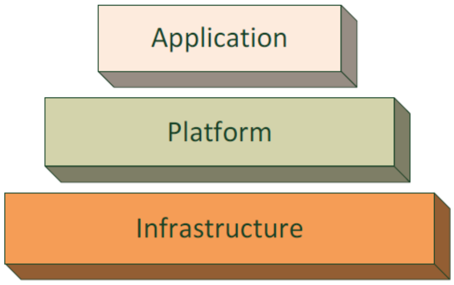

Figure 2.2: Service layers

The most common layers when it comes to cloud computing are displayed in the Figure 2.2, however, in more detail, it is also possible to subdivide the base layer into three, namely a computational (IaaS), a storage (DaaS) and a communications (CaaS) sublayer. At the base of the pyramid lies the Kernel Software and the true base would be the Firmware/Hardware (HaaS). [Youseff et al., 2008]

There are as well approaches of more detailed representation of other possible layers [Rimal et al., 2009], and ultimately, the taxonomy of Everything as a Service (XaaS) may be represented in cloud computing [Lenk et al., 2009].

2.4.1 Application

The enterprises better comprehend software as a service (SaaS) in contrast with the traditional on-premises applications. External providers deliver a set of common code and data definitions, on a pay-for-use or subscription basis witch they own and remotely manage the software. This can potentially save time, costs and resources over the traditional approach of deploying packaged applications for automating functions supporting prospects, customers, internal staff and partners. [Desisto, 2011]

This layer of the Cloud Computing pyramid basically contains applications that offer web-based software as a service (SaaS). Applications in this layer offer the end-user nothing more than what the application itself can do, consequently removing the need to install and run the application on the client's own computers and thus simplifying support and maintenance. Within this part of the pyramid, users are truly restricted to only what the application is and can do. Some of the notable companies here are the public email providers (Gmail, Hotmail, Quicken Online, etc.). This group can include various Software as a Service (SaaS) providers. Most retail consumers use the services within this Cloud layer. A user gets only pre-defined functionality and they cannot go much further than that. Applications are designed for ease of use and GTD (getting things done). SalesForce, a huge Cloud Application/SaaS provider that has led the way for hosted software [Natis et al., 2010], falls into this category as well, however, their force.com product does not. Even online banking offerings could be lumped into this group.

The Table 2.1 represents the SaaS main characteristics. SaaS Cloud providers can be integrated in different categories such as Business Management, Tools, Vertical Apps, Cloud Security and CRM where some known companies like Gmail and SalesForce can be integrated in these groups.

Table 2.1: SaaS Characteristics

Strengths	Weaknesses
• Sometimes free • Easy to use • Lots of different offerings • Easy to access • Good consumer adoption • Proven business models	• The application can be used as far as what it is designed for • No control of underlying technology • No knowledge of underlying technology

2.4.2 Platform

Platform as a Service (PaaS) refers generically to a set of middleware (application infrastructure) functions provided as a cloud service. [Natis, 2011b] In this layer of cloud computing can provide several strategic technologies, including application servers, database management systems (DBMSs), enterprise service buses (ESBs), business process management platforms, user experience platforms among others. [Natis et al., 2011a]

Developers are allowed to build applications to run on the platform provider's infrastructure in middle layer of the Cloud Computing Pyramid. This frees the developers from concerns about environment configuration and infrastructure scaling, but offers limited control. Cloud platform services, also known as Platform as a Service (PaaS), deliver a computing platform and/or solution stack as a service, often consuming cloud infrastructure and sustaining cloud applications. It enables deployment of applications without the cost and complexity of purchasing and managing the underlying hardware and software layers. In the center layer of the Cloud Pyramid resides the Cloud Platform layer where users obtain increased flexibility and control in comparison to the Cloud Application layer, however this is still somewhat constrained as to what a user can or cannot do. This layer also grants an increased level of complexity in terms of managing the frameworks therein (e.g., Java or Ruby). Products and companies like Google App Engine, Heroku, Mosso (now the Rackspace CloudSites offering), Engine Yard, Joyent or force.com (SalesForce development platform) fall into this layer. [Natis et al., 2009]

The Table 2.2 represents the PaaS main characteristics. PaaS Cloud providers can be integrated in different categories such as Deployment and Testing, Application Platform, Database, Integration and General where some known companies like Google App Engine, Microsoft Azure and Force.com (SalesForce) can be integrated in these groups.

Currently, PaaS is the least-developed and least-understood layer in the cloud computing architec-

Table 2.2: PaaS Characteristics

Strengths	Weaknesses
• Developers can upload a tightly configured applications and it simply “runs” within the platform’s framework • Provides more control than a Cloud Application	• Restricted to the platform’s ability • Complicated to work "outside the box" • Sometimes dependent on the Cloud Infrastructure providers

ture, compared with system infrastructure services (IaaS) and application services (SaaS). It is also the fastest emergent in terms of new vendor investments and leap of innovation. [Natis, 2011a] It is expected that during the next five years, the adoption of PaaS in most midsize and large organizations will not lead to a wholesale transition to cloud computing. In its place, it will be an extension of the use patterns of on-premises application infrastructures to hybrid computing models where on-premises application infrastructures and PaaS will interoperate, integrate and coexist. [Natis et al., 2011c]

2.4.3 Infrastructure

The infrastructure layer is the base foundation of the Cloud Pyramid and its where cloud platforms and applications are built on. Cloud Infrastructure providers allow users to create completely virtualized (or hybrid cloud/hardware) IT configurations, giving the user complete control of their environments. Cloud infrastructure services, also identified as Infrastructure as a Service (IaaS), provide computer infrastructure - typically a platform virtualization environment - as a service, along with raw (block) storage and networking. [Leong, 2011] Instead than purchasing servers, software, data-center space or network equipment, clients opt to buy those resources as a fully outsourced service. Suppliers normally bill such services on a utility computing basis, the amount of resources consumed (and therefore the cost) will typically reflect the level of activity. At the bottom of the pyramid are the Cloud Infrastructure providers like Amazon Web Services, GoGrid, and the Rackspace Cloud (specifically Rackspace CloudServers). Companies providing infrastructure enable Cloud Platforms and Cloud Applications. Many companies within this segment operate their own infrastructure [Leong, 2009], enabling them to provide more features, services and control than others within the Cloud Pyramid. Typically, vendors within the Cloud Infrastructure layer deliver Windows or Linux servers that are virtualized so that they can be provisioned dynamically and on demand. Billing for these virtual machines (VMs) is typically done based on usage or configuration and on an hourly or sometimes monthly basis. Other types of common infrastructure components like hardware-based load balancing or firewalls and other services like storage (frequently

Table 2.3: IaaS Characteristics

Strengths	Weaknesses
• Provides full control to a company's infrastructure • Not restricted to "containers" or "applications" or more restrictive instances	• Sometimes comes with a price premium • RackCan be complex to build, manage and maintain (based on provider)

scalable on-demand cloud storage) exist within this layer, providing a full assortment of network and infrastructure devices for an end-to-end solution.

The Table 2.3 represents the IaaS main characteristics. IaaS Cloud providers can be integrated in different categories such as Cloud Management, Virtualization, Networking, Storage, Content Delivery Networks and Computing where some known companies like GoGrid, RackspaceCloud and Amazon Web Services can be integrated in these categories.

2.5 Deployment models

The clouds can be implemented in different ways to better suit the requirements and needs of each user. Each of the deployment models are described in the Figure 2.3 and they can be:

- Public
- Hybrid
- Private

2.5.1 Public cloud

Public cloud defines cloud computing in the traditional mainstream sense, whereby resources are dynamically provisioned to the general public on a fine-grained, self-service basis over the Internet, via web applications/web services, from an off-site third-party provider who charges on a fine-grained utility computing basis.

2.5.2 Hybrid cloud

A hybrid cloud is a composition of two or more clouds (private, community or public) that stay distinctive entities but are bound together thus offering the full benefits of several deployment models.

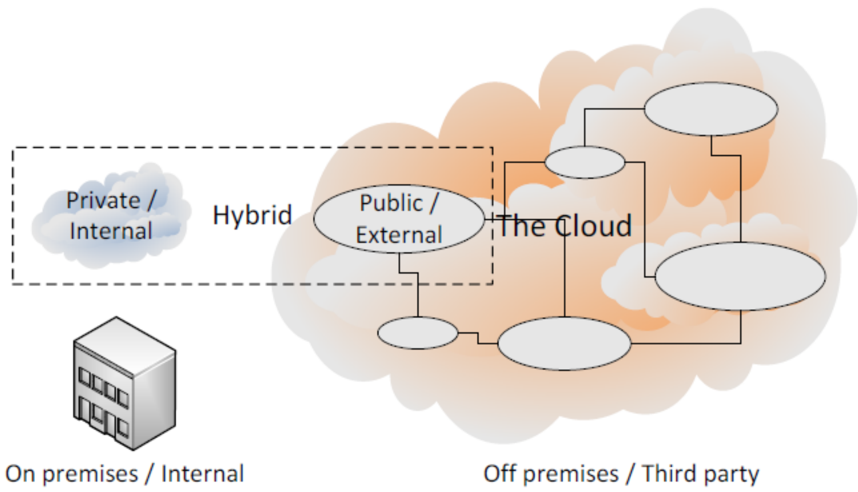

Figure 2.3: Cloud computing types

2.5.3 Private cloud

Private cloud is an infrastructure operated and managed exclusively for a single organization, either if it is managed internally or by a third-party and hosted internally or externally. They have attracted criticism because users still have to buy, build, and manage them and thus do not benefit from lower up-front capital costs and less hands-on management, essentially lacking the economic model that makes cloud computing such an interesting concept.

2.5.4 Community cloud

Community cloud shares its infrastructure amongst several organizations from a specific community with common concerns (compliance, security, jurisdiction, etc.) and it can be managed internally or by a third-party entity and it can be hosted internally or externally. The costs of this scenario are spread over fewer users than a public cloud (but more than a private cloud), so only some of the benefits of cloud computing are realized.

2.6 Drivers and Inhibitors

Private cloud services are eye-catching to IT organizations because they enable self-service ordering of frequently requested services as well as dynamic provisioning, thus, increasing the agility and reducing the cost of delivering IT services. [Scott, 2010] Most of private cloud computing implementations are focused on a particular service or use case, a few are focusing their private cloud computing initiatives on top-down strategic IT architectures. Both can be considered valid methods to the technology, with the last expecting to have significantly greater ROI (if successful), and the first focusing on a experimental approach to private cloud computing, with plans to implement it more broadly based on the results of the initial implementation. [Scott, 2011]

To select the right cloud solution approach requires in-depth analysis regarding the level of security, control, customization and support that a business requires. [Internap, 2011a]

Internap published an article containing a survey that exposes the main drivers (Figure 2.4) and inhibitors (Figure 2.5) when facing the adoption of a cloud technology.

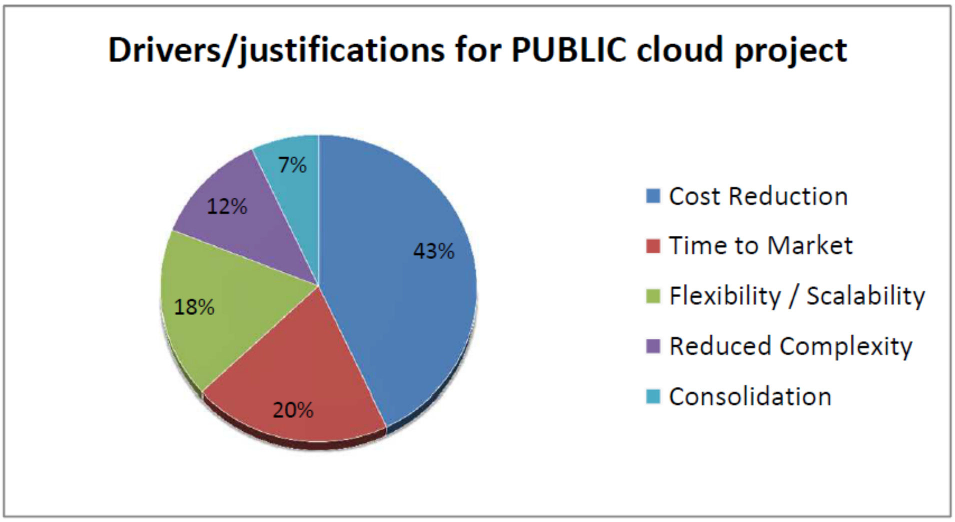

Figure 2.4: Drivers/justifications for PUBLIC cloud project

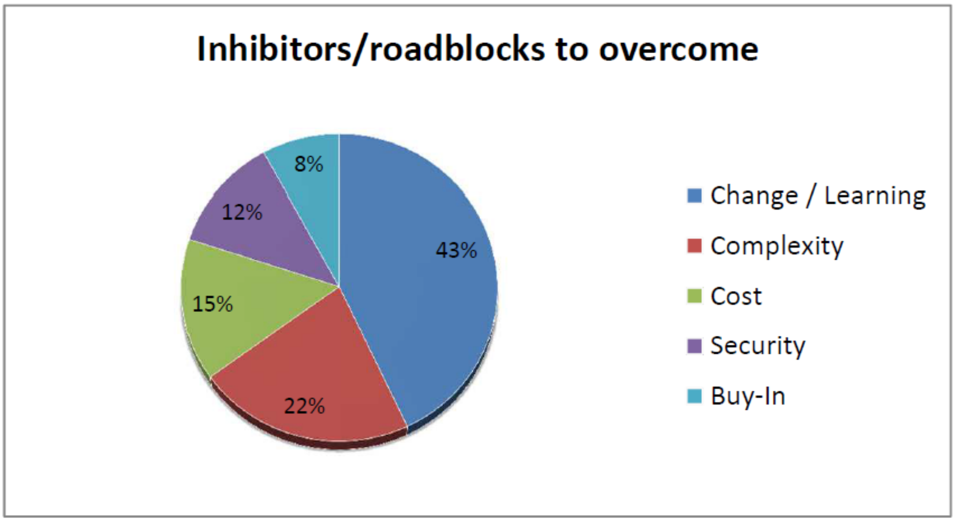

Figure 2.5: Inhibitors/roadblocks to overcome

One of the inhibitors of cloud computing, not shown in this survey results, is the performance issue making difficult for companies to adopt this technology for transaction oriented and other data-intensive applications. [Leavitt, 2009]

Table 2.4: Public vs. Private Cloud

Private Cloud	**Public Cloud**
More Security	Less Security
More Control	Less Control
More Customization	More Standardization
More Support	More Self-Service

Also, the practice of cloud security risk assessment, along with the variety and maturity of the necessary security mechanisms, will continue a rapid evolution over the next two to five years and beyond. For the near future, potential buyers and current users are justified in concerns about placing private or sensitive data within the public cloud, just as they are justified in concerns about any external service provider. Increased levels of regulatory enforcement, and pending legislative actions in multiple jurisdictions, are compounding existing reluctance to take full advantage of private clouds. [Heiser and Cearley, 2011]

Public and private cloud technologies raise management and security challenges, which neither is impossible to meet. To efficiently and effectively secure the use of cloud computing, it is mandatory to match threats and business demands with the right management security approach. [Pescatore, 2011] Three different approaches to this issue of securing public and private clouds can be used, to depend on the cloud infrastructure for security, to run its own security controls in the cloud or to keep security separated from the cloud. Some hybrid alternatives can also be implemented. [Pescatore, 2010]

The end-user expectations and demands for performance and reliability isn't effectively dealt in cloud computing. This is why the discussion of cloud latency has to shift away from IT defined acceptable levels of latency to end-user behavioral judgments as to what latency is acceptable. After all, end-users will penalize applications and websites based on the smallest performance delays or downtime. [Internap, 2010]

2.7 Private versus Public Cloud

Cloud users must select a cloud hosting model after a careful evaluation of each application or workload individually, as the security and customization requirements may vary greatly. For instance, web content for public distribution may not need the same levels of control and security as a company's ERP applications. Cloud users should leverage both private and public cloud solutions for different initiatives or applications. Independently of the particular deployment, it is imperious to select a solution that delivers superior availability and an enhanced end-user experience. [Internap, 2011b]

Cloud computing potential advantages are very well documented. If designed and provisioned properly, cloud deployments can lower capital and operating costs, increase flexibility and reduce complexity.

It is required to weigh the benefits gained from using one type of solution over another. [Ridder and Rold, 2010] Although the Table 2.4 shows some aspects of the differences between private and public clouds, when it comes to making the decision between private and public, it is useful to consider the following questions:

2.7.1 Security

Private cloud solutions typically run highly sensitive, mission-critical, proprietary applications, while public clouds usually host development or test environment applications. With certain private clouds, the applications reside within a private dedicated hosting environment, assured with firewall and multi-layered security features such as role-based security and intrusion detection. Within the multi-tenant environment of a public cloud, the rules of the shared community are followed. Even though public cloud security rules are stringent, they usually cannot be changed to meet the specific needs of individual companies. [Sengupta et al., 2011] If an organization has certain security policies that dictate how mission-critical applications are classified, deployed and accessed, then it would possibly be better suited to use a private cloud. If it needs support for traffic peaks on non-classified material, a public cloud is most likely to be a better choice for cost conscious organizations. Therefore, application and traffic trend considerations would help direct the decision between an enterprise-grade private cloud and a public cloud solution.

As enterprises move from the virtualization of their data centers to build private cloud-computing infrastructures, security issues must also progress to support this change. [MacDonald and Bittman, 2010]

It may not be easy to prognosticate what challenges lie ahead in security and privacy for cloud computing, but its certain that several will arise As in every other market and industry that adopts inter-networked computing, security challenges always arise, and almost always security is "bolted on" after the fact, rather than designed in from the start. [Ghosh and Arce, 2010]

2.7.2 Control and customization

The IT organizations may require highly customized environments for each distinct application. If so, keep in mind that certain configuration flexibilities may not be available in a pre-defined, commoditized public cloud. For example, configurations that oblige complicated networking designs or specialized hardware components may be difficult to accommodate in a public cloud. A private cloud runs multiple applications for one customer, whereas a public cloud supports a variety of applications for multiple customers, thus restricting choices for those public users. Furthermore, a private cloud is delivered through a managed service model, whereas a public cloud is often delivered through highly automated means. Consequently, the ability to be prescriptive about build and run requirements for your applications will be more robust in a private cloud. The private cloud user chooses the environment that can be configured to meet different specific needs. This level of control and customization does not exist within a public cloud. As a result, a private cloud solution may be an easier first step into a cloud environment, especially if you are not ready to redesign applications suited for a public cloud.

2.7.3 Support services

In most public clouds, customer service typically does not exist in a 24/7 type model, but may be available at an additional cost. Though, public cloud users are usually comfortable with this reduced level of support and do not expect highly interactive dialogues with their cloud provider. In a private cloud environment, users can access enterprise-grade support and a dedicated account team. As a result, for complex environments with sensitive applications, an enterprise-grade private cloud with a higher support model is the most appropriate solution. The support team in this scenario will be more intimately involved with your infrastructure and network personnel, providing hands-on support when you need it most.

2.7.4 Cost and flexibility

Choosing a public cloud solution can be less expensive due to its commoditized attributes, whereas the customized traits of a private cloud can increase costs. A public cloud is highly economical for variable workloads, while a private cloud may be economically attractive for stable, predictable application uses. Small business without the need for an expansive data center footprint or unneeded overhead, the public option may be the best choice. Nevertheless, it is important to contemplate all relevant costs in order to make an educated decision. For example, support costs are often overlooked in defining the true cost of moving into a public cloud.

A public cloud can offer superior flexibility with no contracts or long-term commitments. Furthermore, a public cloud typically offers more elasticity or "burstability" - giving users the ability to rapidly scale their solutions to meet planned or unplanned peak loads. A private cloud also has the ability to quickly scale, but is limited by the build-out of the existing environment. Both options are superior from a scalability perspective than in-house deployments that may require procuring, deploying and configuring new hardware to meet scalability demands.

2.7.5 Availability and performance

Once the customization and support models are defined, you should review the solution's overall performance and availability. It is relevant the fact that the network beyond the cloud characterizes the greatest risk to an enterprise; performance and availability is only as good as the network. Consequently, the provider must ensure that the solution does not end at the cloud, but instead comprehends an end-to-end approach from the cloud to the end user.

2.8 Datacenter

The physical infrastructure that serves as support to a system of cloud computing must meet certain minimum requirements. The ability to offer SLAs to users depends not only on the technology used and its potential, but also the hardware and conditions existing in the datacenter where it resides.

Datacenters in organizations today are the core of most business processes, applications residing therein, databases, storage and communication systems that ensure the continuity of the same, and whose ideal is to be available 24x7x365. Recent years have witnessed an exponential growth in data volume, which has required more processing power, bandwidth and storage space, so in many cases the capacity of data centers has run out quickly.

2.8.1 Framework

Technological development in terms of electronic components allows today, through the blade technology, a higher processing concentration and simultaneous a reduction of space, but in return a significant increase in required power, and consequently of the heat dissipated by the rack. On one hand it is reduced the need to occupy space, but in return points of high heat dissipation are created and an obstacle to be overcome at the risk of compromising the datacenter's efficiency.

Currently, with the blade technology and multicore x86 64bits processors the processing capability is increased, reducing power consumption and the cost of system management. Using virtualization tools, and provided that the applications permit it, it is possible to share resources at the level of processing, storage and networking in an architecture called cloud, thereby reducing the total cost of ownership and increasing the availability.

The increase in demands on the bandwidth requires that the infrastructure of copper wiring and fiber optic data centers are conveniently sized and use components that ensure the future growth of the datacenter, to facilitate its daily management and minimize the possible risks of stopping.

It is required for the datacenter infrastructure to be efficient, reliable, and scalable, able to adapt easily to growing demand of business processes and technological developments.

For this reason, the design and draft of a datacenter, today is a challenge that requires multidisciplinary expertise with a global vision of long-term, in order to develop an infrastructure that can maintain levels of performance for the longest period of time.

On the other hand the design and dimensioning of the technical infrastructure to support the functioning of a Data Processing Center have been evolving very significantly in recent years, establishing a new paradigm supported, fundamentally, the concepts of availability and energy efficiency.

It should be noted, though, because of its importance, the use of modularity and scalability criteria that allow optimizing the financial balance of the investment throughout its development. The physical infrastructure supporting the Data Processing Center includes the following areas:

- Wiring
- Physical Security

- Acclimatization
- Racks
- Energy
- Management of the infrastructure

2.8.2 Considerations

With the aim of maintaining a high availability network, should be considered on four base topics:

- People
- Processes
- Technology
- Network-Critical Physical Infrastructure

These four layers are mutually dependent, as is evident in Figure 2.6. If any of them fails, the business processes that depend on the network can be compromised. Thus, we can consider that the four layers represent a chain whose availability is equal to the weakest.

The NCPI is the support of this whole chain, consisting of several components:

- Physical infrastructure (safe room type)
- Energy (electrical, UPS and distribution)
- Cooling (production, distribution of cold air and remove hot air)
- Racks and Accessories
- Wiring and remaining physical infrastructure support
- Security, fire detection and fire protection
- Management and monitoring
- Services (installation, engineering, etc.)

2.8.3 Aggregation

Aggregation should be made taking into account the needs of the horizontal distribution of the datacenter and the necessary redundancy of connections required to ensure connectivity to the servers. This aggregation allows the existence of two levels of patching, which are extremely important to the organization of the connections of servers or other production equipment installed in the datacenter. Thus access to the backstage where the servers or other production equipment reside is only needed for installation or maintenance issues and not to modify existing connections.

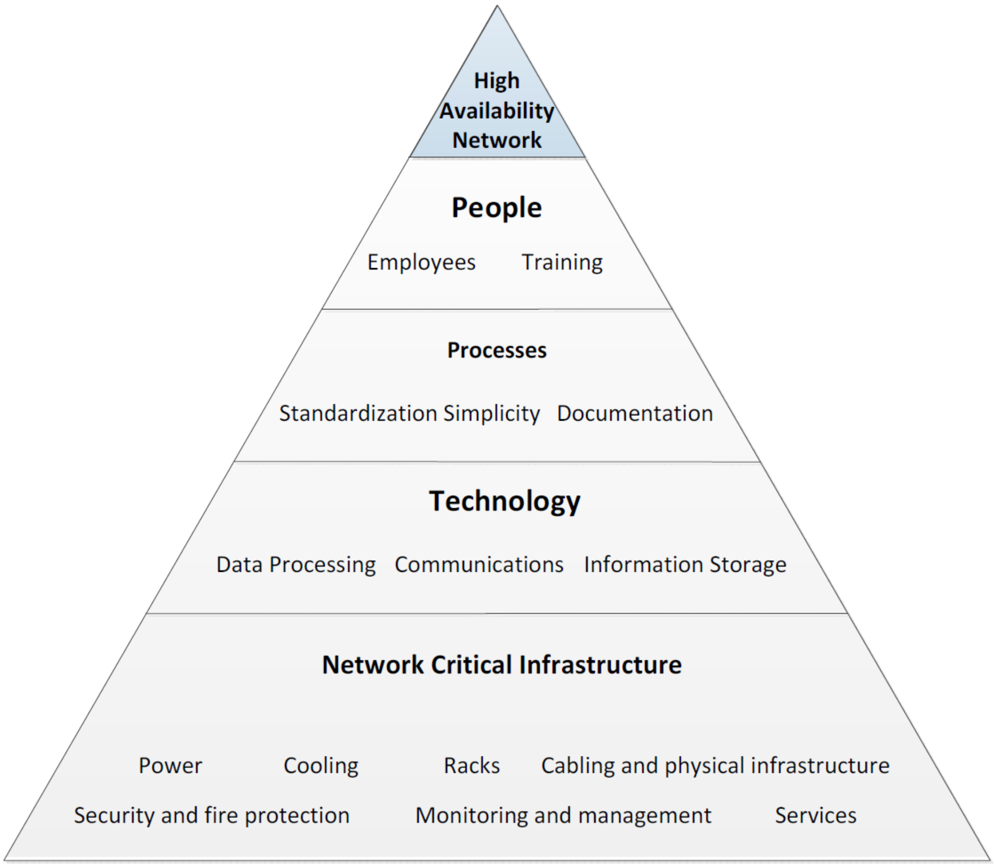

Figure 2.6: Network-Critical Physical Infrastructure

2.8.4 Evaluation criteria

The planning of the physical infrastructure of the datacenter must have regard to the business and its specific objectives through described below: **Information protection** - The implementation of solutions for physical security and redundancy that reduce the level of inoperability of the services; **Resources optimization** - Modular architectures that respond in time to business requirements through standard solutions reducing training costs; **Maintenance of the infrastructure** - Design and implement simple solutions, flexible and reliable with low maintenance and operation costs.

The fundamental principles in the planning of the infrastructure of the Data Center (Simplicity, Flexibility, Scalability, Modularity and High Availability) must be mirrored in the functional requirements of the physical infrastructure of the datacenter.

The main technical criteria that should guide the choices to make:

- Performance;
- Scalability;

- Flexibility;
- Availability.

The main criteria of financial nature are:

- Cost of acquisition;
- Cost of implementation;
- Costs of technological non-technical and functional or obsolescence;
- Costs of operation, management and maintenance.

Chapter 3

State of the art

An introduction to cloud computing was addressed in the previous chapter exposing its fundamental essence by comparing it to other forms of computing, detailing its characteristics, reviewing its history and by studying the layers, deployment models and architecture. In this chapter the main platforms of cloud computing focusing on the Microsoft Azure, Google Apps, Cloud Foundry, Delta Cloud and Open Stack are exposed. Other platforms that are emerging are also included therefore providing a broader glance to the technological solutions available.

Cloud computing main idea is to shift the IT complexities to providers, thus focusing on acquiring services. The private cloud is a real over marketed phenomenon but done well, can help IT better support to its customers with speed and quality. The core of private cloud computing can be considered the intersection of a maturing virtualization trend and the rapidly growing cloud-computing trend. [Bittman, 2010b] Considering Gartner's Magic Quadrant for x86 Server Virtualization Infrastructure, VMware, Microsoft and Citrix Systems are the leaders, therefore the cloud derivate of these technologies will be addressed in the state of the art. [Bittman et al., 2011]

3.1 Microsoft Azure

Windows Azure is an operating system in the cloud distributed around the globe throughout datacenters in North America (Chicago and San Antonio), Europe (Dublin and Amsterdam) and Asia (Hong Kong and Singapore). Controlled environment in these datacenters provide the end user hassle free features like upgrades, routers, patching, memory environments, OS installing, zero downtime, hard drive failure, storage, server acquisition, load balancers, . . .

The main components comprised in Windows Azure, embedding Microsoft SQL Azure and Windows Azure Platform (App Fabric):

- Cloud Fabric
- Roles (Web and Worker)

- Storage (Including SQL Azure)
- Dev Fabric
- App Fabric

The Cloud Fabric, Roles and Storage are described ahead. The Dev Fabric and App Fabric are much similar to the Cloud Fabric implementing, each in their own context (application development and application execution respectively) implementing access control and the service bus. Furthermore, the SDK, the Windows Azure Platform Appliance and the Windows Azure Services Platform are also explained.

3.1.1 Cloud Fabric

The Windows Azure components are resting on a software layer that controls, manages and supports the infrastructure. The Windows Azure Cloud Fabric provides the following characteristics:

- Multiple virtual instances;
- Easy provision of applications;
- Detect failures
- Replacement of failed instances by spinning up new ones;
- How many instances and what role the will play;
- Load balancers and DNS;
- Service elasticity (by scaling down/up number of instances).

3.1.2 Roles

Windows Azure has two main roles, a web role for web application programming in ASP.NET or PHP and a worker role for performing work on behalf of the web role. The worker role may be written in .NET, Ruby, Java and it is used to improve scalability by increasing parallelism and asynchronicity and for background work (see Figure 3.1).

In more detail these two main roles, a web role providing a single HTTP endpoint and a single HTTPS endpoint for external clients and a worker role providing internal endpoints for HTTP and TCP (may also receive work from Windows Azure Storage Queues). These roles can make outbound HTTP/S or .NET Framework class library socket connections to Internet accessible resources and can also access Windows Azure Storage services via REST APIs or the Windows Azure Storage Cliente Library.

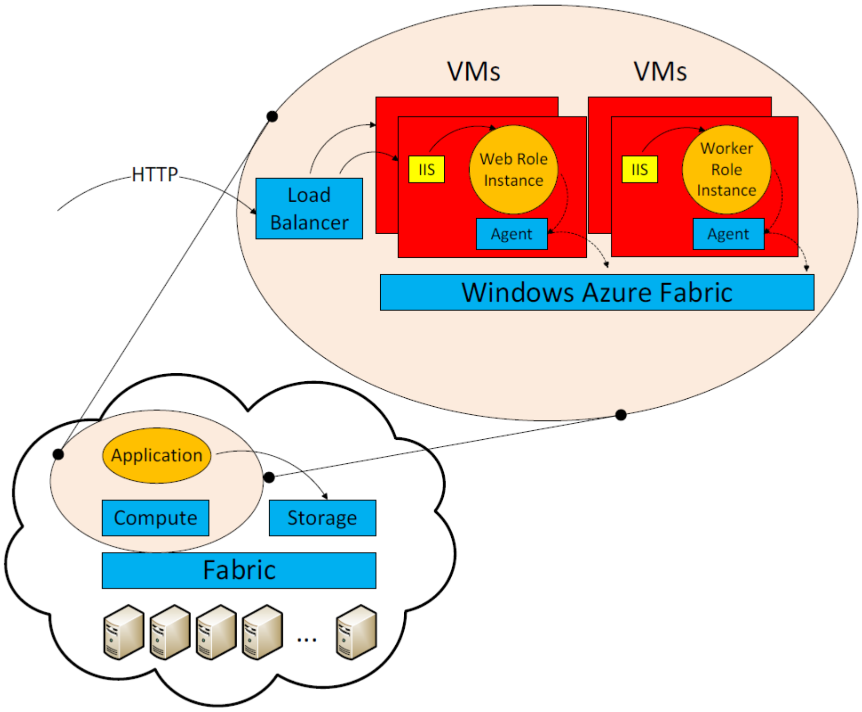

Figure 3.1: Windows Azure Roles

3.1.3 Storage

Windows Azure Storage comprises the following components as illustrated in Figure 3.2:

- Blobs: Large Data Store
- Queues: Background work processing
- Tables: Very Fast / Scalable Storage
- Drives: NTFS Formatted Page Blobs
- SQL Azure: Relational SQL in the Cloud

3.1.4 SDK

The Windows Azure SDK enables developers to benefit from the Azure full potential. This is accomplished by using specific project templates from the Visual Studio (with Visual Basic or C sharp languages), thus straightforwardly moving the applications to the cloud. Some of the available features are the Windows Azure Cloud Service solution with multiple roles, the tools to manage and configure roles, the Local Development Fabric, the Local Development, the Storage services, the Local Cloud

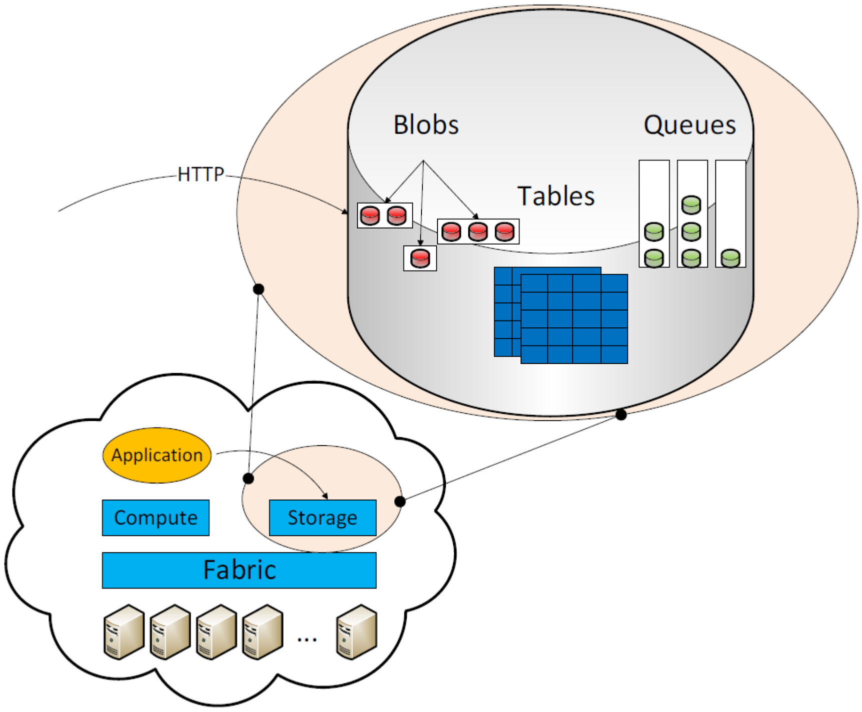

Figure 3.2: Windows Azure Storage

Service debugging support, the Cloud Service package and the deployment builder. SQL Azure does not require a SDK.

3.1.5 Platform Appliance

Windows Azure Platform Appliance (WAPA) is a fusion of a hardware package made available from a Microsoft OEM partner and software that is installed and updated by Microsoft with its Windows Azure set of platform services. It will be sold to service providers, enterprises and government customers. [Bittman and MacDonald, 2010]

With WAPA, Microsoft strongly specifies the architectural requirements (with certified storage, network and server hardware, similar to what it was made with the "chassis" model with Windows phones OEMs). The customer (whether a service provider or an enterprise that wants to host its own platform services) is responsible for the acquisition and day-to-day maintenance of the hardware architecture. What's really different with WAPA is that Microsoft maintains the software (OSs, a fabric controller, services like SQL Azure, etc.). The same software that runs in Microsoft's Windows Azure platform data centers runs in every WAPA installation, and the updates are deployed in the same way at the same time.

3.1.6 Services Platform

Microsoft's effort taking on Amazon and Google with a developer environment that runs on the Web. It means apps scale according to users customer needs, making possible to tap into Microsoft data centers instead of expanding your users.

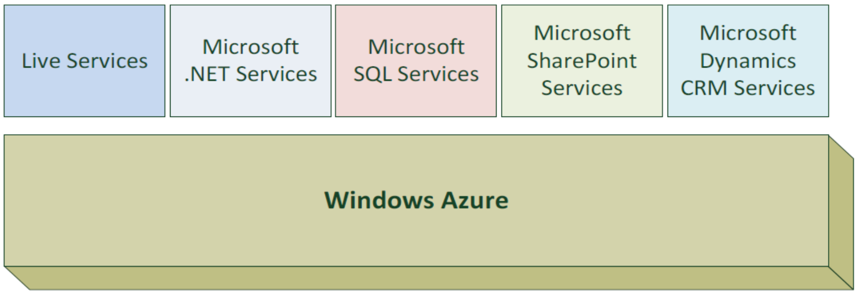

Figure 3.3: Microsoft Azure Services Platform

There are five main components in Microsoft Azure represented in Figure 3.3: Live Services (data and user resources), SQL Services (for a distributed and relational Web database), the security and communication between apps is assured by .NET framework, plus SharePoint and Dynamic CRM services (for tracking customers).

The main advantage to Azure is the platform integration ability with current developer workflows. It can be used directly with Visual Studio 2008 to write apps, deploy, test, and maintain them. Azure is also a good platform for testing because it means companies can create test apps and see how they perform in the real-world without having to build the infrastructure to support them.

The one missing element to all of this is the pricing. There are scarce details about how much it will cost to scale apps. Without that info, it's difficult to know how Microsoft has positioned Azure. It can be regarded primarily as an enterprise product or for Web 2.0 start-ups or even for a free-lance developer that don't have the hardware infrastructure for hosting. In comparison with Google Apps, it's clear that Google has the SMB (small and medium business) market in mind, as they recommend and exemplify a small insurance company that is growing quickly and needs to continuously add corporate services to remote offices.

One other unknown issue is that while Google Apps supported Python out of the box, Windows Azure only supports Visual Studio. Although it is promised Python support along with PHP and Eclipse support, down the road.

3.2 Google Apps

Google Apps is an example of cloud computing accessible to small businesses making possible to reduce IT costs and allow employees to do more by integrating their applications with Gmail, Google Docs, Google Sites and more. Google apps allows companies to optimize it's time and reduce cost of license.

When most users think of Google Apps, the first thing that comes to mind is the mailbox in Gmail or the ability to work on documents with other users in real time and collaboratively. For the majority, they do not stop to think about what happens behind the scenes. Google's data centers for cloud computing (cloud computing) offers customers scalability and reliability in all our products and websites, with support for millions of businesses using Google Apps and more than a billion searches that take place daily in internet. Google's pure and proven cloud offers Google Apps customers a great protection for the data, with a level of security that would be difficult for customers to achieve on their own. This is exactly the same infrastructure Google uses to carry out its own business.

The machines that run the applications in data centers have been developed according to Google's specifications, including those focusing on security. The hardware is restricted to that necessary to run the applications and eliminates unnecessary components such as connectors for video cards or peripherals. Similarly, the software that runs on these machines is a specialized and simplified version of the Linux operating system, which allows to avoid unnecessary software code such as device drivers. This focus helps deliver a computing environment less prone to vulnerabilities, compared to systems deployed at customer sites (systems on-premise), the so-called private clouds or hybrid IT environments.

The services that Google offers are mainly based applications and platforms on the Internet. Some Google employees have published materials on some of our key underlying technologies, such as BigTable, the protocol SPDY, Google File System (GFS) and MapReduce. These last two were the inspiration for Hadoop, the open-source framework from Apache that supports many leading applications that run in the cloud or that manage very high volumes of data. An interesting book about some of the pockets of Google was written by Urs Hölzle and Luiz André Barroso, addressing the topic of data center design and respective warehouse scale machines [Barroso and Hölzle, 2009].

High network traffic can be produced by the presence of a large number of users which provides significant advantages in terms of safety. For example, the Gmail spam filter quickly gets visibility into emerging threats or has evolved from viruses and spam, which helps block the vast majority of them. This type of internet infrastructure on a large scale also offers greater protection against attacks by DoS. It can also detect malicious traffic and helps protect users against malware.

But this unprecedented global scale would not have mattered without the ability to deliver reliably, business-critical services. This is another important feature of the technology and discipline processes in Google. The platform was developed to be able to withstand hardware failures, software-based and highly automated processes to ensure compliance with service level agreement (SLA) of 99.9% uptime of systems without downtime scheduled for maintenance. In 2010, Gmail's uptime was 99.984%, and the first half of 2011 surpassed the average of 99.99%. This is a goal that would hardly be achieved with technological systems "on-premise."

Some of the most productivity and collaboration tools available to the users by Google Apps are considered to be the following:

- GMail for business: Mail Management with a simple and friendly interface, very easy to use that also includes instant messaging and chat ..
    - Google Calendar: A web calendar that lets you optimize the schedules of employees, signifi-

cantly reducing the time spent organizing meetings or events.

- Google Docs: Your documents in a Web application where available from anywhere, anytime, on any device.
  - Google Groups: Groups can use Google Apps as a distribution list to share documents, calendars, websites, videos, etc ... from a number of users or other groups.
  - Google Sites: Google Sites is the easiest way to create web pages or intranet type of project, both internal and external.
- Google Video: A simple and secure way to host your corporate videos, training or product.

The Google Apps are available in two packages aimed at different target groups. One version for businesses and one for education. Google Apps for Business offer 100% web, reducing IT costs, maintenance and management, minimizes and accelerates the initial configuration. Google Apps for Education in the other hand, enables the approach of teachers and students thus causing a significant change on campus and around the world.

3.2.1 Google Apps for Education

Google Apps for education adds value to educational institutions, faculty, staff and students at any level will benefit from this wide range of communication and collaboration tools.

Tools such as Google Sites and Google Groups are really needed in the education world because it allows the exchange of information and interaction with any member of the organization and not just for schools, Google Apps for education is used by a wide variety of organizations nonprofit.

Any nonprofit organization with fewer than 3,000 employees will be entitled to use Google Apps for Education at no cost that can benefit the following advantages:

- Free: Google Apps is free (and ad-free) for students, teachers and school employees eligible.
- Security: ISAE 3402 certificate of type II. Your personal information will remain secure this, the service is operating at 99.9% of the time and has the best disaster recovery at no cost.
- Anytime, anywhere: Because all processes are based on the browser, students and teachers can easily work on any computer without having to buy the software.
- Protection: set of filters to limit e-mail messages to the control center and set up their sites and documents to be shared only within the center.
- Collaboration: Increasing participation and appreciation of the class using the collaboration tools in real time, allowing students to work in the same document at the same time, from anywhere in the world.
- Open Standards: the use of numerous API and support for open industry standards facilitates the integration of Google Apps into existing IT systems.

3.2.2 Google Apps for Business

Google Apps offers a 100% web, reducing IT costs, maintenance and management, minimizes and accelerates the initial configuration, focusing on the importance of business and not on the mail server. Google Apps for Business offers a powerful suite of messaging and collaboration based on Cloud Computing platform offering a 100% web, reducing IT costs, maintenance and administration, enabling greater investment in the business.

IT departments can be released from the endless cycle of revisions and upgrades of servers and computers and software to manage tasks.

IT programmers can stop worrying about backups, scalability, reliability and performance, and start collaborating with the company in designing and creating solutions that generate real competitive advantage. Thus, no longer require large investments based on an estimate of future needs, but an investment based on usage.

- E-mail, IM, voice and video chat: Each user has 25 GB storage for email and instant messaging (50 times the industry average).
- Access anytime, anywhere: Gmail works via the web safely, so you can work from your office computer, at home, when you scroll through the phone and even unrelated.
- Sync with Android, iPhone, iPad and BlackBerry: Enjoy the benefits of Google Apps in the major mobile platforms.
- Research and immediate detection of e-mails: Spend less time organizing e-mail. Quickly find the message you want thanks to Google search to your inbox.
- Less spam: The effective spam filter Gmail helps you focus on messages that matter to you. Postini filters gives you the ability to customize the spam.

3.3 Cloud Foundry

Cloud Foundry, was released by VMWare (NYSE: VMW) as an open source project that delivers platform as a service (PaaS), enabling users to decide on what cloud is more appropriate for their needs, developer frameworks and/or application services. Launched in April 2011 with the source code available by the Apache license, Cloud Foundry is an application platform from the cloud era to simplify and speed up the development, implementation and operation of applications. [Natis et al., 2011b]

Cloud Foundry can support multiple frameworks, multiple cloud providers, and multiple application services all on a cloud scale platform as it can be verified in Figure 3.4.

This platform provides the agility needed to the developers have resulting in a set of performance and productivity gains.

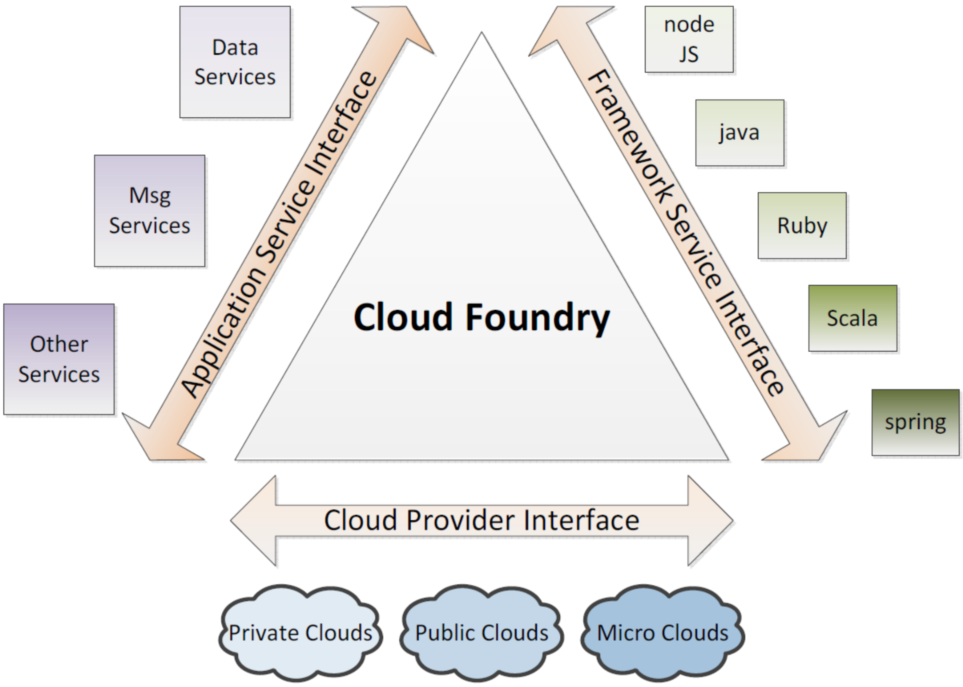

Figure 3.4: Cloud Foundry Overview

- Friction-free way to develop, test and deploy applications. Cloud Foundry lets developers focus on writing their applications and not meddling with middleware and infrastructure. While providing self-service access to a choice of high productivity frameworks and application services, developers can quickly build and test next generation applications on their own laptop and scale to the cloud with no code modifications.

- It is possible to achieve optimized software delivery, to assure portability without alterations (either in development, test or production) and to deploy into private or public clouds.

- Developers can write the application once, test it, scale it and deploy it to production without changing the code. Cloud deployment options can be preserved across private, public and hybrid clouds. Cloud Foundry enables applications architects and operation teams to reduce time to market by shortening the software delivery pipeline and reduce application backlog.

Developers have also the liberty to choose from various development frameworks, cloud types, and services.

- Enables to choose from what kind of cloud is used for deployment, across private, public and hybrid clouds (CloudFoundry included).

- Developers can choose from any industry standard frameworks including Ruby for Rails and Sinatra, Spring for Java, node.js, Grails, Scala on Lift and more via partners (for example Python, PHP).

- The decision of the application services to use is also configurable, including MySQL, MongoDB, Redis and RabbitMQ and vFabric PostgreSQL from VMware and more from third parties and the open source community.
- Developers and architects have at their disposal an extensible architecture that enables to "future proof" their organizations for rapid cloud innovation.
- A Community open-source project [1] allows any developer to access, evaluate and contribute. It is also possible to include integration with other frameworks, adding application services and deploying to other cloud infrastructure.

3.3.1 Micro Cloud Foundry

Micro Cloud Foundry ™ is the industry first downloadable PaaS to run on a developer laptop, and dozens of partner clouds running on private or public infrastructures including vSphere/vCloud, AWS, OpenStack, Rackspace, Ubuntu and more.

3.4 Delta Cloud

Apache Deltacloud is a REST-based (HATEOAS [2]) cloud abstraction API, that enables management of resources in different IaaS clouds using a single API. A series of back-end drivers 'speak' each cloud provider's native API and the Deltacloud Core Framework provides the basis for implementing drivers as needed for other/new IaaS cloud providers.

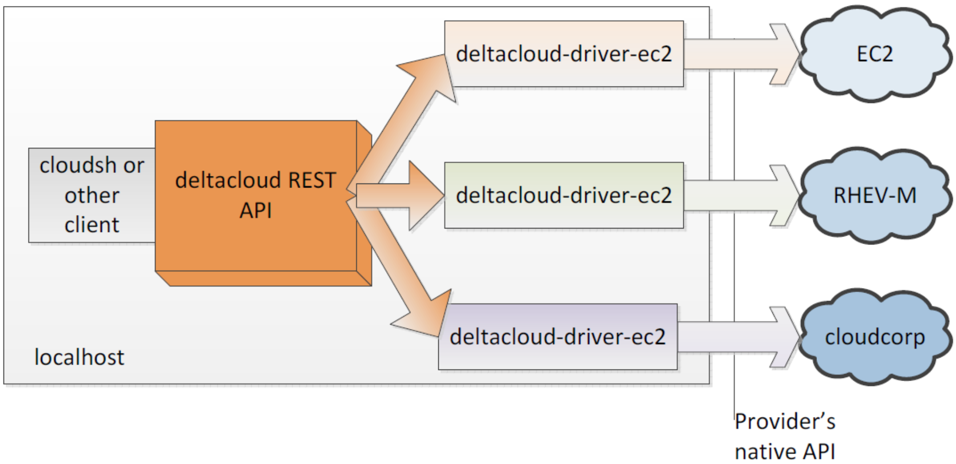

Figure 3.5: DeltaCloud Cloud API Management

[1] www.cloudfoundry.org
[2] Hypermedia as the Engine of Application State

The Apache Deltacloud project empowers its users in avoiding lockin to any single cloud provider. Deltacloud provides an API abstraction that can be implemented as a wrapper around a large number of clouds (see Figure 3.5), freeing users of cloud from dealing with the particulars of each cloud's API.

According to RedHat, the perspective on the future of open cloud computing requires more than open source, with all its importance. [RedHat, 2011] It also requires, among other attributes, APIs that are open, pluggable and extensible, allowing users to add resources, suppliers and technologies from various manufacturers or other sources. It is critical that the APIs are not under the control of a specific vendor or tied to a specific implementation, but they are under the auspices of an organization that allows third-party contributions and extensions of open and transparent way.

This is the exact description of the approach adopted by the Project Apache Deltacloud. The Deltacloud was started by Red Hat and is now a project housed by the Apache Software Foundation. The community that grew around the Deltacloud develops governed by the principles of collaborative and meritocratic ASF. The primary objective is to eliminate the Deltacloud the differences between clouds using a variety of techniques that offer consistency while making it possible to take full advantage of the functionality present in different clouds. For example, a technique called introspection allows cloud providers to display optional features, which can then be used when present. All this helps to avoid the problem of "lowest common denominator" that can be a drawback in multiplatform APIs.

The goal of Deltacloud, therefore, is not an API to be more cloud providers, but meet the needs of users who want to have the ability to operate a hybrid cloud environment as perfectly as possible. (Red Hat CloudFormsTM is a product management and hybrid cloud that leverages open Deltacloud as one of its upstream projects).

The Deltacloud is also an architecture with a high degree of modularity, offering a wide range of choice both in the way users communicate and cloud providers which it supports. With version 1.0, the Deltacloud offers a variety of front-ends, ie interfaces through which users make a request to a cloud provider, brings a new front-end Amazon EC2 basics, that can use the widespread Amazon Web Services interface to communicate with a diversity of cloud providers. And it is primarily intended for users seeking to migrate between EC2 and other cloud provider.

The new interface EC2 is part of the front-end that implements RESTful Cloud Infrastructure Management Interface (CIMI) from the Working Group of the DMTF Cloud Management. The CIMI defines a logical model for resource management in the field of Infrastructure-as-a-Service (IaaS). These standardization efforts garnered broad support from industry, the model and protocol provisional list of 60 employees or more authors. For those interested in the nitty-gritty details about CIMI, a range of specification documents work in progress can be found on the DMTF.

There is also a front-end RESTful "classic" that implements an API developed before the availability of labor with the DMTF in the CIMI.

The back-end Deltacloud, ie one who talks to the cloud providers themselves, is also modular. The code specific cloud vendor is encapsulated in what is called a "driver". This provides support for new cloud vendors are implemented independently of other components Deltacloud. Among other advantages, it is easier for cloud providers write their own drivers if they want to do it.

The list of driver Deltacloud already well expanded. A full array of support is available at the project site and includes suppliers of computing and / or storage for cloud providers like Amazon, Rackspace, IBM Smart Business Cloud, Fujitsu Global Cloud Platform, VMware vSphere, OpenStack and Red Hat Enterprise Virtualization. (In this context, "cloud provider" means any infrastructure that can provide resources for hybrid cloud - including platforms for virtualization management).

3.5 Eucalyptus

The platform implements the Eucalyptus cloud computing model based on private and public. In addition, Eucalyptus implements the model of service Infrastructure as a Service (IaaS) that includes the services offered in the layer of infrastructure (servers, routers, storage systems and other computing resources). It is also responsible for providing the necessary infrastructure for the Software as a Service (SaaS) and Platform as a Service (PaaS).

The Eucalyptus project (Elastic Utility Computing Architecture Linkin Your Programs To Useful Systems) started in the Computer Science Department at the University of California, Santa Barbara researcher RichWolski and his team, in order to investigate the area of High Performance Computing (HPC) more specifically from the project Virtual Grid Application Development Software project (VGrADS) funded by the National Science Foundation. To complete the final testing of the VGrADS on supercomputers, was chosen to Amazon's public cloud. However, VGrADS always been a joint project between the University and Research Laboratories with private data and it required a more detailed site investigation on the behavior of different applications. Thereafter, in early February 2008, began the development of open source Eucalyptus platform. The first version was released on May 29, 2008 only supported the EC2 and in December 2008 was included to interface with the S3 service. In 2009, the team created the company Eucalyptus Eucalyptus Systems Inc. to commercialize the Eucalyptus Enterprise.

Eucalyptus is an infrastructure of open source software that uses multiple computational resources available to researchers, such as workstation clusters and farms. In order to promote the exploitation of the scientific community systems of cloud computing, Eucalyptus was developed completely modularized, allowing researchers to perform experiments of security, scalability, reliability and implementation of interfaces. This was successful achieved by the Cloud-Enabled Space Weather Platform (CESWP) project bringing the power and flexibility of cloud computing to space weather physicists [Satchwill et al., 2011]. The development of Eucalyptus was guided to achieve two main goals:

- Being able to be deployed and run on hardware and software environments that are not under the control of its creators;
- Extensibility: must be modularized to allow easy maintenance of a component or even replacement.

Extensibility is guaranteed to build an architecture using Web Services, in which interfaces are well defined using WSDL documents and the authentication is done via WS-Security mechanisms. At the same time, the Eucalyptus supports an interface with the cloud popular Amazon AWS. Eucalyptus is compatible with various Linux distributions, including Ubuntu, Red Hat Enterprise Linux (RHEL), CentOS, SUSE Linux Enterprise Server (SLES), OpenSUSE, Debian and Fedora and a variety of virtualization technologies, like VMware, Xen and KVM hypervisors.

Furthermore, Eucalyptus was developed with enterprise level technology solutions running on top of the open source core software. Thus it is possible to provide features such as end-user customization, support legacy applications, Service-level agreements (SLA) custom monitoring of clouds and supports automatic scaling. As the Eucalyptus implements IaaS service model, then, undoubtedly, it shares some characteristics related to infrastructure, such as a single interface for management, an Application Programming Interface (API) for interacting with other hosts, switches, routers, and support for adding new equipment in a simple and transparent.

Eucalyptus has two versions: Enterprise Edition and Eucalyptus Eucalyptus Open-Source. Organizations in the Enterprise version has the possibility to implement private clouds with its own policies, fairly intuitive using the entire apparatus developed by the Eucalyptus platform. In the Open-Source version, which also includes the entire apparatus of the platform, any configuration can be viewed by the whole community, and the following restrictions are imposed in order, according to the Eucalyptus team, one of the main goals is to minimize the negative effects of experiments gone bad by maximizing the number of community members who can benefit from the system.

- An instance has a maximum duration of 6 hours;
- A maximum of 4 instances can run simultaneously per user (when more than four instances, the instances are randomly closed until only four remain);
- A static IP allocation has the maximum duration of 6 hours;
- The maximum duration of a "bucket" Walrus is 3 weeks;
- The maximum size of a "bucket" Walrus is 5GB;
- The maximum lifetime EBS is an instance of 3 weeks;
- The maximum size of an EBS instance is 5GB.

According to the Eucalyptus team, in case any is open-source project with affected productivity due to these restrictions, it is possible to request the provision of an adequate infrastructure, and for this purpose all is needed is to send an email requesting intervention.

3.5.1 Architecture

A private cloud based on Eucalyptus supports on a virtualization infrastructure where the components that make up the cloud typically exist as virtual machines. System management is performed internally, and the communication with the outside (via HTTP or API) is performed through a node responsible to act as a firewall. The Figure 3.6 illustrates an overview of the architecture of Eucalyptus and its communication interfaces.

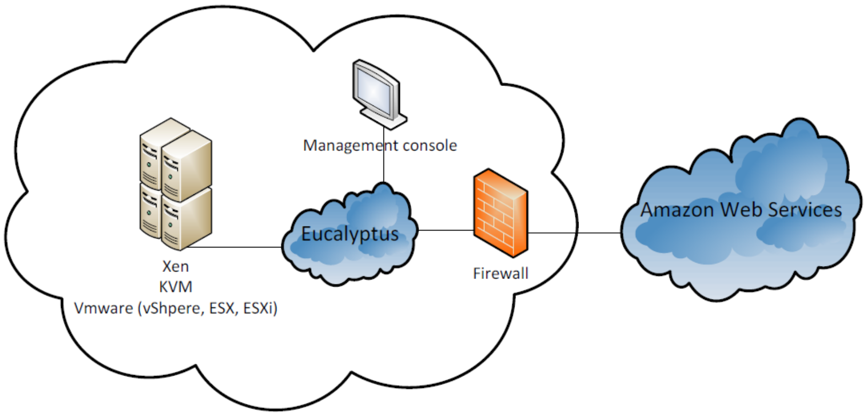

Figure 3.6: Eucalyptus Architecture Overview

Eucalyptus users have access to various resources, such as small clusters, and several set desktop workstations or server machines. Because IP addresses are scarce and public access to the entire branch of the cloud internal hosts may be dangerous, usually network administrators elect a node (known as frontend node) responsible for making a routing of public information for internal hosts. These internal hosts can only communicate with their respective front-end node or other internal hosts via a private network.

The components that make up the architecture of Eucalyptus are: Node Controller (NC), Cluster Controller (CC), Cloud Controller (CLC), and Walrus Storage Controller. The Figure 3.7 illustrates this architecture and components.

3.5.2 Node Controler

The Node Controller is the component that performs the physical resources of a virtual machine. Thus, we can only allocate a Node Controller on a physical machine and several other Node Controllers running on virtual machines. Each Node Controller is responsible for controlling the inspection, enforcement and termination of the respective instance. Although this scenario is very suitable to generate a bottleneck in the architecture of Eucalyptus, this assumption is virtually ruled out due to the monitoring performed by the Cluster Controller. In each Node Controller is a WSDL file that defines the interface to perform various operations, such as runInstance and describeInstance who are responsible, respectively, by starting an instance and get a report of an instance of physical resources (number of cores , memory and disk capacity).

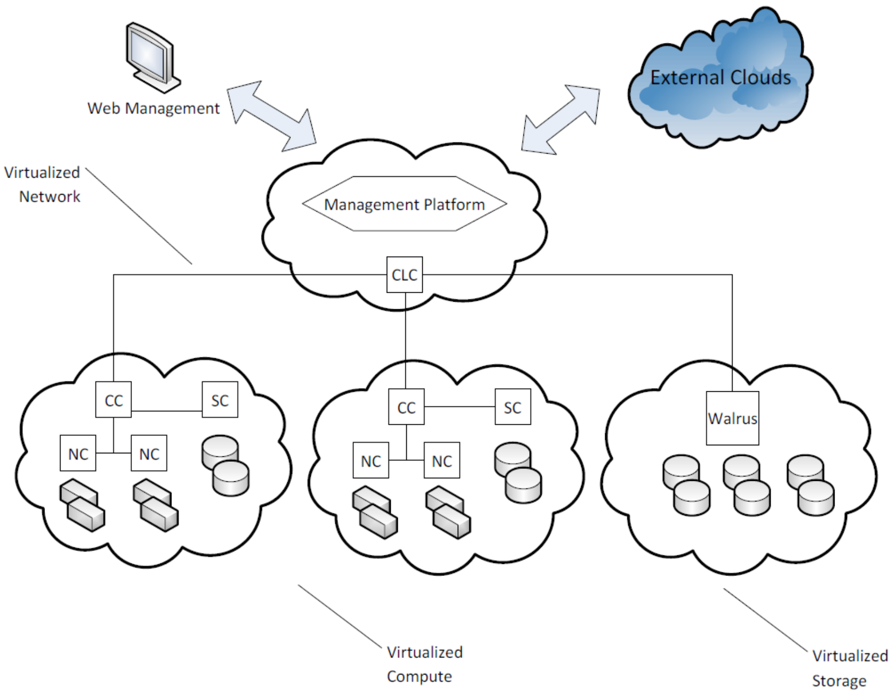

Figure 3.7: Hierarchical architecture of resources used by the Eucalyptus

3.5.3 Cluster Controler

The Cluster Controller is the component responsible for performing the routing of packets between networks virtualized external (public) with the internal (private). Thus, all manipulations performed on the physical resources allocated to the Node Controllers must necessarily be managed by the Cluster Controller. To this end, the Cluster Controller stores a WSDL file with all definitions of interfaces to perform the necessary operations. This WSDL file is very similar to the Node Controller, but this enables you to perform operations both for a single Node Controller and for a set of nodes Controllers. When a Cluster Controller receives a request from runInstance, it evaluates how best Node Controller to perform this new instance based on criteria such as amount of available resources and energy consumption of the physical machine associated with the Node Controller.

3.5.4 Cloud Controler

The Cloud Controller is the gateway to the cloud for administrators, developers and end users. The Cloud Controller is responsible for consulting the control node for information about the physical resources, make decisions of high-level programming and implement them by means of requests to the controllers of the clusters. The Cloud Controller is composed of a collection of services that handle user requests, authentication, persistence system and user metadata (eg, images of virtual machines and ssh key pairs). These services are configured and managed by an Enterprise Service Bus (ESB). To ensure the extensibility that level of granularity, the components that make up (including virtual machines, SLA and user and administrator interfaces) are mutually isolated and the interfaces are well defined internal communication, in which all this orchestration is performed by the ESB. Thus, the implementation of the Cloud Controller enables you to work with Amazon EC2 using Web-Services and Query Interfaces.

3.5.5 Storage Controler

Storage controller implements the access to network storage blocks, such as Amazon Storage Elastic Block (EBS) and is able to interact with storage systems, such as NFS, iSCSI, etc.. An elastic block can be attached to a virtual machine with Linux, but all traffic is sent to disk a remote storage site. An EBS volume can not be shared between instances, but allow a snap-shot to be created and stored in a central storage system, as in the Walrus.

3.5.6 Walrus

The Walrus allows users to store data and organize them into buckets. Furthermore, it is possible to insert Walrus access policies buckets and stored data. Each bucket is compatible with Amazon S3 and Amazon Machine Image (AMI), which provides a mechanism for access to user data by the different instances of virtual machines.

3.6 Open Stack

A global collaboration of developers and cloud computing engineers originated OpenStack, thus producing the ubiquitous open source cloud computing platform for private and public clouds. The project goal is to provide solutions for all types of clouds that is massively scalable, and feature rich and yet simple to implement. [OpenStack, 2011] The technology consists of a series of interrelated projects delivering various components for a cloud infrastructure solution.

OpenStack was founded by NASA and Rackspace Hosting and it has grown to be a global software community of developers collaborating on a standard and extremely scalable open source cloud operating system. Its mission is to permit that any organization to create and offer cloud computing services running on standard hardware.

This project has three major components (discussed below):

- Compute
- Object Storage
- Image Service

In addition to the three existing core projects above, two new projects were incubated with the Diablo release and will be promoted to ‘core’ for the forthcoming Essex version.

3.6.1 Architecture

The Figure 3.8 represents the OpenStack components with each of the daemons running and the communication network that each of the components use.

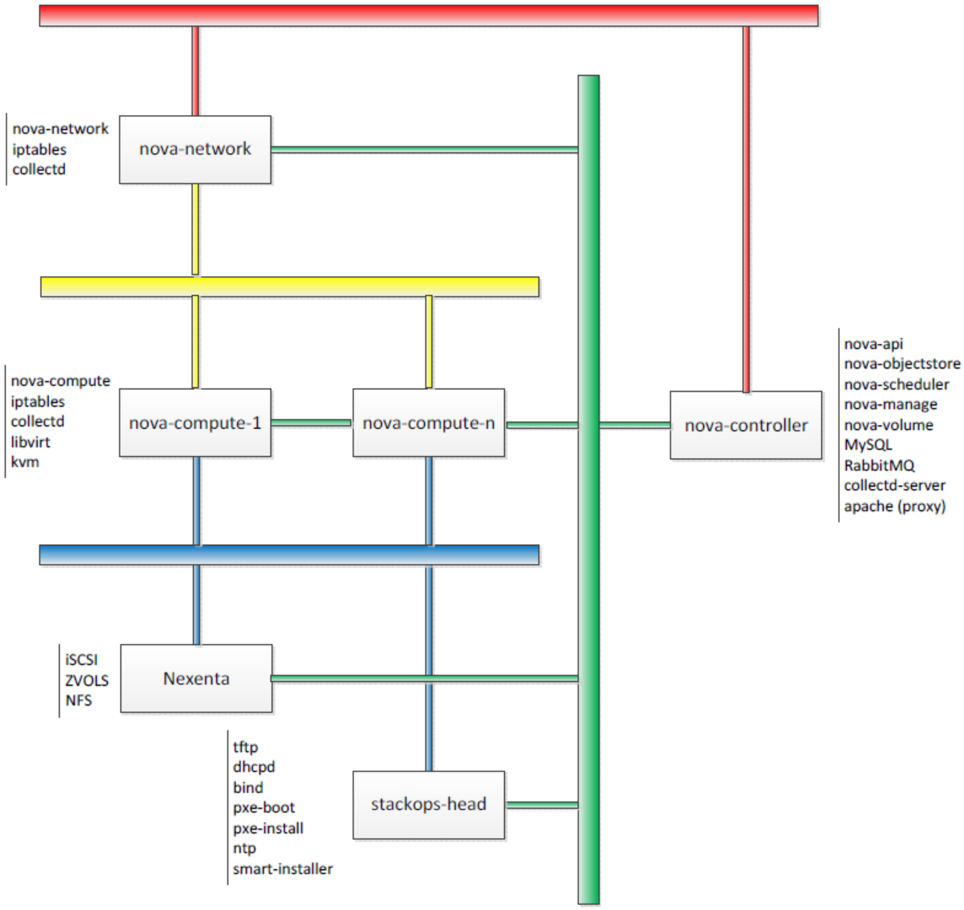

Figure 3.8: StackOps architecture

3.6.2 OpenStack Compute

OpenStack Compute is open source software that is conceived to provision and manage large networks of virtual machines, consequently creating a scalable and redundant cloud computing platform. It delivers the software, control panels, and APIs required to orchestrate a cloud, including running instances, managing networks, and controlling access through users and projects.

The Table 3.1 describes the main features and benefits of the OpenStack Compute.

OpenStack Compute endeavors to be both hypervisor and hardware agnostic. It currently supports seven major hypervisors and a variety of standard hardware configurations.

Popular Use Cases

- Service providers that offer IaaS compute platform
- IT departments that provision compute resources to teams and projects
- Computer processing of large data with tools (like Hadoop)
- Scaling (up or down) compute availability to meet the demand for web resources and applications

3.6.3 OpenStack Object Storage

OpenStack Object Storage (code-named Swift) is open source software for creating redundant, scalable object storage using clusters of standardized servers to store petabytes of accessible data. It is neither a real-time data storage system nor a file system, but instead a long-term storage system for a more long-lasting type of static data that can be updated if needed, recovered and leveraged. This type of storage model can be applied to certain data types, such as virtual machine images, photo storage, email storage and backup archiving. Having no need for a central processing server ("brain") or master point of control this system provides greater scalability, redundancy and permanence.

Objects are written to multiple hardware devices in the data center, with the OpenStack software responsible for ensuring data replication and integrity across the cluster. The storage clusters can be scalable horizontally by the addition of new nodes. In case of a node failure, OpenStack enables required mechanisms to replicate its content from other active nodes. Because OpenStack uses software logic to ensure data replication and distribution across different devices, inexpensive commodity hard drives and servers can be used in detriment of more expensive equipment.

The Table 3.2 describes the main features and benefits of the Object Storage.

Popular Use Cases

- Service providers offering IaaS storage platform
- Integrates with OpenStack Compute to store server images
- Document storage
- Back-end for Microsoft SharePoint
- Archive platform with long retention periods (for log files and other data)
- Store web images and thumbnails

3.6.4 OpenStack Image Service

The OpenStack Image Service, also known as Glance (code-name) provides services for the virtual disk images to be discovered, registered and delivered. The API server of the Image Service offers a standard REST interface for querying information about virtual disk images stored in a range of back-end

Table 3.1: Features and benefits of OpenStack Compute

Feature	**Benefit**
Manage virtualized commodity server resources CPU, memory, disk, and network interfaces	Racks of commodity servers as pools of computing Improved utilization and automation of resources for greater cost efficiencies
Manage Local Area Networks (LAN) Flat, Flat DHCP, VLAN DHCP, IPv6	Programmatically allocate IPs and VLANs VLANs (for rapid provisioning of network capabilities and security features) Flexible networking models to suit needs of each user group and/or application
API with rate limiting and authentication	Designed for automation and security (to make it easy for you to manage who has access to compute resources and prevents users from impacting each other with excessive API utilization)
Distributed and asynchronous architecture	Massively scalable and highly available system (for increased assurance of system uptime)
Virtual Machine (VM) image management	Easily store, import, share, and query images (to make it easy for you to spin up new standardized VMs)
Live VM management (Instance) Run, reboot, suspend, resize, terminate instances	Increase productivity with lifecycle management (from a single user interface and using the APIs)
Floating IP addresses	Keep IPs & DNS correct when managing VMs
Security Groups	Flexibility to assign and control access to VM instances by creating separation between resource pools
Role Based Access Control (RBAC)	Ensure security by user, role and project
Projects & Quotas	Ability to allocate, track and limit resource utilization
VNC Proxy through web browser	Quick and easy CLI administration
Advanced Scheduler	Scheduler decision framework for more efficient mgmt./provisioning
Federated Auth with Zones	Allows to control permissions b/w public and private zones

Table 3.2: Features and benefits of Object Storage

Feature	Benefit
Store and Manage files programmatically via API	Automates resource management/provisioning
Create Public or Private containers	Better control. Allows to share data publicly or keep it private
Leverages Commodity hardware	No lock-in, lower price/GB
HDD/node failure agnostic	Self-healing Reliability, Data redundancy protecting from failures
Unlimited Storage	Huge & flat namespace, highly scalable read/write access Ability to serve content directly from storage system
Multi-dimensional scalability (scale out architecture) Allows to scale vertically and horizontally-Distributed storage	Backup/Archive large amounts of data with linear performance
Account/Container/Object structure no nesting, not a traditional file system	Optimized for scale Allows to scale to multiples Peta-bytes, billions of objects
Built-in Replication (N copies of accounts, container, objects) 3x+ data redundancy compared to 2x on RAID	High Availability
Easily add capacity unlike RAID resize	Elastic data scaling with ease
No central database	Higher performance, No bottlenecks
RAID not required	Allows to handle lots of small, random reads and writes efficiently
Built-in Mgmt. utilities	Acct. Management: Create, add, verify, delete users Container Management: upload, download, verify Monitoring: Capacity, Host, Network, Log trawling, cluster health
Drive auditing	Allows to detect drive failures preempting data corruption
VNC Proxy through web browser	Quick and easy CLI administration

stores, including the OpenStack Object Storage. With the Image Service, clients are able to register new virtual disk images with the, query for information on publicly available disk images and use the Image Service's client library for streaming virtual disk images.

OpenStack Image Service is a multi-format image registry allowing uploads of private and public images in a variety of formats, including:

- Raw
- Machine - AMI (kernel/ramdisk outside of image)
- VHD (Hyper-V)
- VDI (VirtualBox)
- qcow2 (Qemu/KVM)
- VMDK (VMWare)
- OVF (VMWare, others)

The Table 3.3 describes the main features and benefits of the Image Service.

Popular Use Cases

- Service providers providing an IaaS virtual machine image registry
- Enterprises creating a low cost Disaster Recovery/Business Continuity Planning (DR/BCP) platform to back up their virtual server images
- Independent Software Vendors (ISVs) providing pre-configured and optimized images of their software

3.6.5 OpenStack Identity

The OpenStack Identity Service (also known with the code-name Keystone,) offers unified authentication across all OpenStack projects and integrates with existing authentication systems.

3.6.6 OpenStack Dashboard

Dashboard is a graphical user interface (GUI) that allows administrators and users to access and provision cloud-based resources through a self-service portal.

Table 3.3: Features and benefits of Image Service

Feature	Benefit
Image-as-a-service	Store and retrieve virtual machine images at scale
Multi-format/container support	Compatible with all common image formats
Image status	Provides visibility and availability structure
Scalable API	Image Services scales with OpenStack
Metadata	Store and retrieve information about the image
Image Checksum	Ensure data integrity
Extensive Logging	Provide audit and debugging capability
Integrated testing	Verify functionality of the virtual machine
Back-end store options	Greater flexibility with Swift, Local, S3 or HTTP
Version control	Provides structure and control
CLI access	Administrative options
Built-in Mgmt. utilities	Acct. Management: Create, add, verify, delete users Container Management: upload, download, verify Monitoring: Capacity, Host, Network, Log trawling, cluster health
Drive auditing	Allows to detect drive failures preempting data corruption
VNC Proxy through web browser	Quick and easy CLI administration

3.7 XEN Cloud Platform

The open source Xen Cloud Platform (XCP) is an enterprise-ready server virtualization and cloud computing platform that delivers the Xen Hypervisor supporting a range of guest operating systems including Windows and Linux, with network and storage support, management tools in a single, tested installable image, which is also called XCP appliance. [Xen, 2011]

The requirements of cloud providers are addressed by XCP, hosting services and data centers by merging the isolation and multi-tenancy capabilities of the Xen hypervisor with network virtualization technologies and enhanced storage and security offering a rich set of virtual infrastructure cloud services. The user requirements for availability, security, performance and isolation across both public and private clouds are also addressed by the platform.

XCP consolidates server workloads, power savings, cooling, and management costs are consolidated with the use of this system consequently contributing to an increased ability to adapt to ever-changing IT environments, environmentally sustainable computing, an optimized use of existing hardware, and an improved level of IT reliability.

3.7.1 Xen hypervisor

The Xen hypervisor, is a dominant open source industry standard for virtualization. It provides a powerful, efficient, and secure feature set for virtualization of x86, x86_64, IA64, ARM, and other CPU architectures. The Xen hypervisor supports a wide variety of guest operating systems including Windows, Linux, Solaris, and various versions of the BSD operating systems.

Xen muscles many hosting services and most public cloud services, such as Amazon Web Services, Rackspace Hosting and Linode. Oracle VM and XenServer (commercial virtualization products) are built on top of Xen, in addition to desktop virtualization solutions such as XenClient and Qubes OS. From Linux 3.0 forward, all of the code necessary to run Linux as the Xen management OS and as a Xen guest, is part of the Linux kernel.

With Xen virtualization, a thin software layer known as the Xen hypervisor (less than 150,000 lines of code) is interleaved between the operating system and the server's hardware (see Figure 3.9). This provides an abstraction layer that allows each physical server to run one or more virtual servers, effectively decoupling the applications and the operating system from the underlying physical server.

The Xen hypervisor is a distinctive open source technology, developed collaboratively by the Xen community and engineers at over 50 of the most innovative data center solution vendors, including (among others) AMD, Cisco, Dell, Fujistu, HP, IBM, Intel, Novell, Red Hat, Samsung, Sun, Unisys, Veritas, and Citrix. The open-source paravirtualization technology provides a platform for running multiple operating systems in parallel on one physical hardware resource, while providing close to native performance. [Xen, 2011]

- XAPI (also knwon as XenAPI): is an administration stack that controls and configures resource pools and its Xen-enabled hosts and also manages resources within the pool. XAPI exposes the

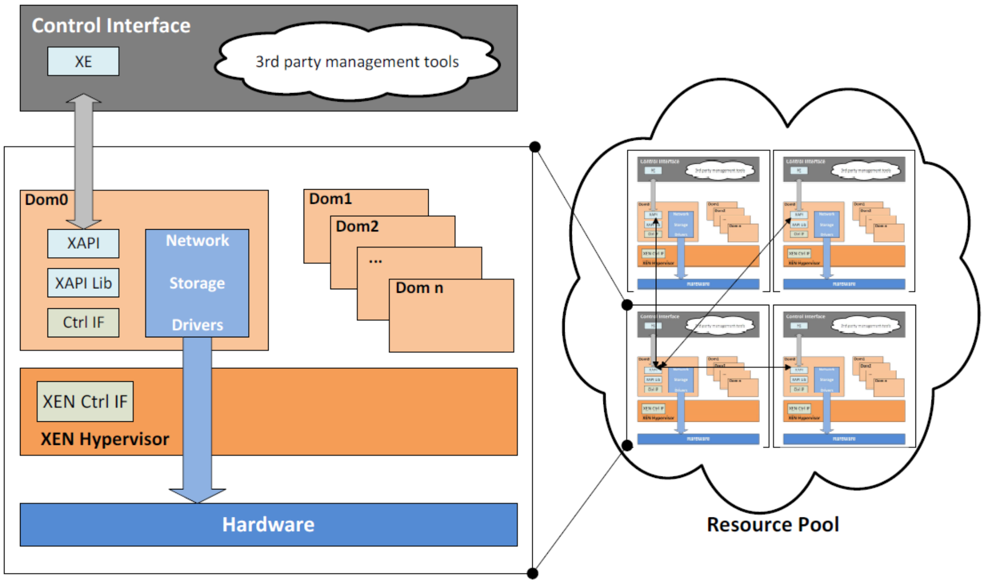

Figure 3.9: Xen Cloud Platform architecture

XenAPI interface (xe).

- XAPI Lib: it's a set of various libraries and utilities.
- Control Interfaces: XCP comprehends the XE command line tool to control resource pools and Xen virtual machines. Various graphical user interfaces that work with XCP are also made available by commercial vendors and open source projects.
- Xen Hypervisor: XCP comprises the Xen Hypervisor.
- Dom0 privileged domain including support for storage, network and drivers: XCP contains the Dom0 Linux kernel including drivers, support for cloud storage infrastructures and rich virtual networking capabilities via Open vSwitch.
- Guest Operating Systems: the XCP binary distribution is distributed with a wide range of Windows and Linux guests.

Chapter 4

Current implementation

After covering some of cloud computing platforms in the previous chapter, in this chapter the implementation IT scenario of the new building housing two of the organizational units of the University of Oporto - Biomedical Science Institute of Abel Salazar and Faculty of Pharmacy and its private cloud architecture is addressed.

The internal organizational structure of these two faculties remain independent, although both of the IT staff agreed to merge efforts into developing a new solution that would fit the new organizational and technological needs, considering the fact that some of the services can obtain greater value if engaged as a common resource.

Private cloud computing, often called a Virtual Private Cloud or Private Data Center, is considered to be total IT system which has been virtualized and hosted within a data center facility.

The new datacenter and its specifications and requirements were studied and designed to support the elasticity needed to deploy successfully a private cloud infrastructure.

4.1 Technology

In this chapter the main technologies used in this project will be addressed, bearing in mind the characteristics of the underlying hardware explained further ahead.

4.1.1 Citrix XEN Server

Citrix XenServer enables businesses to have Windows and Linux virtual machines deployed rapidly and easily manage them and their related assets like storage and networking from a single management console easy to handle. It is the product we use to virtualize our operating systems, whether for office or server environments.

XenServer is a product that Citrix acquired by buying part of XenSource (the leader in virtualization through a technology called Xen hypervisor, open source technology). Therefore the new XenServer

Citrix product that uses virtualization with Xen hypervisor is based on a thin layer of software called a hypervisor is installed directly on the hardware or Baremetal and separates the hardware virtual machines, the innovation of this technology is the use of para-virtualization to achieve high performance and secure environment. Taking advantage of the hardware technologies that exist today as Intel VT and AMD-V.

Virtualization can occur into 2 different forms, the full virtualization and the para-virtualization. KVM, Microsoft® Virtual Server and VMware® ESX Server™ software (among others) are examples of full virtualization. Nevertheless, full virtualization may experience a performance penalty. The VMM (virtual machine monitor) must provide the VM with an image of an entire system, including virtual BIOS, virtual memory space, and virtual devices. The VMM must also create and maintain data structures for the virtual components, such as a shadow memory page table. For every corresponding access by the VMs involves an update to the data structures. In contrast, para-virtualization, used in Xen hypervisor, presents each VM with an abstraction of the hardware that is similar but not identical to the underlying physical hardware. In order to use para-virtualization techniques to the guest operating systems that are running on the VMs must be subject to modifications. As an outcome, the guest operating systems are aware that they are executing on a VM-allowing for near-native performance. [Dell, 2005] However, tests indicate up to the known interval of 5% gain for intensive computations. [Evoy et al., 2011] More recent tests show that for the processing and network access, VMware has superiority over Xen. Xen, in turn, showed superiority in accessing the hard disk, and both had instability in memory. [Grosmann et al., 2011]

Citrix XEN Server 6 also provides a feature (available since version 5 Enterprise) that enables the migration of virtual machines from one physical server to another in the same pool represents a great a benefit in achieving high availability (HA). [Marchi et al., 2010]

4.1.2 Microsoft Virtual Desktop Infrastructure

Microsoft provides a desktop delivery model allowing users to access desktops running in the datacenter entitled Virtual Desktop Infrastructure (VDI). Many organizations are currently exploring whether virtual desktop infrastructure (VDI) is a solution that can help them address some of the desktop management challenges.

VDI Is a desktop delivery model that allow users to access full desktop environments including operating systems, applications, user settings and data witch reside in virtual machines running in centralized servers in the datacenter. This can be achieved by the implementation of an infrastructure as shown in Figure 4.1.

It is important to understand that VDI is just one form of desktop virtualization. As such it can be ideal for some scenarios but may not be a perfect fit for others. Successful scenarios includes deliver secure desktops to unmanaged devices, centralizing desktops for remote, offshore, outsourced or branch offices, helping highly-regulated industries maintain compliance with security requirements and providing enhanced desktop personalization in traditional task workers scenarios that is possible through session virtualization. In use cases such as these, VDI can provide significant benefits such as

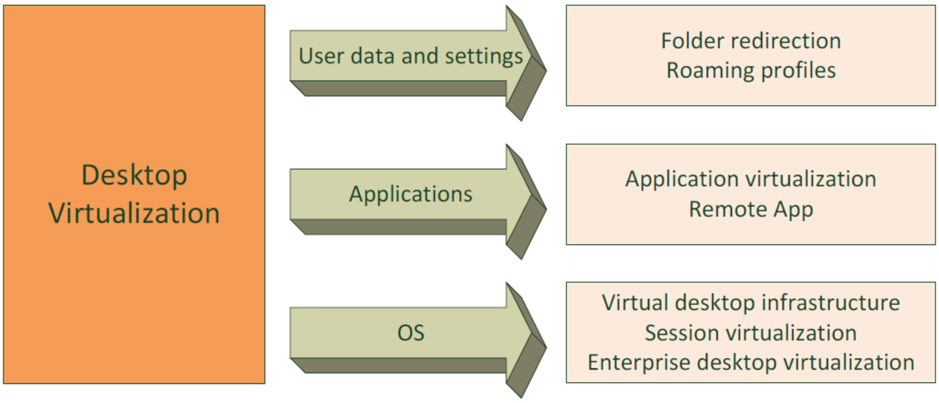

Figure 4.1: Microsoft Virtual Desktop Infrastructure

delivering a secure desktop to virtually any connected device, since desktop, applications and user data all reside in the datacenter.

Centralized management and instant provisioning of desktops to new users, departments or offices and minimize disruption due to hardware failures, through automated desktop backup and recovery. VDI from Microsoft is an integrated cost-effective solution that consists of the underlined virtualization platform, desktop delivery, application delivery and end-to-end management.

Although not used in the current implementation, Hyper-V, the Microsoft virtualization platform, helps optimize hardware resources by consolidating multiple client operation systems on a single server. It offers compelling features in performance, scalability, availability and security. Remote Desktop Services (RDS), provides the complete remote desktop delivery infrastructure for VDI. It includes a basic connection broker suitable for lower complexity, lower scale deployments witch acts like a desktop delivery service that enables end users to connect to virtualized desktops. For larger scale environments, customers can leverage third party partner's solutions such as Citrix XenDesktop.

Even within a hosted virtualized desktop environment, managing the delivery of applications is critical to success. Microsoft offers two complementary technologies that can help address this issue. Application Virtualization (AppV) decouples the Application layer from the operating system and enables them to run as network services. Removing the application footprint from the operating system helps reduce application conflicts, bring patch and upgrade management to a central location and accelerates the deployment of new applications and updates. Remote App provides users access to applications running on a separate RDS host server. It suppresses the desktop and presents only the user interface of the applications on to a local or desktop image.

To tie it all together, Microsoft System Center is a comprehensive integrated management suite that allows IT organizations to manage the physical and virtual desktop as well as servers from a single familiar management console. This can significantly reduce the cost and complexity of management multiple disparity environments, especially for those who invested in System Center as part of their IT infrastructure. The VDI solution from Microsoft is available through purchase of one of the two Microsoft

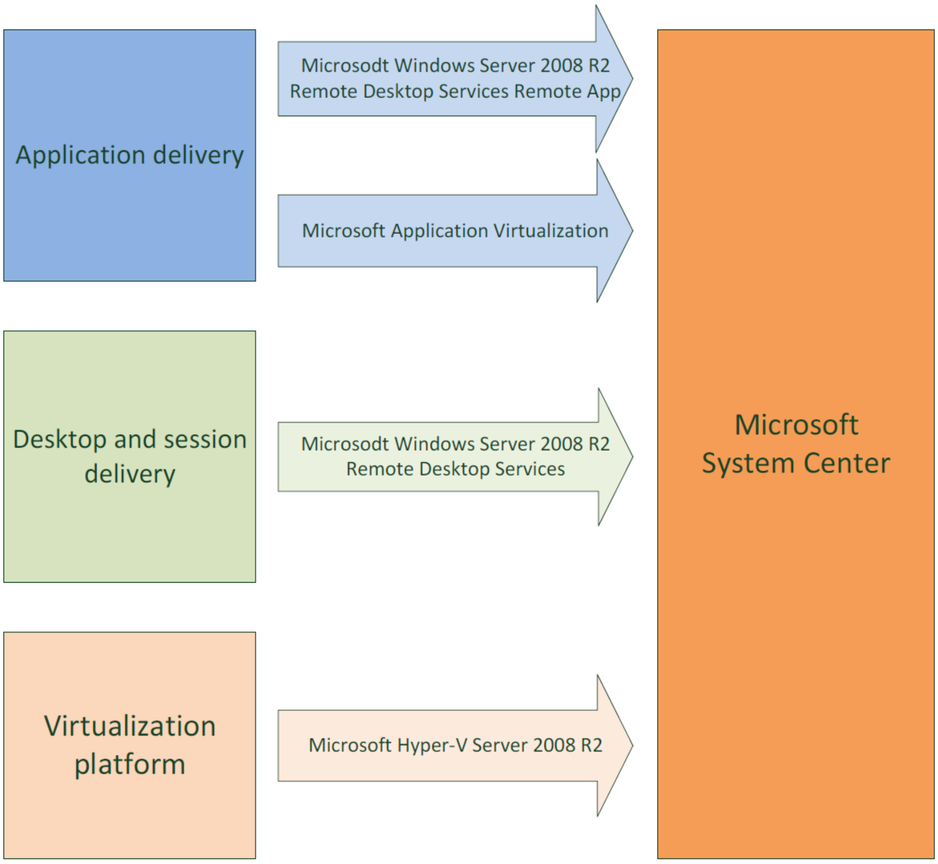

Figure 4.2: Microsoft VDI Architecture

VDI Suites which includes the virtualization platform, desktop delivery, application delivery and core management technology (see Figure 4.2).

4.1.3 Remote Desktop Services

In the new building was decided to use a solution of remote desktops to students and staff and departments/laboratories, using the Microsoft Remote Desktop Server (RDS). The use of shared desktop environments based on Microsoft, for the specific case of the Faculties, takes advantage of the following advantages:

- Benefit from the use of the MS Campus Agreement and License CAL existing agreement. These licenses are in number equal to the FTE's (Full-time equivalent) contemplating both server and client licenses, allowing customers to use them, based on any OS including Linux;
- Centralization of the investment, management and migration of jobs in IT's customers for the Datacenter;

- Gain of energy savings resulting from the use of the barebones.

The Microsoft RDS service comprises on the Microsoft Server 2008 R2 version, shared work environments (RD Session Host - formerly Terminal Services) and VDI's (RD Virtualization Host). Both the RD Session Host as the RD Virtualization Host are roles that are added to servers, according to the service you want to provide. In the case of the RD Virtualization Host, is required to install Microsoft Hyper-V, it is not possible to use another hypervisor to manage the virtual machines. RDS in addition to these two roles still have the RD Gateway, RD Licensing, RD Web Access and RD Broker.

- RD Connection Broker is responsible for controlling connections to the RD Session Host servers and RD Virtualization Host. It provides load balancing and allows the recovery of an open session. To maintain this role in a structure of high availability, it is necessary to use a cluster service (Failover Cluster).
- RD Session Host: Server for sharing desktop environments.
- RD Virtualization Host: Integrated with Microsoft Hyper-V allows provisioning virtual machines to users.
- RD Web Access: Allows the connection to remote environments and applications through a web browser (IE).
- RD Gateway: Enables remote users to access applications installed on the server's session hosts by tunneling RDP in HTTPS communications, not involving the use of VPN's.
- RD Licensing: It manages the licensing of RDS. There is a way of making this role highly available. How to contour can be activated two license servers and licenses are divided between them.

Clients

The jobs to be used on clients are one of the most pertinent issues that arise in this type of solution: thin clients, fat clients or barebones. Then there is the matter of using Windows or Linux OS. Of the tests we performed we conclude the following: The thin client manufacturers like IGEL, Dell, HP, devonIT, Wyse, have the advantage of having a centralized management generally a software provided by each manufacturer. They also have very low power consumption, typically between 7 and 11 W. The disadvantage of such solutions is the purchase price which exceeds the price of a fat client and using a proprietary firmware which can be discontinued, as well as management systems for independent manufacturer. The fat clients allow reusing the existing hardware and enable the installation the OS if necessary. The energy costs are very high, since a standard PC has power supplies typically from 350W and above. The barebones computers or the eeepc like computers position themselves between the previous solutions. Have an acquisition cost in the house of 150-250 euros and power between 30 and 40W, enabling to install any OS.

For the operating system the use of Windows or Linux was weighted. Linux solutions (open thin client or thin station) have the price advantage, but in the other hand, they lack customization for each hardware and do not have all the features of RDP 7.1, including the ability to use RemoteFX or USB redirection. The use of Windows, including Windows Thin PC version, specific to computers with low resources, takes advantage of all features of RDP allowing you to use old or barebones PC's with less advanced processors (atom like). Furthermore, it allows to be managed by the domain controller in order to allow centralized management of client workstations. Thus, we choose the solution at this stage of relocation has been to reuse existing hardware, and acquire barebones where possible, using Windows thin PC. The idea is to use RDP in classrooms, for administrative, for the library, lecture theaters and gradually introduce the solution in the teachers whenever possible or to move these to a VDI solution.

With this approach, the idea is to withdraw investment of workstations and redirect to the datacenter, reduce energy consumption by adopting barebones, allowing a more centralized management, increase mobility and increase the productivity of IT services, freeing their elements of the work on local posts, since all data is centralized.

RDS Farms

The implemented solution is based on a simplified scheme that comprises the Remote Desktop Connection Broker (RDCB) with RD Licensing to manage the connections to Remote Desktop Session Hosts (RDSH) and a pool of session RDSH with the Desktop Experience feature enabled to have the look and feel of windows 7 as shown in Figure 4.3.

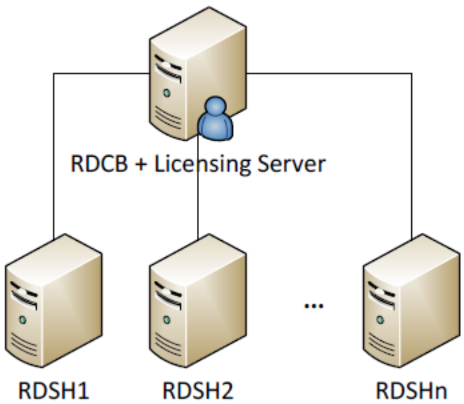

Figure 4.3: Remote Desktop Services

It was decided not to install the RD Web Access role, since some of the functionalities are limited to IE. Furthermore, considering the integration of the RDS in the checkpoint, and it is intended to promote access via VPN, the installation of this role was avoided.

To check the logs and the broker services to verify that brokers are active and operational. This is possible through the Server Manager's Diagnostics Event Viewer and by consulting the Applications and Services Logs of the Microsoft Windows Terminal Services Session Broker.

Defining a number of farm members to deploy depends directly of the use given of the group that uses the farm and its resources. In the solution explored, three major groups were identified that use the same resources: Staff, Departments and Students. Each of these groups was assigned a deployed server farm in with specific software and resources. The farm is elastic, so it can increase or decrease in number of members depending of the resources used.

To add a new Remote Desktop Session Host (RDSH), it is required to add the Remote Desktop Services Role in the Server Manager snap-in. It is worth mentioning that the type of authentication used was the less secure (allowing connections from computers running any version of Remote Desktop) The option Require Network Level Authentication is more secure, since it prevents DoS attacks. In this type of authentication is not necessary to open a full RDS session to validate users. However this option is only implemented from RDP 6.1 that came with Windows Vista. In Windows XP, SP3 if any, you can also enable this feature, but it is necessary to tweak the registry to make this possible.

In this sense, since it is intended that students can access from different devices and OS to the RDS, it was decided to allow any version of RD client to access. The issue of DoS is partially resolved, since access to this service will only be on the internal network or VPN, is not available directly outside.

Users and groups that are eligible to access RDS can be managed in the Local Users and Groups in the Computer Management snap-in. For the RDS clients to profit from the Windows 7 ambience, the client experience configuration must be activated, so that the design that appears to remote users is the Windows 7. This feature can also be activated at a later stage by adding the Desktop Experience feature in Server Manager.

The RemoteFX was also activated, and it is divided into two parts: the activation server session hosts and activating in the clients workstations (to enable redirection for RemoteFX USB). RemoteFX is a new technology included in Windows Server 2008 SP1 R2/Windows 7 designed to improve the visual capabilities of Windows 7 Remote Desktop clients connected through the Hyper-V role to a Windows 7 virtual machine. The improvements achieved with the use of the RemoteFX include many 3D graphics capabilities, full motion video, some OpenGL, and support for Hyper-V's Windows 7 virtual machines USB redirection. To enable this feature, it is necessary to use gpedit.ms to navigate to Computer Configuration - Administrative Templates - Windows Components - Remote Desktop Services - Remote Desktop Session Host - Remote Session Environment and enable RemoteFX Configuration.

The RDP supports two types of redirection USB: USB high level redirection and RemoteFX USB redirection. RemoteFX USB redirection permits the use of many types of USB devices to be used in the remote VDI session. In addition to the existing high-level device redirection common mechanisms in RDP (printers, USB drives, smart cards, PnP devices, audio, etc.), RemoteFX USB redirection enables the use of devices such as multifunction printers, scanners, webcams, and more to be used on the virtual machine.

To activate of using RemoteFX USB redirection on the clients, it is necessary to enable the client workstation by changing the Group Policy of the Remote Desktop Services of the Windows Components in the Administrative Templates within the Computer Configuration. In the Remote Desktop Connection Client and by selecting Remote USB Device Redirection it is possible to configure the users that are capable to redirect devices.

There are different device support status Redirection methods available:

- RemoteFX USB Redirection for all-in-one printer support
- Supported Printer Easy Print
- RemoteFX USB redirection Supported Scanner
- Biometric Supported while in a session
- Not supported RemoteFX USB Redirection During logon
- Play Device Redirection and PTP Camera Supported Plug
- Plug and Play Device Redirection and MTP Media Player Supported
- RemoteFX USB Redirection for Webcam Support (LAN only)
- RemoteFX USB Redirection for VoIP Telephone / Headset Support (LAN only)
- Audio Redirection for Audio (not a USB composite device) Support
- Drive Redirection for CD or DVD drive Support for read operations
- Drive Redirection for Hard Drive or USB Flash Drive Support
- Smart Card Redirection for Smart Card Reader Support
- RemoteFX USB Redirection for USB-to-Serial Support
- Blocked USB Network adapter (also includes some personal digital assistants) (N/A)
- Blocked USB Display (N/A)
- Input Redirection for USB Keyboard or Mouse Support

For the use of smartcards using the citizen card using RDP, you must install on the remote server Session Hosts drivers card reader and software citizen card. We experienced problems with the installation of software on the citizen card in 64-bit version 1.24. so a downgrade was inevitable and the version 1.23 was used.

The installation of software in the RDS must be made in installation mode, it is necessary to place the RDS this state, so that the application is properly installed:

- change user / install -> Place the RDS in installation mode
- change user / execute -> Place the RDS in execution mode
- change user / query -> Knowing which way the RDS is

RDS Broker

After installing the Remote Desktop Connection Broker (RDCB) role, selected from the Remote Desktop Services, it is required to specify the RD Session Host Servers that constitute the farm to the Session Broker Computers group of the RDCB. To accomplish that task, in the Configuration option in the Server Manager snap-in, in the Groups within the Local Users and Groups, it is possible to add new members by selection the Session Broker Computers in the center pane.

In each RD Session Host server it is mandatory the configuration of the Broker that it will be mediated by. To configure this it is necessary to edit the settings area, under RD Connection Broker by double-clicking Member of farm in RD Connection Broker in the Remote Desktop Session Host Configuration.

To determine which servers are members in the same RD Session Host server farm, the RD Connection Broker uses a farm name. It is required the use the same farm name for all the servers that are in the same that load-balanced RD Session Host server farm. If a new farm name is typed, a new farm is created in RD Connection Broker and the server is joined to the farm. If an existing farm name is typed, the server joins the existing farm in RD Connection Broker. Although the name farm in RD Connection Broker does not have to be registered in Active Directory Domain Services, it is recommend the use of the same name that is used in the Domain Name System (DNS) for the RD Session Host server farm. The RD Session Host server farm name in DNS name that represents the virtual clients will be used to connect to the RD Session Host server farm.

Still in the Remote Desktop Session Host Configuration, the selection of the Participate in Connection Broker Load-Balancing check box is active. Optionally, it is possible to change the default value of each of the farm member's server's Relative weight (that is 100). The server weight is relative. Therefore, if you assign one server a value of 50, one and a value of 100, the server with a weight of 50 will receive half the number of sessions. IP address redirection is enabled (as configured by default).

RDS Licensing

We chose to install the licensing service along with the broker, since both are light weight. To install the Licensing manager, the Remote Desktop Licensing role must be added in the Server Manager. To enable Server Licenses it is required to activate the server in the Remote Desktop Licensing Manager, select automatic connection (recommended) and fill out the licensing information along the contract number provided by the deanery.

The final step is to place the license server in Remote Desktop Server Hosts, for this is necessary to configure the RD Session Host servers indicating which server handles the licensing. This is accomplished by editing the settings area, under Licensing, in the Remote Desktop Session Host Configuration snap-in.

4.1.4 Microsoft Desktop Optimization Pack

Microsoft provides a suite of technologies accessible as a subscription for Software Assurance customers designated Microsoft Desktop Optimization Pack (MDOP). This package provides tools to: manage and restore.

Manage:

- MBAM: Microsoft BitLocker Administration and Monitoring
- AGPM: Advanced Group Policy Management

Virtualize:

- MED-V: Microsoft Enterprise Desktop Virtualization
- App-V: Application Virtualization
- UE-V: User Experience Virtualization

Restore:

- DaRT: Diagnostics and Recovery Toolset

Application Virtualization

Microsoft Application Virtualization (App-V) transforms applications into centrally managed services that are never installed and don't conflict with other applications.

This is the technology that Microsoft acquired a few years and was called SoftGrid. The concept of this type of technology is very interesting, but despite taking the name are quite different virtualization models that met in Hyper-V and MED-V. This product is composed by the App-V Sequencer and the App-V Manager.

The version in use is the Microsoft Application Virtualization 4.6 (Service Pack 1). App-V is very easy and fast to use, sequence and deploy applications. IT staff have at their disposal several utilities and documentation such as the App-V Package Accelerators, Sequencer Templates, Packaging Diagnostics and built-in best practices. The Virtual Desktop Infrastructure (VDI) and Remote Desktop Services (RDS) environment is therefore optimized and disk storage is better availed.

This method is extremely flexible and transparent to the end users, applications are available simply by clicking to launch them, anytime. Applications are also available to users transversely in the domain, a user can logon to a new PC (even with another version of Windows) and the applications are also available. There is no need to wait for applications to install and reboot to start using them. The updates are also simplified to deploy and update the packages.

Traditional applications can be transformed with App-V into virtual and centrally managed applications. The sequencer for App-V makes it simple to create virtual applications. It is possible to create standard MSI-based packages that are ready to deploy with existing infrastructure (either by group policy or scripting), or take advantage of built-in deployment and management.

To deploy a virtualized application, it is necessary to "sequence" an application and create a package. This process involves the monitoring and recording (in the file system, registry keys) of everything that happens when an application is installed. This package file is then configured, to enable deployment in different OS versions and architectures and file types associated, icons used and security settings.

IT can use the built-in management in App-V to stream applications to users on demand and allow users to update applications without disconnecting or rebooting. It is also possible to assign applications to users (or security groups) using Active Directory, and retire applications quickly by removing the assignment (without the need to uninstall). Tools to keep track of application licensing and meter usage are also available.

System Center Configuration Manager (SCCM) can also integrate with App-V to enable central management of virtual and physical applications, sideways with hardware and software inventory, operating system and patch deployment, and more.

App-V is a proven technology with real business results and considered to be a mature and verified application virtualization solution. This solution is integrated with the RDS and all the other domain computers enabling a centralized software management.

4.1.5 Microsoft Deployment Server

Windows Deployment Services is the updated and redesigned version of Remote Installation Services (RIS), it is a role (function) of Windows Server, which is available from the Windows Server 2003 SP2 (or SP1 after installing an update). Basically, WDS distributes images via network to the target computers, including:

- Boot image (WinPE), if your computer does not have an operating system installed
- Installation image. For this, we have 2 options: Windows installation (default setup) or an image already installed containing, for example, applications, settings, updates, etc.
- Image capture. Basically it is a modified WinPE, which contains instructions for making custom image capture, which has gone through Sysprep.

As prerequisites for use with this role, it is necessary to have:

- Active Directory: the WDS server must be part of an Active Directory domain or have that role.
- DHCP: need a DHCP server (or have that role) with an active scope on the network because Windows Deployment Services uses PXE, which in turn depends on the IP address by DHCP performed.
- DNS: a functional DNS server on the network is required for name resolution.
- NTFS Partition: in which to store the images (image store).
- Credentials: Membership in "Local Administrators" on the WDS server. To install an image, you need to be part of the group "Domain Users".

- Windows Server 2003 SP2 or later

It is possible to use only WDS to create and distribute images, and still have a certain degree of automation using Unattend.xml answer file, for example. However, to obtain better flexibility and management Microsoft recommends using it in conjunction with Microsoft Deployment Toolkit (MDT) and/or System Center Configuration Manager (SCCM) to allow greater ease in automating the process, from image creation, capture, distribution and installation. Thus, we have:

- MDT to streamline the creation of standard image, including installation, configuration and capture
- The WDS to distribute the boot image (WinPE), and image installation (Windows 7, for example)

In short, the MDT helps in creating and capturing an image, and WDS serves to distribute the image via network to the target computers.

4.2 Architecture

To address the common areas requirement, with multiple resources that can be used by both users (from ICBAS and FF) and to manage the identities independently from the separate organizations, a Windows Active Directory parent and children domains organization was implemented.

Creating the child domains in the UP forest tree ensures a two-way, transitive parent and child trust establishment. Although trust relationship is not defined between ICBAS and FF domains, as both trusts UP domain, they also trust each other as illustrated in Figure 4.4.

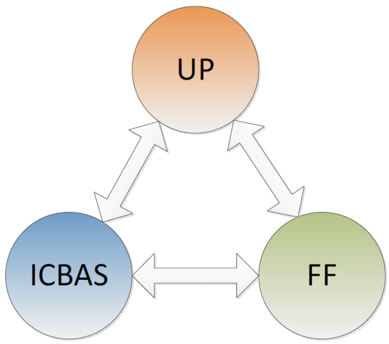

Figure 4.4: Domain trust relationships

Microsoft suggests a private cloud architecture implementation based in Microsoft Hyper-V and Microsoft System Center to enable IaaS [Microsoft, 2010]. As we have a successful experience with XEN hypervisor since the ICBAS datacenter update, based on Pinhal's project within the discipline of Project / Internship at ISEP-IPP [Pinhal, 2008], the adopted XEN hypervisor proved itself to be reliable, fast, robust and cloud proven. Therefore a hybrid solution was encountered to make the transition as smooth as possible using the XEN virtualization technology combined with Microsoft's approach to VDI (see Figure 4.5).

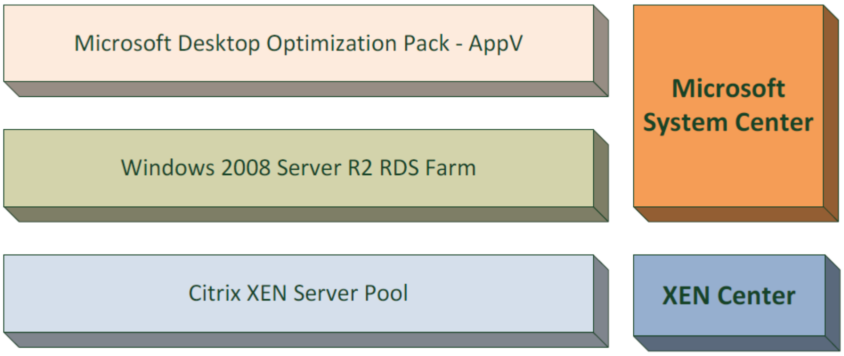

Figure 4.5: ICBAS|FF-UP Private Cloud Architecture

4.3 Hardware

The choice of XEN Hypervisor using Citrix XEN Server, requires certification and compatibility of the equipment used for this platform. This influenced the choice of the hardware.

4.3.1 Storage

The EMC®VNX®series unified storage systems deliver uncompromising scalability and flexibility for the mid-tier while providing simplicity and efficiency to minimize total cost of ownership. [EMC, 2012]

Based on the Intel Xeon-5600 processors, the EMC VNX implements a modular architecture that integrates hardware components for block, file and object with concurrent support for native NAS, iSCSI, Fibre Channel and FCoE protocols. This series can provide file (NAS) functionality via two-to-eight X-blade data movers and block (FCoE and FC, iSCSI) storage via dual storage processors leveraging full 6Gb SAS disk drive topology. The unified configuration includes various rack-mounted enclosures such as:

- Disk processor enclosure (holds multiple disk drives) or storage processor enclosure (requires a disk drive tray) plus a standby power system to deliver block protocols;
- One or more data mover enclosure to deliver file protocols (required for file and unified configurations);
- A control station (required for file and unified configurations);

The VNX5300 meets all the requirements for the solution adopted, allowing connectivity to 10Gb which is intended for data transfer and it provides a simple interface to configure and manage the storage. VNX / VNXe Capacity Calculator provides a comprehensive way to build unified SAN and NAS, in the disk configurations (see Figure 6.2) allowing a better perspective of the possible configurations.

Table 4.1: Usable space in GB

	Fast (Demand)		RAID5	
EFD	5%	600	6%	700
SAS	30%	3600	35%	4200
NL-SAS	65%	7800	59%	7100
Total	100%	12000	100%	1200

Management

EMC Unisphere facilitates VNX systems management from anywhere with a simple user interface and integrated distributed storage environments. The Unisphere control panel consists in a single screen, providing management and reports with speed and simplicity, which enables administrators to obtain instant knowledge about what happens in any environment.

Capacity

Integrated features help ensure that redundant data, inactive or anticipated not consume precious storage resources. The compression block, designed to LUNs relatively inactive (such as backup copies, and static data repositories) automatically compresses data, enabling recovery and reduce capacity by 50% of the physical data. The compression / file-level deduplication reduces the disk space used by 50% and deduplicating selectively compressing inactive files. Because these features operate as background tasks, the overhead on system performance is minimal.

In this solution a multi-tier disk approach was adopted in the storage configuration in order to achieve the best performance results as indicated in Table 4.1.

Physical Specifications

The Table 4.2 shows the VNX physical specifications, whereas file components are described in Table 4.3 and Table 4.4 describe other characteristics.

Connectivity

Flexible connectivity options are available to the VNC series via the UltraFlex IO modules for both block storage processors for FC and iSCSI host connectivity and the file X-blades for NAS connectivity. Each storage processor connects to one side of each of two or four redundant pairs of four-lane x 6 Gb/s SAS buses, providing continuous drive access to hosts in the event of a bus or storage processor fault. VNX models require a minimum of four “vault” drives (SAS our Near-line SAS) and support a maximum of 1000 disk drives in up to 67 disk expansion chassis. Each vault drive consumes approximately 200 GB for the VNX operating environment software and data structures.

Table 4.2: VNX5300 Physical Specifications

Min/Max Drives	4/125
Array Enclosure	3U Disk Processor Enclosure (Holds 15x3.5” or 25x2.5” SAS/Flash drives)
Drive Enclosure Options (DAE)	25x2.5” SAS/Flash drives - 2 U 15x3.5” SAS/Flash drives - 3 U
Standby Power System	1 U 1.2KW
RAID Options	0/1/10/3/5/6
CPU/Memory per Array	Intel Xeon 5600 / 16GB
Max Block UltraFlex™IO Modules per Array	4
Embedded IO Ports per Array	8 FC ports and 4 SAS ports (2 BE SAS buses)
Max Total Ports per Array	24
2/4/8 GB//s FC Max Ports per Array	16
1 GbaseT iSCSI Max Total Ports per Array	8
10 GbE iSCSI Max Total Ports per Array	8
Max FCoE Total Ports per Array	8
6 Gb/s SAS Buses (4 Lanes per Bus) for DAE Connections	2

Table 4.3: VNX5300 File Components

File X-Blades	1-2
Control Stations	1-2 x 1U Server
X-Blade: CPU/Memory	Intel Xeon 5600 / 6Gb
Max File UltraFlex™IO Modules per X-Blade	3
Min/Max 2/4/8 Gb/s FC Ports per X-Blade	4
Max IP Ports per X-Blade	8
Max 10 GbE Ports per X-Blade	4

Table 4.4: VNX5300 Other Characteristics

Management	LAN 2x 10/100/1000 Copper GbE
Max Raw Capacity	360 TB
Max SAN Hosts	2.048
Max Number of Pools	20
Max Number of LUNs	2.048
Max LUN Size	16 TB (Virtual Pool LUN)
Max File System Size	16 TB
Maximum Usable File Capacity per X-Blade	256 TB
OS Support	Block OS's Plus File OS's see E-Lab Navigator and NAS Support Matrix on Powerlink

4.3.2 Servers

Ideal for virtualization, consolidation of work load and databases of medium size, the Dell PowerEdge R810 server is a 2U high rack density and with two or four sockets, which combines advanced management capabilities with features with the best price performance. [Dell, 2012]

- More processing power and memory capacity to maximize the load of the work;
- Integrated management systems. Lifecycle Controller and built-in diagnostics to help maximize uptime;
- Technology FlexMem Bridge - more memory per processor solves the growing needs of demand for memory.

4.4 Network

In a higher logical abstraction level the ICBAS|FF-UP network architecture is be represented in Figure 4.6. The network implementation is designed to improve the IO data bound with a 10Gbps switch to maximize the data transfers rates between the storage and the XEN pool servers.

Traffic to and from the server's pool and storage from the switches stacks are directed compulsorily to the checkpoint cluster for filtering and enforcement of management policies and ACLs. Although the core and backbone switches support routes, only a few were defined, thus delegating the control to the checkpoint cluster.

Redundant paths are shown in Figure 4.7 with the checkpoint cluster in the middle. The checkpoint cluster has a virtual IP address that is made available by a mechanism similar to Linux Virtual Server (LVS) where the members are checked by a hearbeat and implemented in an active-passive mode. A network interface connecting to the WAN is available in each of the cluster members.

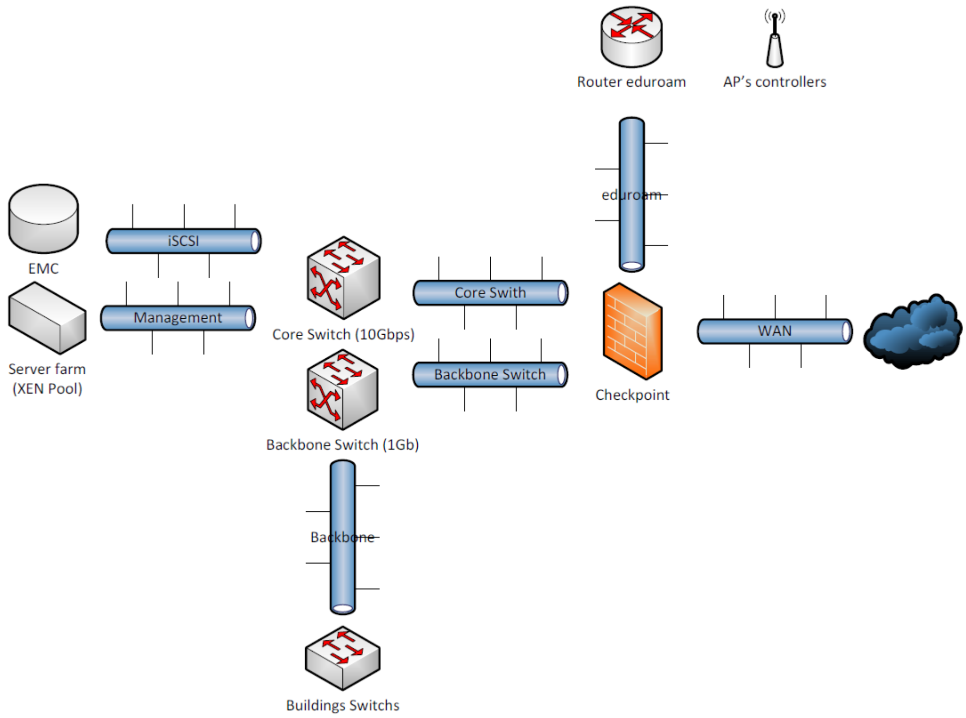

Figure 4.6: ICBAS|FF-UP Network Architecture

The Alcatel stack has two switch carts of 1Gb represented as the backbone switch stack that connects to each of the buildings stacks by a fiber link aggregation (LAG), and the core Dell stack is composed by two 10Gb switches. As the EMC storage has two disks array controllers, each one provide two network interfaces making possible to connect to the storage by any of these. It would be possible to aggregate these links maximizing the available bandwidth, but regardless of the IO bound software, the bottleneck is most likely to be the backbone switch.

Both the Xen Server pool servers and the EMC storage are connected to both of the core switches by 4 network interfaces transporting exclusively iSCSI traffic enabling multipath between the pool and the storage. A different network interface is also connected to provide access for management of the resources. In this schema, as the Alcatel stack is made of a blade chassis, it represents the only single point of failure (SPOF) in case of hardware malfunction.

Although the building switches stacks are only of 100Mbps, contrary to what we'd been expecting, these are sufficient to implement the technological solution. However, we would expect nowadays to have available Gb switches that surely would offer a smoother user experience, particularly in the RDS and in the delivery of virtualized applications.

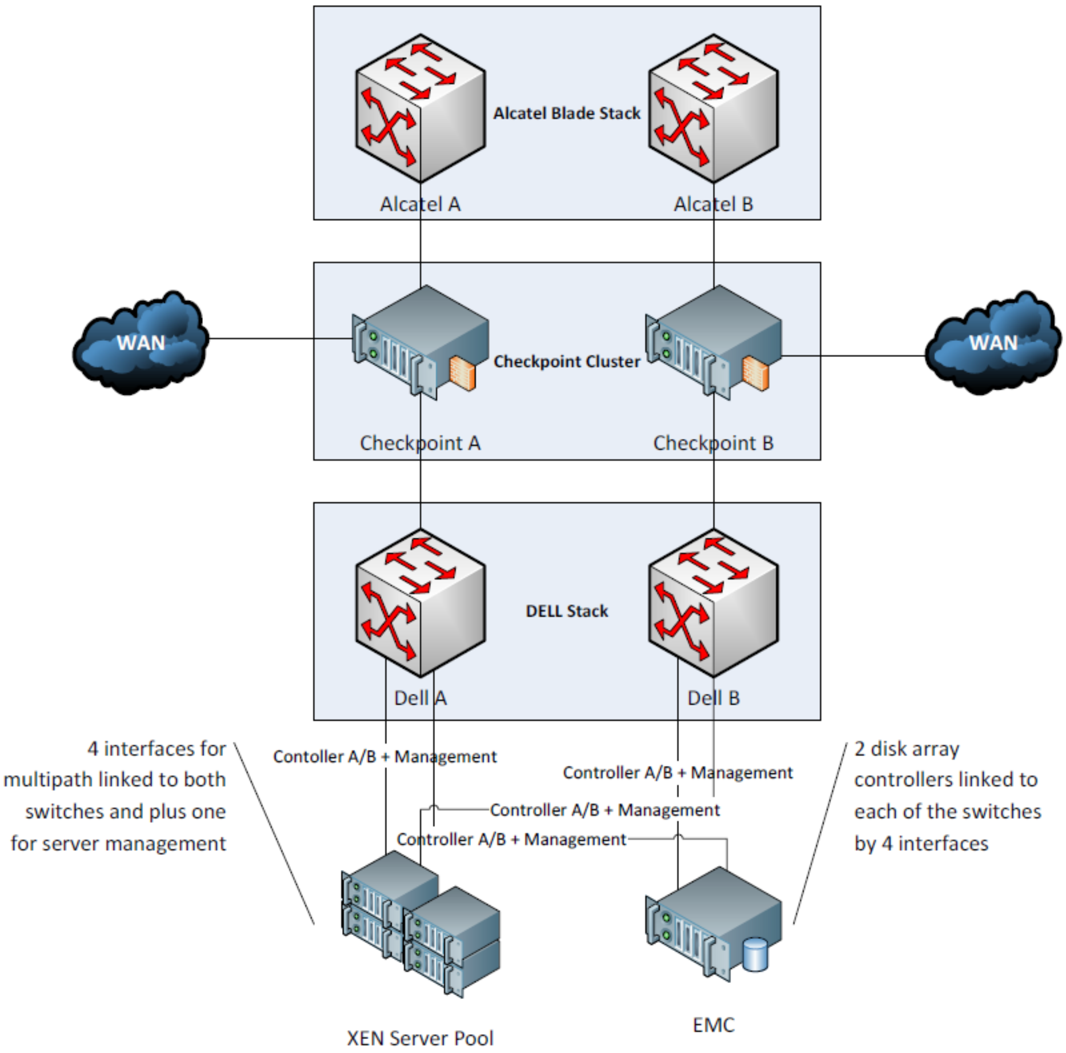

Figure 4.7: Checkpoint cluster

4.5 Organization

The structure engineered is to have 3 separate domains, a top domain up.intranet and two child domains icbas and ff (representing each of the organizational unit). This approach ensures not only the trust relationships between each other, but also the user and computer policy enforcement needed to this organization.

As can be understood from the RDS farms, they belong to the top level domain (UP) for having shared resources that can be accessed by any member of the organization ICBAS|FF. It is possible in case of significant growth to be scaled as necessary. A template was created of each kind of server

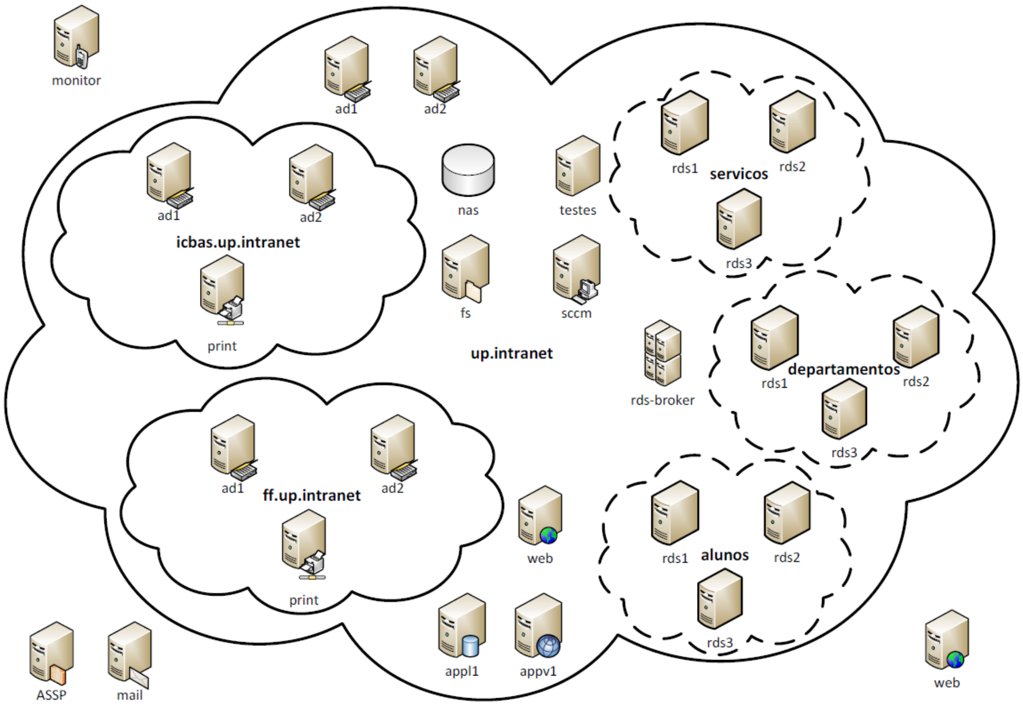

Figure 4.8: ICBAS|FF-UP Servers Architecture

(with appropriate roles and including all necessary software to deliver virtualized applications) to enable a simple escalation. Whenever needed to add another server to a particular farm, after the virtual machine created, it is added to the domain and the respective farm (broker).

The broker redirects requests for each of the farms depending on the desired connection RDS (students, services or departments), and access to each of the farms is restricted with security groups, thus limiting unauthorized access to the farms.

From the Figure 4.8 it is also possible to infer the existence of other servers (not all being represented) in each of the child domains, and others outside the top level domain that represents the DMZ.

Besides the web servers (only represented one of them) and the email along the Anti-Spam-SMTP-Proxy (ASSP) the monitor is a Nagios server enforcing surveillance of the servers (either physical or virtual) along some of the critical services and other devices.

Chapter 5

Proposed evolution

In the previous chapters several cloud technologies were studied, either open and paid, public and private, providing IaaS, PaaS or SaaS. A wide view of the cloud world can be glanced and its potentials can be evaluated. A case study of a private cloud scenario deployment and its available resources, used technologies and requirements were discussed. In this chapter some technologies to deploy and integrate the case study scenario are studied in order to fulfill some of the IT requirements and to enable systems agility.

In this chapter it's implemented an OpenStack architecture in a virtualized server within a contained test environment, from the simpler form to the more complex deployment. It is further addressed an alternative technology to the RDS and App-V, the Ulteo to deliver remote desktops and applications. As for the virtualization platform, is also studied Microsoft Hyper-V for further integration with the existing systems.

Architecting private cloud services requires leadership, vision and a focus on a wide-ranging array of changes, as well as a living strategic plan to make the private cloud initiative a success. [Bittman, 2011a]

Private cloud computing holds tremendous promise to the improvement of the speed of delivery and the flexibility of an infrastructure, sourcing for the right workloads. Not all services require self-service and rapid provisioning, and not all workloads require the ability to migrate or flex into an external cloud domain. The private cloud might simply become a more automated, more standardized infrastructure delivery mechanism therefore making most of advertised features in the private cloud vision might be overkill. It's important to have your own vision for how the enterprise wants to deliver infrastructure in the future. [Bittman, 2010c]

The test of these technologies provides a great opportunity to leverage private cloud infrastructure. The ability to dynamically provision a set of resources with a specific infrastructure configuration enables the sharing of hardware resources, the testing of multiple configurations and the ability to see the effects of change on scalability. [Murphy, 2010]

The hardware underneath the test implementation scenario is a SuperServer 6015 series in a high-density 1U form-factor. This server supports the maximum frequency quad-core Xeon® 5400/5300 and dual-core Xeon® 5200/5100 Series processors.

This server had a specific role that required a large amount of processing capability. It mostly executed calculations with a chemistry algorithm.

With the following specifications:

- 2 x Intel(R) Xeon(R) CPU E5410 @ 2.33GHz
- 8 GB Memory
- 2 x 1Gb NIC's
- 2 x 160Gb SATA HDD (RAID)

In an early stage of testing some platforms, XEN Cloud Platform 1.0 was installed in this server, and as it is viable to manage with XenCenter through its API. This hypervisor version was used to test other technologies (see Figure 6.3).

5.1 StackOps

StackOps is a private cloud platform that is based on the Open Stack. It comes with a ready to install ISO for deployment on the machines that will compose the cloud.

There are several organizations that support OpenStack, including Rackspace, NASA, Dell, Citrix, Cisco, Canonical and 50 others, having evolved into a worldwide software community. OpenStack is a set of open source technologies to deliver elastic cloud operating systems.

It may be assumed that the distribution of OpenStack is the simplified Linux cloud operating system. It aspires to be a reference distribution to IT professionals and departments due to the combination of the best open source tools combined with a simple deployment interface, thus minimizing costs for IT professionals and departments and obtaining higher profits for providers that wish to sell Public Cloud services.

5.1.1 Architecture

The implementation of a system based on StackOps could be addressed using different architectures, each serving its purpose. It is even possible to have a solution tailored to meet specific needs.

Single Node

For testing and assessment purposes, all of OpenStack's components can be installed in just one node (server) as represented in Figure 5.1. It is also possible to install the distribution on a virtual machine using QEMU emulator to the detriment of the KVM hypervisor, thereby achieving an ideal environment for an assessment system. Important to note that this would not be the production environment since it would not take advantage of the KVM hypervisor, only being feasible to a first approach to StackOps. This method of implementation has only a requirement of one network interface.

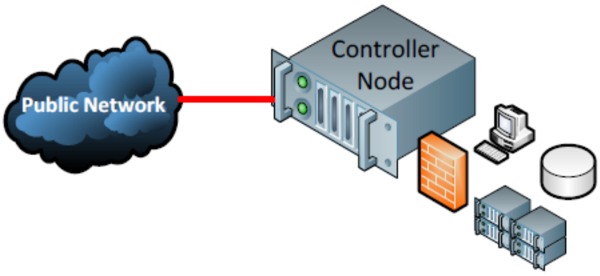

Figure 5.1: StackOps single node architecture

Dual Node

It is recommended for a minimal installation, in a production environment, a system with two different nodes (servers). The Compute node is installed in one server and the Controller node that contains all of the other OpenStack's components is installed on the other as illustrated in Figure 5.2. In this scenario it is possible that both nodes can be executed in a virtualized environment, however, it is recommended that at least the Compute node is installed on the bare metal. It is also possible, and whenever necessary, to add additional Compute nodes. This method of implementation requires at least two network interface in each server in order to successfully run this configuration and the server's hardware must meet the minimum hardware requirements.

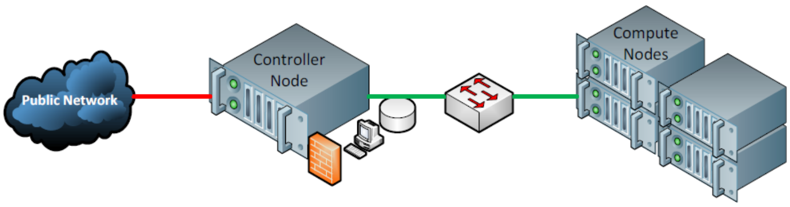

Figure 5.2: StackOps dual node architecture

In this architecture a single server is has the roles of Controller, Network and Volume, and N other servers have the Compute roles. Two different networks exist between the control node and the server nodes:

- **Management Network** is used for the communication of Nova components among them.
- **Service Network** is used for the communication of Network node with the net of guest virtual machines.
- **Public Network** (Optional) if used for the communication of internet/intranet with Network node.

Multi Node

In a production environment the architecture to adopt would be to multiple nodes where OpenStack components are dispersed on different nodes in a minimum of four physical servers (see Figure 5.3). This would be the ideal setting to begin implementing a private cloud based with StackOps. The addition of compute nodes is feasible whenever necessary, allowing the scalability of the system. The complexity in this architecture is more pronounced, and the requirements to implement this system are also very demanding, both from the hardware and the network capabilities.

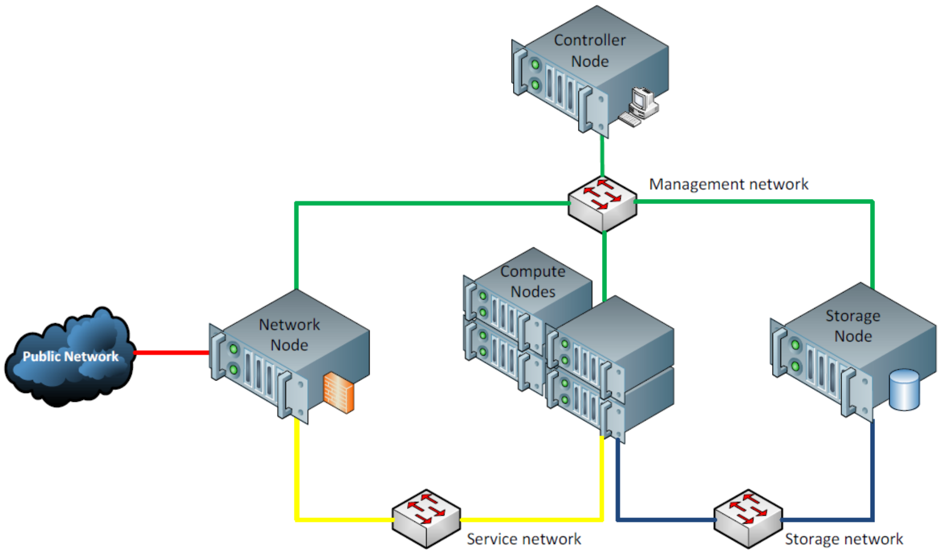

Figure 5.3: StackOps multi node architecture

In this architecture all the servers are distributed and with well-defined roles, among them three different networks are set up between the control node and the server nodes, where:

- **Management Network** is used for the communication of Nova components among them.
- **Service Network** is used for the communication from the Network node with the guest virtual machines network.
- **Storage Network** (Optional) assures the communication of all Compute nodes with Volume node.
- **Public Network** is used for the communication of internet/intranet with Network node.

5.1.2 Deployment

In this test scenario using the XCP, a simple node architecture (as shown in Figure 5.1) was adopted, where all the components of OpenStack reside on the same server,to assess the system's functionality and features. The main limitation, besides de memory and disk space, was the lack of vSwitch in order to enable Cross-Server Private Network, a feature in XenServer that is only available in XenServer with Feature Pack 1 or later Enterprise or Platinum Edition or later.

The system installation is facilitated through a web interface, that depending on the selected options, generates an installation script that the end of it, you get the features opted. As shown in Figure 6.8, and in this particular case all components have been installed in only one node.

After checking the system it is possible to access the web platform available to users (see Figure 6.10) and administrators (see Figure 6.9) to login as used.

5.2 Ulteo

Ulteo stemmed from the creator of Mandriva Linux, Gaël Duva, who developed a Virtual Desktop infrastructure open source project based on Debian and Ubuntu. [Ulteo, 2012] With this open VDI project, it is possible to allow users to run Linux applications from within a Microsoft Windows environment, thus allowing corporates to deploy virtualized GNU/Linux and/or Windows desktops through the Open Virtual Desktop.

Through Ulteo's Open Virtual Desktop (OVD) technology, corporations are capable to quickly install and deploy an OVD Session Manager with one or more OVD Application Servers. Subsequently to the definition of the users and the creation of a "publication", end users can access the applications delivered through Open Virtual Desktop using different possible modes:

- Linux or Windows Desktop with a mix of remote Linux and/or Windows applications.
- Application Portal, to run remote Linux and/or Windows applications from web links.
- Application Publishing, to get remote Linux and/or Windows applications seamlessly integrated in the local end-user desktop.

5.2.1 Architecture

Two different examples of how Ulteo can be deployed, Open Virtual Desktop V3.0 core architectures Desktop mode and Application mode (Web portal or Desktop integration) respectively represented in Figures 5.4 and 5.5.

In the Desktop mode, the client connects to the Ulteo Session Manager server(s) and establishes a RDP over SSL connection to the Linux ApS server(s) to retrieve the desktop environment as shown in Figure 5.4. If an external application is invoked, the RDP connection is bounced to the Windows ApS

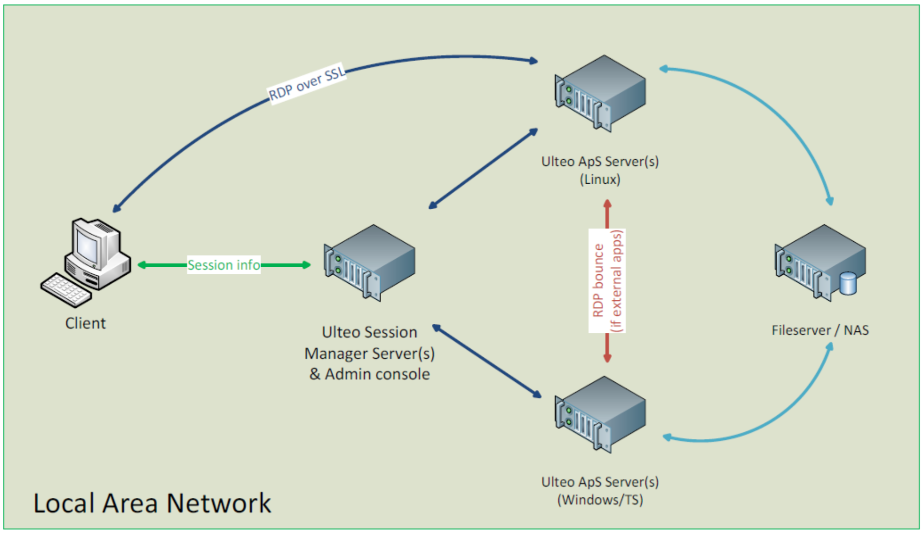

Figure 5.4: OVD - Desktop mode

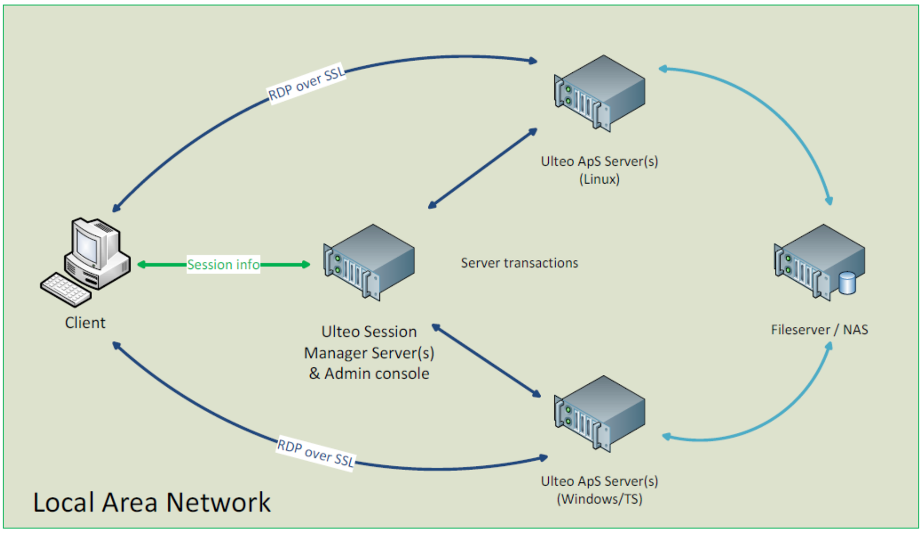

Figure 5.5: OVD - Application mode (Web portal or Desktop integration)

server(s) (Windows/TS). In each of the RDP connections the Fileserver/NAS is made available for the ApS server(s).

In the Application mode, either in Web portal or Desktop integration, both of the Linux and Windows ApS server(s) can have client RDP over SSL connections depending on which application the user opens (see Figure 5.5). As always, the Fileserver/NAS is available for the ApS server(s).

5.2.2 Infrastructure

Ulteo OVD is all about mixing various applications sources into a consistent stream that can be delivered to users, depending on their needs. It's also been intended to be integrated in heterogeneous environments and capable to inter-operate with various technologies. The Figure 5.6 illustrates the Ulteo infrastructure overview where clients can access from any device to a public or private cloud.

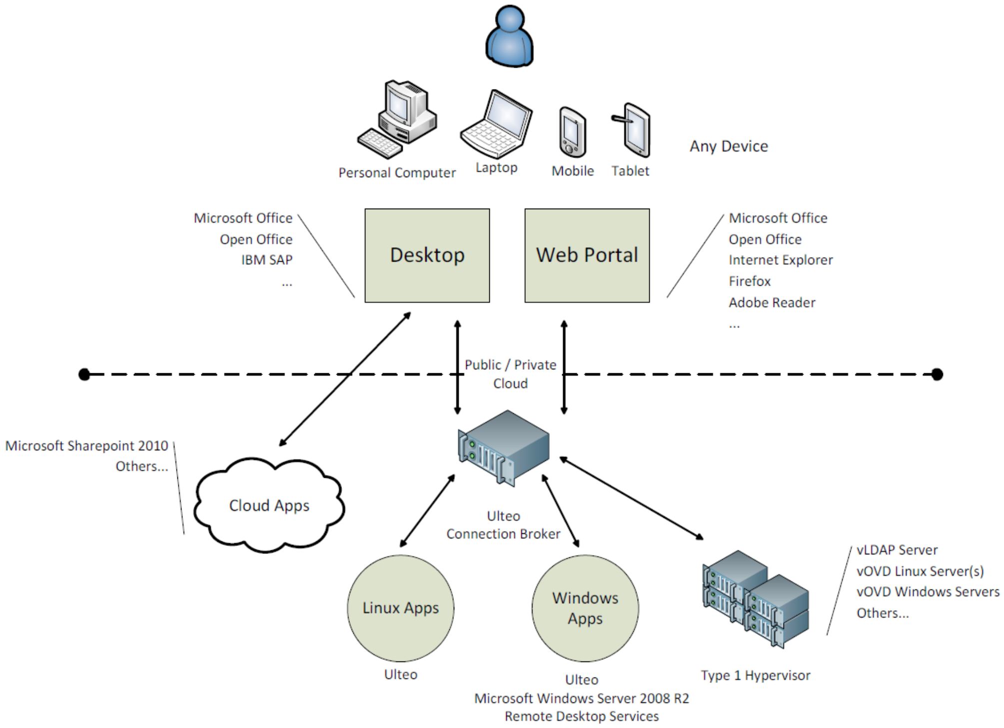

Figure 5.6: OVD - Infrastructure overview

5.2.3 Features

Windows and/or Linux application delivery

Ulteo's Open Virtual Desktop delivers applications that are hosted on Linux servers or via Windows Terminal Services to any Java-enabled web browser, to any desktop (Windows, Linux, MacOS, thin client), that are available anywhere. Windows applications are made available through Windows Server 2003 or 2008/2008R2 servers.

Ulteo provides a default "base" set of Linux applications including the OpenOffice productivity suite, Firefox web browser, Thunderbird email client, Pidgin instant messaging client, and several others. Supplementary Linux applications can be installed and deployed to meet any particular and specific needs by the system administrator.

Multiple choices for end-user interface

Applications can be delivered to users by way of:

- A Remote desktop (Windows 7 like or Linux): Users can log into a full-featured remote desktop to access applications "published" to the user (or security group) by the administrator.
- Client-server model: Client-server computing refers broadly to any application that is distributed which differentiates between service providers (servers) and service requesters (clients).
- Integrated: The dynamic integration of applications in the end user's local desktop, making them available as if they were locally installed.

User clients include:

- a web-browser (the web-browser needs a Java plugin to perform).
- a dedicated Ulteo client software for Windows, Linux or thin clients.
- an Android or iOS or tablet (only available in desktop mode).

Ulteo Open Virtual Desktop enables many features and characteristics such as:

- Local desktop integration: Network shares, local drives, local printers, USB sticks, sound support, copy-paste are provided.
- Centrally managed administrator web console:
  - A full-featured administrator web console is provided for configuring server settings, managing users, publishing applications, and for monitoring servers.
  - The administrator is guided through a wizard that walks through the process of publishing applications. The OVD administration console enables conveniently the installation of new applications.
  - It is possible to delegate administration privileges through the administration console in order to specific users are able to administer the system.
- LDAP and authentication:
  - User authentication can be simplified using Windows Active Directory and also with any LDAP server.
  - Authentication against CAS servers is also supported.
- File Servers
  - Windows or Linux (Samba) fileservers
  - WebDAV embedded fileserver

- Scalability: OVD has been designed to serve applications to thousands of users at the same time. Comprehensive tests and tuning have been performed to ensure excellent Linux and Windows® application servers load handling and directory management up to 50,000 users and more.
- Bandwidth usage: New settings and better cache management are available to reduce bandwidth usage.
- Monitoring, logging, reporting and load-balancing: Detailed monitoring, logging and reporting of Windows and Linux application servers are provided to the OVD administrator. Applications Servers can also benefit of customizable load-balancing configuration.
- Security and remote access:
  - Wan Internet AND LAN accesses secured by SSL.
  - OVD Gateway module simplifies WAN access: one IP address, one port.
- Other
  - The system is also compatible with Amazon's EC2 virtualization infrastructure.
  - Session recovery and reconnection supported as well as the support for several languages and keymaps.
  - The Application Server and Session Manager may be installed on a single server by using a dedicated DVD ISO image.
  - SSO-ready through the API.

System recommendation

Ulteo OVD recommends the following server system configurations:

- OVD Application servers for Linux applications: x86 servers with multi-core or quad CPU with 1GB or more RAM is the minimum configuration. For each 15 concurrent end-users a minimum of 1Gb memory should be added. Supported Host OS (32 or 64 bit): Ubuntu 10.04.*, RHEL 5.5, 6.0, Centos 5.5, 6.0, SLES 11SP1 and OpenSuSE 11.2, 11.3.
- OVD Application servers for Windows applications: Windows 2003 (32 bit), 2008R2 (64bit) Server + Active Directory and Terminal Services on any hardware. 1GB or more RAM.
- Servers for OVD Session Manager: x86 server w/512MB or more RAM. Host OS (32 or 64 bit): RHEL 5.5, 6.0, Centos 5.5, 6.0, Ubuntu 10.04.*, SLES 11SP1 and OpenSuSE 11.2, 11.3.

Note about OVD ApS and SM servers: It is possible to install both OVD components in only one server, but for evaluation purposes. It is not recommended this configuration for production environments.

Ulteo OVD recommends the following client system configurations:

- Web Client: Oracle Java 1.6 (32 bit only) enabled browser: Firefox 2+, Internet Explorer 7+, any platform. Safari 5+ on MacOS. 512 MB RAM or more (1GB or more is recommended).
- Native Client: Windows® XP/7 or Linux platform. Minimum of 512 MB RAM and 1GB or more is recommended
  - iOS 4.3 or 5.x minimum client for iPad or iPhone
  - Android 2.3+ Client for tablets and phones

Other recommended environmental characteristics:

- Network: at least 10Mbps (minimum recommended) or more LAN
- User directory servers: Active Directory, LDAP and e-Directory/ZenWork servers are supported.
- Fileservers: CIFS (Windows and/or Linux), embedded Ulteo WebDAV and CIFS servers

Minimum hardware requirements

For a Session Manager:

- Evaluations/POCs: Minimum requirements are 1 CPU core and 512MB RAM
- Production environments: Recommended a minimum of 2GB RAM and 2 CPU cores

For an Application Server (Windows or Linux):

- Evaluations/POCs: Minimum requirements are 1 CPU core and 1GB RAM
- Production environments: Recommended a minimum of 4GB RAM and 2 CPU cores

For a WebClient:

- Evaluations/POCs: Minimum requirements are 1 CPU core and 512MB RAM
- Production environments: Recommended at least 1GB RAM or (preferentially) more
- If the same machine is used to host the Session Manager, the Web Client and the Linux Application Server:
  - Evaluations/POCs: The minimum requirements are 1 CPU core and 1GB RAM
  - Production environments: Ulteo does not recommend the use of this installation method for production environments.

5.3 Hyper-V

This section describes the procedure of setting up the bare minimum components to demonstrate a Private Cloud environment using current release versions of Microsoft products and technologies. This test scenario is not meant for nor is it an ideal configuration to use in a production environment.

Microsoft provides in depth documentation in to how to build a full production private cloud using a Hyper-V Cloud according to the best practices, enabling the full potential of the technology such as Hyper-V Cloud Fast Track. [Microsoft, 2011a]

Once the installation and configuration are complete, it is possible to demo the use of System Center Virtual Machine Manager and the SCVMM Self Service Portal 2.0 for a Private Cloud to be built and managed. The Microsoft System Center (MSC) Technologies enables full cloud management and administration by monitoring, reporting, change management, deployment and more and even include Xen and VMware pools. More software and hardware resources would be easily integrated in the implementation as it would be facilitated though the use of the MSC.

It is required to have at least a basic understanding of the roles and services in Windows 2008 R2, a superficial knowledge of how to install SQL Server 2008 R2, and a fair knowledge of how the System Center Virtual Machine Manager works. It is worth reiterate the fact that for a production environment, the careful installation of SQL server is required to ensure the proper functioning of System Center.

The resulting implementation of this configuration does not provide for any failover or redundancy and is intended solely as a lightweight demo/test/learning environment. The concepts addressed here can be applied as a template to install a production Private Cloud.

5.3.1 Architecture

Microsoft Hyper-V Cloud foundation is built on the Windows Server platform with the System Center end-to-end service management capabilities, the Hyper-V virtualization capability, and the Windows Server Active Directory identity framework. In order to the infrastructure as a Service can be consumed, it is needed the pooling, allocation, provisioning, and usage tracking of datacenter resources that is brought by the new System Center Virtual Machine Manager Self-Service Portal 2.0 that simplifies these operations. [Microsoft, 2010]

For this test scenario, two machines were added to the existing infrastructure. As Hyper-V role needs hardware virtualization support, one physical computer performed this task and installed with Windows Server 2008 R2 server core (recommended for Hyper-V). The existing domain controller, and DNS server were used along with the existing SQL Server and one other virtual machine was created with the Web Server (IIS) role. This system architecture is illustrated in Figure 5.7.

The Windows 2008R2 servers with the following roles were required:

- Active Directory Domain Services
- DNS Server

- Hyper-V
- Web Server (IIS)

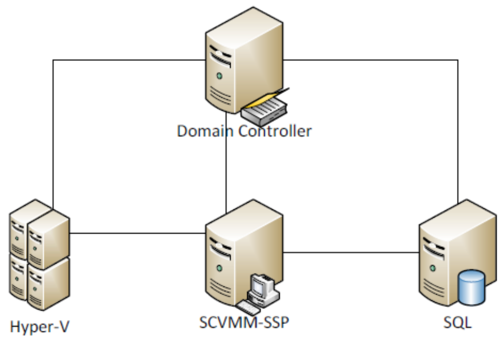

Figure 5.7: Hyper-V Architecture Configuration

5.3.2 Hardware

A desktop computer with a AMD Quad-processor 3.4GHz CPU, 8Gb RAM and two 500Gb 7200rpm SATA drive for host operating system and guest VM storage was used to install Hyper-V. Both hard drives could be striped as RAID-0 for additional performance or they could be formatted independently and place guest VM's on separate spindles. For this experimental scenario both drives separately were used to guarantee the maximum the usable space for snapshots (undo disks) and templates. This implementation lacks a backup solution that would have to be engineered, to ensure the business continuity in case of hardware failure. All the other involved are virtual machines running in the ICBAS|FF-UP XEN Server pool.

5.3.3 Software

The necessary software was downloaded in the UP software repository, however, anyone who wishes, can use free trial versions (in lack of a Technet or MSDN subscription). For testing purposes, one needs to be aware of the individual of when each product will expire. Using the components below, from the day the Host machine OS is installed (Hyper-V), 180 of working days will be assured.

The software needed for this infrastructure to work:

- Windows Server 2008 R2 with Service Pack 1 (SP1)
- Microsoft SQL Server 2008 R2
- Microsoft System Center (MSC) Virtual Machine Manager (VMM) 2008 R2 Service Pack 1 (SP1)
- Microsoft System Center (MSC) Virtual Machine Manager (VMM) Self-Service Portal (SSP) 2.0 with SP1

5.3.4 Installation

The two new servers needed to be configured from scratch, the Hyper-V and the VMM Self-Service Portal. As the all the other stakeholders are virtual machines already existing in the ICBAS|FF-UP XEN Server pool.

Hyper-V Host

The installation of the Windows 2008 R2 server core is rather simple, and after added the Hyper-V role, several necessary updates are required.

Pre-install hardware configuration is obligatory - ensure that you have enabled virtualization support in the BIOS of your computer. The form to configure/enabled depends on the hardware manufacturer and the BIOS used. It is also recommended that the Data Execution Prevention (DEP) is active.

Virtualization extensions provide support for privileged instructions (Intel-VT and AMD-v) with support for MMU virtualization that enables PCI pass-through. It is somewhat comparable to XEN paravirtualization and enables the guest virtual machines to directly use peripheral devices, such as Ethernet, accelerated graphics cards, and hard-drive controllers, through DMA and interrupt remapping.

The main characteristic of DEP is to aid inhibition of code execution from data pages. Normally, code is not executed from the default heap and the stack. Hardware-enforced DEP identifies code running from these locations and raises an exception when execution occurs. Software-enforced DEP can help avoid malicious code from misuse the exception-handling mechanisms in Windows.

Windows Server 2008 R2 is the foundation up which the entire private cloud is built. The built-in Hyper-V hypervisor is leveraged for virtualizing the servers, clients and their applications that can then be served up through the self-service portal. It is absolutely critical that the base server is installed properly and is 100% stable.

Web Server (IIS) Role

The Internet Information Services (IIS) is required by the Self Service Portal (SSP) 2.0. The web portal also needs specific Web Server (IIS) role services and the Message Queuing Feature to be enabled.

With the role Web Server IIS, the following role services are required:

- Static Content
- Default Document
- ASP.NET
- .NET Extensibility
- ISAPI Extensions
- ISAPI Filters

- Request Filtering
- Windows Authentication
- IIS6 Metabase Compatibility

The server feature Message Queuing is also necessary for the Self Service Portal.

SQL Server

After the Windows Server 2008 R2 Foundation is complete, the Windows Server 2008 R2 and Hyper-V host is now complete, some few additional optional steps are suggested just for precaution. The installation of Windows Server Backup Features and the execution of a Bare Metal Recovery Backup to external storage using Windows Backup are recommended. The SQL Server 2008 R2 is used for storing configuration information for System Center Virtual Machine Manager and the SCVMM Self-Service Portal. Considering a full production SQL Server R2 was already up and running in the ICBAS|FF-UP infrastructure, installing a new server was not required.

System Center Virtual Machine Manager

For the installation of the System Center Virtual Machine Manager R2, some prerequisites are needed for the system to be in compliance including security updates, and AIK (Windows Automated Installation Kit). The SQL Server instance access is required at this stage by providing the user credentials for this purpose.

After the installation of the Virtual machine manager Administrator console, it will become the principal interface used when interacting with the virtualization infrastructure. Occasionally it may be needed to go back to the standard Hyper-V MMC, but for the majority of tasks, the SCVMM Administrator console is more than adequate and easy to use.

Self-Service Portal

The Microsoft Private Cloud is represented mainly by the Self-Service Portal. Administrators can, through this portal, create resource pools consisting of networks, storage, load balancers, virtual machine templates and domains. It is the possible for administrators to create and manage business units who can use the self-service portal to requests these pools of resources and create them on demand.

In the SSP 2.0 setup, after selecting VMMSSP Server Component and Website Component, and clearing the necessary prerequisites, once again the SQL Server instance is required for access.

Though, a Private Cloud is more about the how you use the infrastructure to create value, provide self-service, reduce overheard, automate resource creation and ultimately resulting in cost decrease.

5.4 Future Work

IT agility, by rapid deployment, decommissioning and elasticity can be achieved as direct benefits of Private Cloud Computing [Bittman, 2011b]. It is possible to implement different approaches to the way IT is delivered to the users (teachers, researchers, staff and students). Enabling self-service access to virtualized systems can empower testing and development environments to these users whenever necessary.

Following this guideline, there is still much work to accomplish in this field, including a StackOps multi-node architecture to be implemented in various physical servers (which may be distributed physically) as represented in Figure 5.8. This architecture approach, although complex, decouples the roles of each component of the system enabling more flexibility associated to the network separation providing more security. It is more ambitious project than the multi node because it provides failover capabilities assured by the duplicate (or more) servers in each of the roles.

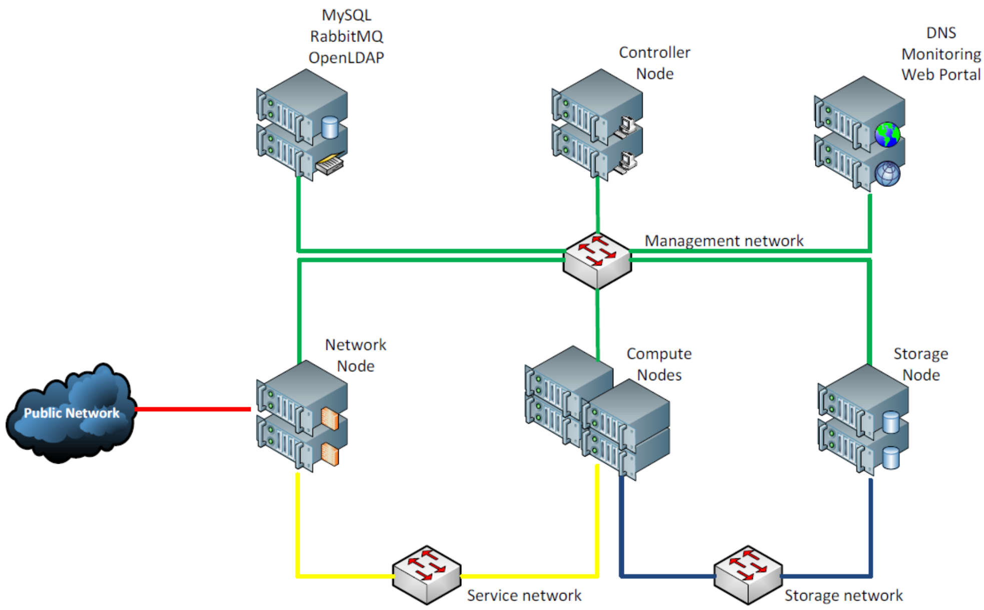

Figure 5.8: StackOps advanced multi node architecture

There are several projects that can be implemented in an integrated manner with one of the solutions covered including:

- Enhance the monitoring system [1]
- Optimizing the virtual desktop solution [2]
- Replica of the UP LDAP core [3]

[1]By scaling up Nagios or adopting other strategy
[2]Either providing RD or alternatively using Ulteo
[3]for enabling these users to login to the common computers

- Implementation of Federation authentication with shiboleth [4]
- Change the email server [5]

Fine-tuning the Nagios server considering the virtualization technology. Change the monitoring strategy based on introspection considered to be more sophisticated. In this approach, the variables are collected from outside the virtual machine by a system installed on the physical machine's operating system. These systems inspect the memory footprint of the virtual machine, and discover the value of these variables. This approach is being used for intrusion detection. [de Carvalho, 2010]

Nagios can also be used in conjunction with other monitoring solutions, providing a full integrated metric monitoring system. Gonçalves [2011]

The ability, on demand, to provision instances of virtual machines provided to teachers, investigators, and students, is being considered at this time, therefore taking advantage of the available hardware resources resulting of an implementation.

A requirement for future implementation is the availability of remote working, motivated by the outbreak of swine flu in 2009. That can be accomplished with the checkpoint VPN connection and the existing RDS farms. As part of the Disaster Recovery Management (DRM), the work-at-home, providing the users desktop from a remote location, is a model that is currently taken into account when planning and implementing Business Continuity Management (BCM). [Morency and Witty, 2011]

An alternative scenario would be an implementation of a Ulteo Open Virtual Desktop configuration that would permit a approach to this solution, thus providing Linux and windows applications to the users (see Figure 5.9) making it an asset with respect to the heterogeneity of systems which may exist.

The full RDS framework is achieved with Microsoft Windows Server 2008 R2. The suggested complete architecture has five core components as shown in Figure 5.10.

- Remote Desktop Web Access (RDWA)
- Remote Desktop Gateway (RDG)
- Remote Desktop Connection Broker (RDCB)
- Remote Desktop Session Host (RDSH)
- Remote Desktop Virtual Host (RDVH)
- Licensing Server
- Active Directory

Each of the components shown plays a very specific role with its own set of features. As a hole, the delivery of Terminal Services applications, remote desktops, and virtual desktops, are accomplished.

[4] for authenticating users of the LDAP core to access IT resources
[5] either using Microsoft Exchange or Google Apps or both

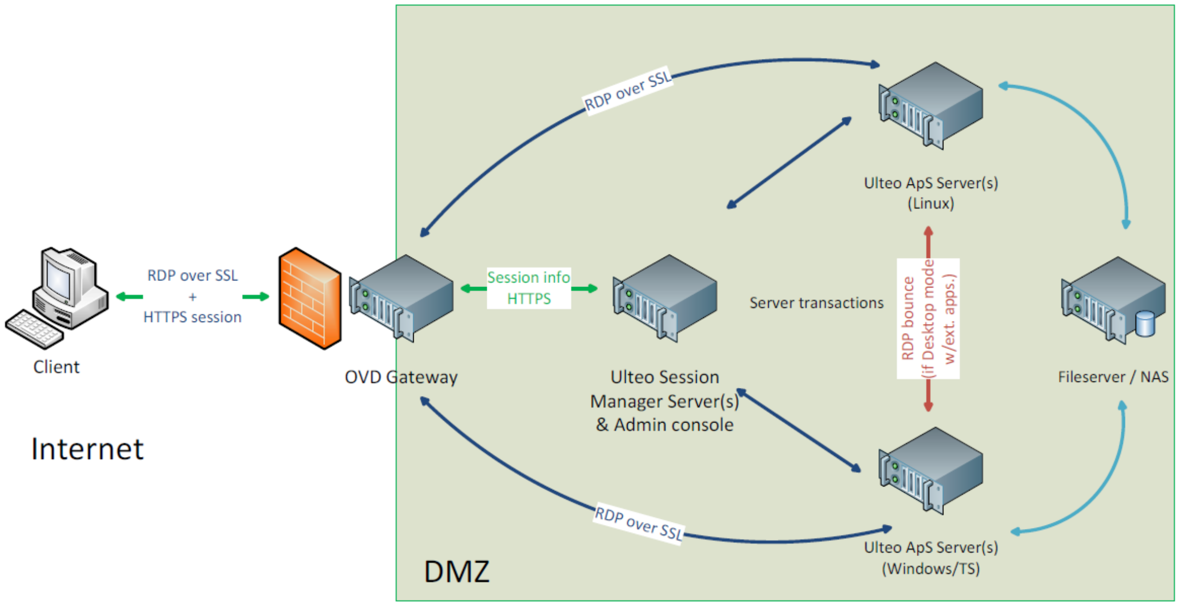

Figure 5.9: OVD - Application mode or Desktop mode with WAN access through OVD Gateway

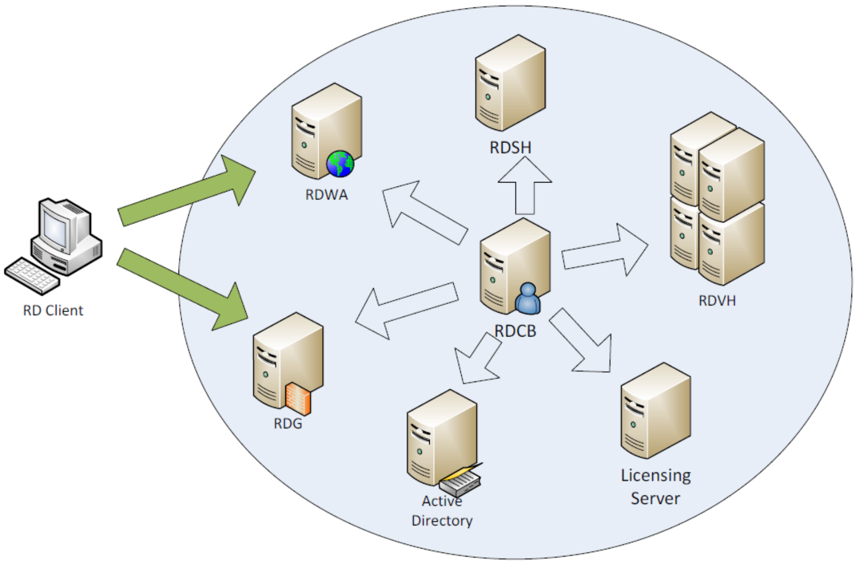

Figure 5.10: Microsoft Remote Desktop Services suggested architecture

An interesting project in hand is also a replica of the LDAP core of UP in read-only mode to be created (based on OpenLDAP) on the campus. Therefore, common resources in the ICBAS|FF-UP campus can be used with any user from any organization unit within UP. A script will be developed to synchronize the users to the top domain bearing in mind the difference between the LDAP core schema used in the UP

Table 5.1: Windows Server Licensing

Windows version	Licensing mode	Number of free virtual machines
Windows Server 2008 R2 Enterprise	Per server	4
Windows Server 2008 R2 Datacenter	Per processor	Unlimited

and the Active Directory schema.

The migration of the email to the cloud using google apps for education (in the universities edition), thereby allowing a critical seamless service and at the same time enabling the access to a range of collaboration tools so far unused. Success case studies are well documented. [Google, 2011]

5.5 Suggested implementation

Considering the exposed technologies in the case study and in the implementation chapter, different architectures can be suggested. Either open or paid technologies can be adapted to the used scenario. A full Microsoft based implementation requires a structural change in the used virtualization technology and it may be complicated to deploy since some of the systems are non-Windows.

Although it would be my approach to the design of the system, since there are utilities to convert XEN virtual machines into Hyper-V and considering that most of the systems are Microsoft based, it would guarantee a more integrated system. Taking advantage of the licensing benefits that the University of Porto has, and considering that the Hyper-V Server (server) license offers up to 4 Windows Server free licenses in the Enterprise edition (see Table 5.1), this approach is also viable in terms of licensing costs.

VMware, although a major player, will not be considered due to high costs in licensing, which can be up to 16 times higher than that of Microsoft, beyond the fact of the compulsory certification of the existing hardware to be VMware compliant. [Microsoft, 2012]

The use Hyper-V Server, which can be made downloaded directly from Microsoft, comes with no interface for managing VMs in advanced mode, which is already included in VMware and Citrix XenServer. A very basic management can be done from any Windows installing RSAT (Remote Server Administration Tools), but that goes for environments with few VMs. An environment with many VMs to facilitate the management will need to purchase: System Center Virtual Machine Manager (VMM) 2008 R2 SP1 - Enables centralized management of IT infrastructure, physical and virtual, greater server utilization and dynamic resource optimization across multiple platforms for virtualization. It includes end to end capabilities from the planning, the deployment, the management, and optimization of the virtual infrastructure. Licensing of this tool can is explained in Table 5.2.

Microsoft recommends for a private cloud solution the use of the Datacenter edition, not only because of the unlimited licensing of virtual machines [Microsoft, 2011c], but also because of the extended

Table 5.2: Virtual Machine Manager Licensing

Environment	Product	Licensing mode
Big	Virtual Machine Manager 2008 R2 Server ML Enterprise	Management server
Big	Virtual Machine Manager 2008 R2 Client ML	License for each managed virtual server
Small	Virtual Machine Manager 2008 R2 Workgroup	Manages up to 5 physical hosts without the need of virtual server licensing

Table 5.3: System Center Essentials 2010 Licensing

Product	Licensing mode
SCE 2010 Management Server Software License	Management server with SQL Express (recommended up to 15 servers)
SCE 2010 Management Server Software License with SQL Server Technology	Management server with SQL Standard
SCE 2010 Server ML	Licence for each managed server (physical and virtual)

support to this technology such as Cloud Jumpstart [Microsoft, 2011b].

Microsoft System Center Essentials 2010 provides a unified experience for managing physical and virtual to IT professionals of midsized organizations. It is more comprehensive compared to SCVMM, it allows secure, update, monitor and troubleshoot more efficiently from a single console, so that you can manage your IT environment proactively and effectively. Its limitation is the management of up to 50 servers among physical and virtual. The licensing is described in the Table 5.3.

The newer version, System Center 2012 will be licensed only by the physical processor being managed, not the current Server license plus Management License model. [Buchanan and O'Brien, 2012]

Migrating all the systems to Microsoft could be more or less complex (ie VPN or eMail), but in the long run the management could be more simpler, for example when creating users in the active directory, they could immediately have a VPN access account (ie via checkpoint that is already integrated with the domain controller) and an active email account with Microsoft Exchange.

For the email service, in order to migrate that service to the cloud, Google Apps for Education (University edition) is a choice of weight, thereby releasing, human and technological resources, for targeting to the private cloud. In addition to the 25GB offered by Google, is a free, secure, familiar (to a large number of users) and with access to collaborative tools.

Thus, being a hybrid structure would be achieved in order to have resources of a private cloud in conjunction with a public cloud (see Figure 5.11) and using both open and payed technologies. Users can access ICBAS|FF-UP private cloud, through the VDI solution or the private cloud (or both) and also access the Google Apps public cloud.

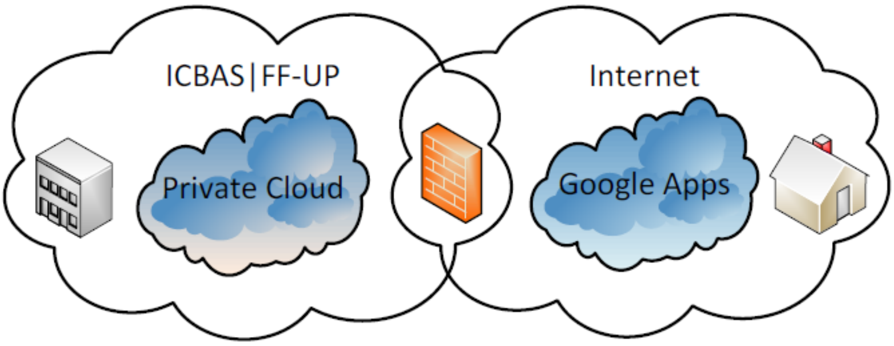

Figure 5.11: ICBAS|FF-UP Suggested Cloud Architecture

Users from home, could access their school/work desktop with their installed applications and share, as well as an email service that resides in the public cloud. It would be also possible to use provision virtual machines instances for users to use in the best way they find appropriate.

Chapter 6

Conclusions

After the analysis of cloud computing technologies and study the existing implementation (in relation to the technologies used) on the campus of ICBAS | FF-UP, a suggested evolution of the used systems is presented to evaluate its functionality and thus maximize existing the existing resources, making them even more agile and consequently adding more value to the users.

The technological approach made in the previous chapter was intended to comprise alternative solutions to the detriment of proprietary (paid) solutions considering the existing infrastructure without compromising the initial requirements and that would be as transparent as possible to the end user. Combined with this evaluation, a private cloud of IaaS implementation has been considered, to create value added to the academic community, either for testing purposes or for pursuing R&D goals therefore making more agile the way IT is delivered.

As shown in the Table 6.1, all components used can have (a free or paid) alternative, being demonstrated in this work the feasibility of changing the virtualization infrastructure, the remote desktop and application delivery.

Nevertheless not always open technologies are more suitable, for in this case, an implementation of Microsoft Windows Server 2008 R2 Hyper-V with a Microsoft VDI solution, complemented with Ulteo OVD (to deliver Linux applications) can also be achieved.

In a more closed vision, email could always be replaced with Microsoft Exchange Server, and the system integration would be complete since the infrastructure would be all focused on Microsoft. Thus the system management with System Center might be even more profitable by taking advantage of virtually all its features. Despite the system integration advantages, adopting a public cloud service like Google Apps, offers more than the eye can meet. By relocating IT staff resources to other technologies, and freeing up storage space, it turns out that it just might be a perfect solution in order to consolidate and optimize the existing IT resources.

Implementing a private cloud is not just limited to the installation and configuration of software (open or paid) on any server that meets the minimum requirements in order to promote the services and features of the infrastructure. It is a multidisciplinary subject that involves several areas of computer

Table 6.1: Alternative technologies

Technology	Open	Proprietary
Virtualization	XEN Hypervisor	Microsoft Hyper-V with With Virtual Machine Manager
Identity Management	OpenLDAP	Microsoft Active Directory
Remote Desktop	Ulteo	Microsoft Remote Desktop
Application Delivery	Ulteo	Microsoft App-V
Email	Cyrus	Microsoft Exchange
Productivity Tools	Open Office	Microsoft Office
Monitoring	Nagios	Microsoft system center configuration manager
Private Cloud	OpenStack	Microsoft Hyper-V with With Self Service Portal

engineering applied to all the aspects involved, so that the final architecture of the system is suitable for a production environment that is efficient, reliable, secure and elastic.

Cloud Computing allows a number of important benefits for their users, and the major cost reductions, greater flexibility and mobility, and above all, be a catalyst for innovation. These advantages are some of the most recognized, however certain factors are also inhibitors, such as loss of control, security and privacy. [Zhou et al., 2010] Some of the providers of cloud computing were studied, and after the presentation of the current implementation at ICBAS-UP, some of the available platforms were tested regarding the requirements and matching them to this paradigm.

Although there is still some distrust in public clouds, private clouds will tend to evolve because they are in controlled environments, so keeping some level of security and data integrity.

The cloud computing platforms are intrinsically linked to virtualization, will therefore be necessary more hardware to support a virtualization platform able to have adequate performance in a production model. The open source alternative is currently a viable alternative to consider. The existing open platforms are now easier to deploy, configure and use and with proven reliability. [Chang et al., 2010]

In a first instance, OpenStack, is the solution to adopt given the nature of the specificities for this scenario of cloud computing. Microsoft's solution, although relatively simple to implement and fully integrated across the existing architecture, has the cost associated with the licensing. At this stage, we were able to provide a test pilot for the ICBAS|FF-UP direction to ponder the investment decision in this type of technology.

In the context of this research study, and considering that licensing through the Microsoft Campus Agreement (MCA), by the Rectory of UP, is quite attractive, since the technologies to use are liable to be included in the MCA. The benefit to the IT team is huge, since the administration of systems is performed in an environment where everyone is already somewhat familiar (saving time in early learning and adaptation to the platforms). Besides DreamSpark Premium, renamed the Microsoft Developer

Table 6.2: Hypervisor licensing cost comparison

	VMware vSphere 4.1 and below	VMware vSphere	Microsoft Hyper-V 2008 R2 SP1	Citrix Xenserver 6
Licensing	CPU	CPU	Free [1](Without limiting processor and memory)	By hypervisor functionality (Without limiting processor and memory)
Cores for processor and memory	6 cores for each license Standard and Enterprise, ESS, ESS + or 12 cores for each license Advanced and Enterprise Plus	Unlimited	Unlimited	Unlimited
Total Virtual Memory per processor Licensee	Not applicable	In accordance with vSphere version: 32GB VRAM Essentials Kit; 32GB VRAM Essentials Plus Kit; Standard VRAM 32GB; 64GB VRAM Enterprise; Enterprise Plus 96GB VRAM	Not applicable	Not applicable

Network Academic Alliance (MSDNAA) included in the MCA, is also included access to TechNet SA for obtaining online technical support on products and technologies, enabling a privileged communication channel for the resolution any problems and technical difficulties.

Although to operate a Microsoft integrated environment, it is required to adopt Microsoft System Center 2012. Microsoft recently changed the licensing terms for this product as a way to improve private cloud licensing to promote viable comparisons with VMware or Citrix editions. [Buchanan and O'Brien, 2012]

Still, with the Microsoft UP campus agreement, these licensing costs are somewhat more friendly than the others (see Table 6.2).

Having this architecture deployed enables the background requirements to enable a Microsoft based private cloud solution managed centrally with System Center.

The consideration of all factors intrinsic to the technology of cloud computing and the needs and requirements existing in the current context lead to approaches that can be so different in philosophy, but on the other hand, so close to achieving in its functionality and provision of services.

In a purely economic perspective, the open source technologies lead to distinct advantage. Existing

[1] The Hypervisor is totally free, but not its professional management.

solutions are now sufficiently mature and stable, with great level of support from the online community projects and providing a range of functionality very close to proprietary versions purchased. In the case of the XEN hypervisor, many features not available in free versions can be created in scripts through its API, many of them are also available in communities.

It should be noted that despite the inherent cost benefits of using these technologies requires a high degree of specialized knowledge (or knowledge acquisition) for that technology to be properly implemented and integrated.

On the other hand, the paid solutions (in this particular case the Microsoft) have a form of implementation easier either via the graphical user interface (GUI) or by various wizards available to perform the most diverse tasks. However, this approach also implies a high degree of expertise and adapt to the used GUI. In the case of Microsoft System Center Configuration Manager, the wizards are practically nonexistent, so it is (almost always) necessary to consult the Microsoft Developer Network (MSDN) to get to perform any kind of operation, not being at all intuitive.

In a broader perspective, and taking into account the organizational structure of the UP, it could be considered a solution for implementation on a larger scale (focusing the organizational units). The Rectory of the UP would provide an academic cloud that would provide their services to the organizational units via a real-time infrastructure (RTI). This type of service could also be scalable to the other could layers (PaaS, DaaS, SaaS, ...) thus unifying all the services of IT and standardizing and streamlining the processes and the organizational management.

This solution could be integrated with existing private clouds in each of the organizational units. In Figure 6.1 are represented the ICBAS and FF, but the intention would be to include all 14 faculties, and still being capable of extending the services to other institutions if necessary.

6.1 Final considerations

With the substantial advances in Information and Communications Technology (ICT) over the last half century, there is a growing perceived general opinion that computing will become the 5th utility (after telephony, water, electricity and gas). This form of computing utility, like all the other four existing utilities, will deliver the basic level of computing service that is considered the essential to meet the everyday necessities of the consumers. [Buyya et al., 2009]

Private cloud computing is emerging from the hype and has turned into very real implementations; moreover, it's expected implementations to double in 2011. Technology solutions are still relatively incomplete and immature, and clients should be cautious to focus on the fundamental business model, process changes, management changes, cultural changes and customer relationship changes first or at least in a correspondent manner with small technological deployments. [Bittman and Scott, 2011]

Considering that today hybrid cloud computing is mostly aspirational, visionary enterprises and service providers are choosing technologies and building on-ramps. During 2012, it is expected to start to determine vendor winners, emerging best and bad enterprise practices. Private cloud will have successful deployments based on business requirements where agility and speed help the business grow thus

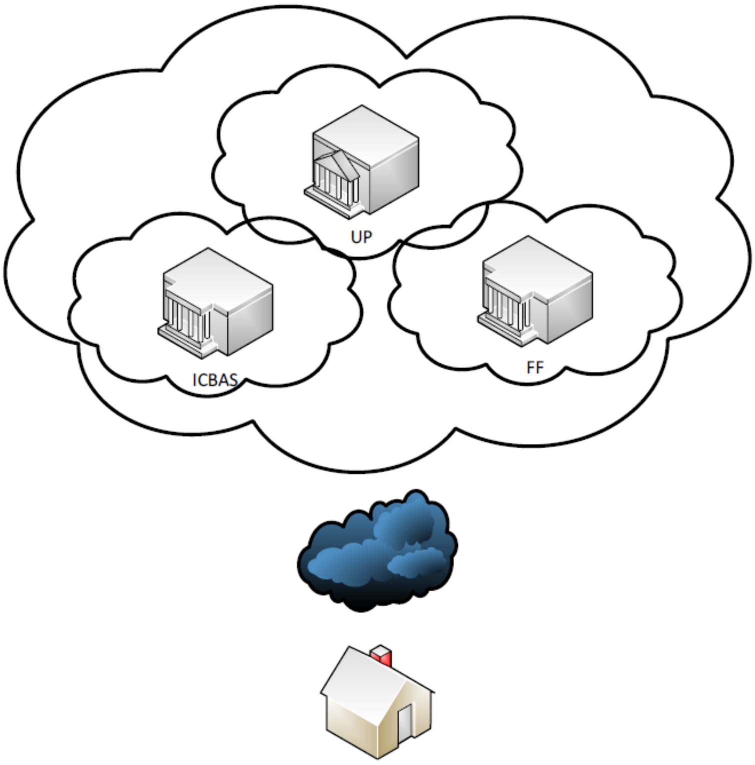

Figure 6.1: UP Academic Cloud

finding its appropriate position in many companies. [Bittman, 2012]

The approach of software and hardware architects must now be conscious to the end-to-end systems in order to obtain optimal solutions designs. The individual and independent applications or single server applications must be put aside keeping in mind the physical and economic mechanisms involved in a datacenter. [Barroso and Hölzle, 2009]

In order to consider cloud computing's energy consumption it must be taken into account the processing, the storage and the transport hence in some cases, it has been shown that cloud computing is more energy-efficient when major computing tasks are of low intensity and frequency. However, under some circumstances more energy can be consumed, even applying technologies such as virtualization and advanced cooling systems. [Baliga et al., 2011] It seems a contradiction in fact, evolving to a paradigm where the computing power is remote instead of the common local PCs, not always represent a reduction in energetic costs.

Cloud computing initiatives should be planned using three organizing principles these being the definition of concepts, the implications and scenarios and the technologies and vendors. [Cearley and Reeves, 2011]

- Concepts defined: The exploration of any subcategories and the initiative scope in this dynamic area of innovation.

- Implications and scenarios: The identification of scenarios where the innovation probable will affect current IT. Identify the impact on business risks and goals. The exploration of how to use the innovation, including the definition of a start point, the organization's readiness and strategies to adopt the innovation.

- Technologies and vendors: The identification of future and current technologies that are eligible for this innovation. Exploration of the most significant vendor offerings.

Considering that each case has its specific implementation issues, it is pertinent to take into account the optimization than can be addressed by the allocation of resources in a private cloud such that cost to the provider is diminished (through a growth of resource sharing) while endeavoring to meet all client application requirements as specified in their respective Service Level Agreements (SLA). [Ghanbari et al., 2012]

Some cloud advantages achieved by adopting cloud platforms are improved enterprise agility and risk management, decreased IT costs, and shifted IT's focus to the business. [Gomolski, 2011]

When done well, private cloud computing can enable significant improvements in speed of service delivery and responsiveness, reduced operational costs and quality of service matching business requirements. [Bittman, 2011d]

Private cloud computing is quickly becoming the preferred deployment option for enterprises' most sensitive information due to the flexibility, scalability and control derived from the private cloud's dedicated environment. [Bittman and Scott, 2011]

The computing paradigm is at a turning point and cloud computing is the main driver for change.

Bibliography

Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., and Zaharia, M. (2009). Above the clouds: A berkeley view of cloud computing. Technical report, Electrical Engineering and Computer Sciences University of California at Berkeley.

Baliga, J., Ayre, R. W. A., Hinton, K., and Tucker, R. S. (2011). Green cloud computing: Balancing energy in processing, storage and transport. *Proceedings of the IEEE 2011*, volume: 99 issue: 1:149–167.

Barroso, L. A. and Hölzle, U. (2009). The datacenter as a computer: An introduction to the design of warehouse-scale machines.

Bittman, T. J. (2010a). Private cloud computing an essential overview. Technical report, Gartner.

Bittman, T. J. (2010b). Private cloud computing: Clearing the air. Technical report, Gartner.

Bittman, T. J. (2010c). Vmware and private cloud computing. Technical report, Gartner.

Bittman, T. J. (2011a). The 10 fundamentals of building a private cloud service. Technical report, Gartner.

Bittman, T. J. (2011b). The drivers and challenges of private cloud computing. Technical report, Gartner.

Bittman, T. J. (2011c). Evaluating the benefits of private cloud computing. Technical report, Gartner.

Bittman, T. J. (2011d). Key issues for private cloud computing, 2011. Technical report, Gartner.

Bittman, T. J. (2011e). Private cloud computing: Emerging from the mist. Technical report, Gartner.

Bittman, T. J. (2012). Top five trends for private cloud computing. Technical report, Gartner.

Bittman, T. J. and MacDonald, N. (2010). Microsoft's windows azure platform appliance a major experiment. Technical report, Gartner.

Bittman, T. J. and Scott, D. (2011). Private cloud computing ramps up in 2011. Technical report, Gartner.

Bittman, T. J., Weiss, G. J., Margevicius, M. A., and Dawson, P. (2011). Magic quadrant for x86 server virtualization infrastructure. Technical report, Gartner.

Buchanan, S. and O'Brien, F. (2012). Plan now for microsoft system center 2012 licensing changes. Technical report, Gartner.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., and Brandic, I. (2009). Cloud computing and emerging it platforms - vision, hype, and reality for delivering computing as the 5th utility. *Elsevier*, Future Generation Computer Systems Volume 25 Issue 6:599–616.

Cearley, D. W. and Reeves, D. (2011). Cloud computing innovation key initiative. Technical report, Gartner.

Chang, B. R., Tsai, H. F., Huang, C.-F., and Huang, H.-C. (2010). Private small-cloud computing in connection with linux thin client. *PCSPA '10: Proceedings of the 2010 First International Conference on Pervasive Computing, Signal Processing and Applications*, Pervasive Computing Signal Processing and Applications, Sept. 2010:pp.82–87.

Commission, E. (2012). Europe2020. http://ec.europa.eu/europe2020. accessed 15-April-2012.

de Carvalho, M. B. (2010). Adaptação da ferramenta nagios para o monitoramento de servidores virtuais. Master's thesis, Universidade Federal do Rio Grande do Sul.

Dell (2005). An overview of xen virtualization. Technical report, Dell.

Dell (2012). Dell. http://www.dell.com/. accessed 7-December-2011.

Desisto, R. P. (2011). Hype cycle for software as a service, 2011. Technical report, Gartner.

EMC (2012). Emc. http://www.emc.com/. accessed 28-November-2011.

Evoy, G. V. M., Schulze, B., and Garcia, E. L. M. (2011). Performance and deployment evaluation of a parallel application on a private cloud. *Wiley Online Library*, Concurrency and Computation: Practice and Experience, 2011, Vol.23:pp.2048–2062.

Fenn, J. and LeHong, H. (2011). Hype cycle for emerging technologies, 2011. Technical report, Gartner.

Ghanbari, H., Simmonsa, B., Litoiu, M., and Iszlai, G. (2012). Feedback-based optimization of a private cloud. *Future Generation Computer Systems January 2012*, Volume 28, Issue 1:104–111.

Ghosh, A. and Arce, I. (2010). Guest editors' introduction: In cloud computing we trust - but should we? *Proceedings of the IEEE 2011*, volume: 99 issue: 1:149–167.

Gomolski, B. (2011). Case study: D-link makes aggressive move to the cloud. Technical report, Gartner.

Gonçalves, B. O. (2011). Gerência e monitoramento de uma nuvem privada. Master's thesis, Universidade Federal de Santa Catarina.

Google (2011). Google apps for education. http://www.google.com/a/help/intl/pt-BR/edu/university.html.

Grosmann, D., de Sousa, N. F. S., Diniz, F. A., and da Silva, T. R. (2011). Estudo comparativo sobre o uso do vmware e xen server na virtualização de servidores. Master's thesis, Universidade do Estado do Rio Grande do Norte – UERN.

Heiser, J. and Cearley, D. W. (2011). Hype cycle for cloud security, 2011. Technical report, Gartner.

Internap (2010). Latency: The achilles heel of cloud white paper. Technical report, Internap.

Internap (2011a). Cloud buyers guide.

Internap (2011b). Cloud hosting: Private or public? - decision brief.

Knipp, E., Smith, D. M., Cearley, D. W., and Clark, W. (2011). Client-cloud applications: The rebirth of client-server architecture. Technical report, Gartner.

Leavitt, N. (2009). Is cloud computing really ready for prime time. Technical report, IEEE.

Lenk, A., Klems, M., Nimis, J., Tai, S., and Sandholm, T. (2009). What's inside the cloud? an architectural map of the cloud landscape. *CLOUD'09, Vancouver, Canada*, CLOUD '09 Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing:Pages 23–31.

Leong, L. (2009). How to select a cloud computing infrastructure provider. Technical report, Gartner.

Leong, L. (2011). Adopting cloud infrastructure as a service in the "real world". Technical report, Gartner.

Loutas, N., Peristeras, V., Bouras, T., Kamateri, E., Zeginis, D., and Tarabanis, K. (2010). Towards a reference architecture for semantically interoperable clouds. *Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on 2010*, 2010 IEEE Second International Conference on Cloud Computing Technology and Science, Nov. 2010:pp.143–150.

MacDonald, N. and Bittman, T. J. (2010). From secure virtualization to secure private clouds. Technical report, Gartner.

Maio, A. D. (2011). Cloud computing in government: The reality behind the hype. Technical report, Gartner.

Marchi, A. A., Foscarini, E. D., Scheeren, F. M., Fialho, F. F., and Rey, L. F. (2010). Implementação de solução de virtualização e consolidação de servidores no cpd da ufrgs usando o citrix xenserver. Master's thesis, Universidade Federal do Rio Grande do Sul.

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., and Ghalsasi, A. (2011). Cloud computing: The business perspective. *Decision Support Systems*, 51:176–189.

Microsoft (2010). Microsoft hyper-v cloud datasheet. Technical report, Microsoft.

Microsoft (2011a). *Hyper-V Cloud Fast Track White Paper*. Microsoft.

Microsoft (2011b). *Hyper-V Cloud Jumpstart Datasheet*. Microsoft.

Microsoft (2011c). *Microsoft Private Cloud*. Microsoft.

Microsoft (2012). *Microsoft Private Cloud*. Microsoft.

Morency, J. P. and Witty, R. J. (2011). Hype cycle for business continuity management and it disaster recovery management, 2011. Technical report, Gartner.

Murphy, T. E. (2010). Test and quality management: The first frontier for private cloud. Technical report, Gartner.

Mutavdžic', R. (2010). Cloud computing architectures for national, regional and local government. In *MIPRO 2010*.

Natis, Y. V. (2011a). 2011 hype cycle for cloud application infrastructure services (paas). Technical report, Gartner.

Natis, Y. V. (2011b). Gartner reference model for paas. Technical report, Gartner.

Natis, Y. V., Knipp, E., Jones, T., Malinverno, P., Skybakmoen, T., Murphy, T. E., Smith, D. M., and Lheureux, B. J. (2011a). Cool vendors in platforms as a service, 2011. Technical report, Gartner.

Natis, Y. V., Knipp, E., Smith, D. M., and Pezzini, M. (2011b). Cloudfoundry.com vmware as a paas provider. Technical report, Gartner.

Natis, Y. V., Knipp, E., Valdes, R., Cearley, D. W., and Sholler, D. (2009). Who's who in application platforms for cloud computing: The cloud specialists. Technical report, Gartner.

Natis, Y. V., Lheureux, B. J., Pezzini, M., Cearley, D. W., Knipp, E., and Plummer, D. C. (2011c). Paas road map: A continent emerging. Technical report, Gartner.

Natis, Y. V., Pezzini, M., Thompson, J., Iijima, K., Sholler, D., Knipp, E., Valdes, R., Lheureux, B. J., Malinverno, P., and Driver, M. (2010). Magic quadrant for application infrastructure for systematic soa-style application projects. Technical report, Gartner.

OpenStack (2011). Openstack cloud software. http://www.openstack.org/. accessed 22-October-2011.

Parkhill, D. F. (1966). *The Challenge of the Computer Utility*. Addison-wesley.

Pescatore, J. (2010). Three styles of securing public and private clouds. Technical report, Gartner.

Pescatore, J. (2011). 2011 key issues for securing public and private cloud computing. Technical report, Gartner.

Pinhal, N. (2008). Virtualized high availability server cluster. Technical report, DEI-ISEP.

Português, G. (2012). Agenda digital 2015. http://agendadigital.gov.pt/. accessed 18-May-2012.

Qian, L., Luo, Z., Du, Y., , and Guo, L. (2009). Cloud computing: An overview. *Queue, June 2009*, Vol.7(5):pp.3–4.

RedHat (2011). Redhat. http://www.redhat.com/.

Ridder, F. and Rold, C. D. (2010). Comparing infrastructure utility services and private clouds. Technical report, Gartner.

Rimal, B. P., Choi, E., and Lumb, I. (2009). A taxonomy and survey of cloud computing systems. *Fifth International Joint Conference on INC, IMS and IDC*, 2009 Fifth International Joint Conference on INC, IMS and IDC, Aug. 2009:pp.44–51.

Satchwill, B., Rankin, R., Shillington, J., and King, T. (2011). An internationally distributed cloud for science: the cloud-enabled space weather platform. *SECLOUD '11: Proceeding of the 2nd international workshop on Software engineering for cloud computing*, SECLOUD '11 Proceedings of the 2nd International Workshop on Software Engineering for Cloud Computing:pp. 1–7.

Scott, D. (2010). Building private clouds with real-time infrastructure architectures. Technical report, Gartner.

Scott, D. (2011). Critical success factors for private cloud computing. Technical report, Gartner.

Sengupta, S., Kaulgud, V., and Sharma, V. S. (2011). Cloud computing security - trends and research directions. *IEEE World Congress on 2011*, 2011 IEEE World Congress on Services, July 2011:pp.524–531.

Stevens, H. and Pettey, C. (2008). Gartner says cloud computing will be as influential as e-business (special report examines the realities and risks of cloud computing). http://www.gartner.com/. accessed 28-September-2011.

Sultan, N. (2010). Cloud computing for education: A new dawn? *International Journal of Information Management*, 30:109–116.

Sultan, N. A. (2011). Reaching for the cloud: How smes can manage. *International Journal of Information Management*, 31:272–278.

Teneyuca, D. (2011). Internet cloud security: The illusion of inclusion. Technical report, Information Security Technical Report.

Tomas, G. H. R. P. (2011). Eucalyptus: Uma plataforma de cloud computing para qualquer tipo de usuário. Master's thesis, Universidade Federal de São Carlos, Campus Sorocaba.

Truong, H.-L. and Dustdar, S. (2010). Cloud computing for small research groups in computational science and engineering: current status and outlook. *Springer-Verlag*, 2010:75–91.

Ulteo (2012). Ulteo - open source enterprise virtual desktop and application delivery solutions. http://www.ulteo.com/. accessed 16-October-2011.

UMIC (2012). Iniciativa nacional grid. http://www.umic.pt/. accessed 15-April-2012.

Weiss, G. J. and Chuba, M. (2011). Hype cycle for server technologies, 2011. Technical report, Gartner.

Xen (2011). Xen. http://xen.org/. accessed 19-November-2011.

Yang, J. and Chen, Z. (2010). Cloud computing research and security issues. *Computational Intelligence and Software Engineering (CiSE), 2010*, International Conference on 2010:1–3.

Youseff, L., Butrico, M., and Silva, D. D. (2008). Toward a unified ontology of cloud computing. *2008 Grid Computing Environments Workshop, Nov. 2008*, 2008 Grid Computing Environments Workshop, Nov. 2008:pp.1–10.

Zhang, L.-J. and Zhou, Q. (2009). Ccoa: Cloud computing open architecture. *IEEE International Conference on Web Services*, Web Services, 2009. ICWS 2009. IEEE International Conference on 2009:607–616.

Zhang, S., Chen, X., Zhang, S., and Huo, X. (2010). Cloud computing research and development trend. *Second International Conference on Future Networks*, Future Networks, 2010. ICFN '10. Second International Conference on 2010:93–97.

Zhou, M., Zhang, R., Xie, W., Qian, W., and Zhou, A. (2010). Security and privacy in cloud computing: A survey. *Semantics Knowledge and Grid (SKG), 2010 Sixth International Conference on 2010*, 2010 Sixth International Conference on Semantics, Knowledge and Grids, Nov. 2010:pp.105–112.

Zhu, W., Luo, C., Wang, J., and Li, S. (2011). Multimedia cloud computing. *Signal Processing Magazine, IEEE 2011*, volume: 28 issue: 3:59–69.

Glossary

AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
AWS	Amazon Web Service
CaaS	Communication as a Service
CAL	Client Access License
CRM	Customer Relationship Management
DaaS	Data as a Service
DoS	Denial of Service
DMZ	DeMilitarized Zone
DNS	Domain Name System
EC2	Elastic Compute Cloud
EFD	Elastic Compute Cloud
ERP	Enterprise Resource Planning
FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
FTE	Full-time equivalent

HaaS	Hardware as a Service
IaaS	Infrastructure as a Service
IIS	Internet Information Services
iSCSI	Internet Small Computer System Interface
LAN	Local Area Network
LUN	Logical Unit Number
NAS	Network Attached Storage
NL-SAS	Nearline SAS
Qos	Quality of Service
REST	REpresentational State Transfer
ROI	Return On Investment
SAN	Storage Area Network
SaaS	Software as a Service
SAS	Serial Attached SCSI
SLA	Service Level Agreement
SOA	Service Oriented Architecture
PaaS	Platform as a Service
S3	Simple Storage Service (Amazon)
VMM	Virtual Machine Manager
VPN	Virtual Private Network

Appendix

6.2 EMC VNX

Figure 6.2: VNX5300 - Capacity Calculator

6.3 XEN Center

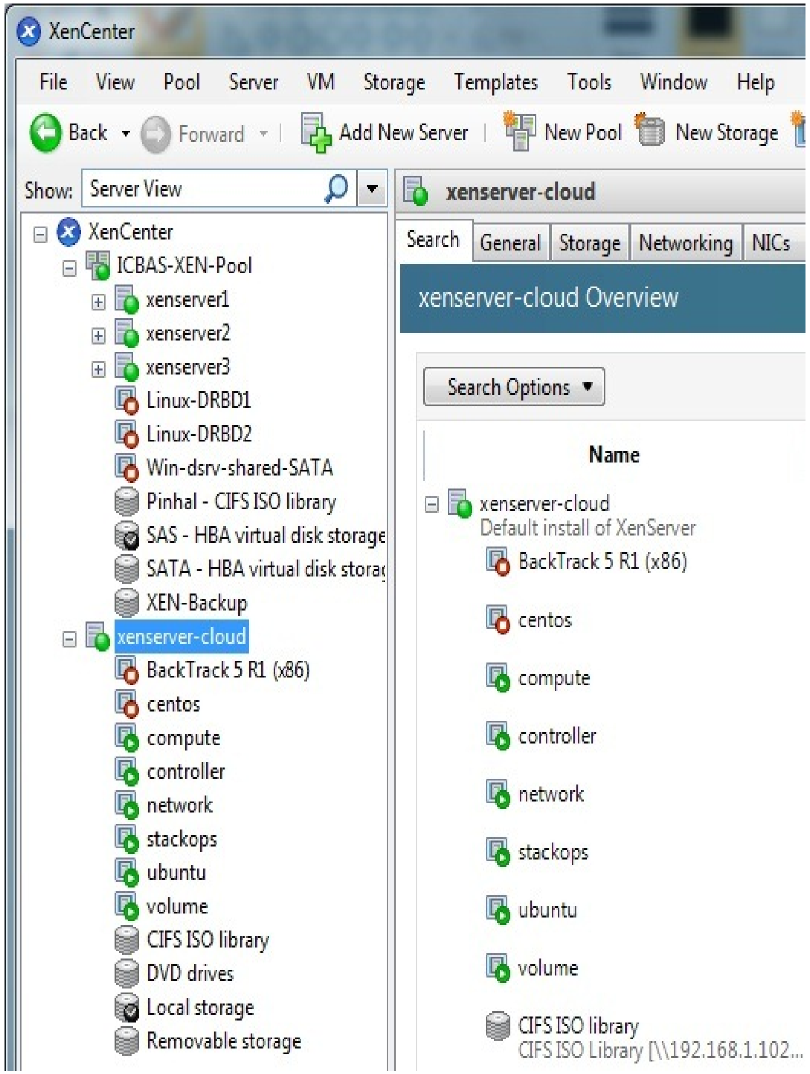

Figure 6.3: XenCenter - XCP

6.4 Microsoft Application Virtualization

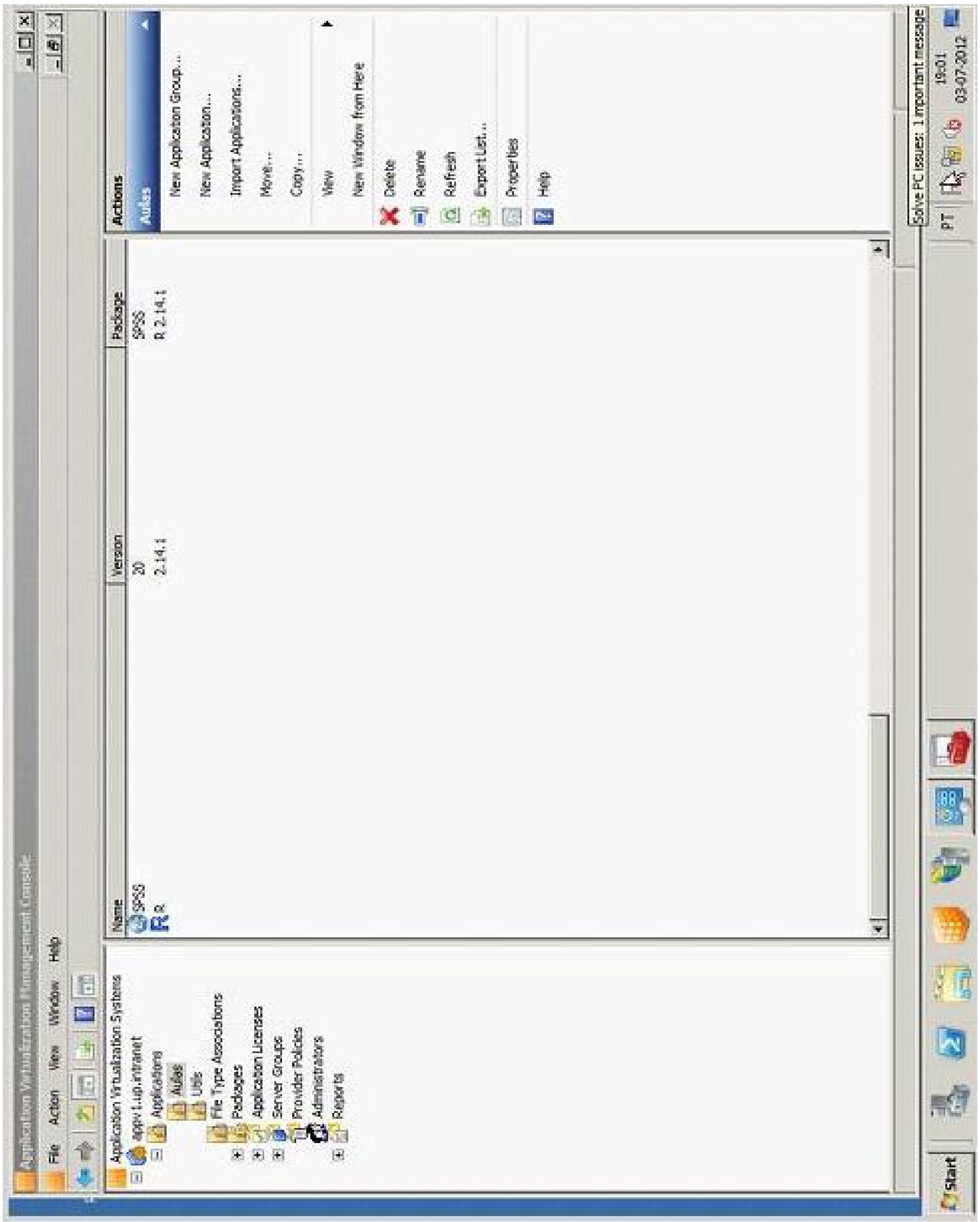

Figure 6.4: Microsoft Application Virtualization Server - Applications

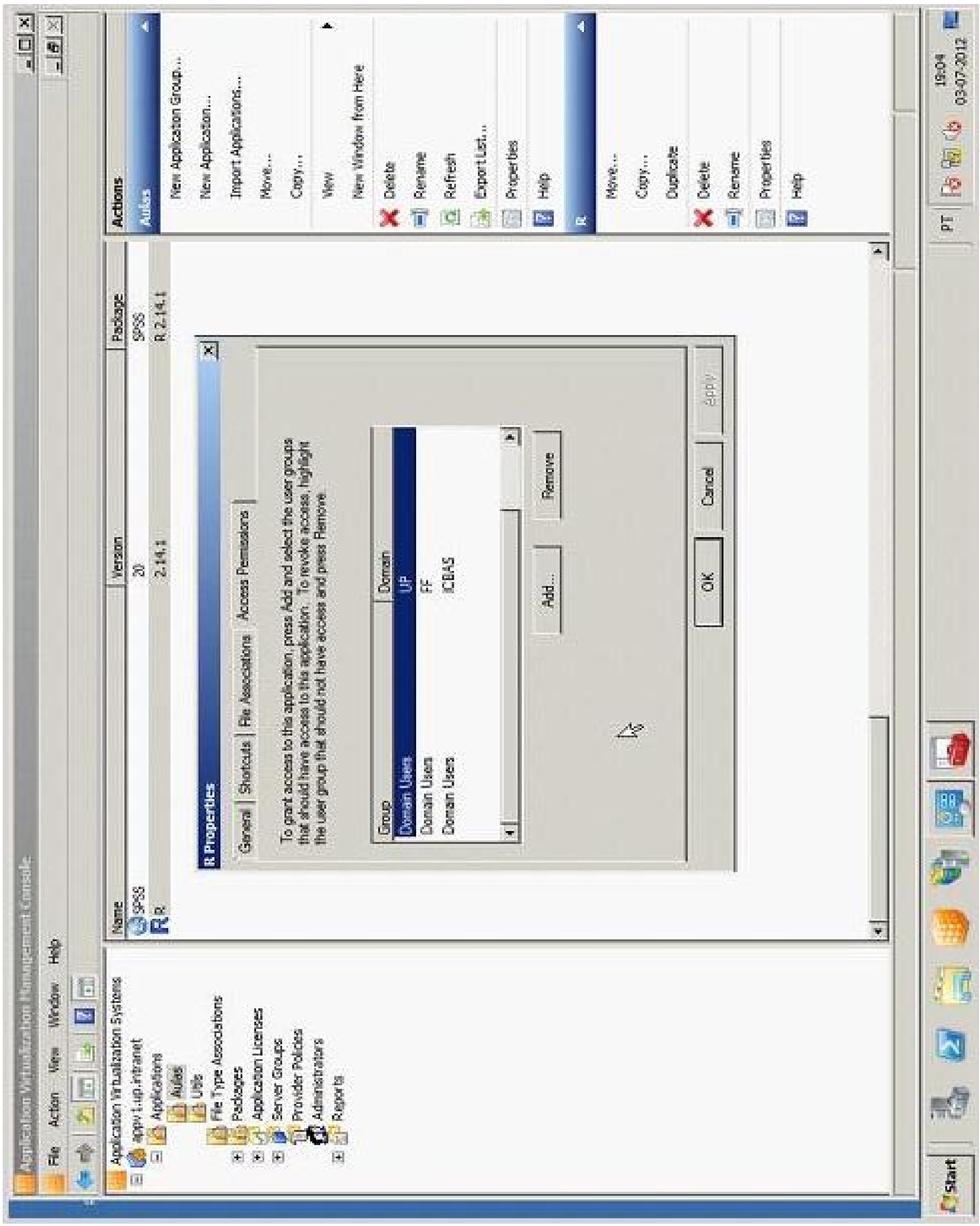

Figure 6.5: Microsoft Application Virtualization Server - Access Permissions

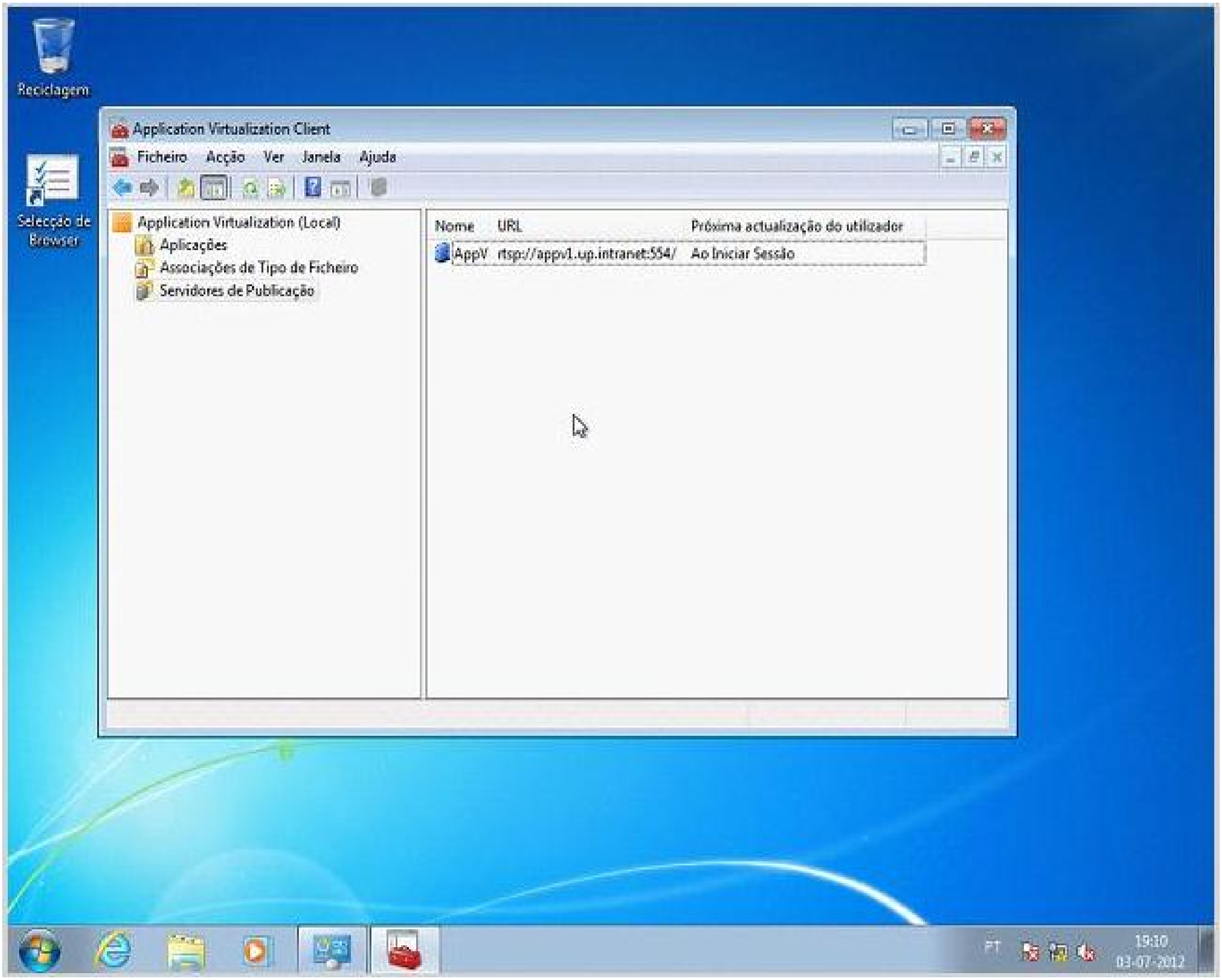

Figure 6.6: Microsoft Application Virtualization Client - Configuration

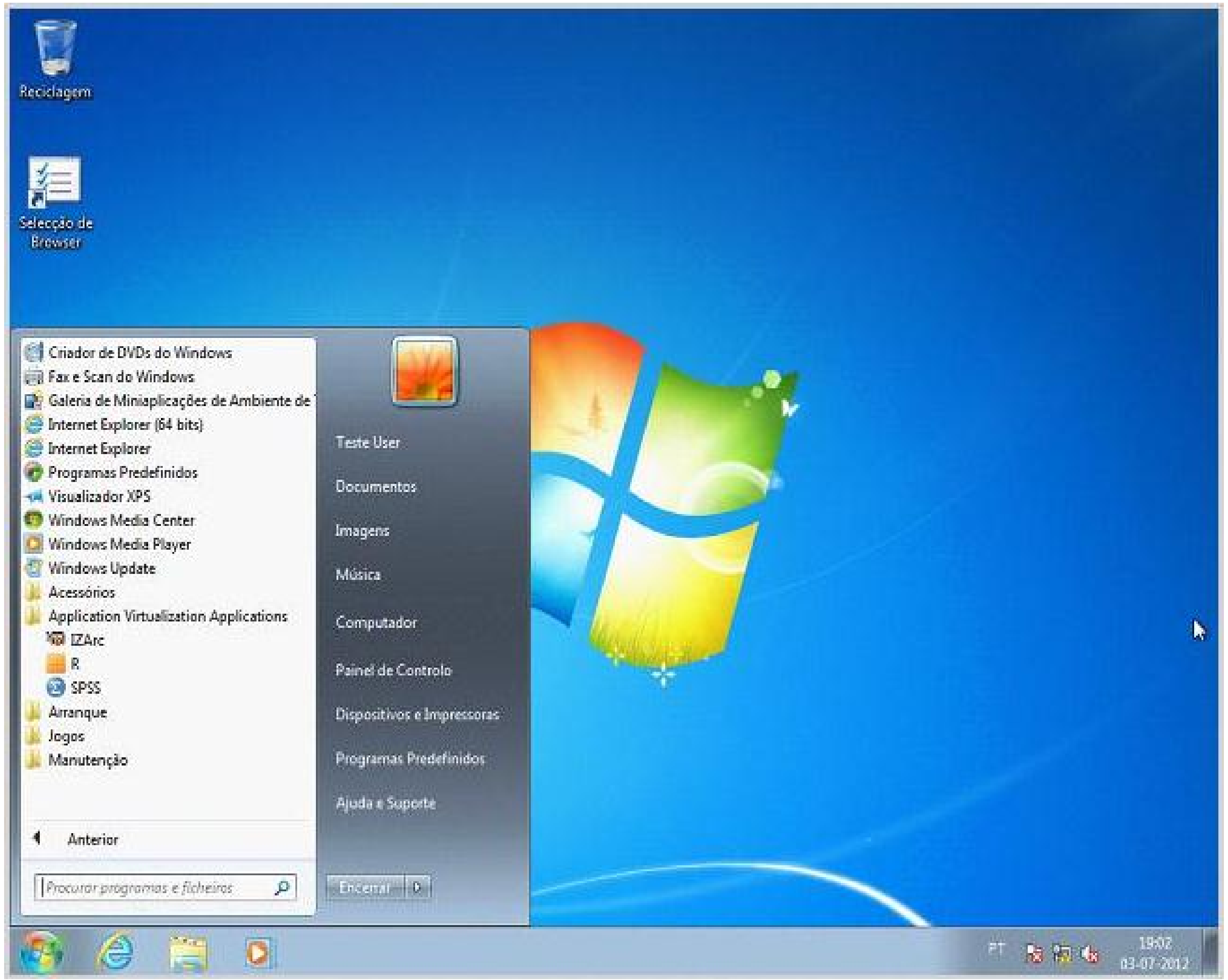

Figure 6.7: Microsoft Application Virtualization Client - Published applications

6.5 StackOps

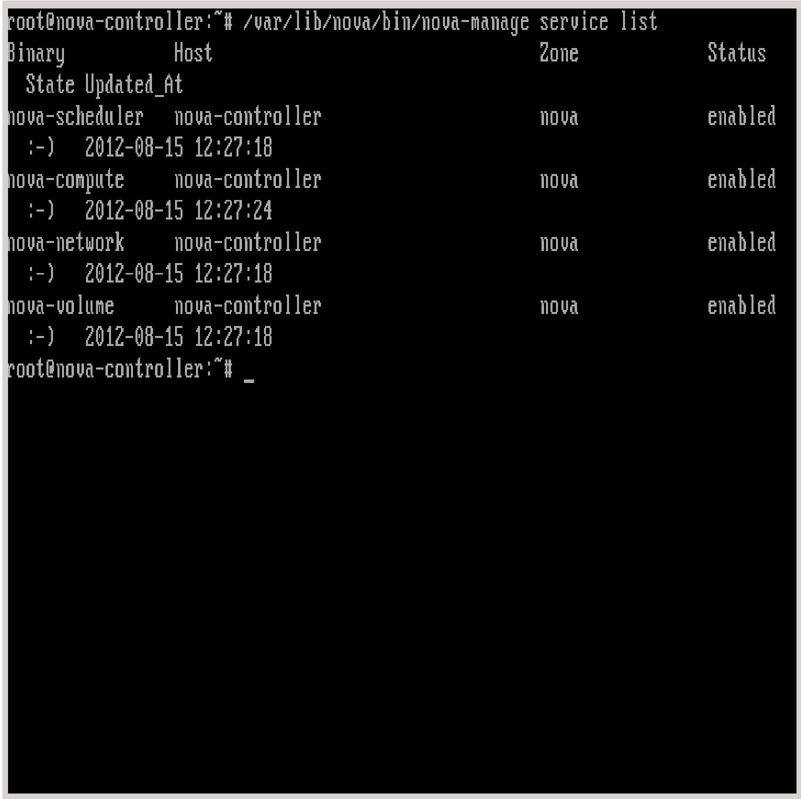

Figure 6.8: StackOps - Console

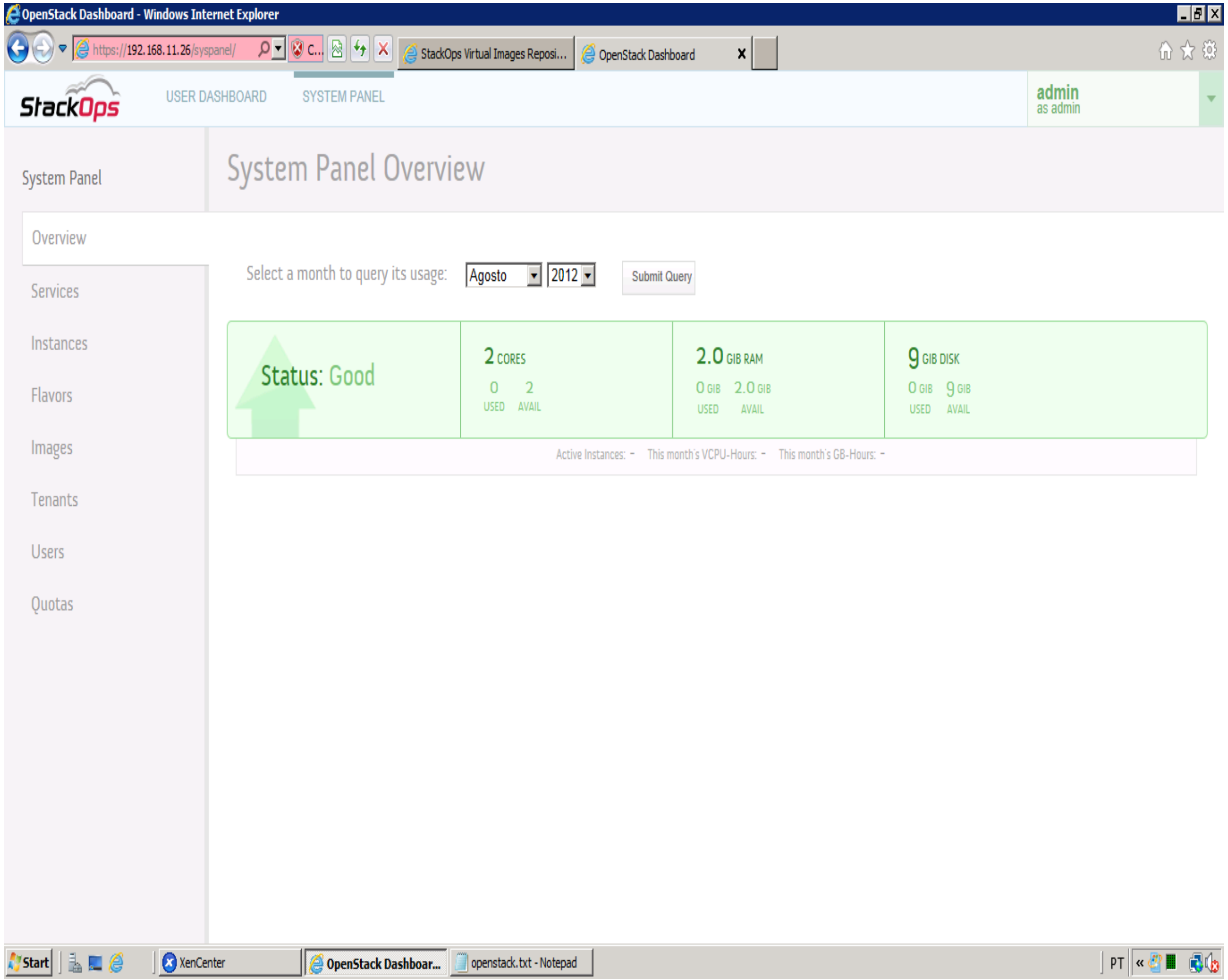

Figure 6.9: StackOps - Administrator

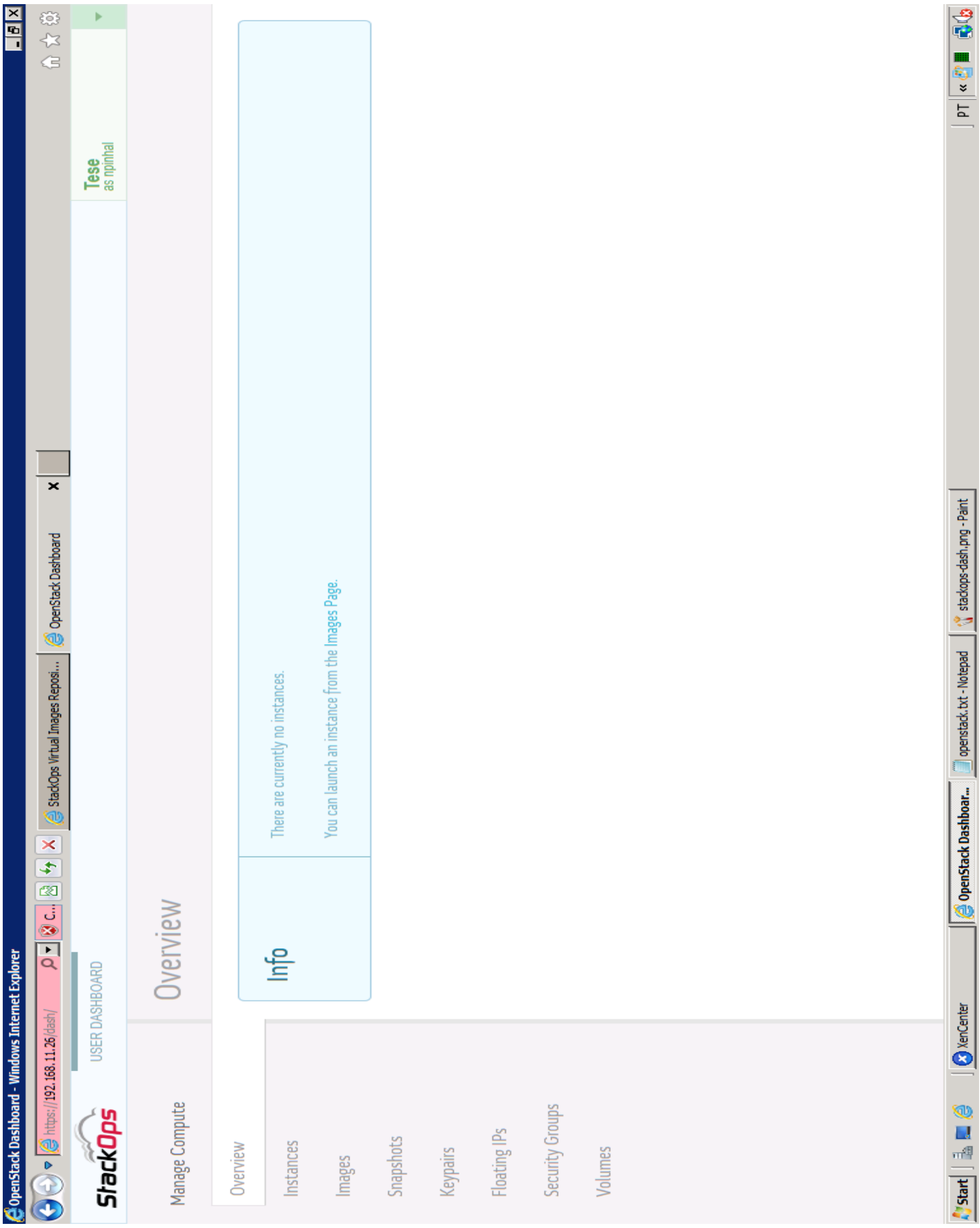

Figure 6.10: StackOps - User

6.6 Ulteo Open Virtual Desktop

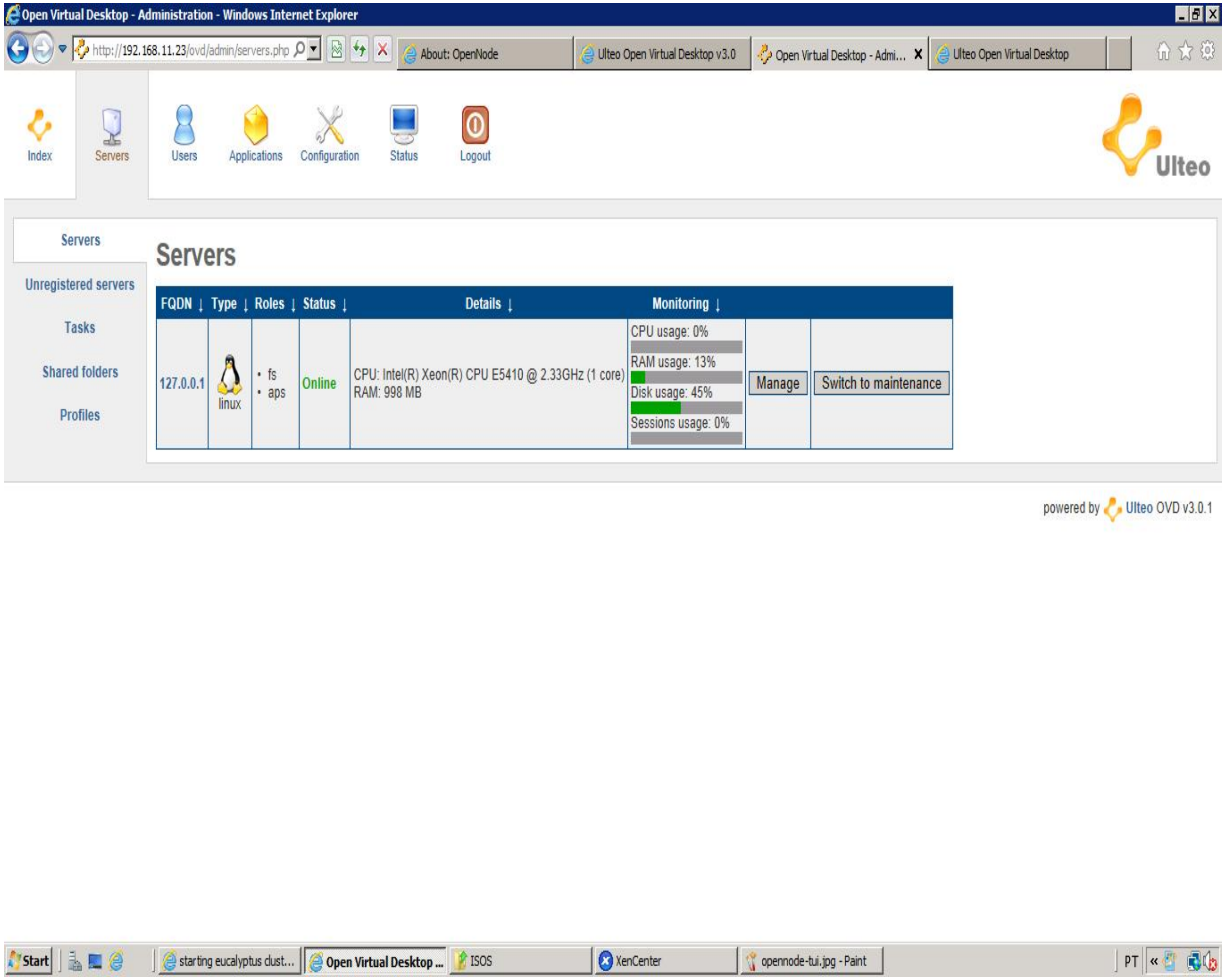

Figure 6.11: Ulteo Open Virtual Desktop - Administration

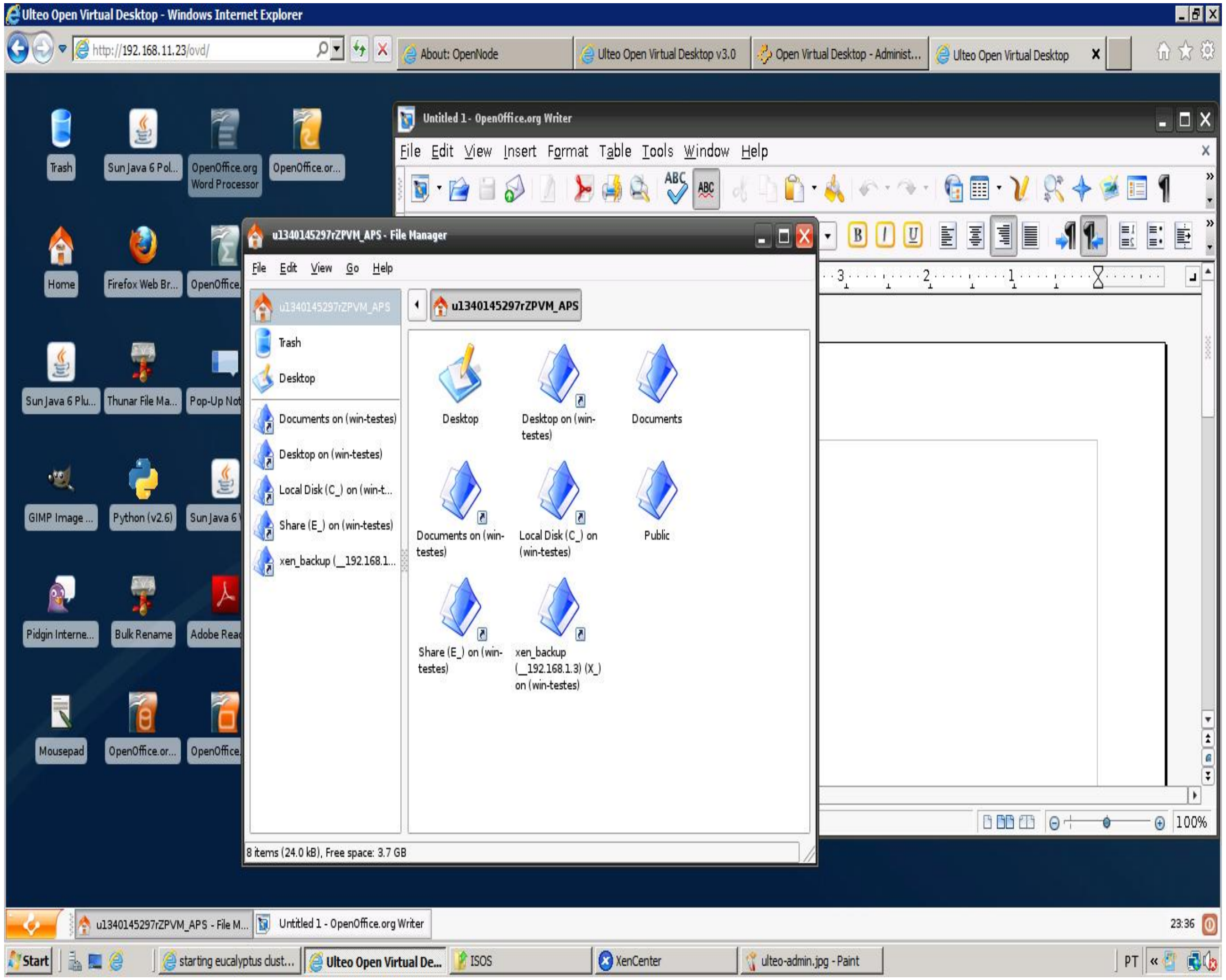

Figure 6.12: Ulteo Open Virtual Desktop - Desktop

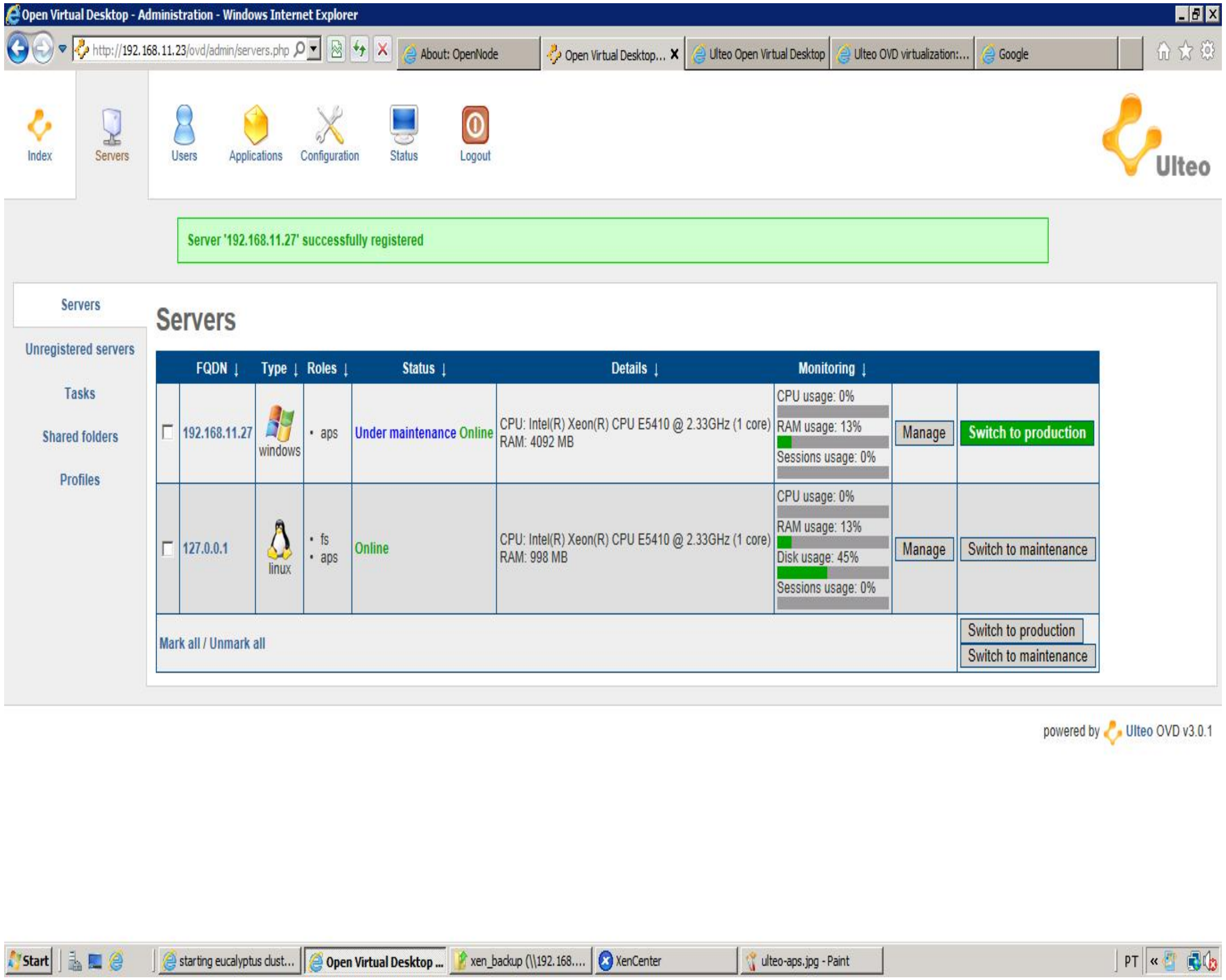

Figure 6.13: Ulteo Open Virtual Desktop - Application Server

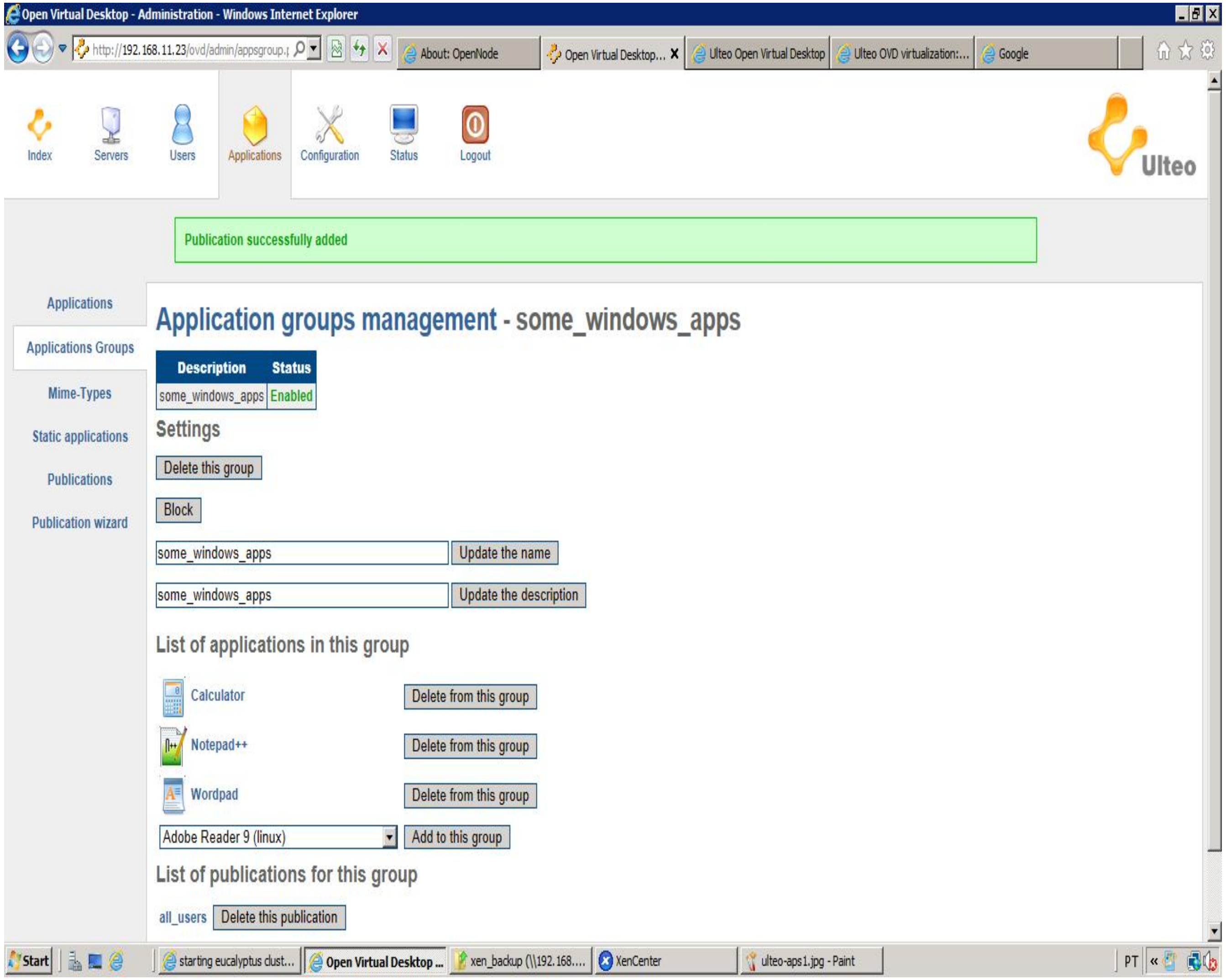

Figure 6.14: Ulteo Open Virtual Desktop - Application Server Management

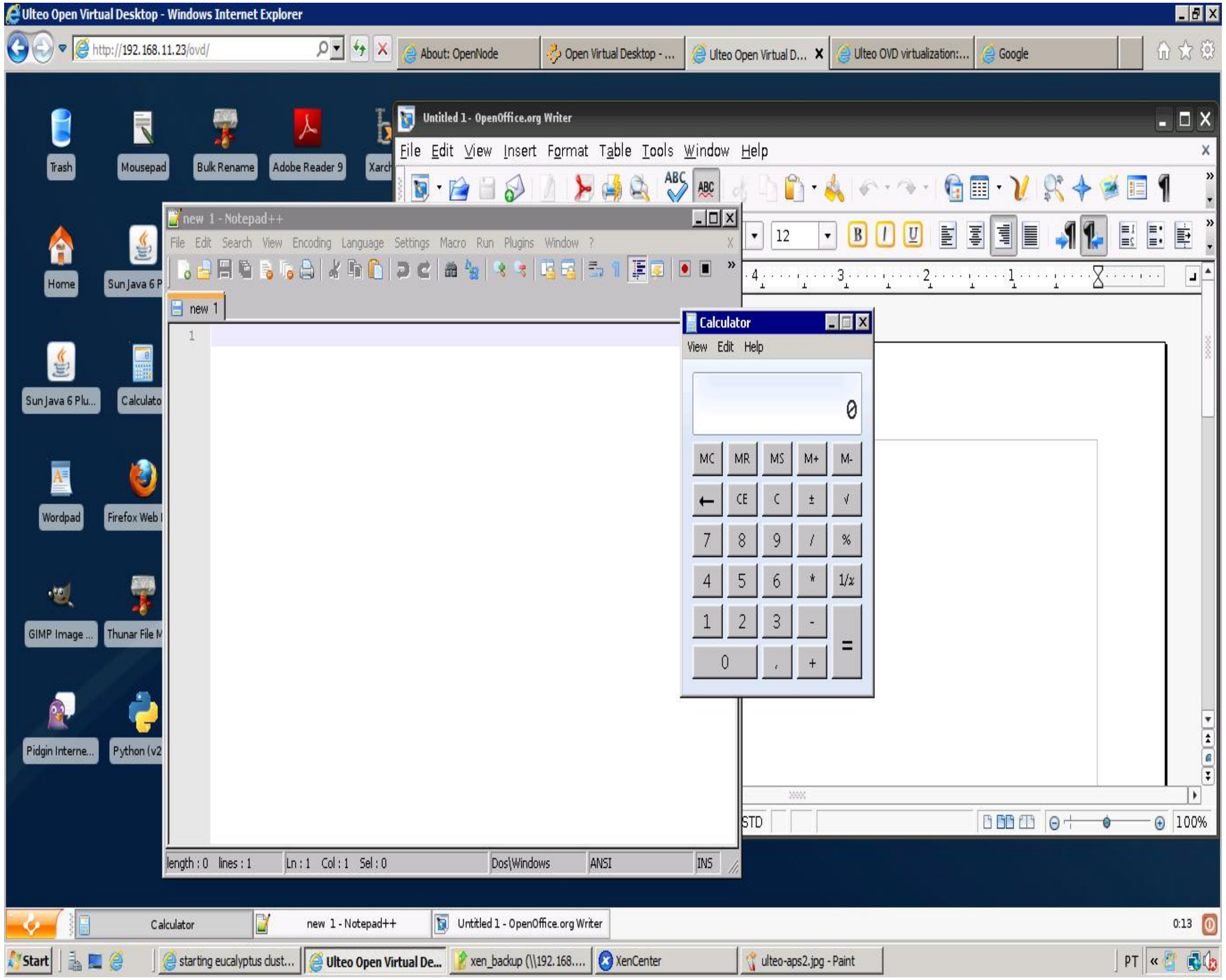

Figure 6.15: Ulteo Open Virtual Desktop - Desktop (Linux and Windows)

6.7 Hyper-V

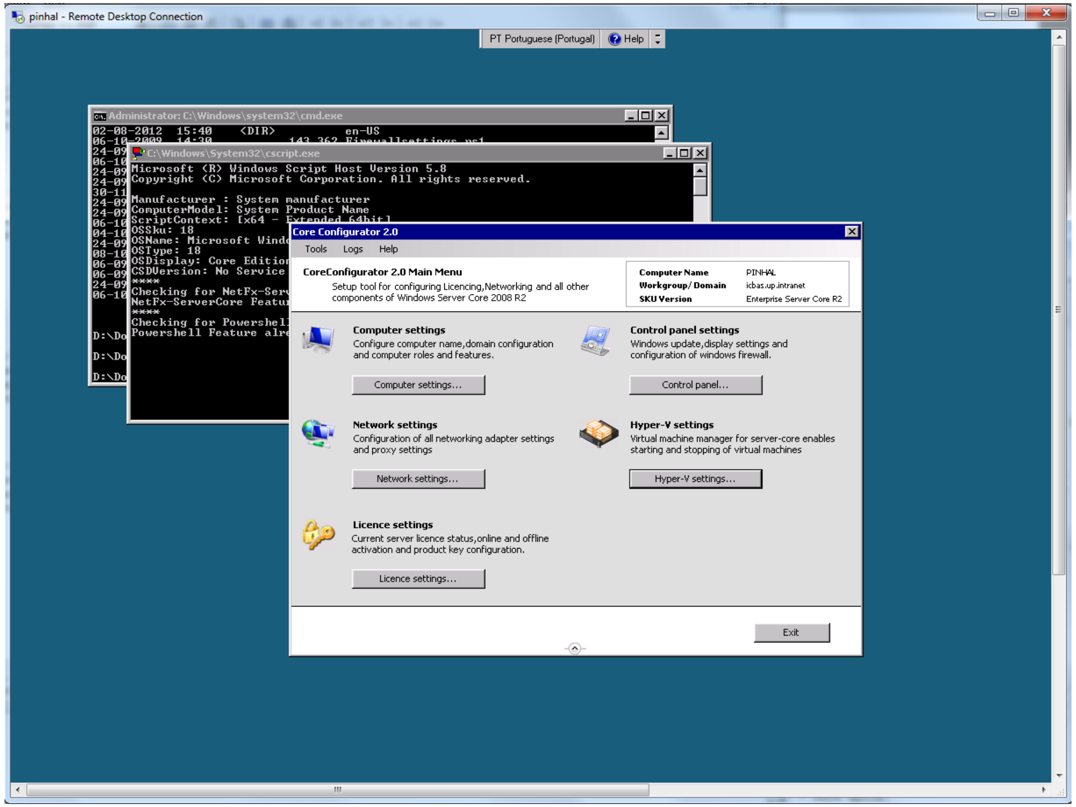

Figure 6.16: Hyper-V Server Core - CoreConfig

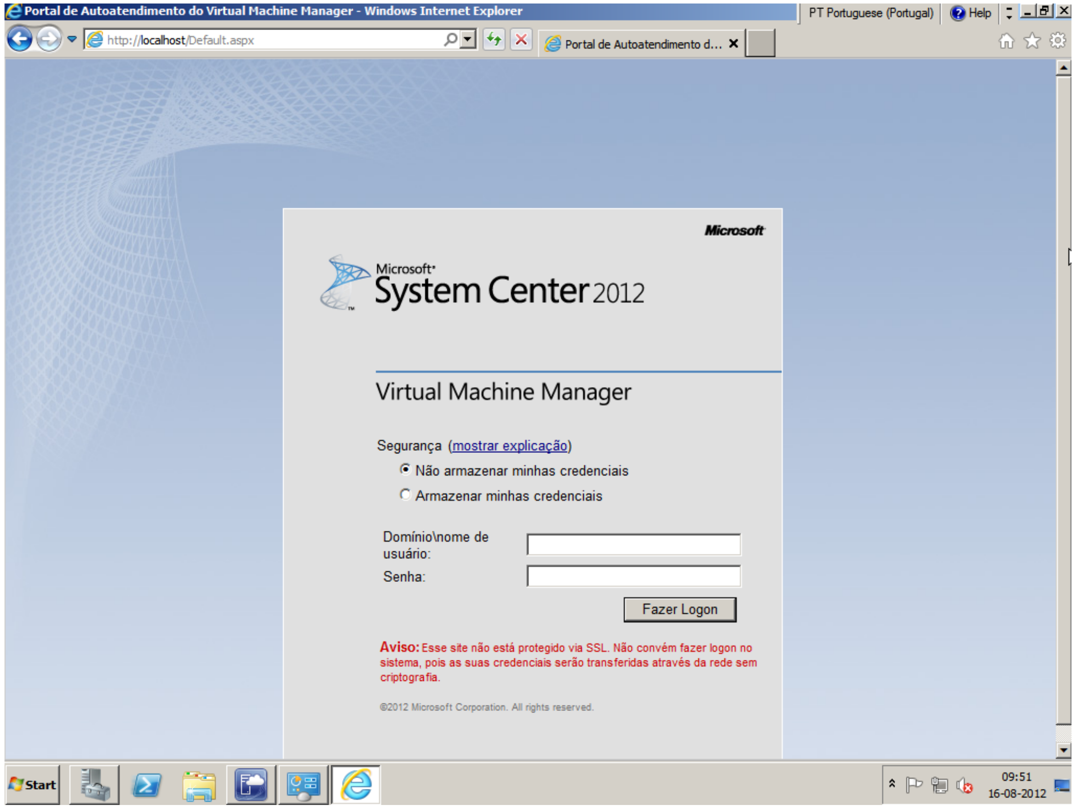

Figure 6.17: Hyper-V Self Service Portal - Login

Figure 6.18: Hyper-V Self Service Portal - Configuration