Survey on Data Security Issues in Cloud Computing

[1]Jayesh Dabhi , [2]Prof. Zisan Y. Noorani
LD College of Engineering
[1]djaysh.r@gmail.com, [2]er_zishan@yahoo.com

***Abstract*— As we all know that every coin has two sides, cloud computing is nothing different. Despite of having many advantages and enterprise applications there are many disadvantages. Cloud computing allows its user to store data in large amount. User can access those data as and when required from anywhere because Cloud computing is rest on internet. Because of that it faces many security issues such as privacy, data security, confidentiality, authentication etc. This paper describes succinct but all-round survey on data security issues related to Cloud computing across all levels of data life cycle.**

***Keywords*— Cloud Computing, Access control, Cloud Computing security, Data security**

I. INTRODUCTION

We can define Cloud Computing as "Internet computing" because everything in cloud computing resides on internet and user can access those data or services online and pay according to whatever they use. In other words, it is a model consisting of information processing, storage and delivery where physical resources are given to user as and when demanded. User doesn't have to purchase any networking equipment in real because they can get these resources from a cloud provider whenever they want as an outsourced service. It can be defined as "managing the available resources and information as services over the internet whenever user demands for them." [1]

II. BASIC CLOUD COMPUTING THEORY

Cloud Computing

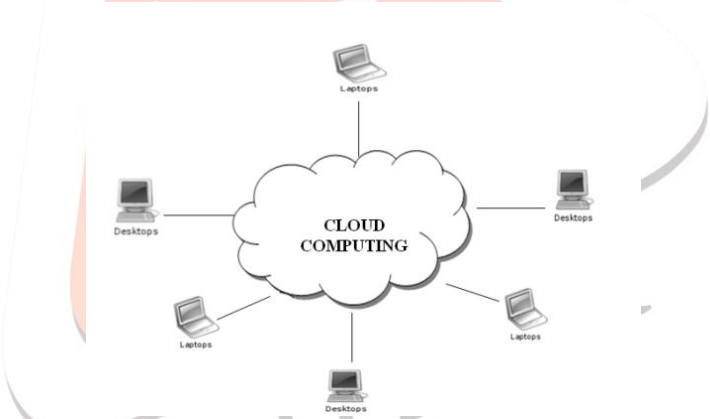

Figure (1) Cloud Computing [1]

Characteristics

Following are some of the characteristics of cloud computing:

- **Shared Infrastructure:** The environment of this cloud uses an imposing software model. That model allows cloud to share its physical services and storage and networking capabilities among various users. The cloud infrastructure is to find out most of the available infrastructure across multiple users.
- **Network Access**: User can access these Cloud services over a network by the use of standards based APIs from devices like computers and mobile devices.
- **Handle Metering:** Providers of these Cloud services always do store the information of their clients because they have to manage and optimize the service and also to provide reporting as well as billing information. Because of this, customers are payable only for those services which they have actually used during that billing period.[1]

Service Models

There are three types of models do exist for providing services of cloud. These three models are also known as the "**SPI Model** (Software, Platform and Infrastructure) ".

- **Software as a Service (SaaS):** Customers obtain the facility to access and use an application or service that is hosted in the cloud. As an example 'Salesforce.com', where necessary information for the interaction between the consumer and the service is hosted as part of the service in the cloud.

- **Platform as a Service (PaaS): To** gain access to the platforms Customers have to enable them to organize their own software and applications in the cloud.
- **Infrastructure as a Service (IaaS):** The facility provided to the customer is to lease some fundamental computing resources such as processing, storage etc. The customer has control over operating systems, storage, deployed applications but it does not manage or control the basic cloud infrastructure.[1]

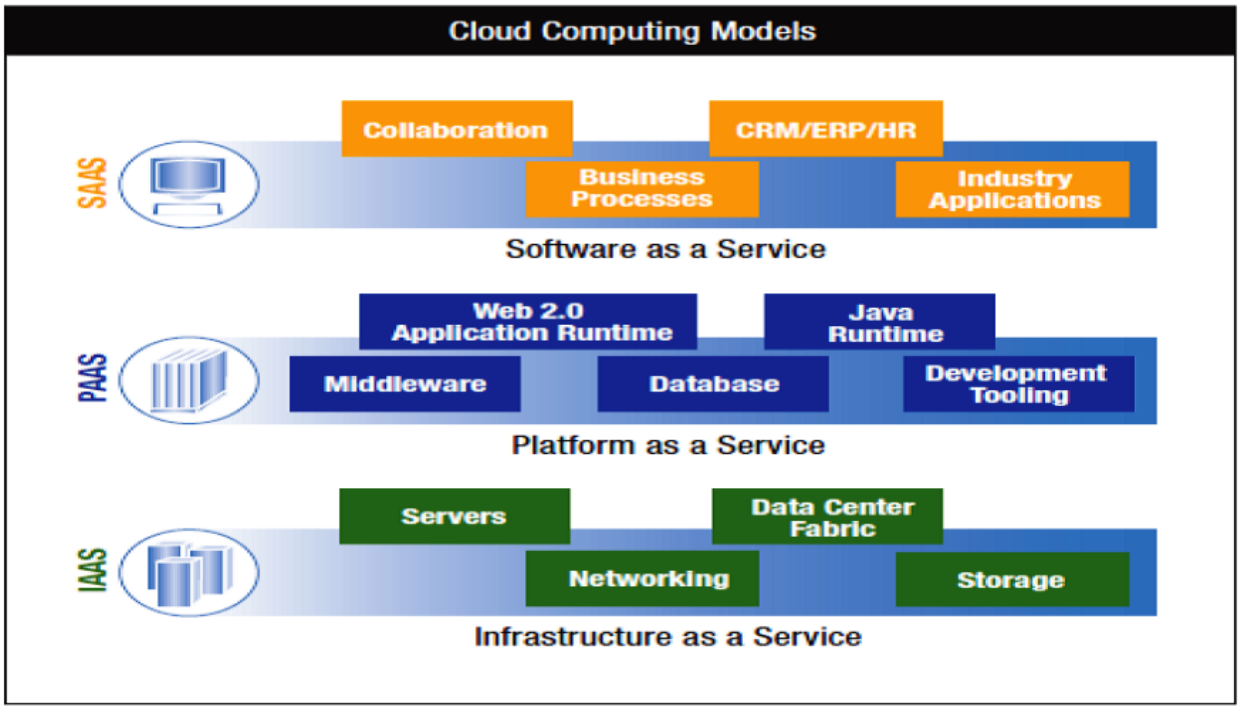

Figure (b) Computing Service Delivery Models[2]

Deployment models

There are three primary ways in which cloud services can also deployed which are described below:

- **Public cloud**: In Public cloud, customers can access web applications and services over the internet. Each individual customer has its own resources which are dynamically provided by a third party vendor (cloud providers).These providers facilitate multiple customers from multiple data centres, manage all the security measures and provides hardware and infrastructure for the cloud customers to operate. The customer has no idea about how the cloud is managed or what infrastructure is available. Customers of Public Cloud services are considered to be untrusted.
- **Private cloud**: In private clouds customers has complete control over that how data is managed and what security measures are in place while data processing in cloud. The customers of the service are considered "trusted." Trusted customers of service are those who are considered to be part of an organization including employees, contractors, & business partners.
- **Hybrid Cloud**: Hybrid Clouds are a combination of public and private cloud within the same network. Private cloud customers can store personal information on their private cloud and use the public cloud for handling large amount of processing demands.[1]

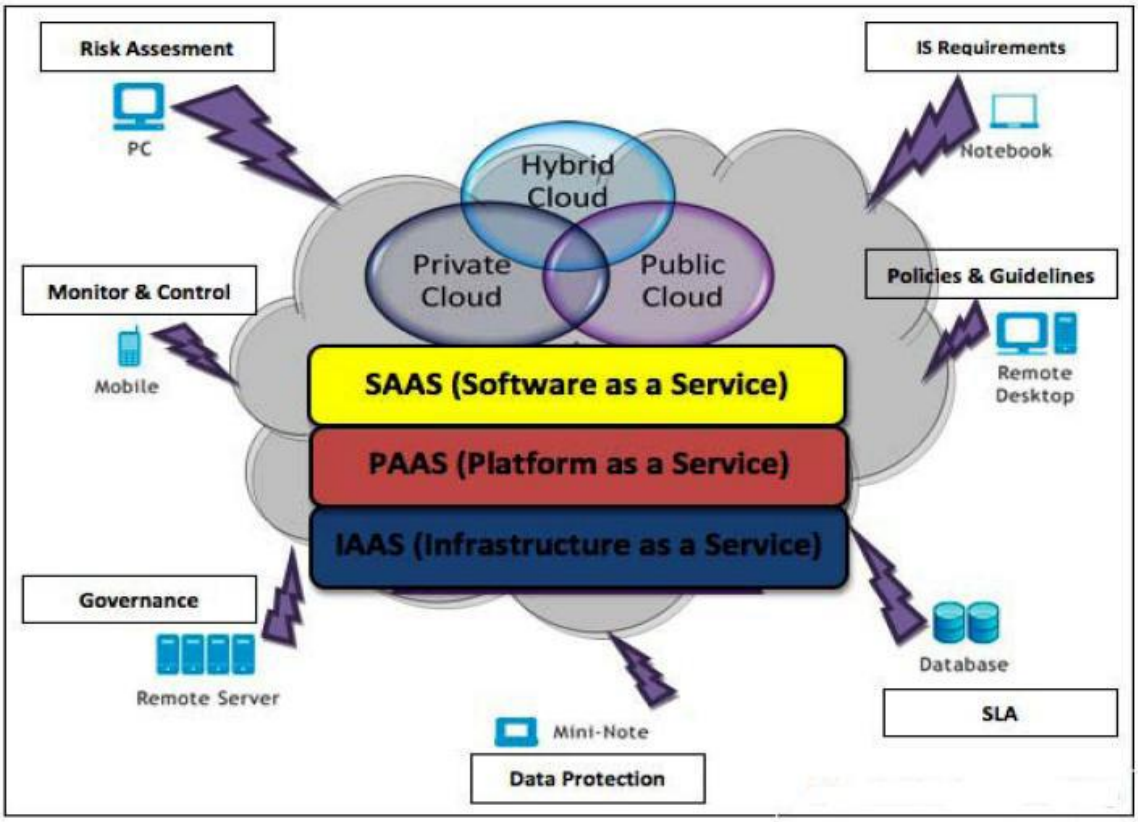

Figure (c) Cloud Deployment Model[2]

Security control measures in cloud are similar to ones in traditional IT environment. As multi-tenant characteristic, service delivery models and deploy models of cloud computing, compared with the traditional IT environment, however, cloud computing may face different risks and challenges. [3]

Traditional security issues are still present in cloud computing environments. But as enterprise boundaries have been extended to the cloud, traditional security mechanisms are no longer suitable for applications and data in cloud. Due to the openness and multi-tenant characteristic of the cloud, cloud computing is bringing tremendous impact on information security field[3]:

1. Due to dynamic scalability, service abstraction, and location transparency features of cloud computing models, all kinds of applications and data on the cloud platform have no fixed infrastructure and security boundaries. In the event of security breach, it's difficult to isolate a particular physical resource that has a threat or has been compromised.
2. According to the service delivery models of cloud computing, resources cloud services based on may be owned by multiple providers. As there is a conflict of interest, it is difficult to deploy a unified security measures;
3. As the openness of cloud and sharing virtualized resources by multi-tenant, user data may be accessed by other unauthorized users.
4. As the cloud platform has to deal with massive information storage and to deliver a fast access, cloud security measures have to meet the need of massive information processing.

This paper describes data security and privacy protection issues in cloud. This paper is organized as follows: Section II gives a brief description of what exactly cloud computing security-related issues are. Section III discusses data security and privacy protection issues associated with cloud computing across all stages of data life cycle. Section IV shows current solutions for data security and privacy protection issues in cloud. Section V summarizes the contents of this paper. Section VI describes future research work.[3]

III CLOUD COMPUTING SECURITY ISSUSES

A. *Cloud Computing Security*

Cloud Computing Security's definition as per Wikipedia [4] "*Cloud computing security (sometimes referred to simply as "cloud security") is an evolving sub-domain of computer security, network security, and, more broadly, information security. It refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.*" Here we need to clear that cloud computing security is not cloud-based security software products such as cloud-based anti-virus, anti-spam, anti-DDoS, etc.

B. *Security Issues Associated with the Cloud Computing*

There are so many security issues associated with cloud computing which can be grouped into any number of dimensions.

According to Gartner [5], before making a choice of cloud vendors, users should ask the vendors for seven specific safety issues: Privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support and long term viability. In 2009, Forrester Research Inc. [6] evaluated security and privacy practices of some of the leading cloud providers (such as Salesforce.com, Amazon, Google, and Microsoft) in three major aspects: Security and privacy, compliance, and legal and contractual issues. Cloud Security Alliance (CSA) [7] is gathering solution providers, non-profits and individuals to enter into discussion about the current and future best practices for information assurance in the cloud. The CSA has identified thirteen domains of concerns on cloud computing security [8].

S. Subashini and V. Kavitha made an investigation of cloud computing security issues from the cloud computing service delivery models (SPI model) and give a detailed analysis and assessment method description for each security issue [9]. Mohamed Al Morsy, John Grundy and Ingo Müller explored the cloud computing security issues from different perspectives, including security issues associated with cloud computing architecture, service delivery models, cloud characteristics and cloud stakeholders [10]. Yanpei Chen, Vern Paxson and Randy H. Katz believed that two aspects are to some degree new and essential to cloud: the complexities of multi-party trust considerations, and the ensuing need for mutual auditability. They also point out some new opportunities in cloud computing security [11].

According to the SPI service delivery models, deployment models and essential characteristics of cloud, there are security issues in all aspects of the infrastructure including network level, host level and application level.

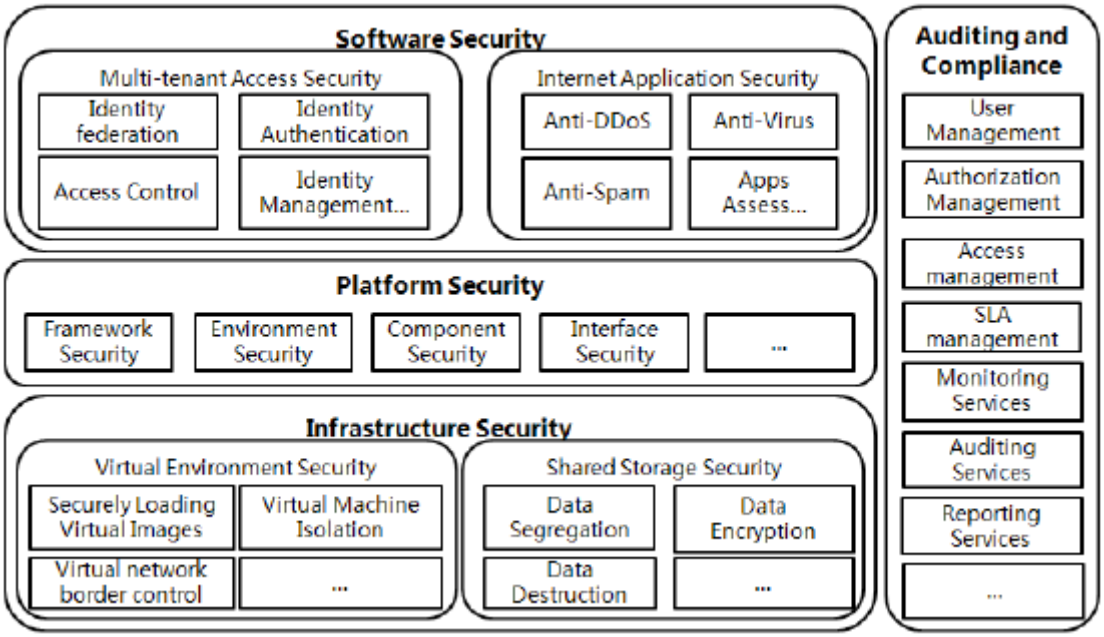

Figure (d) Cloud computing security architecture[3]

IV DATA SECURITY ISSUES

The content of data security in cloud and the traditional data security are similar. It is also involved in each stage of the data life cycle. Characteristics such as openness and multi-tenant of the cloud, these content of data security has its particularities.

The next sections analyze data security issues in cloud around the data life cycle[3].

1. DATA LIFE CYCLE

The entire process from generation to destruction of the data is called Data life cycle. This cycle is divided into seven stages as shown in the figure below:

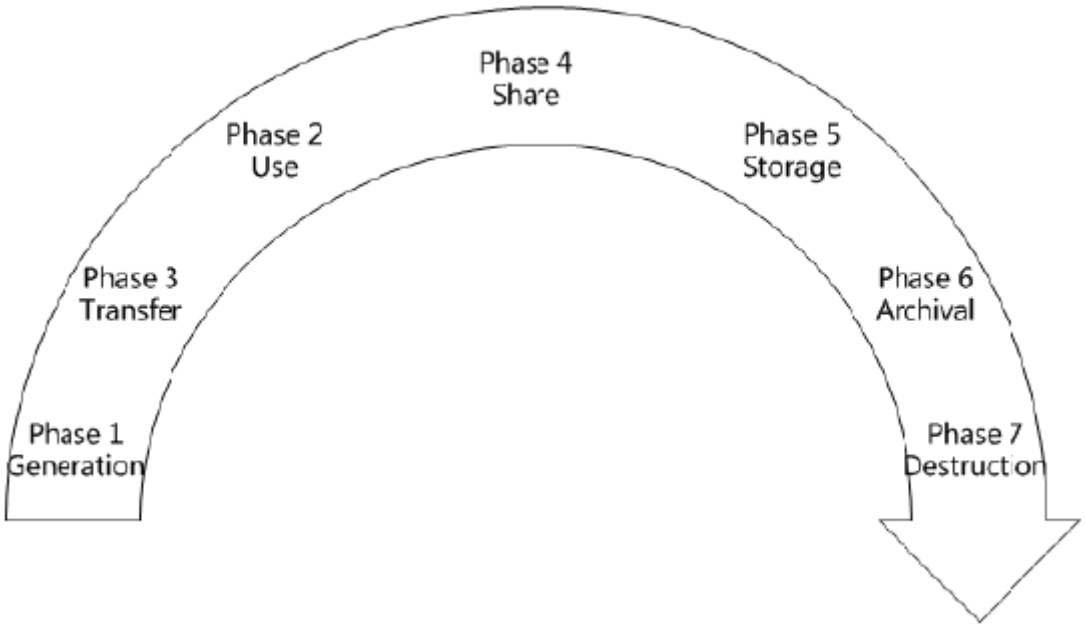

Figure (e). Data life cycle

2. Data Generation

It is involved in the data ownership. In the traditional IT environment, generally users or organizations own the data and manage them. But while migrating data into cloud, we should consider that how to maintain the data ownership. For personal private information, data owners are entitled to know what personal information being collected, and in some cases, to stop the collection and use of personal information.

3. Transfer

Within the enterprise boundaries, data transmission usually does not require encryption, or just have a simple data encryption measure. For data transmission across enterprise boundaries, both data confidentiality and integrity should be ensured in order to prevent data from being tapped and tampered with by unauthorized users. In other words, only the data encryption is not enough. Data integrity is also needed to be ensured. Therefore it should ensure that transport protocols provide both confidentiality and integrity. Confidentiality and integrity of data transmission need to ensure not only between enterprise storage and cloud storage but also between different cloud storage services. In other words, confidentiality and integrity of the entire transfer process of data should be ensured.

4. Use

For the static data using a simple storage service, such as Amazon S3, data encryption is feasible. However, for the static data used by cloud-based applications in PaaS or SaaS model, data encryption in many cases is not feasible. Because data encryption will lead to problems of indexing and query, the static data used by Cloud-based applications is generally not encrypted. Not only in cloud, but also in traditional IT environment, the data being treated is almost not encrypted for any program to deal with. Due to the multi-tenant feature of cloud computing models, the data being processed by cloud based applications is stored together with the data of other users. Unencrypted data in the process is a serious threat to data security. Regarding the use of private data, situations are more complicated. The owners of private data need to focus on and ensure whether the use of personal information is consistent with the purposes of information collection and whether personal information is being shared with third parties, for example, cloud service providers.

5. Share

Data sharing is expanding the use range of the data and renders data permissions more complex. The data owners can authorize the data access to one party, and in turn the party can further share the data to another party without the consent of the data owners. Therefore, during data sharing, especially when shared with a third party, the data owners need to consider whether the third party continues to maintain the original protection measures and usage restrictions. Regarding sharing of private data, in addition to authorization of data, sharing granularity (all the data or partial data) and data transformation are also need to be concerned about. The sharing granularity depends on the sharing policy and the division granularity of content. The data transformation refers to isolating sensitive information from the original data. This operation makes the data is not relevant with the data owners.

6. Storage

The data in the cloud may be divided into: (1) The data in IaaS environment, such as Amazon's Simple Storage Service; (2) The data in PaaS or SaaS environment related to cloudbased applications. The data stored in the cloud storages is similar with the ones stored in other places and needs to consider three aspects of information security: confidentiality, integrity and availability. The common solution for data confidentiality is data encryption. In order to ensure the effective of encryption, there needs to consider the use of both encryption algorithm and key strength. As the cloud computing environment involving large amounts of data transmission, storage and handling, there also needs to consider processing speed and computational efficiency of encrypting large amounts of data. In this case, for example, symmetric encryption algorithm is more suitable than asymmetric encryption algorithm.

7. Archival

Archiving for data focuses on the storage media, whether to provide off-site storage and storage duration. If the data is stored on portable media and then the media is out of control, the data are likely to take the risk of leakage. If the cloud service providers do not provide off-site archiving, the availability of the data will be threatened. Again, whether storage duration is consistent with archival requirements? Otherwise, this may result in the availability or privacy threats.

8. Destruction

It is must that data should be completely destroyed after there is no more requirement of that data. The physical characteristics of storage medium are such that the data may still exist and can be restored even though it was deleted. This may result in inadvertently disclose of sensitive information

V. CURRENT SECURITY SOLUTIONS FOR DATA SECURITY AND PRIVACY PROTECTION

To get rid of these data security issues various encryption algorithm and mechanisms are used. People try to find the best way to solve these issues by using different combination of these algorithm and mechanisms to provide better security to the data in the cloud.

In June 2009, IBM developed a fully homomorphic encryption scheme in which allows data processing is done without being decrypted [12]. Roy I and Ramadan HE applied decentralized information flow control (DIFC) and differential privacy protection technology into data generation and calculation stages in cloud and put forth a privacy protection system called airavat [13] which can prevent privacy leakage without authorization in Map-Reduce computing process. Key management is a key problem for data encryption solutions. On the one hand, the users are not capable enough to manage their keys. While the cloud service providers have to maintain a large number of user keys. There are some organisations like The Organization for the Advancement of Structured Information Standards (OASIS) and Key Management Interoperability Protocol (KMIP) which try to solve such issues [14]. About data integrity verification, the users cannot download data and verify its correctness and then upload the data, because of data communication, transfer fees and time cost. The data is dynamic in cloud storage that's why traditional data integrity solutions are no longer suitable. NEC Labs's provable data integrity (PDI) solution can support public data integrity verification [15].

About data destruction, U.S. Department of Defense (DoD) 5220.22-M (the National Industrial Security Program Operating Manual) shows two approved methods of data (destruction) security, but it does not provide any specific requirements for how these two methods are to be achieved [20].

VI. CONCLUSION

As we all know cloud computing has many advantages but still there are many disadvantages that need to be solved. There are models like service delivery models and deployment models. There are some other essential features also. We can conclude that we need to solve these data security issues. These issues exist in all levels in service delivery models as well as all seven stages of data life cycle. As per the analysis for issues above, expectation is to have an integrated and comprehensive security solution to meet the needs of defence in deep. As we mentioned in this paper there are some solutions available to overcome these kind of issues but still we need to find some more proper and perfect solution up to some extent.

REFERENCES

[1] Ayesha Malik, Muhammad Mohsin Nazir," Security Framework for Cloud Computing Environment: A Review" Journal of Emerging Trends in Computing and Information Sciences VOL. 3, NO. 3, March 2012

[2] Mohsin Nazir "Cloud Computing: Overview & Current Research Challenges" IOSR Journal of Computer Engineering (IOSR-JCE) ISSN: 2278-0661, ISBN: 2278- 8727Volume 8, Issue 1 (Nov. - Dec. 2012), PP 14-22

[3] Deyan Chen and Hong Zhao, "Data Security and Privacy Protection Issues in Cloud Computing" 2012 International Conference on Computer Science and Electronics Engineering

[4] Cloud computing security, http://en.wikipedia.org/wiki/Cloud_ computing_security.

[5] Gartner: Seven cloud-computing security risks. InfoWorld. 2008-07-02. http://www.infoworld.com/d/security-central/gartner-seven-cloud- computing-security-risks-853.

[6] Cloud Security Front and Center. Forrester Research. 2009-11-18. http://blogs.forrester.com/srm/2009/11/cloud-security-front-and- center.html

[7] Cloud Security Alliance. http://www.cloudsecurityalliance.org.

[8] Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing, V2.1, http://www.cloudsecurityalliance.org/ guidance/csaguide.v2.1.pdf.

[9] S. Subashini, V.Kavitha. A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications 34(2011)1-11

[10] Mohamed Al Morsy, John Grundy, Ingo Müller, "An Analysis of The Cloud Computing Security Problem," in Proceedings of APSEC 2010 Cloud Workshop, Sydney, Australia, 30th Nov 2010.

[11] Yanpei Chen, Vern Paxson, Randy H. Katz, "What's New About Cloud Computing Security?" Technical Report No. UCB/EECS-2010-5. http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-5.html

[12] "IBM Discovers Encryption Scheme That Could Improve Cloud Security, Spam Filtering," at http://www.eweek.com/c/a/Security/IBM- Uncovers-Encryption-Scheme-That-Could-Improve-Cloud-Security- Spam-Filtering-135413/.

[13] Roy I, Ramadan HE, Setty STV, Kilzer A, Shmatikov V, Witchel E. "Airavat: Security and privacy for MapReduce," In: Castro M, eds. Proc. of the 7th Usenix Symp. on Networked Systems Design and Implementation. San Jose: USENIX Association, 2010. 297.312.

[14] "OASIS Key Management Interoperability Protocol (KMIP) TC", http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=kmip.

[15] Zeng K, "Publicly verifiable remote data integrity," In: Chen LQ, Ryan MD, Wang GL, eds. LNCS 5308. Birmingham: Springer-Verlag, 2008. 419.434. [