International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

An Efficient Approach to Forensic Investigation in Cloud using VM Snapshots

Mr. Mahesh Dhumal and Prof. Monika Rokade

Department of Computer Engineering, Sharadchandra Pawar College of Engg. Dumbarwadi,

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

Cloud computing has recent times arisen as a technology to allow users as well as clients to access infrastructure, storage, software as well as deployment Environment based on payfor-what-use model. Conventional digital forensic can't be investigated due to some technical challenges like environmental as well as technical. The vibrant nature of cloud computing provides massive opportunities to identify malicious request using various security algorithms in cloud environment. Proposed research work identifies the current issues and provides solutions to reduce the challenges of digital forensics in the cloud environment and some challenges. In this paper system proposed forensic investigation of cloud security for trusted and untrusted environments. System illustrated the various machine learning algorithms for eliminate the malicious request, and investigate the malicious user also. Proposed method generate the user log base snapshot during the active session and manual investigator can verify all logs and identify the malicious user. We offer a skilled approach to forensic examination in the cloud using virtual machine (VM) snapshots.

Keywords: *Software as Services, Snapshot generation, Cloud Computing, VM, Cloud Service Provider.*

Introduction

Cloud is an emerging technology and cloud-based storage is a newly adopted idea that not only allows users to upload data to the web, but also allows quick access to the available resources and data sharing with anyone at any time is. But Cloud is a technique that creates a challenge for the person who is investigating and detecting forensic evidence that can help in forensic analysis, because the data stored on the cloud is from any system and any system can be accessed from and the scarf remains in very small quantities.

The 21st century is called the age of the digital world. There has been adopted computers to a great extent. Today without computers and Internet one cannot survive as we are dependent on these machines for almost all our work. Taking into consideration starting from home to education till banking and even corporate functioning everything has now been automated to computers. Computers contain all our important data in the digital format. With this the need to store the digital data has increased and virtual environment has replaced the physical storage for storing all our credentials as shown in Fig. 1. The most destructive challenge of the cloud is to prevent the unauthorized extinction of the data stored on the cloud, because anyone can easily remove the stuff without any proper authority. Removing data on the virtual machine removes nodes pointing to some information is completely dependent on deletion.

VMs are rapidly gaining popularity due to the simulation of computing environments, separating users, restoring previous states, and supporting remote initiation. All of these features have positive security side effects. VM's hardware abstracts and isolation limits the scope of the attack and formulate it much complicated for external attacker to use not permitted data and resources on the physical machine. VM state restoration enables clients to come back to a state preceding assault or information calamity provides an easy way to remove malware and data protection. By allowing users to start and stop VM remotely, the attackers have short-time windows in which they should be prepared and attacked. This is a surprisingly effective security measure. Since the hypervisor runs out of Virtual Machine, Its having a ability to monitor malware. Due to such reasons, VM Infrastructure has the ability to secure than physical server infrastructure.

Literature Survey

According to [1] Cloud computing has recently arisen as a technology to allow users to access infrastructure, storage, software as well as deployment environment based on a usability of user what users have been used model. The vibrant and multi-tenant environment of

traditional digital forensic cloud environments cannot handle nature because it has to face numerous procedural, authorized and directorial challenges specific to the cloud system. The dynamic nature of cloud computing provides enormous opportunities for enabling digital check in the cloud environment. It has been addressed in the untrusted cloud situation to ease the challenges of digital forensics and some of the existing solutions. We offer a skilled approach to forensic examination in the cloud using virtual machine (VM) snapshots.

Identification of digital forensic in the cloud can add a new dimension to the process of creating confidence in the cloud in [2]. But Lots of cloud features such as transparency, virtualization, lack of legal issues etc., Challenges for the Cloud Forensics Whether it is a traditional digital forensic or cloud forensic, collecting comprehensive data for analysis is a major challenge in the investigation. Data gathering in exceptionally virtualized conditions like cloud is very tedious. The final goal of proof collection and analysis is to prove the official courtroom that they are forensic sound. We can use introspection techniques because they will not corrupt the source of evidence while collecting necessary data.

According to [3] Content is often repeated, modified or modified on primary storage systems, and users lose control over its dispersion on the system. The content identified with a specific venture from the framework in this way turns into a work escalated errand for the client. In this work system illustrates, a system that helps the user easily remove project interconnected content, but this does not require change in user behavior or any system component, Such as file system, kernel or application IRCUS is transparently integrated inside the client's framework, works in client space and stores the subsequent metadata with files. This work system describes evaluation of system and showed that its overhead and accuracy is acceptable for practical use and deployment.

Cloud computing systems illustrates [4] an prototype to the distributed dispensation of digital data. Digital forensic investigations associated with such systems area unit doubtless to involve a lot of complicated digital proof acquisition and analysis. Some public cloud computing systems could embrace the storage and process of digital knowledge in several courts, and a few organizations could value more highly to encode their knowledge before getting into the cloud. together with cloud design, these two factors will build rhetorical examination of such systems a lot of complicated and long. There are not any established digital rhetorical tips that specifically address the investigation of cloud computing systems. during this letter we tend to examine the legal aspects of the digital forensic investigation of the cloud computing system.

System [5] proposed the cloud automatic data processing system hosts most of today's industrial business applications, which provides it high revenue that makes it the target of cyber attacks. This emphasizes the necessity for a digital rhetorical system for the cloud surroundings. standard digital forensics cannot be directly given as a cloud forlantic answer as a result of its thanks to virtualization of multi-tenancy and resources within the cloud. whereas we have a tendency to do cloud forensics, information cloud element logs, virtual machine disk pictures, volatile memory dumps, console logs and network capture area unit to be inspected. during this letter, we've go together with a foreign proof assortment and preprocessing framework victimization Straits and Hadoop distributed filing system. the gathering of VM disk pictures, logs etc. is triggered by a pull model once triggered by the investigator, whereas the cloud node sporadically pushes network capture to HDFS. Pre-processing steps like bunch of logs and correlation and VM disk pictures area unit done through mahout and VICA to implement track analysis.

According to [6] Cloud computing is the computing paradigm which modify getting resources like code, hardware, services over the net. Most of user store their knowledge on cloud for knowledge security and integrity ar prime connected. this text encompasses a downside to confirm the integrity and data storage in cloud computing. to confirm the accuracy of the info, the operate of permitting Third Party Auditor (TPA) to be accustomed highlight the danger of cloud storage services by Cloud consumer to verify knowledge integrity hold on within the cloud Take it. This paper focuses on knowledge security, we provide implement Correction code in file distribution to produce redundancy and guarantee knowledge dependency. By mistreatment homomorphic token with distributed verification of erasure coded knowledge, our theme succeed storage correctness in addition as error localization. intensive security analysis shows that the planned arrange is very economical and versatile against the failure of Byzantine, malicious data repatriation attacks and even the server collision attacks.

Problem Statement

This work focuses the challenges of digital forensics in the cloud. In recent times, cloud environments are used by many customers for the storage and distribution of illegal information. A digital forensic structure dedicated to the cloud environment is required. The proposed research work carried an competent method to forensic examination in the cloud using the Virtual Machine (VM) Snapshot. It will collect VM snapshot logs from different users' sessions, whose reliability cannot be compromised. This approach should be executed for many VMs.

Objectives of System

The objectives of this research work include the following:

1. Explore the challenges and requirements of forensics in the virtualized environment of cloud computing

2. Design a digital forensic structure for the cloud computing systems from the view point of investigator and/or cloud architecture
3. Address the issues of dead/live forensic analysis within/outside the virtual machine that runs in a cloud environment
4. Using digital forensic triage in the examination and partial analysis phase of cloud forensics

IV. PROPOSED METHODOLOGY

In the proposed research work to design and implement a system that can provide the security to data, in cloud environment and provide the security from insider attacks like collusion attack, bruited force attack as well as SQL injection attack.

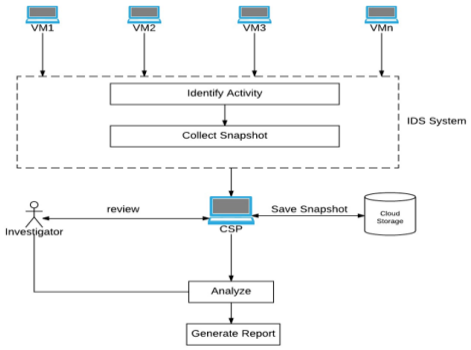

Figure 1 : Proposed System Architecture

We propose a protected information sharing plan for element individuals. Initially, we suggest a protected path for key dissemination with safe correspondence channels, as well as the clients can securely acquire their private keys from collecting chief. In our proposed system we use three different entities data owner, group manager, cloud server and attacker is untrusted entity. In this module first data owner upload the data file to cloud server using cryptography algorithm once data has store into database, owner gets the notification about file storage successfully. The data owner having a full access of specific data file he can share or access, so data owner can share the any file to any group manager then it will automatically access to all group members. The shared group members can access each file to anytime by cloud server. In first phase if data owner revoke any user from access the file then he can't access such file. If he can try to generate any collusion attack using SQL injection queries, even our system will system will prevent such attacks. Second data owner can share and revoke file to individual user to specific group, and third once any user revoke system will automatically generate proxy key generation that means existing keys will expired. The overall approach improves the system efficiency as well security on drastic level.

The Role Base Access Control (RBAC) Data share respective file to different number of users, the data owner can set the specific role to respective user for read, write, delete etc. The particular user can view the files which are shared by data owner in access control tab. According to the given credentials specific user can download the desired file. The data owner also eliminates file access of particular user using revocation function. The revocation function which removes file access to those users who watch the existing authenticated user. The same time system expired the existing case this strategy should be eliminate collusion attacks.

V. SYSTEM ANALYSIS

Algorithm: Elgamal Encryption Scheme Key Generation phase

Input: Random input data textMetadata

Output: returns the keys {a,b,p,g}

Step 1 : Initialized the random text input using textMetadata

Step 2: ResultData[] = GetRandomP (textMetadata.getbyte).bit length according to the probable prime number. **Step 3:** p= ResultData [0] g= ResultData [1] **Step 4:** produce a using P a=GenerateA(p)

Its generates like p.bitLength()-1,Random. **Step 5:** Calculate b= calculateb(g, a, p); so, b= g.modPow(a, p);

Step 6: Key generation done.

Encryption

Input: Text data d,p,b,g Output cipher as C1,and C2.

Step 1 : initialize BigInteger [] rtn = {null, null};

Step 2 : message=d.getbytes();

Step 3 : [] result= ElGamal.encrypt(message, p, b, g);

Step 4 : k = ElGamal.getRandomk(p);

Step 5 : C1 = g.modPow(k, p);

Step 6 : C2 = m.multiply(b.modPow(k, p)).mod(p);

Decryption Input : input c1 and c2 as cipher a and p as private keys

Output: Plain text d.

Step 1: m = C2.multiply (C1.modPow (a.negate(), p)).mod (p); **Step 2:** return m.

SHA 256 Algorithms

Input: string required to ascertain the SHA score.

Output: SHA score of string

Step 1: Padded with the length in such way that the result is various in least 512 piece long.

Step Step 1: Initialize the C

Step 2: Shascore= SHA256(C)

Step 3: Return Shascore

Step 4: give back the H(i) SHA score of given string.

Machine learning dynamic attack query pattern Weight Calculation Algorithm also applicable for SQL injection

Input: Query generated from user Q, each retrieved list L from webpage.

Output: Each list with weight.

Here system has to find similarity of two lists: $\vec{a} = (a_1, a_2, a_3, \ldots)$ and $\vec{b} = (b_1, b_2, b_3, \ldots)$, where $a_n$and b_ne the components of the vector (features of the document, or values for each word of the comment) and the n is the dimension of the vectors:

Step 1 : Extract all the features from Test set using below

ReceiveCommand = (T[j])

Step 2: Read all features from Trainset using below

PolicyList = (T[k])

Step 3: Read all features from Trainset using below

Step 4 : Generate weight of both feature set

= (ReceiveCommand , PolicyList)

Step 5 : Verify Threshold

SelectedInstance= result = W > T ? 1 : 0; Add each selected instance into L, when n = null

Step 6 : Return L.

Mathematical Model

First we consider a

A={A1,A2,A3……..An} each set holds the specific module activity of system. A1={file uploading phase or file sending phase}

A2={data encryption and re-encryption phase}

A3={Share and Access control for delegates}

A4={Revocation and proxy key re-generation} A1 define the first module which is user the upload the multiple documents

Data[ ] = [ ] + (a1, a2 … … . . an)

d[k] ⍰ {Att1,Att2……Attn} each documents contains the set of attributes

keys[] ⍰ Keygen(RandomText)

Enc[c1] [c2] ⍰ encryption(Data, keys[])

DecData ⍰ decryption ([c1] [c2], keys[]) Role base access control for each ith user has been defined using below formula

U[i]⍰file(x) = u[ ][read,write, update, delete]

User revocation has done using below formula

U[i] ⍰ Revoke(F) : Data_Owner

Software Requirements

1. **System interfaces:** Windows Operating System
2. **User interfaces:** User interface using Jsp and Servlet
3. **Hardware interfaces**

Processor:- Intel R-Core i3 2.7 or above

Memory:- 4GB or above

Hard Disk :- 500 GB

4. **Software interfaces**:

Front End: Jdk 1.7.0, Eclipse

IE 7.0/above

Back-End: Mysql 5.1.

5. **Communications interfaces**

System will use HTTP as well as SMTP and SOAP protocol for establishing connection and transmitting data over the network.

6. **Services:** Amazon EC2 as Public cloud Environment

Results and Discussion

For the system performance evaluation, calculate the matrices for accuracy. The system is executed on java 3-tier architecture framework with INTEL 2.8 GHz i3 processor and 4 GB RAM with public cloud Amazon EC2 consol. For the system evaluation we create 2 machines on physical environment with Wi-Fi and 10 VM with Amazon EC2 as public cloud environment. After implementing some part of system we got system performance on reasonable level. In the first experimental we have calculated the accuracy of attack detection module using various number instances.

Table 1: System performance

Test Instances	Accura cy	Precisio n	Recall	F-Measu re
10	0.90	0.91	0.94	0.95
20	0.91	0.92	0.95	0.96
50	0.89	0.90	0.93	0.94
100	0.90	0.93	0.92	0.95

In second experimentation system show the user verification time with different approaches. In current system we consider as four different uthorities for runtime verification. The below Fig. 2 shows the performance measures using some existing

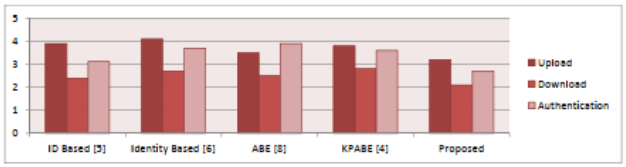

Fig. 2 : System Performance Measures proposed vs Existing approaches

Graph Comparison

In second experiment Figure 3 shows data encryption performance which works to display that the data it will encrypt in exactly how much period in seconds. Suppose there is a 100kb data is encrypted in 150 second so the outcome will show inevitably in that time of encryption data from the operators.

Figure 3 : Data encryption performance based on data size

Conclusion

Proposed system provides the highest security from different type of attack in cloud environment to end users confidentiality data. In other hand AES 128 encryption algorithm also maintain the robust security mechanism. Access control and revocation maintain the security and efficiency of system. The system achieves Role Base Access control in single as well as multi cloud environment with this approach. The current architecture is very efficient for security purpose, but sometime it's utilized multiple resources. When such system allocates multiple resources it will generate a lot of dependencies. For the next update we can focus on minimum resource utilization with system flexibility like power, VM's, network, memory etc. The proposed work also describes the efficiency of system in cloud environment, that able to detect the runtime investigation as well as malicious attacks. The secure revocation in RBAC module, provides the defence from collusion attacks as well as enhance the efficiency of system. Finally, system able to work smoothly in trusted or untrusted cloud environment due to drastic supervision of detection algorithms.

References

[1] Mr. DigambarPowar, Dr. G. Geethakumari "Digital Evidence Detection in Virtual Environment for Cloud Computing" ACM, 2012.

[2] Saibharath S, Geethakumari G "Cloud Forensics: Evidence Collection and Preliminary Analysis" IEEE, 2015

[3] Mr. Chandrashekhar S. Pawar, Mr. Pankaj R. Patil, Mr. Sujitkumar V. Chaudhari "Providing Security and Integrity for Data Stored In Cloud Storage" ICICES, 2014.

[4] Deevi Radha Rani, G. Geethakumari "An Efficient Approach to Forensic Investigation in Cloud using VM Snapshots" International Conference on Pervasive Computing (ICPC), 2015.

[5] BKSP Kumar RajuAlluri, Geethakumari G"A Digital Forensic Model for Introspection of Virtual Machines in Cloud Computing" IEEE, 2015.

[6] Hubert Ritzdorf, NikolaosKarapanos, SrdjanCapkun "Assisted Deletion of Related Content"ACM, 2014.

[7] Liang X, Zhao J, Shetty S, Liu J, Li D. Integrating blockchain for data sharing and collaboration in mobile healthcare applications. InPersonal, Indoor, and Mobile Radio Communications (PIMRC), 2017 IEEE 28th Annual International Symposium on 2017 Oct 8 (pp. 1-5). IEEE.

[8] Manoj R, Alsadoon A, Prasad PC, Costadopoulos N, Ali S. Hybrid secure and scalable electronic health record sharing in hybrid cloud. In2017 5th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud) 2017 Apr 6 (pp. 185190). IEEE.

[9] Khan SI, Hoque AS. Privacy and security problems of national health data warehouse: a convenient solution for developing countries. InNetworking Systems and Security (NSysS), 2016 International Conference on 2016 Jan 7 (pp. 1-6 IEEE.

[10] Shrestha NM, Alsadoon A, Prasad PW, Hourany L, Elchouemi A. Enhanced e-health framework for security and privacy in healthcare system. InDigital Information Processing and Communications (ICDIPC), 2016 Sixth International Conference on 2016 Apr 21 (pp. 7579). IEEE.