FEAA
FACULTATEA DE ECONOMIE
ŞI ADMINISTRAREA AFACERILOR
ŞCOALA DOCTORALĂ DE ECONOMIE
UNIVERSITATEA "AL. I. CUZA" IAŞI

Security of Electronic Document Management Systems in Cloud

PhD thesis summary

Scientific coordinator:

PhD Univ. Prof. Mircea Radu GEORGESCU

PhDs:

Natalia Suicimezov

Research field:

Economic Informatics

IAŞI, 2015

Introduction

We are in an uncertain, insecure era which raises many questions about the structure of society and of the individual, about how the aspects of life and of the work in organizations are treated. As consumers, we often adopt a heterogeneous mass of personal and business needs which have to be defined correctly, efficiently and reliably to yield predictable results, adapted to the working environment and information flows to and from it. In this context, technology enabled the emergence of some systems specific to human needs that best meet current challenges. The traditional "breathing space" has been replaced by an agile development of Cloud services, in which the road of information from the source to the issuer was shortened and grew into a unit within a much larger and well developed project. Cloud technology has deepened the belief that the IT sector is crucial and strategically linked to the performance of an organization, be it of personal nature (family), societal (human groups) or related to work (companies). We witness a shift in the way in which information is delivered to the final consumer. From concept to widespread use, congestion of information, which take the form of huge data centres, once dispersed, and with useful but irrelevant information links are translated into an integrated infrastructure, a "global car" that trades the so-called "Big Data". The high level performance means Cloud services, and analysts warn that a big storm is heading toward us. Similar to their forecasts and by drawing an analogy with the more common forecasts found in everyday environment - the weather ones - Grab your umbrellas and face the changes of the new Hyper-Cloud[1].

In search of the right instruments, to guide all IT activities in an organization towards an efficient and relevant forecast in the long term, we turned to workflows in the Cloud, referring to document management and analyzing the security of the entire set as a prerequisite and criticism for its explanation. We initiated this scientific approach in order to unify into a coherent structure the security of electronic document management systems in Cloud , covering the concepts listed above and developing a model of analysis of Cloud solutions in the pre-implementation phase and a plan to improve IT infrastructure capacity, starting from the back-up data security using virtualization solutions and SAN.

To determine the methodological framework corresponding to the content of this paper we consulted several studies on the *methodology of scientific research*, based on which we established *the research project.*

After drafting the project, the research continued with *documentation* phase, achieved both through *bibliographic documentation* and *through direct documentation of reality*. The current study is the result

[1] Georgescu, M., Suicimezov, N., Emerging Markets Queries in Finance and Business, Elsevier, IT governance in Cloud (2013), p. 4

of synthesizing a vast research literature, and in the first case, *the criteria* adopted for consultation and reading bibliographic sources were: actuality of the material, relevance and usefulness of the data for the research topic and credibility of the source. The selection and examination of the most current and relevant resources enabled the creation of a correct, complete theoretical foundation, but without an exhaustive character of research.

The retrieved and interpreted data were obtained from sources available at national and international level. The *documentation* was made based on current and representative scientific articles published in academic journals, in ISI indexed journals, in databases such as: Elsevier, Springer, ProQuest, IEEE Xplore Digital Library and Wiley. In addition, we studied online publications developed by providers of Cloud computing, document management and information security, as well as reports of research of world institutions of technology consultancy, such as Pierre Audoin Consulting, IDC, Gartner, Forrester, CSA . Also, for this paper, we have considered the conclusions and recommendations of the HPC and Cloud presentations, supported by international specialists in international conferences held in the US (HPC Convention Salt Lake City, HP Power Grid) or Europe. Materials included in this paper and which have contributed to the elaboration of theoretical chapters were properly cited during the thesis.

Simultaneously to the bibliographic documentation we also made a *direct documentation of reality* through participatory direct observation. The proposed objective was to correlate and confront the theoretical framework previously developed on the topic of security of electronic document management systems in Cloud with concrete situations existing in specialty practice or in the scope of these technologies. Following the analysis, configuration, implementation and use of Cloud and security services and of the direct work with document management solutions, we had access to valuable information for the current thesis, applicable in the achievement of the reference framework proposed in the last chapter or generation of recommendations and conclusions presented in the paper. Since Cloud computing and the workflow system in Cloud are evolving at a rapid pace with the changing needs of the customers, this paper reflects the author's opinion based on the experience in the field, the current situation and demonstrates the progress of research and its results.

Analysis techniques used for the *processing and analysis of information* obtained from documentation are: the study of documents, comparative analysis, content analysis, case study and systematization by grouping, making tables and graphical representation. The present research has both *theoretical and practical* character and can be the starting point of building a plan for the adoption of a Cloud solution or of development and personalization of the one presented in Chapter IV.

The structure of the paper follows the following objectives:

1. Objective 1: Analysis of the need to use new Cloud technologies – as a requirement of modern organizational management;

2. Objective 2: Worflow systems trend analysis, from classic to cloud;

3. Objective 3: Identification and description of the components of workflow system architecture in cloud;

4. Objective 4: Identification of dimensions of Quality of Services of workflow systems in cloud by studying the similarities and unique indicators of cloud services;

5. Objective 5: Analysis of security factors of information systems in Cloud;

6. Objective 6: Determination and description of the features of traditional information systems vs the ones based on Cloud Computing, through the specifics of each model;

7. Objective 7: Development of a reference framework for the implementation of a security service in cloud;

8. Objective 8: Proposing a hardware platform model that can be an implementation support for a cloud system, for the future migration of services in Cloud.

This paper can be used as a reference for the practitioners interested in the analysis, development and implementation of workflow systems principles in Cloud and for the operational management area interested in the proper execution of workflows at a business level. It can also be used as a reference in applying the principles of security and the security area in Cloud mainly focus on the SaaS model.

Electronic Document Management Systems

The workflow systems are rooted in automation of tasks, starting with 1970 to streamline organizational management. Lately, these systems were geared to automating complex business processes on a global scale and scientific applications. Many systems have been implemented on HPC infrastructures, clusters, peer-to-peer and grid computing.

One of the engines of this phenomenon is the increasing demand for large-scale applications, using modeling or forecasting both in e-business area and in the area of top sciences. Examples include the process of exchange of securities on the capital market, the online ticket booking in a travel agency, the process of identifying the pulsar in astrophysics, weather forecasting process in meteorology. These workflow applications require the resource support of a strong infrastructure (like HPC with advanced CPU units, vast memory space and network speed).

However, the issue of scalability and resource elasticity still exists in the IT conventional paradigm. Since many of these resources are independent and organized in a heterogeneous manner, the degree of scalability is very low. The phenomenon entails a high cost, almost impossible to allocate in a short period of time, in case of resource syncopes at peak times that would require attracting foreign aid, whatever its nature - human resources, material or knowledge. Simultaneously, since in the current paradigm of computing, workflow systems must maintain up and running its own resources, rather than use them to or

from third party providers, resource elasticity is poor. This is why most processing resources during off peak are inactive, leading to a low ROI - Return of Investment and a huge consumption of energy.

Impact of Cloud Computing in Organizational Management

In an attempt to identify an answer to the question "what is cloud computing?" online environment will provide a range of topics and content included in studies such as "white paper" on technical blogs, forums of IT specialists or commercial presentation sites. This wealth of confusing, heterogeneous, poor quality information only raises further questions on the topic. Transformation of information technology services into a commodity[2] entitles us to refer to holistic concept of 'cloud' formulated by Larry Ellison; in his view, the company "redefines the cloud to include everything that we already do." [3]

A summary of the latest market forecasts and estimates reveal that public investment (CAGR - compound annual growth rate) in the Cloud field will increase by 26.4% from 2011 to 2016, a figure that represents 17% of total IT spending in the forecast period[4]. At the end of this period, 46% of all global IT spending will be allocated to applications and platforms targeting the cloud. In the same vein, all Cloud technologies will converge and will reach the threshold of 107 billion dollars in public area, estimates IDC[5].

In the current economic and highly competitive environment, IT provides support and leads to dynamic growth of any organization. Current models used by business applications are expensive: the user should become the owner of the hardware and software infrastructure, while the technical assistance to users is difficult. Capital and operating costs are high, unpredictable and uncontrollable and cannot be linked dynamically and agile with the activities actually carried out, in line with market requirements. Many skills for the support of these processes are not part of the core of skills needed to conduct a business[6].

Information Security

Information security is a process to protect information and information systems from unauthorized access, use, disclosure, disruption, modification or destruction of data. The terms "information security", "cybersecurity" and "provision of information" are often used interchangeably.

[2] US Department of Commerce, National Institute of Standards and Technology. (2012, June) *Cloud computing: a review of features, benefits and risks, and recommendations for secure, efficient implementation (ITL bulletin for june 2012),* http://csrc.nist.gov/publications/nistbul/june-2012_itl-bulletin.pdf, p. 2, accessed on May 15, 2014

[3] Farber, D., *Oracle's Ellison nails Cloud computing*, http://www.cnet.com/uk/news/oracles-ellison-nails-cloud-computing/, accessed on December 15, 2014

[4] http://www.prweb.com/releases/2013/11/prweb11341594.htm, accessed on May 21, 2014, 9 pm

[5] http://www.idc.com/getdoc.jsp?containerId=prUS24298013, accessed on May 21, 2014

[6] Rimal, B. P., Jukan, A., Katsaros, D., Goeleven, Y.: *Architectural Requirements for Cloud Computing Systems: An Enterprise Cloud Approach*, Journal of Grid Computing, 9(1) (2011), p. 3-26

Information security refers to the confidentiality, integrity and availability of data, in whatever form - electronic or listed.

Essentially, ensuring information systems involves unique identification of threats and challenges that must be addressed accordingly. Finally, identification means selecting the security control and its introduction in a standard process, efficiently integrated, functional and operational with the requirements of the general system (reliability, affordability).

Information security includes three functions: access control, secure communications and personal data protection. Access control includes both the original entry and re-entry of the same participant or access of other participants. Important to remember is that a participant can be an individual or a process. Secure communication includes the transfer of information between any of the participants. Protection of personal data includes storage devices, processing units and even cache memory.

Even if Cloud technology is geared to provide a judicious use of resources - through virtualization technologies - and make the customer's work easier, they have security gaps[7].

All these Cloud's fundamental elements require security, which depends on and varies depending on the model of Cloud used, the way in which it is delivered and the character represented. Some of the fundamental security challenges are posed by data storage environmental security, data transfer, security of applications and security controlled by third parties involved in the project.

Treatment of applicability of IT technologies migrating to Cloud

Cloud Security Alliance (CSA) collects vendors, nonprofits and individuals to discuss current and future practice of providing information in the cloud. Cloud Standards website gathers and coordinates information about Cloud standards developed by global stakeholders. Open Web Application Security Project (OWASP) has created a list of top vulnerabilities of Cloud or SaaS-models, which is updated whenever new threats appear.

Because the clear IT trend of the last period is directed towards Cloud Computing solutions, the majority of small and medium companies analyzed it and they want to adopt it in their own infrastructure. The challenge in choosing the right solution implies the search and testing of adjacent IT solutions, which would work together as a unified solution that provides functionality required by the business environment (eg management, monitoring, configuration, audit, reporting solutions, etc. Of Cloud Computing technologies).

On the other hand, companies providing IT services are put in the situation to innovate, to expand its IT solutions portfolio, depending on market trends and foremost, on customer needs. This is the case for the company we made the first analysis, an analysis of the product in terms of services offered and its

[7] Seccombe A, Hutton A, Meisel A, Windel A, Mohammed A, Licciardi A, et al. "*Security guidance for critical areas of focus în Cloud computing*", v2.1. CloudSecurityAlliance, 2009, p. 25

requirements. This analysis was conducted to discover and describe the skills, methods of implementation and customization possibilities of *Shadow IT* service.

In the current context of a services market oriented to Cloud Computing, adopting a new strategy, which offers customers additional flexibility in Cloud Computing services, allowing managing and securing Cloud environment based on business needs of its clients, is a major advantage.

The suggestion of a *hardware platform model – future support for the implementation of a Cloud system* is a suggestion of advanced architecture for a new SAN solution, Hyper-V Backup, implemented in a Corporate infrastructure. As every project has clearly defined workflows, this proposal is part of the IT infrastructure improvement.

The reference framework, in this case, came after the organization lost valuable information because the only back-up solution they had was using a memory stick, more over, backups were performed on DC servers with the role of File Servers which means that:

• Company Details were not classified in an objective manner;

• IT infrastructure was not properly scaled to the company's requirements;

• No importance was paid to data security.

The plan was based on the current analysis of the infrastructure, identification of risks, challenges, and based on the goals of solving the problems we've outlined a proposal for a stable, sustainable solution, which can be on the medium and long term, based on a Cloud infrastructure.

In the current organizational context, the information is the "treasure" of enterprises, and when we migrated all the treasure in IT advanced technology the commitment that any organization must assume is going at the same pace with innovation.

In any project of implementation of an IT solution in the infrastructure of an organization, they want to retain current infrastructure, hence the challenge of integrating them. Based on the latest information available in the research literature and practice, we have outlined scenarios through which we detailed the implications of replacement of SAN, Hyper-V and Backup solution. Given the central theme of this research, we analyzed based on a model proposed in this respect the particularities of the integration process of these solutions into IT infrastructure.

CONCLUSIONS

Information surely remains today the most valuable resource, as it has been always considered by civilization since the beginning of the first societies. We can move forward with this statement, claiming that information is a genetic fingerprint of all living organisms that supports the development, survival and adaptation to the environment. A critical factor of information is its ability to not be vulnerable, public and accessible to everyone and kept as a strong point of an entity, be it single structure or organization of individuals. In this way, the information retains the intrinsic character of value, by the very fact that only

allows certain players to own and use it. In today's society, information is messages, ideas, documents, people, and its importance only depends on the correct ways of communication and transfer to become effective to operate and maintain.

This translation from micro to macro raised one of the most significant risks to information - security and preservation of the character of being "valuable".

The time of occurrence of Cloud technology marked a change in the way in which information travels, is retained and archived. The widespread use of Cloud technology has brought to the fore the feature of information security, of workflows ethics, of fairness of massive displacement of electronic documents between parties and a safer environment.

This paper can be used as a reference for practitioners interested in the analysis, development and implementation of workflow systems in the Cloud and for the operational management area, interested in the proper execution of workflows at a business level. It can also be used as a reference in applying the principles of security and the security area in the Cloud mainly focus on the SaaS model.

As regards own contribution, the first element is the analysis and delimitation of the conceptual framework. In this regard, we went through a process of selection, confrontation, filtering, synthesizing, processing and reasoning of the information provided by the research literature. Our study involved the description of the document management systems, of Cloud computing technology and of their security in order to develop complete and integrated architecture solutions. Our own way of approaching the topic is, moreover, another original element contributing to the strengthening and development of the current state of knowledge related to the topic addressed.

Also, the theoretical framework presented in this paper is filled with examples and concrete situations in specialty practice. The statistics in different areas of interest, market studies and practical examples confer an applicative character to the thesis. The usefulness of the scientific endeavour is obvious through the timeliness of the paper from a commercial, market perspective, companies turning to Cloud, secure workflows and integrated architecture. In fact, the whole research was developed in response to the questions faced by companies when they want to know the level of security of electronic document management services in the Cloud. On the other hand, the paper can be seen as a tool for analyzing a system for electronic document management in the Cloud from the point of view of security, or as a practical guide for achieving the migration process to Cloud and of integration of classic systems with cloud technology.

The effective work with cloud solutions, their implementation and management allowed us to have the practical and expertise perspective needed to strengthen this scientific material. This practical character, combined with scientific rigor has led to the debate and validation of the findings in specialized conferences, becoming the topic of own scientific articles listed in the next section of the thesis, entitled

"Dissemination of research results". As a conclusion, we emphasize that this information is relevant both for IT researchers and for stakeholders involved from a commercial point of view in document management systems and cloud, as well as for decision makers from companies, users and practitioners of these technologies. The security issue will be for a long time a challenge and an opportunity for the business environment, especially by analyzing the results obtained by Cloud technology and its adoption rate globally.

Regarding the limits of the research, we have not proposed to deal exhaustively with the topic chosen and complementary themes. This research is carried out as a result of an analysis oriented towards the technologies actually used by companies and the problems they face and involved the selection of research papers and information from the current economic reality. We note that this current character of information and of bibliographic source also restricted the research area. The research was focused on its applicability and usefulness in the case of small and medium enterprises and less to theorizing concepts, which can constitute another limitation of research. We can also add that another limitation of the research is given by the access to information, a phenomenon also mentioned in the paper, which made certain study or specialized materials relevant to the research topic not accessible to the end user.