Available online at www.sciencedirect.com

ScienceDirect

IERI Procedia 10 (2014) 112 – 118

Procedia IERI

www.elsevier.com/locate/procedia

2014 International Conference on Future Information Engineering

From Single to Multi-Clouds Computing Privacy and Fault Tolerance

Maha TEBAA*, Said EL HAJJI

University Mohammed V–Agdal, Faculty of Sciences,
Laboratory of Mathematics, Computing and Applications,
Rabat, Morocco

Abstract

Security issues of data hosted in a Cloud Computing provider remain hidden seen excessive marketing that led to a totally unrealistic view of cloud computing security. Although Cloud Computing has not yet reached the level of maturity expected by its customers, and that the problems of confidentiality, integrity, reliability and consistency (CIRC) are still open, the researchers in this field have already considered a future cloud strategy which aims : a better QoS, reliability and high availability, it is the Multi-Clouds, Cloud of Clouds or Interclouds.
This paper will present the security limitations in the single Cloud and the usefulness of adopting rather Multi-Clouds strategy to reduce security risks, through the use of DepSky which is a virtual storage system that ensures better availability and high confidentiality of data.

Keywords: Multi-Clouds Computing, Security, Privacy, DepSky, BFT, ICStore, HAIL, RACS ;

1. Introduction

Cloud computing is reshaping the IT domain, instead of unpacking computers and stack them in a machine room, the cloud can download virtually equipment and associated infrastructure, theconstruction of aDatacenter is now possible in some minutes with minimal technical knowledge and for a fraction of the purchase cost of a single server. Researches in the field of Multi-Clouds don't have much interest as the single

* Maha TEBAA. Tel.: +212-537-771-834; fax: +212-537-774-262.
E-mail address: maha.tebaa@gmail.com.

doi:10.1016/j.ieri.2014.09.099

Cloud. Although to date, little attention has been given to the distribution of cloud's risk, and managing multiple Clouds from a single technology platform, in the real world, by attaching your business to a supplier single is widely regarded as a bad strategy and unacceptable risks, and that philosophy applies to cloud provider with a single cloud, or even a single geographical area, as the case in EC2. To meet a variety of needs, including availability, business continuity and disaster recovery, it is important to use multi-Clouds strategies. Cloud providers themselves can fail, so for the greatest degree of protection, a company can engage in a Multi-Clouds strategy.

In this work we present the security limitations in the single cloud, and improvements to attributes (CIRC) through the use of a storage system in the multi-Clouds based on the consolidation of several clouds to put up the cloud of Clouds, to manage these various clouds, we use DepSky library in order to handle the heterogeneity of interfaces of each cloud provider. The DepSky system guarantees the availability and confidentiality of data by using multiple providers Clouds, by the association of "the algorithm of byzantine failures tolerance, secret sharing and erasure codes cryptographic". Then we compare DepSky with other storage mechanisms used on multi-Clouds, RACS, HAIL and ICStore that is currently under development and addresses the security aspects (CIRC) in a layered architecture.

2. Single Cloud Computing

Cloud computing is a new way of managing its infrastructure, applications and computer data over the Internet (VPN) by delegating administrative tasks of maintenance and supply of material resources to a third party (Cloud Computing provider), the only way to interact with data is an interface from their smartphone or computer, Cloud providers are essentially concerned by delivering services (PaaS, IaaS or SaaS), but ignore the main aspect which is the protection of privacy in the Cloud, and without any responsibility or obligation towards these customers, private data stored in plain cleartext in the virtual hard disk can be accessed by the provider itself or by other clients accessing the same disk.

2.1. Security limitations of the Single Cloud

Security issues of data stored in cloud are still the number one barrier to adoption of cloud computing for companies and government agencies. Security concerns are unavoidable for cloud computing to reach the level of maturity required, as the next generation of IT. Cloud services should ensure data integrity and provide privacy of the data stored in the cloud, but sometimes they lose control over the data stored in their data centers, as is the case in the following examples.

a) Data Integrity

Data integrity is one of the important issues related to Cloud security issues. The transmitted data between the client and cloud providers may be lost or corrupted, as shown in the examples bellow:

October 2009, a loss of all Sidekick users data (directories of contacts, calendars, photos) due to a server malfunction in Danger's data centers (Microsoft), after one year Microsoft has conceived that the majority of lost data cannot be recovered [2].

- January 2009, servers Ma.gnolia have suffered a total loss of data due to a complete failure; the loss of half a terabyte of data has made the process of recovery impossible, making the site essentially dead [3].

For more examples, refer to Cachinet and al [4].

b) Data Confidentiality and Privacy

Protect sensitive data such as bank details or documents healthcare should be among the priorities of cloud providers that are either internal (malicious administrators who work in the Cloud Provider) or external attacks. Various Cloud provider adopt various technology to resolve the problem of data privacy, but the virtualized nature of cloud make the traditional mechanism unstable for handling the privacy risks, and the use of the different encryption technique still limited. Garfinkel [5] gives an example of the loss of confidentiality such as the Amazon Cloud service. This example shows that just by knowing an Amazon account password, the totality of the account's instances and resources can be reached.

c) Data Availability

In Cloud computing the most important character that encourages customers to migrate to a cloud service, is the high availability of services, data, and applications. If we entrust our data to a single cloud provider andit does not contemplate a backup solution or it hosts the data in a single platform or in a same geographical area this may increase the risk of downtime, and it impacts customers who can get stuck for several hours without access to their data. Amazon [6] underlines in its contract that a service may be cut down at any moment.

2.2. Homomorphic Encryption applied to the Single Cloud

Cloud providers use traditional methods to secure their customers' data (they ask customers to encrypt data with a key before sending them to the Cloud). However when a client wants to perform processing on his data, the cloud provider requires the decryption key. In fact, it's impossible to perform any processing on encrypted data without decrypting. Consequently, cloud provider holds all the clear data. See [1] for examples.

a) Definition

Our proposal is to encrypt data before sending them to the cloud providers, but theyshould be decrypted, whenever there is the need to execute calculations on them. It was impossible to hold encrypted data secured while a third party performs calculations on them. So, to allow the Cloud provider to perform calculations on encrypted data without decrypting them, it is necessary to use the homomorphic encryption cryptosystems.

Homomorphic encryption systems are used to perform operations on encrypted data without a secret key (without decrypting); the client still the unique possessor of the secret key [1].

Definition:

An encryption is homomorphic, if: from Enc (a) and Enc (b) it is possible to compute Enc (f (a, b)), where f can be one of the following operations: +, ×, ⊕ and without using the private key.

We distinguish two categories of homomorphic encryption:

- The additive homomorphic encryption (addition on raw data) is the Pailler and Goldwasser-Micalli cryptosystems.
- The multiplicative homomorphic encryption (only products on raw data) is the RSA and El Gamal cryptosystems.

b) Limitations

For homomorphic encryption systems, the speed of slow moving and large encrypted text sent, the size of the result ciphertext after performing operations and the limited bandwidth, which depends on ISP, are the main limitations, Also the management of private keys that totally depends on the client who does not necessarily have a good knowledge on security, since it has delegated the management of its IT to a Cloud Provider, so if the client loses the private key used to encrypt the data hosted in the cloud provider, it can recover the data only in encrypted state, and therefore it can not be reused.

3. Security mechanisms in the Multi-Clouds

Multi-Clouds, Cloud of Clouds or Interclouds [7] are similar terms used to show that despite the limitations discussed earlier and all the security problems of single Cloud, Cloud computing should not be restricted to a single cloud, and that sensitive data should not be entrusted to a single cloud, to avoid dependency on just one Cloud Provider. Switchingthe cloud computing from a single Cloud to multi-Clouds is mandatory to fulfil data security.

3.1. BFT (Byzantine Fault tolerance)

Distributed computing suffers from harmful reliability problems and availability, the deployed services are becoming more open to the world, and a significant load on the servers is increasingly imposed. In addition, the failure in deployed services, network failure that can happen unexpectedly. These questions can easily compromise the reliability and performance of services. Replication is a key technique used to maintain system reliability by having multiple backup servers (or redundant replicas). Replication aims to raise the availability and performance of the system while ensuring its correctness.

The relationship between cloud computing and BFT was already studied and in recent years it has been considered among the most important aspect in distributed systems, although these studies were purely academic.

A BFT [8] protocol is used to manage communication between replicated systems and clients. This protocol requires at least 3f replicas to ensure consistency between replicas of the system, where f represents the replicas that can be byzantine.

3.2. DepSky

Bessani et al. [9] propose a virtual storage system called DepSky is a coexistence of several clouds to build a cloud of clouds. DepSky improves the availability, integrity and confidentiality of information stored in the cloud through the encryption, encoding and duplication of data. This allows mitigating the limitations of individual clouds using several dependability and security techniques. The DepSky system ensures the availability and confidentiality of data stored in different cloud providers by using the multi-clouds architecture and the association of “the algorithm of byzantine failures tolerance, secret sharing and erasure codes cryptographic”. The DepSky architecture consists of four clouds and each cloud uses its own specific interface. The algorithm DepSky presented in client machines as a software library allowing reading and writing data stored in the cloud, the DepSky system consists of two algorithms:

DEPSKY-A (Available DepSky): brings the accessibility and integrity of data by duplicating storage on different clouds using quorum methods.

DEPSKY-CA (Confidential & Available DepSky): the lack of confidentiality is the major drawback of DepSky-A because the storage of data is in cleartext, Depsky-CA algorithm encrypts data before storing them in the multi-Clouds with a symmetric encryption, then the data is divided into block as: f+1 blocks are necessary to recover the original data, f or less block don't give any information about the data stored in the Multi-clouds.

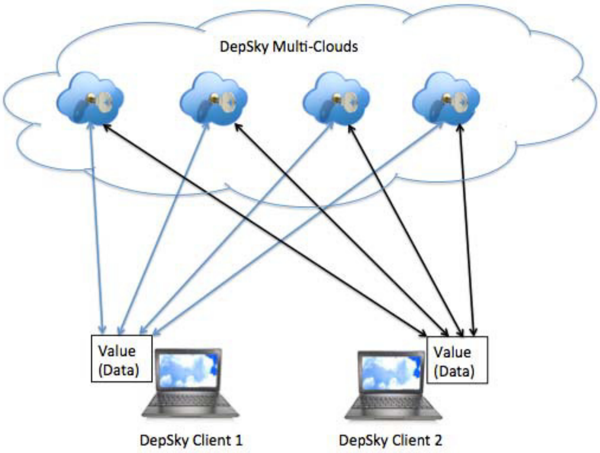

Fig.1. Architecture of DepSky

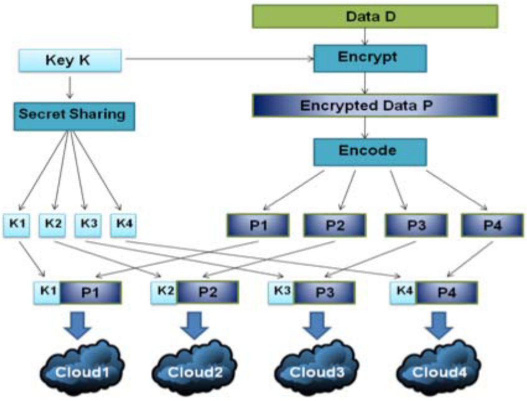

Fig.2. DepSky-CA Algorithm [10]

3.3. RACS (Redundant Array of Cloud Storage)

The RACS [10] system (Redundant Array of Cloud Storage) employs a RAID5-like technique (mainly erasure codes) to implement high-available and storage-efficient data replication on diverse clouds. Generally the purpose of RACS is to avoid vendor lock-in and its associated risks. This problematic occurs when organizations move their data from their data centers to one cloud storage provider and later, even with appealing offers from competition, it's expensive for clients to switch from one provider to another. Typically, storage providers charge clients for inbound and outbound bandwidth. Therefore, the intent behind RACS is to replicate client's data on multiple providers, but in an economical way. Data are spread on multiple providers, which implies redundancy to tolerate possible failures or outages. To sum up, the RACS system is a cloud storage proxy that shares data across multiple cloud storage providers in transparent way.

In contrast to DepSky, the RACS system does not try to solve security problems of cloud storage, but deals with the "economic failures" and vendor lock-in concern. Thus, the system provides no mechanism of detecting data corruption.

3.4. HAIL (High Availability and Integrity Layer)

HAIL [11] (High Availability and Integrity Layer) is a distributed cryptographic system that permits a set of servers to prove to a client that a stored file is intact and retrievable, it's allows checking data retrievability, but if data is deleted by malicious provider, nothing can be done. HAIL provides a software layer to address availability and integrity of the stored data in the Interclouds. It aggregates cryptographic protocols for proof of recoveries with erasure codes to provide a software layer to protect the integrity and availability of the stored data, even if the individual cloud are compromised by a malicious and mobile adversary.

In contrast to DepSky, HAIL does not guarantee the privacy, needs code execution in cloud servers, it does not allow management of different versions of data.

3.5. IC Store (Intercloud Storage)

ICStore [12] allows access to private or public cloud providers to migrate third, save or share files. The advantage of the solution is to protect against downtime, data loss or hacker attacks. Another advantage is filled more easily switch providers migrating all data from one cloud to another.

The ICStore Client, offers to the end client a key value store with simple read and write operations, which is a common base service offered by commodity cloud storage providers. ICStore client consists of three core layers that target different dependability aspects: i) confidentiality, ii) integrity and iii) reliability and

consistency (RC). This layered approach allows individual layers to be switched "on" and "off" to provide different levels of dependability that are to be matched with client's goals, also with performance and possibly even monetary constraints in mind.
In contrast to DEPSKY, ICStore does not use the secret sharing algorithm on the provision of confidentiality. However it is not clear if information- efficient secret sharing or some variant of this technique could substitute the erasure codes employed on these protocols.

Table 1. Existing Security Mechanisms in the Multi-Clouds

	Data Integrity	Service Availability	Privacy
DepSky	√	√	√
IC Store	√		
HAIL	√	√	
RACS			

4. Conclusion and Perspectives

All the mechanisms discussed earlier enable storage, modification or deletion data hosted in the cloud, a client company may also ask provider to perform processing on the data, the homomorphic encryption applied to single cloud allows operations on encrypted data without decrypting.

The swot analysis of the different security mechanisms used in the multi-clouds allows to cloud providers to know what is the mechanism to use to provide better security (Confidentiality, Integrity and Availability) of data stored in their data centers, and the client can understand the limitations of single cloud and the benefits of multi-clouds, DepSky is the most reliable mechanism. The use of multi-clouds computing is not restricted to data storage, but also performing operations on data, Our proposal is to integrate a homomorphic cryptosystem in DepSky algorithm, precisely in the secret sharing scheme, may give better results especially when dealing with sensitive data.

References

[1] Maha Tebaa, Said El Hajji, « Secure Cloud Computing through Homomorphic Encryption», International Journal of Advancements in Computing Technology (IJACT) Volume5, Number16, 2013.
[2] David Sarno, Microsoft says lost sidekick data will be restored to users. Los Angeles Times, 2009.
[3] Erica Naone, Are we safeguarding social data? Technology Review published by MIT Review, http://www.technologyreview.com/blog/editors/22924/, 2009.
[4] C. Cachin, I. Keidar and A. Shraer, «Trusting the cloud», ACM SIGACT News, pp. 81-86, 2009.
[5] S.L. Garfinkel, «Email-based identification and authentication: An alternative to PKI», IEEE Security and Privacy, 1(6), pp. 20-26, 2003.
[6] Amazon, Amazon Web Services. Web services licensing agreement, 2006.
[7] Mohammed A. Al Zainand and al. «A Survey on Data Security Issues in Cloud Computing: From Single to Multi-Clouds», Journal of Software, Vol 8, No 5 (2013), 1068-1078, 2013.
[8] Ali Shokerand and al., BFT Selection, Networked Systems, Lecture Note in Computer Science, Volume 7853, 2013, pp 258-262, 2013.
[9] Bessani and al. «DEPSKY: Dependable and Secure Storage in a Cloud-of-Clouds», EuroSys'11, Salzburg, Austria, 2011.
[10] Fernando Martins André, Thesis «Availability and Confidentiality in Storage Cloud», 2011. [11] Monali Shrawankar and al. «Comparative Study of Security Mechanisms in Multi-clouds Environment»,

International Journal of Computer Applications (0975 – 8887) Volume 77, No.6, 2013.
[12] Christian Cachin and al. «Dependable Storage in the Intercloud», BM Research Report, 2010.