Cryptography and Steganography in Cloud Migration: Challenges and Advances in Data Security

Jasvinder Singh
Department of Computer Science, University of British Columbia, Vancouver, Canada

ABSTRACT

Cloud computing has recently became a widely discussed topic in the IT industry. More and more organizations consider using the Cloud, because it enables an easy and cost efficient way of hosting applications, with dynamic scaling and geographical distribution possibilities. Still, it is not clear how and when cloud computing should be used. Existing application are often written in a way that does not really fit a cloud environment well. Also, certain quality attributes (e.g. performance, security or portability) can be affected. More studies are needed on how existing systems should be plugged into the Cloud and what are the consequences of the migration. Data migration and application migration are one of popular technologies that enable computing and data storage management to be autonomic and self-managing. We examine important issues in designing and developing scalable architectures and techniques for efficient and effective data migration and application migration. The first contribution we have made is to investigate the opportunity of automated data migration across multi-tier storage systems. In this paper, we have discussed various techniques for cloud stegnography and cryptography.

KEYWORDS: cloud computing, public cloud platform, migration, Cloud Stegnography.

1. INTRODUCTION

Cloud computing usually refers to a utility-based provisioning of computational resources over the Internet. Widely used analogies to explain cloud computing are electricity and water supply systems. Like the Cloud, they provide centralized resources that are accessible for everyone. Also, in the Cloud you only pay for what you have used. And finally, it is usually consumed by those who have difficulties to produce necessary resources by themselves or just do not want to do that. Despite the description by analogy, it is difficult to give a unique and precise definition. One of the main ambiguities to define cloud computing is the fact that it is still evolving and taking its shape.

The definitions proposed in the cloud computing community are often focused on different perspectives and do not have common baselines. Analyzing existing sources in order to identify common characteristics, Vaquero et al [7] observed no clear and complete definition in the literature. Nevertheless, the authors proposed three features that most closely describe cloud computing: scalability, pay-as-you-go utility model, and virtualization – and gave the following definition:

"Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs"[1].

Cloud Computing characteristics:

A. **Virtualization** (abstracted infrastructure). Cloud computing became possible through a new evolution of virtualization. Virtualization enables dynamic infrastructure utilization, resource sharing, isolation and security. In contrast to a standard model when processing takes place on specific hardware defined in advance, applications do not have any static computing place in a virtualized cloud environment. Resources are allocated dynamically depending on the demand. Thus, customers do not know the exact place and the type of hardware their applications are running on. Cloud providers can only guarantee minimum performance or storage capacity for the customer [2].
B. **A pay-per-use model.** This is the key characteristic of cloud computing economics. All resources in the Cloud are available on a utility basis, meaning that users are charged based on the quantity consumed by them. This model allows entering the market with no upfront investments into own hardware infrastructure.
C. **On-demand access**. On-demand access means that resources like CPU time or storage can be provisioned automatically when needed without any extra management effort [3].
D. **Elastic scalability.** Elastic scaling signifies that computational resources, used by the application, can be dynamically scaled up or down. In other words, virtualized hardware resources can be resized easily and rapidly on demand. It makes a utility model even more attractive, because consumers use only what they really need.

E. **Resource pooling.** Computing resources of the provider are shared across multiple users. Different resources are pooled in a multi-tenant way so that they can be dynamically assigned and reassigned to serve consumers' needs.

Cloud Data Migration

The critical IO changes of Solid State Disks (SSD) over conventional rotational hard plates makes it an appealing way to deal with incorporate SSDs in layered capacity frameworks for execution improvement. Be that as it may, to coordinate SSD into multitiered capacity framework viably, mechanized information movement amongst SSD and HDD assumes a basic part. In numerous true application situations like saving money and market conditions, workload and IO profile display intriguing qualities and furthermore bear the limitation of workload due date. Step by step instructions to completely discharge the energy of information relocation while ensuring the movement due date is basic to boosting the execution of SSD empowered multi-layered capacity framework. So as to completely exploit the advantages of SSDs in a multi-layered capacity framework with SSDs filling in as the speediest level, it is essential to distinguish the correct subset of information that should be set on this level given the restricted limit of SSD level because of mind-boggling expense per gigabyte. In particular, we need to boost general framework execution by setting basic, IOPS (input/yield activities every second) serious and inactivity delicate information on the quick SSD level through two-way mechanized information relocation amongst SSDs and HDDs. By working with an assortment of big business class stockpiling applications, we watch that numerous square level IO workloads show certain time-subordinate normality as far as access examples and temperature of degrees (hot or icy). For instance, in keeping money applications, IO workloads for account access and credit check are ordinarily heavier amid specific long stretches of multi day. In any case, such examples may change from day-time to evening time, from everyday, from weekdays to ends of the week or from working days to open occasions. Hence, square level IO profiling is the initial step for building a mechanized information relocation framework. The following enormous test is to devise techniques.

2. LITERATURE SURVEY

[1]Issa Khalil, Cloud processing administrations are ending up increasingly mainstream. Be that as it may, the high centralization of information and administrations on the mists make them alluring focuses for different security assaults, including DoS, information robbery, and protection assaults. Moreover, cloud suppliers may neglect to conform to benefit level understanding as far as execution, accessibility, and security ensures. Along these lines, it is of principal significance to have secure and effective instruments that empower clients to straightforwardly duplicate and move their information starting with one supplier then onto the next. In this paper, we investigate the best in class between cloud relocation procedures and distinguish the potential security dangers in the extent of Hadoop Distributed File System HDFS. We propose a between cloud information relocation component that offers better security ensures and quicker reaction time for moving substantial scale information records in cloud database administration frameworks. The execution of the proposed approach is approved by estimating its effect on reaction time and throughput, and contrasting the execution with that of different systems in the writing. The outcomes demonstrate that our approach essentially enhances the execution of HDFS and outflanks its partners.

[2] Ibrahim Ejdayid A. Mansour, Cloud suppliers offer their IaaS administrations in light of virtualization to empower multi-occupant and segregated conditions for cloud clients. Right now, every supplier has its own

restrictive virtual machine (VM) director, called the hypervisor. This has brought about tight coupling of VMs to their fundamental equipment blocking live movement of VMs to various suppliers. Various client driven methodologies have been proposed from both scholarly community and industry to fathom this issue. In any case, these methodologies endure restrictions as far as execution (relocation downtime), adaptability (decoupling VMs from basic equipment) and security (secure live movement). This paper proposes LivCloud to defeat such impediments. An open-source cloud orchestrator, a created transport convention, overlay organize and secured relocation channel are pivotal parts of LivCloud to accomplish compelling live cloud movement. Besides, an underlying assessment of LAN live relocation in settled virtualization condition and between various hypervisors has been considered to demonstrate the movement affect on arrange throughput, organize inactivity and CPU use. The assessment has shown the requirement for advancement inside the LAN condition.

[3] Qingni Shen, with the improvement of distributed computing, cloud security issues have as of late picked up footing in the exploration network. Albeit a great part of the endeavors are centered around securing the task framework and virtual machine, or securing information stockpiling inside a cloud framework, this paper takes an elective point of view to cloud security—the security of information relocation between various mists. To start with, we portray a few dangers when we are doing information movement. Second, we propose a security

component to manage the security issues on information movement starting with one cloud then onto the next. Third, we outline a model to give the instrument a short execution in view of HDFS(Hadoop Distributed File System) and we complete a progression of tests to assess our model. Here, the answers for securing information movement between mists fundamentally include in SSL transaction, relocation ticket outline and square encryption in appropriated document framework and bunch parallel processing.

[4] Sameera Dhuria, Cloud Computing is another registering model in the realm of Information Technology that conveys benefits as utility over the Internet. It has a few focal points when contrasted with conventional registering models like on-request benefits, readiness, versatility, lessened data innovation overhead for the end-client, more prominent adaptability, decreased cost and so on. The points of interest and long haul advantages of this new innovation inspire associations to move their current applications to the cloud. Despite the fact that relocating to cloud gives numerous advantages, there are various difficulties and security issues identified with cloud, that block the procedure of its appropriation by the associations. The present paper plans to examine the real difficulties identified with movement to Cloud Computing.

[5] Virendra Singh Kushwah,Cloud processing is another worldview that joins a few figuring ideas and advancements of the Internet making a stage for more spry and financially savvy business applications and IT foundation. The appropriation of Cloud registering has been expanding for quite a while and the development of the market is consistently developing. Security is the issue most reliably raised as shoppers hope to move their information and applications to the cloud. I legitimize the significance and inspiration of security in the relocation of inheritance frameworks and I do an approach identified with security in movement procedures to cloud with the point of finding the necessities, concerns, prerequisites, angles, openings and advantages of security in the movement procedure of heritage frameworks.

Steganography Techniques

In addition to the general methods of information hiding presented above, many steganography techniques have been proposed during the last few years. These techniques differ in the mechanism or principle being used to hide a secret message or the changes that are taking place during the entire process of embedding. Therefore, there are six categories of steganography techniques:

substitution systems, transform domain techniques, spread spectrum techniques, statistical methods, distortion techniques, and cover generation techniques .

A. Substitution Systems

For a given cover file, it is important to find out some areas or data that can be modified without having any significant effects on this cover file . Therefore, a secret message can be embedded by replacing the redundant or insignificant parts of a cover file with secret message bits, without adding any significant noise to this cover file . Generally, digital covers have a large number of redundant bits (e.g. least significant bits (LSB)). In the

substitution technique of steganography, the bits of the secret message substitute the LSB of the bytes of the cover file without causing a drastic change to this cover file. Moreover, the LSB technique is a spatial domain technique since it embeds the secret bits directly in the cover file. Since LSB substitution technique is relatively quick and easy to use, it is the most common technique used for digital steganography and especially with digital images. However, the embedded information using the LSB technique is highly vulnerable and could be destroyed entirely by applying a slight modification to the stego image such as JPEG compression.

B. Transform Domain Techniques

Unlike spatial domain techniques (e.g. LSB technique), transform (frequency) domain techniques hide secret data in significant parts of the cover file. Therefore, frequency domain techniques are considered more robust to attacks than spatial domain techniques. Hence, most of robust steganographic systems known today rely on frequency domain techniques. There are many transforms used to map a signal into the frequency domain. Discrete cosine transform (DCT), discrete wavelet transform (DWT), and discrete Fourier transform (DFT) are methods used as mediums to embed secret data in digital images. However, when we add a slight noise or secret data to some frequency domain components, it changes the whole image rather than changing only this part of the image. Thus, secret and embedded data will be spread across the entire image and will not be concentrated on one certain area or region.

C. Spread Spectrum Techniques

Marvel et al. (1998) define spread spectrum communication as ”the process of spreading the bandwidth of a narrowband signal across a wide band of frequencies”. In spread spectrum steganography, the frequency domain

of the cover file is considered to be a communication channel and the secret message as a signal that is transmitted through it. Since the secret message is spread through a wide frequency band, this technique is relatively robust against stego file modification or message removal.

D. Statistical Techniques

These techniques embed only one bit of secret data in a cover file. Therefore, it is known as "1-bit" steganography scheme. If "1" is hidden in a cover file, some statistical characteristics (e.g. entropy and probability distribution) of this cover file must be changed significantly to clearly indicate the existence of a message. However, if the hidden bit is "0", the cover file is left unmodified. Therefore, this technique entirely depends on the ability of the receiver to differentiate between changed and intact cover files.

E. Distortion Techniques

Most of the steganography techniques are blind, which means that a receiver does not need the original cover file to extract the hidden message from the corresponding stego file. However, if a distortion technique is used, the receiver requires the original cover file in order to recover the secret message. For a receiver, the embedded message is the difference between the modified cover file received (the stego file) and the original cover file.

Cryptographic techniques for security in cloud

Many security methods for cloud use various cryptographic techniques. Cryptographic techniques have become essential for security in cloud. A key is used for data encryption and decryption. This helps in protecting confidentiality and integrity of data. It ensures security of data being shared in cloud and also allows data to be stored securely.

Cryptography refers to the science of designing ciphers. Encryption refers to the method of converting plain text to secret text (cipher text) which can only be read by owner of secret key. At present various cryptographic algorithms are there which belong to two major categories –

a) Symmetric algorithms such as DES, AES, Triple DES
b) Asymmetric or public-key encryption algorithms such as RSA, Diffie-Hellman, ECC, etc.
The difference is in the way the keys are used. In symmetric key encryption, the person who is sending the data and the person who is receiving the data share a key which is kept secret. This is then used to encrypt and decrypt the messages. In asymmetric key encryption, two keys are involved wherein one is used for encryption (this is publicly available) and the other is used for decryption (this is kept secret).

Identity based encryption -

Identity-based encryption (IBE), is a type of public-keyencryption in which the public key of a user is some unique information about the identity of the user (e.g. a user's email address). It allows any party to generate a public key from a known identity value such as an ASCII string. A trusted third party, called the Private KeyGenerator (PKG), generates the corresponding private keys. This kind of encryption reduces the complexity of the encryption process for both users and administrators.

Attribute based encryption

It is a type of public-key encryption in which the secret key of a user and the cipher text are dependent upon attributes (e.g. the country he lives, or the kind of subscription he has). A user can encrypt a message under a public key and a policy. Decryption will only work if the attributes associated with the decryption key match the policy used to encrypt the message.

Fully Homomorphic Encryption

Homomorphic encryption ensures privacy of data in communication, storage or in use with tools similar to conventional cryptography, but with extra features of computing over encrypted data, searching an encrypted data, etc. Search and manipulation of cipher text was difficult with traditional encryption techniques.

Intelligent Encryption

Conventional public-key and shared-key encryption systems rely on standard protocols and a preestablished public key certification infrastructure(public key infrastructure), allowing people all over the world to use encryption according to standard methods. But in conventional cryptographic methods only one person, i.e., owner of the key can view original data. This creates problem in case of cloud where

number of users are there to access same data. Intelligent encryption works on basis of various conditions. It allows various users to view encrypted data based on certain conditions rather than only single authorized user. It is similar to attribute based encryption but conditions are extended for multiple users.

3. CONCLUSION AND FUTURE SCOPE

In this paper, we have discussed the various techniques for cloud stegnography and cloud cryprography for various environment. It is concluded that, although there are various techniques available for cloud migration, a robust method is required to implement to migrate the data form private cloud to public cloud in a secure way. Hence a two layer data migration is required to be implemented to migrate the data. Out of these two layers one layer encrypt the input data and second layer will be hide the data into another resource.

REFERENCES

[1] Issa Khalil,Ismail Hababeh,Abdallah Khreishah,"Secure Inter Cloud Data Migration",International Conference on Information and Communication Systems(ICICS), 2016

[2] Ibrahim Ejdayid A.Mansour,Kendra Cooper, Hamid Bouchachia,"Effective Live Cloud Migration", 2016 IEEE 4th International Conference on Future Internet of Things and Cloud.

[3] Qingni Shen, Lizhe Zhang, Xin Yang, Yahui Yang, Zhonghai Wu, Ying Zhang, "SecDM: Securing Data Migration Between Cloud Storage Systems",2011 Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing.

[4] Sameera Dhuria, Anu Gupta, R. K. Singla,"Migrating Applications to the Cloud: Issues and Challenges", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 6, June 2015.

[5] Virendra Singh Kushwah, Aradhana Saxena,"A Security approach for Data Migration in Cloud Computing",International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.

[6] Wei Hao, I-Ling Yen and Bhavani Thuraisingham,"Dynamic Service and Data Migration in the Clouds", 2009 33rd Annual IEEE International Computer Software and Applications Conference.

[7] Rashmi Rao, Pawan Prakash, "Improving security for data migration in cloud computing using randomized encryption technique", July 2015 IOSR Journal of Computer Engineering (IOSR-JCE), ISSN: 2278-8727Volume 11, Issue 6.

[8] Rajeshri Vaidya, and Prof. Sumedh Pundkar,"Large Data migration within Cloud Environments using Compression and Encryption Technique", International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 2, July 2015.

[9] Mohammad Manzurul Islam, Sarwar Morshed and Parijat Goswami,"Cloud Computing: A Survey on its limitations and Potential Solutions", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 4, No 2, July 2013.

[10] Chetan M Bulla, Satish S Bhojannavar and Vishal M Danawade, "Cloud Computing: Research Activities and Challenges", International Journal of Emerging Trends & Technology in Computer Science, Volume 2, Issue 5, September – October 2013.

[11] Nirav Shah, Sandip Chauhan,"Survey Paper on Security Issues While Data Migration in Cloud Computing", July 2014 IJIRT | Volume 1 Issue 7 | ISSN: 2349-6002.

[12] Punit K Mendapara, Sandip S Chauhan,"Survey Paper on Secure Live Data Migration in Cloud Computing by maintaining Integrity and Confidentiality", December 2015 | IJIRT | Volume 2 Issue 7 | ISSN: 2349-6002.

[13] Virendra Singh Kushwah, Aradhana Saxena,"A Security approach for Data Migration in Cloud Computing", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.

[14] Zohreh Sanaei, Abdullah Gani,"Heterogeneity in Mobile Cloud Computing: Taxonomy and Open Challenges", IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 16, NO. 1, FIRST QUARTER 2014.

[15] J. Priya Shanthi, Parsi Kalpana,"Migration of Existing Applications to Cloud and Among Clouds", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013.

[16] Y. Ghebghoub, S. Oukid, and O. Boussaid,"A Survey on Security Issues and the Existing Solutions in Cloud Computing", International Journal of Computer and Electrical Engineering, Vol. 5, No. 6, December 2013.