A Secure Approach for Dynamic Investigation using VM Snapshot Logs in Cloud Computing

Sayeda A Rahila[1]
P.G. Student, Department of Computer Engineering
Modern Education Society of Engg, Pune, MH, India[1]

Prof. Shraddha Khonde[2]
Professor, Department of Computer Engineering,
Modern Education Society of Engg, Pune, MH, India[2]

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 pay-for-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 identify 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.

I. 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.

II. LITARATURE SURVEY

Cloud computing systems illustrates [1] 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.

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 describe evaluation of system and showed that its overhead and accuracy is acceptable for practical use and deployment.

According to [4] 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.

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 can not be directly given as a cloud forlantic answer as a result of it's 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.

Fuzzy IBE primitive [6] however using a double tree information structure to record clients characters at leaf hubs. In this way, key-redesign productivity at PKG can be essentially diminished from direct to the tallness of such double tree (i.e. logarithmic in the quantity of clients). By and by, we call attention to that however the double tree acquaintance is capable with accomplish a relative elite, it will bring about different issues: 1) PKG needs to create a key pair for every one of the hubs on the way from the personality leaf hub to the root hub, which brings about unpredictability logarithmic in the quantity of clients in framework for issuing a solitary private key. 2) The measure of private key develops in logarithmic in the quantity of clients in framework, which makes it troublesome in private key stockpiling for clients

In [7], the creators propose a two-contention capacity is figured secretly by two gatherings such that after the calculation, no gathering ought to know anything about the different inputs aside from what he can figure from his own info and the capacity esteem. Some broad relations between the data increase of an ideal convention and the correspondence intricacy of a capacity is likewise said. In this paper, measures for uncovered data required for figuring f have been considered. Chiefly examined the measures given by BarYehuda furthermore, have likewise demonstrated that a few results exhibited by them are incorrect on two-party calculation. They have introduced another definition for the extra data for two gathering conventions and have given a few limits for solid capacities for the extra data. Here they have made utilization of estimations which are troublesome.

In [8], they are examining about Secure Multi-Party Calculation (SMC). Secure Multi-Party Computation (SMC), is a procedure that permits parties with comparable foundation to process results upon their private information, minimizing the risk of divulgence. The exponential increment in delicate information which should be gone upon organized PCs and the development of web has created incomprehensible open doors for agreeable calculation, where parties meet up to encourage calculations and make out determinations that are commonly gainful and at the same time keeps their private information secure. This system is for the most part an augmentation to a formerly proposed convention Encrytpo_Random, which exhibited a basic yet successful way to deal with SMC furthermore set forward an suitably created structural engineering, whereby such an effective convention, including the gatherings that have approached for joint calculations and the outsider who attempts such calculations, can be produced. Through this broadened work an endeavor has been made to advance fortify the current convention and makes utilization of a few layers in structural engineering. These layers are making the entirety framework incredibly confounding. For appointing identifiers to the hubs of a system, proficient calculations are managed such that the identifiers are unknown by making utilization of a circulated calculation without focal power.

In [9], the protected aggregate permits gatherings to work out the aggregate of their individual inputs without uncovering the inputs to each another and it makes a difference to separate the confusions of the safe multiparty calculation. A calculation was introduced for sharing straightforward number data on top of secure entirety and it is utilized by the calculation at all emphases for unknown ID task. However, the safe entirety does not permit to share complex messages.

A protected calculation work generally utilized as a part of the writing of [10], which is secure whole that permits gatherings to register the entirety of their individual inputs without uncovering the inputs to each other. This capacity is well known in information mining applications furthermore clarifies the complexities of the safe multiparty calculation. To separate unknown ID task from mysterious correspondence, consider a circumstance where parties wish to show their information on the whole, yet namelessly, in spaces on an outsider site. The IDs can be utilized to allot the openings to clients, while mysterious correspondence can permit the gatherings to disguise their personalities from the outsider. While looking it all the more carefully its unmistakable can that the information being partaken in remote systems is not exactly simple.

III. 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.

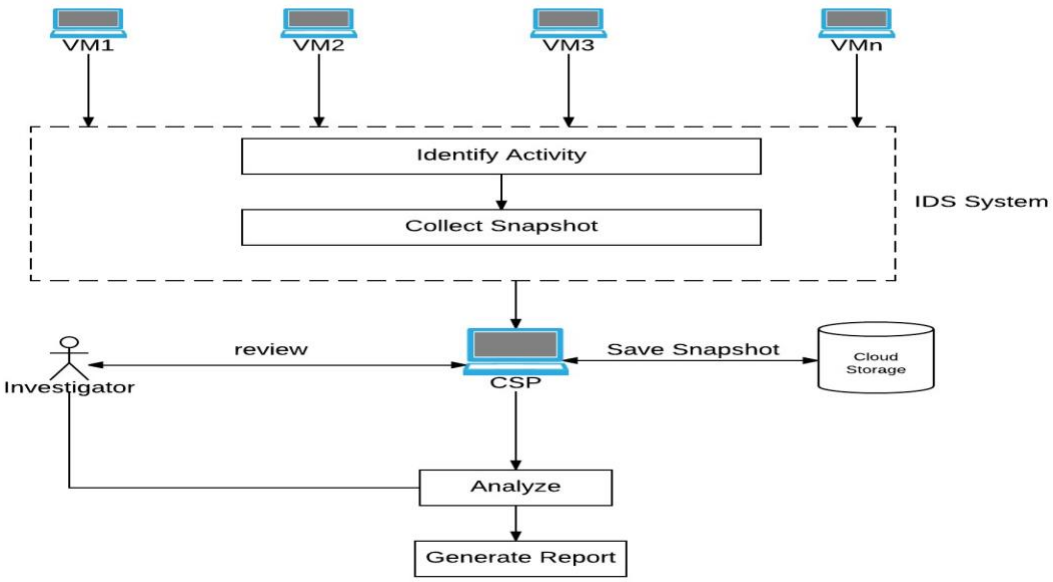

Figure 1 : Proposed System Architecture

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.

IV. 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 and a_n b_n are 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

$$\text{ReceiveCommand} = \sum_{j=1}^{n} (\text{T[j]})$$

Step 2: Read all features from Trainset using below

$$\text{PolicyList} = \sum_{k=1}^{\text{m}} (\text{T[k]})$$

Step 3: Read all features from Trainset using below

Step 4 : Generate weight of both feature set

$$W = (\text{ReceiveCommand} , \text{PolicyList})$$

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

$$\text{Data}[d] = d[k] + \sum_{k=0}^{n} (\text{a1}, \text{a2} \ldots \ldots \ldots \text{an})$$

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

$$\text{U[i]} \leftarrow \text{file(x)} = \sum_{n=1}^{m} (\text{u}_{[n]}[\text{read}, \text{write}, \text{update}, \text{delete}])$$

User revocation has done using below formula

$$\text{U[i]} \leftarrow \text{Revoke(F)} : \text{Data\_Owner}$$

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

VI. 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 WiFi

and 10 VM with Amazon EC2 as public cloud environment. After implementing some part of system we got system performance on reasonable level.

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

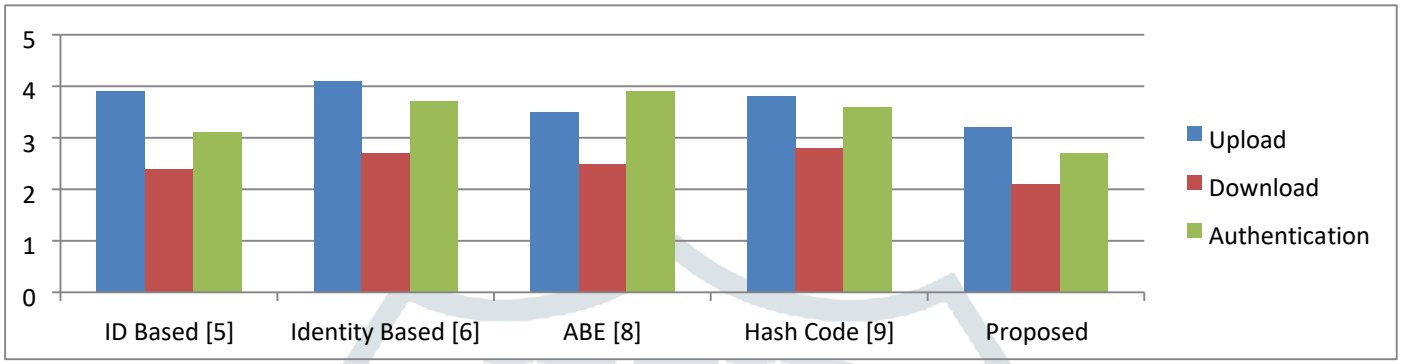

Fig. 2 : System Performance Measures proposed vs Existing approaches

V. CONCLUSION

The secure revocation in RBAC module, provides the defence from collusion attacks as well as enhance the efficiency of system. Proposed system provides the highest security from different type of attack in cloud environment to end users confidentiality data. In other hand Elgamal 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.

REFERENCES

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

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

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

[4] Rahul Reddy Nadikattu. 2016 THE EMERGING ROLE OF ARTIFICIAL INTELLIGENCE IN MODERN SOCIETY. International Journal of Creative Research Thoughts. 4, 4 ,906-911.

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

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

[7] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks, 2008.

[8] RR Nadikattu, 2016 THE EMERGING ROLE OF ARTIFICIAL INTELLIGENCE IN MODERN SOCIETY. International Journal of Creative Research Thoughts. 4, 4 ,906-911.

[9] Andreas Jakoby and Maciej Li´skiewicz (2005), "Revealing Additional Information in Two-Party Computations " , Advances in Cryptology - ASIACRYPT 2005 Lecture Notes in Computer Science Volume 3788, 121-135.

[10] Sikender Mohsienuddin Mohammad, **"IMPROVE SOFTWARE QUALITY THROUGH PRACTICING DEVOPS AUTOMATION"**, International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.6, Issue 1, pp.251-256, March 2018, Available at :http://www.ijcrt.org/papers/IJCRT1133482.pdf

[11] Dr. Durgesh Kumar, Neha Koria, Nikhil Kapoor, Ravish Bahety (2009), "A Secure Multi-Party Computation Protocol for Malicious Computation Prevention for Preserving Privacy During Data Mining", International Journal of Computer Science and Information Security,Vol. 3.

[12] Akheel Mohammed, Sajjad Ahmed Md , Ayesha (2013),"Confidentiality And Anonymity Strengthening in Computational Services", IJRRECS,Volume-1,Issue-6,1006-1011.

[13] Sikender Mohsienuddin Mohammad, **"CONTINUOUS INTEGRATION AND AUTOMATION"**, International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.4, Issue 3, pp.938-945, July 2016, Available at :http://www.ijcrt.org/papers/IJCRT1133440.pdf

[14] Swathi, P.Jyothi, and Anil Kumar(2014), "Assigning Privacy Ids For Each Data That Have Been Sharing In Wireless Networks", International Journal of Communication Network and Security (IJCNS) ISSN: Volume-2, Issue-3.